Contrasting drought legacy effects on gross primary productivity in a mixed versus pure beech forest

Xin Yu[1], René Orth[1], Markus Reichstein[1], Michael Bahn[2], Anne Klosterhalfen[3], Alexander Knohl[3], Franziska Koebsch[3], Mirco Migliavacca[1, 4], Martina Mund[5], Jacob A. Nelson[1], Benjamin D. Stocker[6, 7], Sophia Walther[1], and Ana Bastos[1]

[1]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena, D-07745, Germany
[2]Department of Ecology, University of Innsbruck, Innsbruck, A-6020, Austria
[3]Bioclimatology, University of Göttingen, Göttingen, D-37077, Germany
[4]Joint Research Centre, European commission, Ispra (VA), 21027, Italy
[5]Forestry Research and Competence Centre Gotha, Gotha, D-99867, Germany
[6]Department of Environmental Systems Science, ETH, Zürich, 8092, Switzerland
[7]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, 8903, Switzerland

Correspondence to: Xin Yu (xyu@bgc-jena.mpg.de)

Abstract. Droughts affect terrestrial ecosystems directly and concurrently, and can additionally induce lagged effects in subsequent seasons and years. Such legacy effects of drought on vegetation growth and state have been widely studied in tree-ring records and satellite-based vegetation greenness, while legacies on ecosystem carbon fluxes are still poorly quantified and understood. Here, we focus on two ecosystem monitoring sites in central Germany with similar climate but characterized by different species and age structures. Using eddy-covariance measurements, we detect legacies on gross primary productivity (GPP) by calculating the difference between random-forest model estimates of potential GPP and observed GPP. Our results showed that at both sites, droughts caused significant legacy effects on GPP at seasonal and annual time scales which were partly explained by reduced leaf development. The GPP reduction due to drought legacy effects is of comparable magnitude to the concurrent drought effects, but differed between two neighbouring forests with divergent species and age structures. The methodology proposed here allows quantifying the temporal dynamics of legacy effects at the sub-seasonal scale and separating legacy effects from model uncertainties. Application of the methodology at a larger range of sites will help quantify whether the identified lag effects are general and on which factors they may depend.

1 Introduction

The frequency, intensity, duration, and spatial extent of drought are expected to increase in the next decades due to anthropogenic global warming in many regions (IPCC, 2022). A great number of studies, considering both long-term observations (Schwalm et al., 2010; Zscheischler et al., 2014) and model simulations (Reichstein et al., 2007; Sun et al., 2015) across various spatial scales, have shown that droughts concurrently impact the structure and function of terrestrial ecosystems (Assal et al., 2016; Frank et al., 2015; Lewis et al., 2011; Ma et al., 2015; Orth et al., 2020), potentially turning ecosystems from sinks to temporary sources of carbon (Ciais et al., 2005; Reichstein et al., 2013). Therefore, understanding the impact of droughts on terrestrial ecosystems is a key research question in Earth sciences (Piao et al., 2019).

Drought impacts on terrestrial ecosystems are not limited to concurrent effects, but also include legacy effects during the following seasons and years (Anderegg et al., 2015; Frank et al., 2015; Kannenberg et al., 2020). Legacy effects at tree and/or stand scale can be caused by the higher vulnerability to drought due to previous water depletion of the soil (Krishnan et al., 2006, Galvagno et al., 2013), reduced or delayed leaf development (Migliavacca et al., 2009; Rocha and Goulden, 2010; Kannenberg et al., 2019), drought-induced hydraulic damage of the xylem (Anderegg et al., 2013), adjustments in carbon allocation within the trees (Huang et al., 2021), depletion of non-structural carbohydrates (Peltier et al., 2021) due to reduced carbon availability and adjustments in carbon allocation (Hartman and Trumbore, 2016), tree mortality (Allen et al., 2015), as well as reduced resistance to disturbances (e.g. insects outbreaks) due to depleted non-structural carbohydrates (Erbilgin et al., 2021). However, at the ecosystem level the impact of species and age structures on legacy effects are still less understood (Haberstroh and Werner, 2022, Wang et al., 2022).

Tree-ring records cover periods of decades to centuries and can cover multiple drought events, being therefore widely used to analyze inter-annual legacy effects of drought on tree growth (Anderegg et al., 2015; Huang et al., 2018; Kannenberg et al., 2019). Beyond the level of individual trees, satellite-based observations and model outputs, as expressed through vegetation greenness (Wolf et al., 2016; Wu et al., 2018), canopy backscatter (Saatchi et al., 2013), aboveground carbon stocks (Wigneron et al., 2020), and gross primary productivity (Schwalm et al., 2017, Bastos et al., 2020) have also been used to study seasonal and inter-annual legacy effects of drought. However, studies focusing on carbon fluxes, especially based on eddy-covariance measurements, are still rare (Kannenberg et al., 2020). Eddy-covariance data with hydrometeorological variables measured in parallel have the potential to quantify the timing and magnitude of legacy effects at the sub-seasonal and annual scales, and might provide insights into the mechanisms of legacy effects that might not be fully reflected in vegetation indices and tree rings.

Assessments of drought impacts on the ecosystem carbon fluxes usually focus on direct and concurrent effects (Ciais et al., 2005; Reichstein et al., 2007) without considering legacy effects. This is probably due to the challenge to attribute signals in the observations to a previous drought and hence identify them as legacy effects on ecosystem carbon fluxes (Kannenberg et al., 2020), and the inability of models to reproduce these legacy effects (Bastos et al., 2021). A number of studies consider ecosystems to have ‘recovered’ when the target variable such as gross primary productivity (GPP) and tree-ring width returns to the baseline, which is usually based on pre-drought values of the target variable (Bose et al., 2020; González de Andrés et al., 2021; Zhang et al., 2021). However, this might complicate the detection of legacies since GPP recovery dynamics is affected by hydrometeorological conditions in legacy years, which can either stimulate or slow-down recovery. Here, by estimating potential GPP given hydrometeorological conditions in legacy years, we consider that ‘recovery’ happens when the actual GPP reaches the potential GPP under the given hydrometeorological conditions, rather than the absolute flux.

Therefore, we aimed to develop a novel approach to quantify drought legacy effects on GPP at the sub-seasonal and annual scales. To do this, we followed a residual approach (Beringer et al., 2007) to identify legacy effects as the residuals between actual and potential GPP which is estimated by a machine-learning algorithm (specifically Random Forest regression). Furthermore, it is crucial to understand if the residuals are caused by model uncertainties or can be interpreted as legacy effects. By overlooking model uncertainties, one could misinterpret small residuals as 'legacy effects'. Here we quantified model uncertainties to provide more robust estimates of drought legacies and avoid misinterpretation of results. To test our approach, we used eddy-covariance measurements at two neighbouring sites that experienced similar climate but are characterized by different species and age structures in central Germany. We asked 1) can we detect drought legacy effects on GPP? 2) is the GPP reduction due to drought legacy effects significant compared to the magnitude of drought concurrent effects? 3) how do drought legacy effects on GPP differ at two neighbouring forests with different species and age structures?

2 Data

2.1 Study sites

The two neighboring temperate forest sites studied here, Hainich (DE-Hai, 51°04′46″N, 10°27′07″E) and Leinefelde (DE-Lnf, 51°19′42″N, 10°22′04″E), are located in central Germany, approximately 30 km from each other. These two sites share similar climate conditions, with long-term annual mean of 8 °C for 2-m air temperature and 750 mm of total annual precipitation (Tamrakar et al., 2018). Both sites were affected by the two extreme central European droughts in 2003 and 2018 which reduced gross primary productivity (Fu et al., 2020; Herbst et al., 2015).

The forest at Hainich is an old-growth, uneven aged (1-250 years) mixed forest, dominated by beech (*Fagus sylvatica*, representing approximately 64% of the tree carbon stocks). Ash (*Fraxinus excelsior*, 28%) and sycamore (*Acer pseudoplatanus*, 7%) are co-dominant tree species, and additionally there are few trees of European hornbean (*Carpinus betulus*), Norway maple (*Acer platanoides*), and other deciduous species (Knohl et al., 2003). The forest at Leinefelde can be characterized as a managed even-aged (ca. 130 years) pure beech forest (Anthoni et al., 2004).

2.2 Eddy-covariance and meteorological measurements

Identical eddy-covariance instrumental setups and data acquisition techniques were carried out at the two sites. The methodology of data collection and quality control followed those of Aubinet et al. (2000). The standard processing methods (Pastorello et al., 2020) adopted by the Integrated Carbon Observation System (ICOS) were used to carry out the gap-filling and the partitioning (Warm Winter 2020 Team and ICOS Ecosystem Thematic Centre 2022). The GPP estimated from the nighttime partitioning algorithm (Reichstein et al., 2005) was used for the analysis (GPP_NT_VUT_REF). A detailed description of meteorological data and instrumentation can be found in previous studies (Anthoni et al., 2004; Knohl et al., 2003). We used daily meteorological data alongside carbon and water fluxes, namely GPP, latent heat flux after the energy balance correction (LE_CORR), which was converted to evapotranspiration (ET) using the heat of vaporization, incoming

shortwave radiation (SW_IN), air temperature (TA), vapor pressure deficit (VPD), soil water content at the first layer (SWC_1, 8cm), the second layer (SWC_2, 16cm), the third layer (SWC_3, 32cm), and potential incoming shortwave radiation (SW_IN_POT) for the years 2000-2020 at DE-Hai and 2002-2012, with a gap in 2007-2009, at DE-Lnf.
Additionally, we used daily enhanced vegetation index (EVI) data from the FluxnetEO v1.0 dataset (Walther et al., 2021) for the same years as the eddy-covariance data. EVI was derived from the MCD43A4 product of MODIS with a 500m spatial resolution and we used an average over 2x2 pixels surrounding the tower. We further estimated daily transpiration based on the Transpiration Estimation Algorithm (Nelson et al., 2018).

2.3 Radial increment and net primary productivity of fruits and leaves

Annual radial increment (RI) was calculated from permanent band dendrometers which measures change in stem girth (or circumference) over bark. The effect due to the inclusion of shrinkage and swelling of the bark is a negligible uncerainty for four reasons: 1) we used only the annual increment, 2) the dominant species is beech that has only a thin bark, 3) we recorded the final stem diameter of each year in winter, when the water status of the xylem and the bark is relatively constant, and when stem wood or the bark are not affected by frost or late/early growth or water uptake, and 4) in this study we were interested only in the interannual variability of stem growth, which is less affected by shrinkage and swelling at the described temporal scale than absolute growth rates. The dendrometer trees represented the main species and their respective size classes of the main footprint at DE-Hai for the years 2003 to 2020. Because of technical constraints, damages and a natural dieback of single trees, the number of measurement trees per year varied between 54 and 95. Net primary productivity (NPP) of fruits for the years 2003 to 2020, and NPP of leaves for the years 2003 to 2016 resulted from litter samplings (25-29 traps) within the main footprint area of the flux tower. The high fluctuation of annual fruit NPP is caused by the periodically high fruit production (masting) of beech (*Fagus sylvatica*). In mast years the proportion of beech fruits (nuts and shells) amounted to almost 92% of total fruit mass. At DE-Lnf these data are not available. A detailed description of measurement and processing methods can be found in a previous study (Mund et al., 2020).

3 Methodology

3.1 Data processing

As the first step, we filtered and processed the eddy covariance and meteorological data in the following way:
1) To ensure reliable data for our analysis we used gap-filled daily data for days for which more than 70% of measured and good quality gap-fill data (Reichstein et al., 2005) were available.
2) We only used data during the growing season which was defined as the period when GPP was greater than 10% of maximum of GPP as inferred from a smoothed (centered 7-days moving averages) daily average GPP across all years.

3) We calculated anomalies of all variables by subtracting the mean seasonal cycle and any significant long-term linear trend, detected by the Mann-Kendall test (Kendall, 1948), as these can obscure drought-related signals. We took the mean of each day across all considered years and then used centered 7-days moving averages to calculate the mean seasonal cycle.

4) Furthermore, a 7-days moving average smoothing was applied to the anomaly time series to filter out noise at daily time scales. We expect this to increase the accuracy of our model while preserving drought legacy patterns which rather/better emerge at longer time scales.

As for RI data, we removed for each individual tree any significant long-term linear trend detected using the Mann-Kendall test (Kendall, 1948).

3.2 Water availability index estimation

Soil moisture at the two study sites was measured only at the upper 30 cm and thus does not account for water availability in deeper layers (see Section 5.4). Therefore, we used a bucket model approach based on observed evapotranspiration and precipitation to estimate a vegetation water availability index, WAI (Tramontana et al., 2016), calculated as:

$$WAI_0 = WAI_{wam\text{-}up} \quad (1)$$

$$WAI_t = \min(WAI_{max}, WAI_{t-1} + P_t - ET_t) \quad (2)$$

Where WAI_0 was the initial value of the water availability index (WAI), $WAI_{warm\text{-}up}$ was the end value of WAI from the warm-up of the bucket model (Eq. 1). To warm up the bucket model, we ran it 5 times through the first year before starting the actual computation across all considered years.WAI_{t-1} (mm) and WAI_t (mm) were WAI at time step t-1 and t, respectively, P_t (mm), and ET_t (mm) were, precipitation, and evapotranspiration at time step *t* (Eq. 2). We set the bucket size (i.e. WAI_{max}) as the maximum cumulative water deficit (CWD) at each site. The estimated bucket sizes were 205mm and 191mm at DE-Hai and DE-Lnf, respectively.

Additionally, we calculated the CWD, which was estimated from cumulative differences between observed evapotranspiration and precipitation over periods where cumulative net water loss from the soil (Σ (ET-P)) is positive.

3.3 Drought and legacy years selection

Since legacy effects should result from significant impacts of droughts on ecosystems, we adopted a combined driver and impact-based approach to define droughts. Drought years were defined as those years when both low water availability and a concurrent biospheric response were found, and were evaluated as follows:

1) First, we selected the minimum of negative GPP anomalies relative to the mean seasonal cycle during the growing season (minimum GPP_{anom}) as a proxy to reflect the severity of drought impact on GPP in each year.

2) Then, we calculated the mean WAI anomalies relative to the mean seasonal cycle for days when minimum GPP_{anom} occurred and the previous 14 days (mean WAI_{anom_15}) to reflect the water availability during the development of the GPP anomaly. To

identify drought-related GPP reductions, we considered only years where negative GPP anomalies were associated with dry conditions.

3) Finally, we selected the years with both the lowest minimum GPP_{anom} and mean WAI_{anom_15} (Fig. S1). These were 2003 and 2018 at DE-Hai and 2003 at DE-Lnf (2018 data not available here).

In our data, we define non-legacy years as normal and drought years, while legacy years correspond to the two calendar years following a drought year. Including too few legacy years could lead to an underestimation of legacy effects, and too many legacy years would result in the lack of training data (see Section 3.4). As a trade-off, we selected a legacy period of two years and this choice was justified by the fact that GPP anomalies residuals returned to the range of model uncertainties (i.e. 25th-75th percentiles of model residuals), which is considered as the point when GPP recovers. This happened in 2005 (see Section 4.3) following the 2003 drought at both sites. For the 2018 drought at DE-Hai, data was only available up to 2020.

3.4 Quantification of legacy effects on GPP and transpiration

Here, we followed a residual approach (Beringer et al., 2007) to detect drought legacy effects on GPP. To do this, we fitted a random forest regression model (RF, Breiman 2001) for daily GPP anomalies using the anomalies of hydro-meteorological variables in non-legacy years as predictors. We chose RF because it has the ability to effectively learn 1) the relationship between independent and dependent variables regardless of linear or non-linear relationships; 2) the interactions between independent variables (Ryo and Rillig, 2017). The model was then used to predict GPP anomalies in the legacy years, thereby reflecting the potential GPP anomalies given the climate conditions in that year. Specifically, the approach included the following steps (Fig. 1):

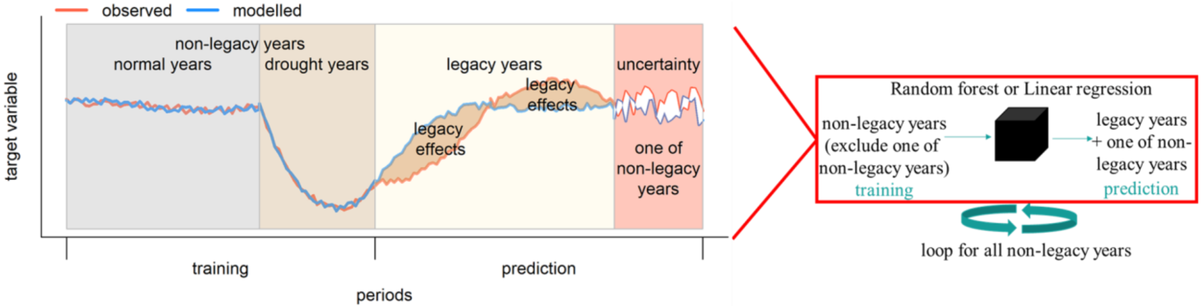

Figure 1. Conceptual diagram of quantification of legacy effects. A random forest (RF) model (or linear regression, represented by the black cube on the right) was used to determine the relationship between the target variable (GPP_{anom} or RI) and hydro-meteorological conditions using a training dataset excluding data in legacy years and one of non-legacy years for each loop. The legacy effects could be quantified as the residuals between observed (red line) and modelled (blue line) target variable (i.e. GPP_{anom}, RI, …) in legacy years. And the residuals between observed and modelled target variable (i.e. GPP_{anom}, RI, …) in all non-legacy years from all loops indicated RF model uncertainties using a leave-one-out approach (see below).

First, all daily data in non-legacy years were used as input for the RF model to determine the relationships between anomalies of GPP (GPP_{anom}) and anomalies of hydro-meteorological variables (SW_IN_{anom}, TA_{anom}, VPD_{anom}, and WAI_{anom}) along with absolute values of SW_IN_POT to capture seasonal variations in the response of ecosystems to hydro-meteorological conditions. These relationships represented long-term controls of climate on GPP, including drought events and near-average or wet conditions. The Out of bag (OOB) scores indicating the prediction ability of RF models were ~0.7 and ~0.8 (where zero indicates no skill and 1 denotes perfect performance) at DE-Hai and DE-Lnf, respectively (Fig. S2). WAIanom is the most important explanatory factor at both sites, followed by SW_INanom at DE-Hai and the phenological stage (given by SW_IN_POT) at DE-Lnf (Fig. S3). The 'randomForest' package in R 4.0.3 was used, and the number of trees, the number of variables randomly sampled as candidates at each split, and the node size of RF were set to 400, 5, and 5, respectively. Tuning those hyperparameters did not significantly change our results.

Based on these relationships and the meteorological anomalies in legacy years, we used the trained RF model to predict the potential GPP_{anom} in the absence of legacy effects and calculated the model's residuals (GPP_{anom} residuals, i.e. observed minus predicted values), which should reflect legacies from the past drought: negative residuals corresponded to more negative or less positive GPP_{anom} than would be expected given the meteorological conditions in that year, indicating negative legacies of drought, while positive residuals corresponded to less negative or more positive GPP_{anom}, indicating beneficial legacies of drought. In order to reduce the noise at the daily scale, daily results were aggregated to the weekly scale.

To account for model uncertainties and evaluate the significance of legacy effects, we used a leave-one-out approach to quantify model uncertainties. In the training phase, one of the non-legacy years was excluded from the training dataset and the trained RF model was then used to predict the GPP_{anom} in that year. This was done for all non-legacy years, and the GPP_{anom} residuals in non-legacy years for each leave-one-out iteration were then considered as model uncertainties. In order to reduce the noise at the daily scale, all the daily results were aggregated to the weekly scale.

In order to infer possible legacy effects due to plant hydraulic damage, the same method was used to quantify legacy effects on transpiration (Tr) estimated by TEA (Transpiration estimation algorithm) approach (Nelson et al., 2018). The TEA approach first isolates the periods when evapotranspiration is most likely dominated by transpiration. Then, a quantile random forest model (Breiman, 2001; Meinshausen and Ridgeway 2006) is trained during the separated periods and transpiration can be estimated at every time step. More detail can be found in Nelson et al., 2018. We did not use evapotranspiration (ET) because given a certain amount of energy even though Tr decreases due to plant hydraulic damage but it could be compensated by increased soil evaporation, and the amount of ET might remainl unchanged.

3.5 Quantification of legacy effects on tree growth

To detect legacy effects on tree growth, we used a multivariate-linear regression instead of RF to develop the relationship between tree growth (detrended radial increment, RI) due to the fewer data points available. We used the following explanatory variables: detrended growing-season mean WAI, detrended growing-season mean VPD, detrended growing-season mean SW_IN, and detrended growing-season mean TA for each species. We detrended the time series of all variables by removing any significant long-term linear trend detected using the Mann-Kendall test (Kendall, 1948). Annual net primary productivity of fruits (fruits-NPP) particularly was added as an additional predictor to only the model for beech since the high fluctuation of annual fruit NPP could be caused by the periodically high fruit production (masting) of beech. We considered fruits-NPP

as a predictor to account for the trade-off between tree growth and reproduction in mast years, which could also cause the change in tree growth in addition to legacy effects from previous droughts (Hacket-Pain et al., 2015).
The strategy to quantify legacy effects and model uncertainties was the same as in the case of GPP. We trained the model in non-legacy years except for each one of them iteratively and predicted potential RI in legacy years and the year additionally excluded. The residuals between observed and potential RI in non-legacy years and legacy years were then considered as model uncertainties and legacy effects, respectively.

3.6 Separation of legacy effects on GPP due to structural and physiological effects

Drought legacy effects on GPP might result from changes in canopy structure (structural effects) and photosynthesis capacity (physiological effects) (Kannenberg et al., 2019). Combining GPP and satellite-based EVI allows separating these structural and physiological effects. To do this separation, we used two model settings: 1) RF, which was the original setting described in section 3.4, included both structural and physiological effects; 2) RF_{EVI}, which added EVI anomalies as an additional predictor to the original model, only included physiological effects because structural effects were already included in the predictor EVI anomalies and GPP_{anom} residuals from this model were expected to be caused by physiological effects. Therefore, physiological legacy effects on GPP were quantified as GPP_{anom} residuals from RF_{EVI} while structural legacies were quantified as the difference between GPP_{anom} residuals from RF and RF_{EVI} (i.e. RF-RF_{EVI}). The same method was used to separate structural and physiological effects of legacy effects on Tr.

3.7 Quantifying concurrent and lagged reduction in GPP from drought

Additionally, we compared the estimated legacy effects on GPP to the concurrent drought-induced GPP anomalies. To compute the concurrent reduction in GPP, we summed up all GPP anomalies over each identified drought period. Here, drought periods were defined as the periods where WAI_{anom} was lower than -1 of standard deviation (WAI_{SD}). WAI_{SD} was calculated for each day of year by using a centered 7-day moving window instead of a single value over the whole time series because WAI_{SD} showed a seasonality. This definition only relied on the water availability without considering biospheric responses because WAI directly indicated the water supply for vegetation while GPP could include other factors in addition to drought in short periods. We quantified the lagged reduction in GPP at the annual scale as the difference between GPP_{anom} residuals in legacy years and the median of the model uncertainties. To compare the reduction in GPP across sites, both concurrent and lagged values were normalized relative to averaged total GPP over the growing season.

4. Results

4.1 GPP time series in drought and legacy years

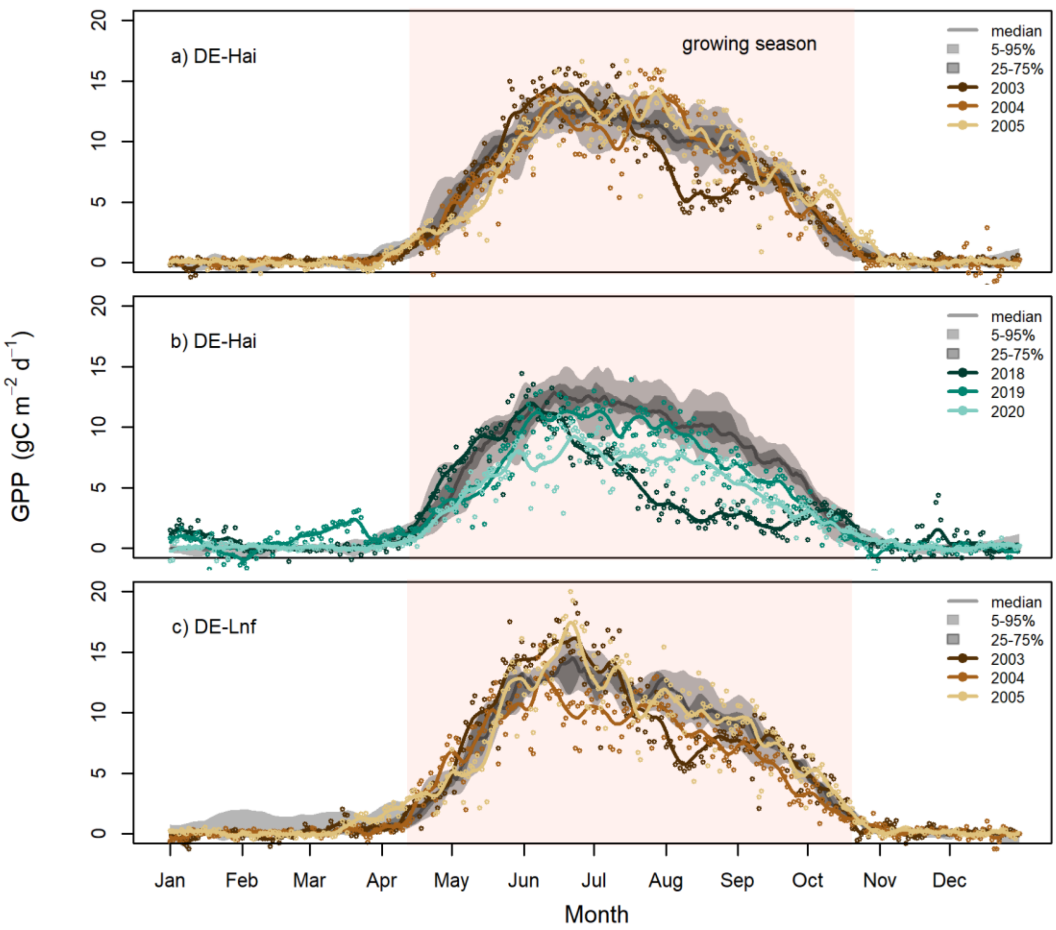

Figure 2. Daily GPP in the selected drought and legacy years at a) DE-Hai 2003, b) DE-Hai 2018 and c) DE-Lnf 2003 showing the droughts and following legacy years, respectively. Colored points and lines showed original and smoothed (7-days average) GPP, respectively, in drought and legacy years. The grey lines and shaded areas showed the median, 25th-75th (dark grey), and 5th-95th (light grey) percentiles of GPP, respectively, over non-drought and non-legacy years. The shaded coral areas indicate the average growing seasons of DE-Hai and DE-Lnf.

In Fig. 2, we show the measured absolute GPP time series in the selected drought (2003 and 2018) and legacy years (2004, 2005, 2019, and 2020) together with the long-term median, 25th-75th, and 5th-95th percentiles GPP at DE-Hai and DE-Lnf. In the drought year 2003, GPP was significantly lower than the baseline, defined as the 25th percentile GPP, during July-September at DE-Hai and July-August at DE-Lnf, respectively. In the post-drought years 2004 and 2005, there was no systematic decrease in GPP at DE-Hai, while GPP at DE-Lnf was slightly lower than the baseline during June-August of 2004. During the 2018 drought, GPP significantly differed from the baseline during June-September at DE-Hai. After the 2018 drought, we could not find any systematic decrease in GPP in 2019, while GPP was consistently lower than the baseline from mid-May to September of 2020 at DE-Hai.

4.2 Drought legacy effects on GPP: seasonal patterns

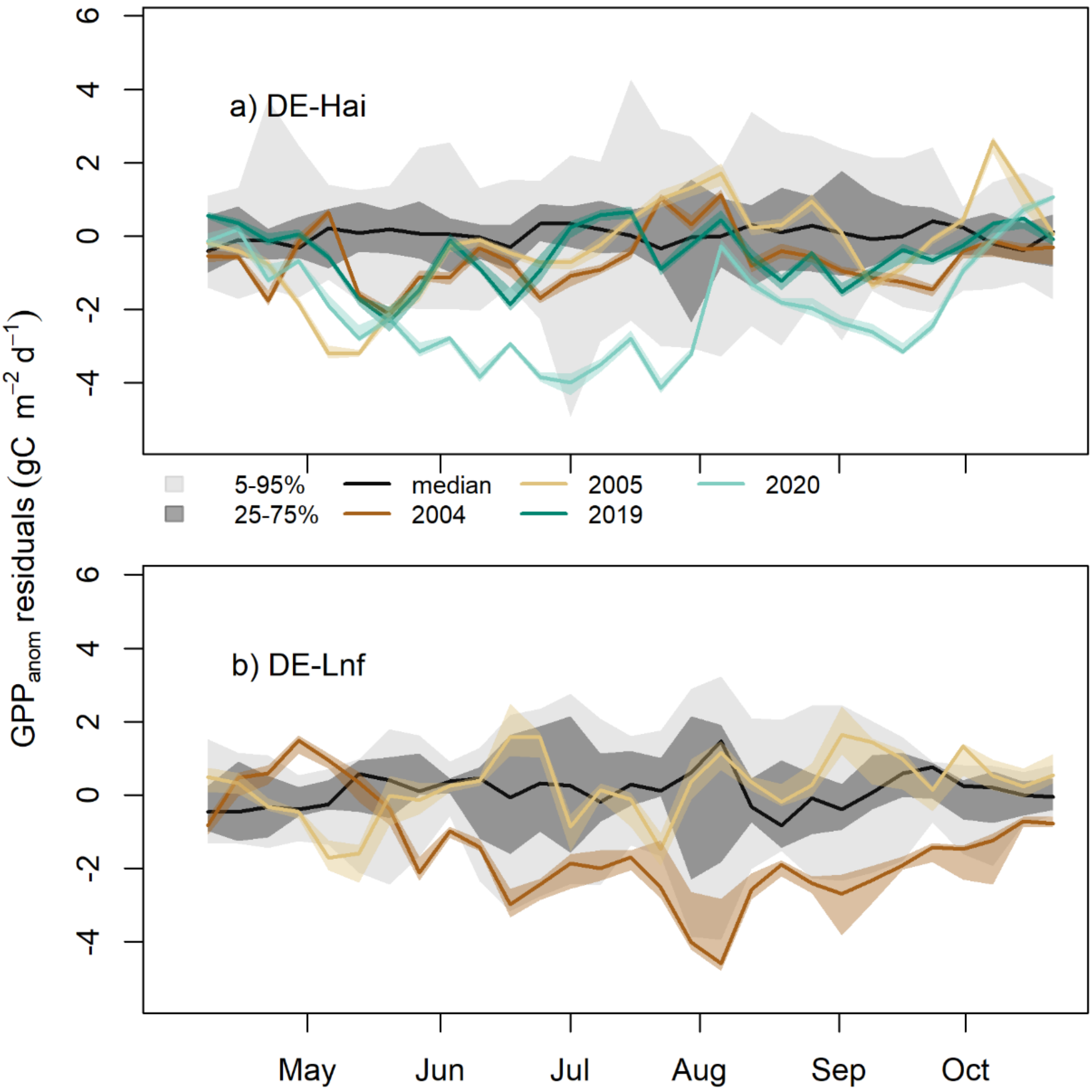

Figure 3. Residuals of GPP anomalies at the seasonal scale in legacy years at a) DE-Hai and b) DE-Lnf. Residuals of GPP anomalies were characterized by observed minus predicted GPP anomalies (GPP$_{anom}$ residuals). The color lines and bands show the median and 5th-95th percentile GPP$_{anom}$ residuals of ensemble model runs (see Section 3.4), respectively. Negative residuals corresponded to more negative or less positive GPP$_{anom}$ than would be expected given the climate in that year, indicating negative legacies of drought, while positive residuals corresponded to less negative or more positive GPP$_{anom}$, indicating beneficial legacies of drought. The model uncertainties (dark and light grey shaded area, respectively) are characterized by the 25th-75th and 5th-95th quantile ranges of GPP$_{anom}$ residuals in non-legacy years. The black line represents the median of GPP$_{anom}$ residuals in non-legacy years. The ticks denote the start of each month.

At the seasonal scale, residuals of GPP anomalies (GPP_{anom} residuals) showed significant departures from model uncertainties at both sites (Fig. 3). After the 2003 drought at DE-Hai, we found negative residuals below the 25th percentile of model residuals in non-legacy years (model uncertainties) during the early and late growing season of 2004 (April-July, September) and May-June of 2005, and below the 5th percentile for short periods, in April and May of 2004 and May of 2005. After June 2005, residuals were mostly within 5-95% of the model residuals. After the 2018 drought at DE-Hai, we found negative residuals (below 25th percentile of model residuals) during May, June, August, and September of 2019. In 2020, residuals showed a persistent decrease from May to July, and generally stayed well below the 5th and 25th percentile of model residuals from mid-May until July and September, respectively.

After the 2003 drought at DE-Lnf, we found persistent negative residuals were below the 25th percentile of model residuals over almost the complete growing season (from May to October) in 2004 and below the 5th percentile of model residuals for periods in June-September. In 2005, residuals remained mostly within 25th-75th percentiles of model residuals.

4.3 Drought legacy effects on GPP: annual patterns

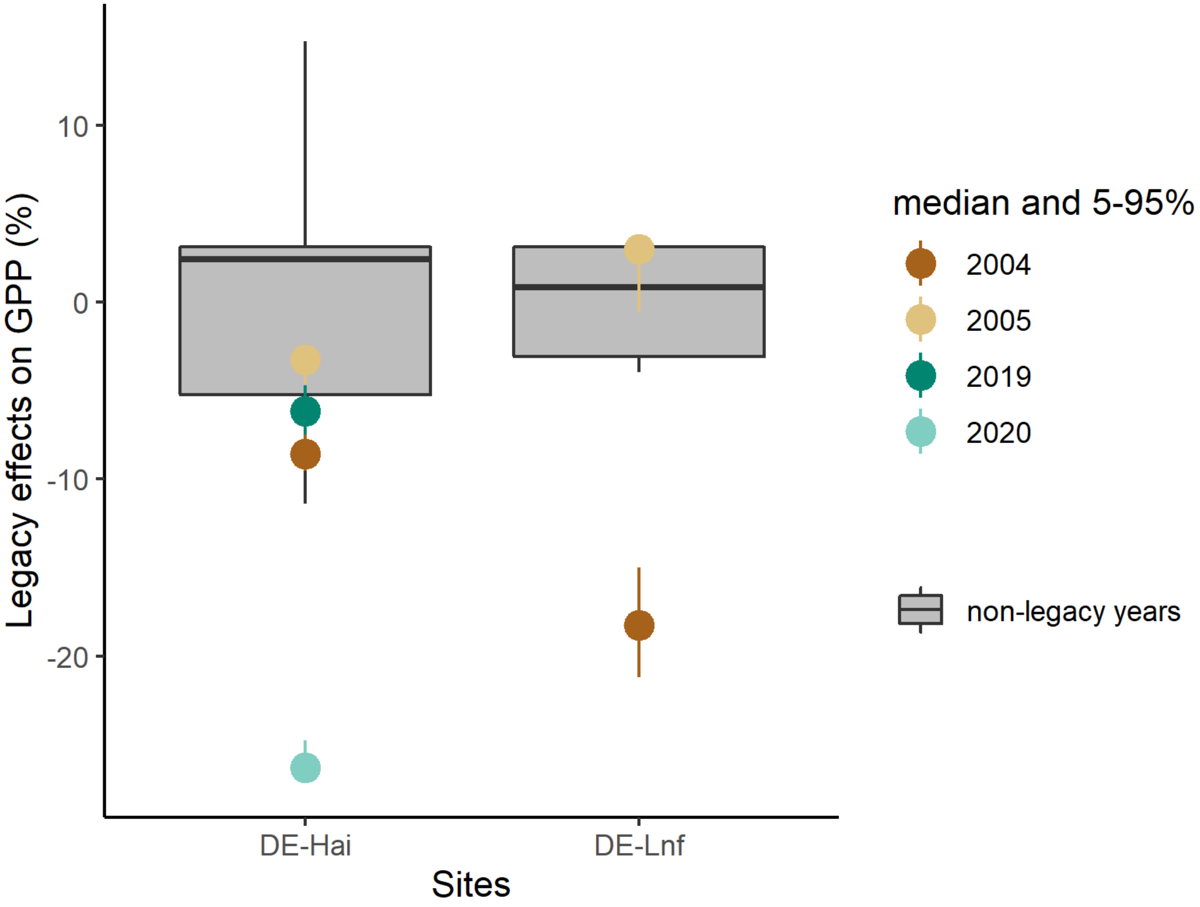

Figure 4. Integrated residuals of GPP anomalies at the annual scale in legacy years at DE-Hai and DE-Lnf. The color points and line ranges show the median and 5-95% percentile integrated GPP_{anom} residuals of ensemble model runs (see Section 3.4), respectively. The model uncertainties (the boxplot) are characterized as the 25th-75th quantile range of integrated GPP_{anom} residuals in non-legacy years.

There were systematic departures of integrated residuals of GPP anomalies in legacy years from model uncertainties at the annual scale (Fig. 4) although the seasonal patterns varied (Fig. 3). After the 2003 drought at DE-Hai, integrated residuals in 2004 were significantly below the 25th percentile of model residuals, while integrated residuals were within the 25th-75th percentiles of model residuals in 2005. After the 2018 drought, integrated residuals in 2019 were near the 25th percentiles of model residuals, while in 2020 integrated residuals were far below the 25th percentile of model residuals.

At DE-Lnf, after the 2003 drought, integrated residuals in 2004 were below the 25th percentile of residuals in non-legacy years, while integrated residuals almost remained within 25th-75th percentiles of model residuals in 2005.

4.4 Drought legacy effects on GPP due to structural and physiological effects

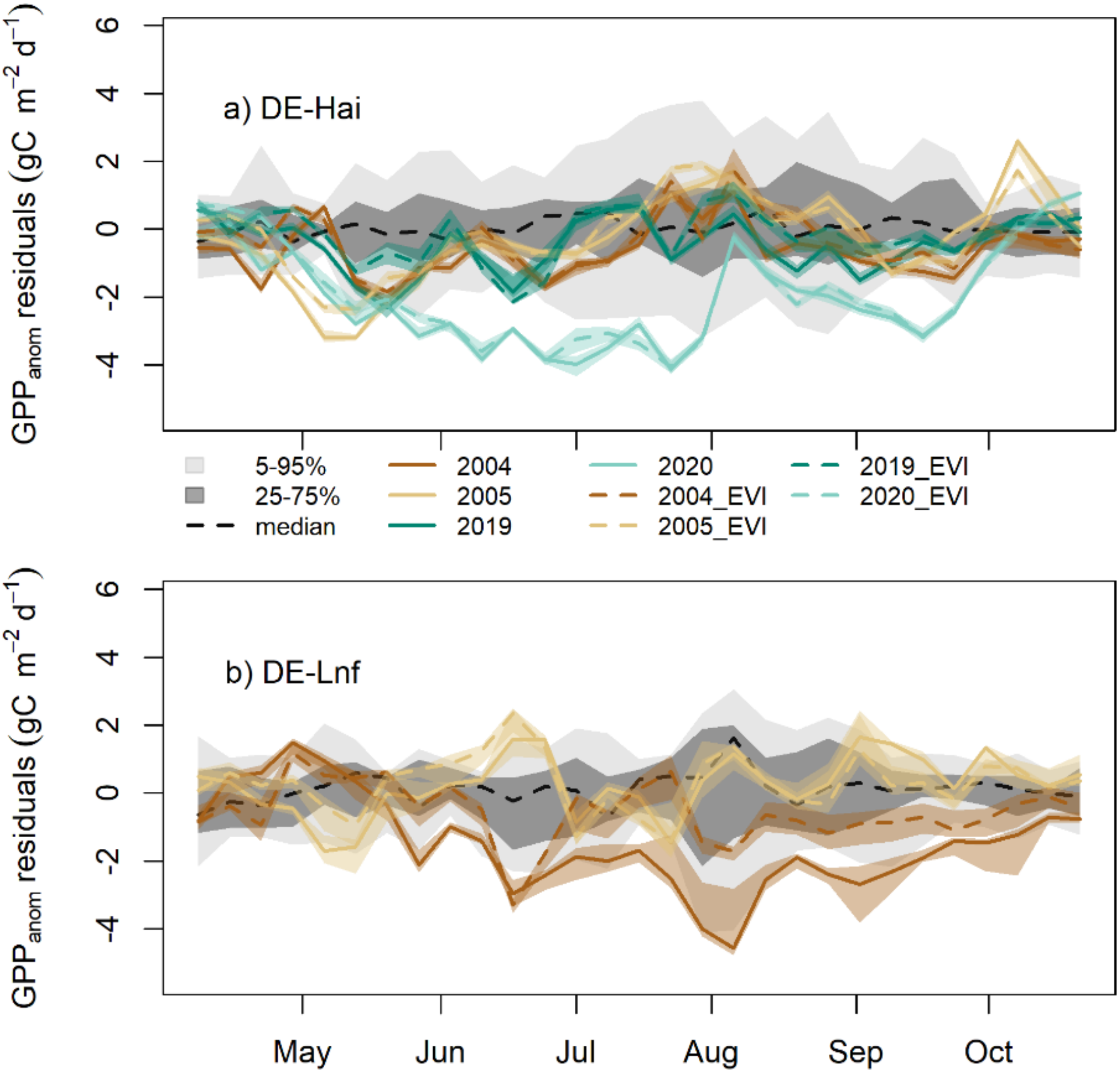

Figure 5. Residuals of GPP anomalies from RF and RF_{EVI} (see Section 3.6) in legacy years at a) DE-Hai and b) DE-Lnf. Residuals of GPP anomalies are characterized by observed minus predicted GPP anomalies (GPP_{anom} residuals). The color lines and bands show the median and 5th-95th percentile GPP_{anom} residuals of ensemble model runs (see Section 3.4), respectively. The solid and dashed lines show the residuals based on RF and RF_{EVI} , respectively. The model uncertainties from RF_{EVI} (dark and light grey shaded area, respectively) are characterized by the 25th-75th and 5th-95th quantile ranges of GPP_{anom} residuals in non-legacy years. The black dashed line was the median of GPP_{anom} residuals from RF_{EVI} in non-legacy years. The ticks denoted the start of each month. Figure S4 shows the results for April-June and August-October at DE-Hai in more detail.

At the seasonal scale, residuals of GPP anomalies from RF_{EVI} (Res_{EVI}) showed significant departures from GPP_{anom} residuals from RF (Res) over some periods at both sites (Fig. 5). At DE-Hai, we found Res_{EVI} was above Res in the early growing season

(April-May) of 2004, 2005, 2019, and 2020, and also in the late growing season of 2004 (August-October) and 2019 (August-September). After the 2003 drought, we found negative $\mathrm{Res}_{\mathrm{EVI}}$ below the 25th percentile of model residuals from $\mathrm{RF}_{\mathrm{EVI}}$ in non-legacy years (model uncertainties) during the early and late growing season of 2004 (May-July, September) and May of 2005, and below the 5th percentile for short periods, in May of 2005. After the 2018 drought, we found negative $\mathrm{Res}_{\mathrm{EVI}}$ (below 25th percentile of model residuals) during June of 2019. In 2020, $\mathrm{Res}_{\mathrm{EVI}}$ showed a persistent decrease from May to July, and generally stayed well below the 5th and 25th percentile of model residuals from mid-May until July and September, respectively. At DE-Lnf, $\mathrm{Res}_{\mathrm{EVI}}$ was below Res from April to mid-May and significantly above Res almost over the growing season of 2004 (from mid-May to September). We found negative $\mathrm{Res}_{\mathrm{EVI}}$ below the 25th percentile of model residuals from $\mathrm{RF}_{\mathrm{EVI}}$ in non-legacy years (model uncertainties) during June, August, and September of 2004, and below the 5th percentile for short periods, in June and September of 2004.

4.5 Drought legacy effects on radial increment

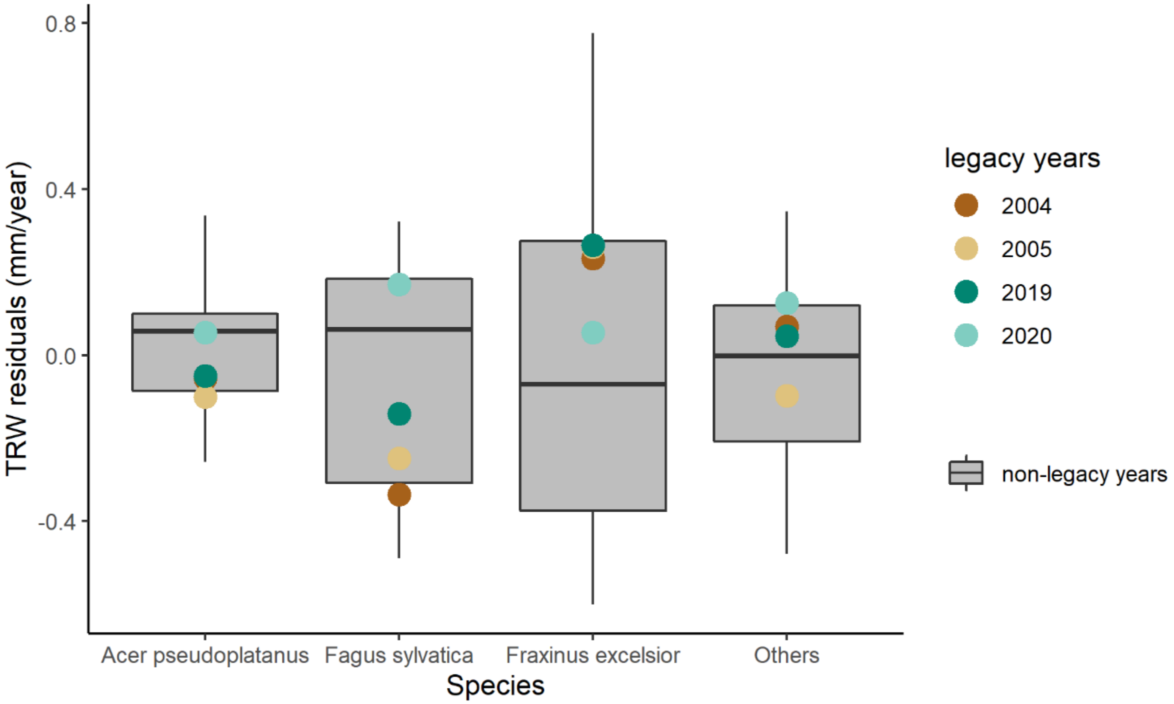

Figure 6. Residuals of RI in legacy years at DE-Hai across species. Residuals of RI are characterized as observed minus predicted RI anomalies (RI residuals). The model uncertainties (the grey area) are characterized as the 25th-75th quantile range of RI residuals in non-legacy years.

To complement the analysis of the legacy effects on GPP at the seasonal and annual scales, we also evaluated legacy effects on tree growth at the annual scale. RI of *Fagus sylvatica* was below the 25th percentile of model residuals in the post-drought year 2004 and returned to the 25th-75th percentiles of model residuals in 2005. For species of *Acer pseudoplatanus*, *Fraxnius excelsior*, and others, residuals of RI were almost within 25th-75th percentiles of model residuals in 2004 and 2005. After the 2018 drought, RI of all species for 2019 and 2020 were almost within or close to 25th-75th percentiles of model residuals.

4.6 Concurrent and lagged reduction in GPP

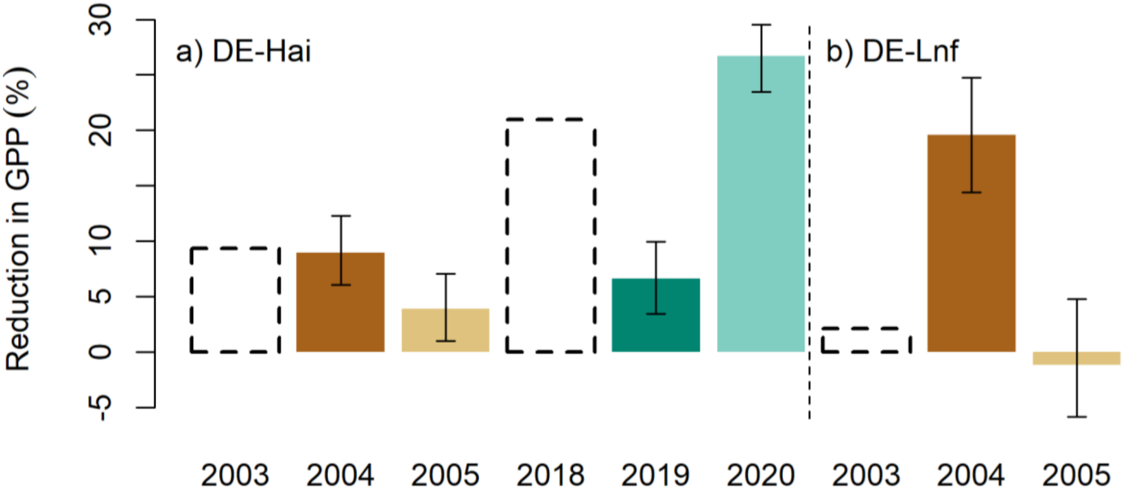

Figure 7. Concurrent (dashed black bars) and lagged (colored bars) reduction in GPP from the 2003 and 2018 droughts at a) DE-Hai and b) DE-Lnf. Concurrent impacts in GPP were quantified as the sum of GPP anomalies over drought periods in drought years relative to averaged total GPP over the growing season (see Method). Lagged impacts in GPP are characterized as the difference between GPP_{anom} residuals in legacy years and median of the model uncertainties relative to averaged total GPP over the growing season. Colored bars and error bars show the median and 5-95%, respectively, of lagged reduction in GPP from ensemble model runs.

Finally, we compared the concurrent impacts on GPP with the lagged impacts due to drought. We found that, at DE-Hai, the concurrent reduction in GPP was 9.4% relative to averaged total GPP over the growing season (hereinafter) in 2003, while 6.1-12.3% indirectly reduced in 2004. And in 2018 concurrent reduction in GPP was 21.0%, while 3.5%-10.0% and 23.5-29.6% indirectly reduced in 2019 and 2020, respectively. At DE-Lnf, the concurrent reduction in GPP was negligible in 2003 (2.2%), while we estimated 14.4-24.8% GPP reduction in 2004, which was higher than the corresponding values at DE-Hai in the same year.

5. Discussion

5.1 A novel methodology to detect drought legacy effects on GPP

There is limited research on discovering legacy effects of drought on the ecosystem carbon cycle using eddy-covariance observations (Kannenberg et al., 2019). Here, we propose a residual-based methodology using a random-forest regression model to detect legacy effects on GPP, and found significant legacy effects on GPP using eddy-covariance data at two forests in central Germany in the similar climate but with different age and species composition. There are three advantages to our methodology: 1) capturing the temporal dynamics of legacy effects at the seasonal scales; 2) separating the influence of meteorological conditions during the post-drought period on recovery rates; 3) estimating model uncertainties to avoid misinterpreting small residuals as ‘legacy effects’.

First, because we used measurements with a high temporal resolution (daily), legacy effects could be determined across different time scales. Previous studies based on tree-ring or satellite-greenness data have mainly focused on legacy effects at the annual scale (Anderegg et al., 2015; Wu et al., 2018) or monthly scale (Bastos et al., 2021), but the legacies can be more ephemeral, for example, if they appear only in critical periods of the growing season, as we have found here. Such temporally confined effects may not necessarily manifest themselves at the annual scale. For example, after the 2003 drought, the annual GPP at DE-Hai in 2005 was close to normal, which was the 25th percentile of model residuals here, but we found short legacies at the seasonal scale (Fig. 3).

Second, recovery is usually considered when the target variable (i.e. GPP, tree-ring width…) returns to the baseline, usually based on pre-drought values of the target variable (Bose et al., 2020; González de Andrés et al., 2021; Zhang et al., 2021). However, meteorological conditions during the recovery period will modulate recovery rates, so that recovery can be delayed e.g. if a drought is followed by other unfavourable climatic conditions. Hence, the evaluation of possible legacy effects should be based on the functional relations between the target variable and meteorological conditions. Our model takes this into account by considering that ecosystems recovered when observed GPP reaches the potential GPP given the meteorological conditions, rather than the absolute flux.

Finally, our approach allows determining the uncertainties in estimated legacy effects. Previous studies (Anderegg et al., 2015; Huang et al., 2018) quantified legacy effects as the residuals between observed and predicted target variables (i.e. tree-ring width, vegetation indices, …) in legacy years, but were not able to consider uncertainties of their trained models. Yet, it is crucial to understand if the residuals are caused by model uncertainties or can be interpreted as legacy effects. In this study, legacy effects are identified only when the model residuals are outside the range of the model uncertainties, so that we are confident that the legacies reported here are significant and avoid interpreting residuals caused by model error as legacy effects. A limitation of our approach is that we have to assume that there are no legacy effects in the climate system because this would potentially bias the interpretation of the residuals.

The methodology we proposed is able to detect the legacy effects of drought on GPP and can be easily applied to other eddy-covariance sites and variables (i.e. evapotranspiration, transpiration, …), in order to improve our understanding of drought legacy effects on the ecosystem carbon cycle at different time-scales.

5.2 Seasonal and annual legacy drought impacts on GPP

We found that residuals of GPP anomalies (GPP_{anom} residuals) in legacy years were significantly larger than model uncertainties at both seasonal and annual scales at both sites, which indicated strong legacy effects of drought on GPP at least in the two years following the drought events.

We found negative legacies on GPP (reduced uptake) in the early growing season of all legacy years (2004, 2005, 2019, and 2020) at DE-Hai. Reduced and delayed leaf development due to physiological effects of the 2003 and 2018 droughts (e.g. metabolic damage, non-structural carbohydrates depletion) could result in reduced ecosystem-level photosynthesis (Migliavacca et al., 2009; Rocha and Goulden, 2010; Kannenberg et al., 2019), and could potentially explain negative legacies on GPP at the start of the growing season. In line with this hypothesis, we found the enhanced vegetation index (EVI, a proxy of leaf area index, Fig. S5 and Fig. S6) at the sites showed lower values than other years in the early growing seasons of 2004, 2005, and 2019 and this delayed spring phenology propagated over the year of 2004 and 2019 with a shift of seasonality. We found consistently lower values of NPP allocated to foliage growth in 2004 than other years (Fig. S7). Furthermore, the detected negative legacies in the early growing season became smaller when adding EVI anomalies as an additional predictor in the random forest model (Fig. 5), indicating that the reduced and delayed leaf development partly explained the estimated legacy effects by the RF model trained with climate predictors only.

Another possible mechanism explaining legacy effects could be hydraulic damage induced by drought (Anderegg et al., 2013), and therefore insufficient ability of water transport limiting sink strength (Körner, 2015) and photosynthetic capacity (Chen et al., 2010), at least until damage is repaired. If this was the case, transpiration fluxes should be reduced. However, we did not find similar negative legacy patterns on transpiration in the early growing season (Fig. S8a). Therefore, hydraulic damage did not seem a likely cause of drought legacies on GPP for these events. Overall, we cannot pinpoint the physiological causes of the detected legacy effects due to limited availability of measurements. This calls for establishing more plant-physiological measurements complementing eddy-covariance and RI measurements to capture sufficient information about plant water relations such as sap flow (Poyatos et al., 2021) and tree water deficit (Nehemy et al., 2021) as well as carbon allocation (Hartmann et al, 2020) to provide a more detailed process understanding of the mechanisms underlying drought legacy effects.

Negative legacies on GPP in terms of lagged reduction in GPP in 2004 at DE-Lnf (14.4-24.8%) were stronger than those at DE-Hai (6.1-12.3%) in the seasonal and annual scales. The persistence of negative legacies throughout the full growing season in 2004 indicates that the 2003 drought likely caused stronger damage, especially reduced leaf development which was supported by largely reduced negative legacies of RF_{EVI} with EVI comparing to RF without EVI (Fig. 5), on the ecosystem at DE-Lnf than that at DE-Hai. From the community-level perspective, the stronger legacy effects found at DE-Lnf compared to

DE-Hai may have been partly related to differences in forest composition between the two sites (Tamrakar et al. 2018, Pardos et al., 2021). Measurements of GPP at tree species level were not available, therefore we relied on the legacies found for RI (reflecting growth), available for individual trees at DE-Hai. It should be noted, though, that the relationship between GPP and growth is complex (Fatichi et al., 2014). Negative legacy effects on RI of *Fagus sylvatica*, dominating at DE-Hai, in 2004, were found, whereas other co-dominating species (*Acer pseudoplatanus* and *Fraxinus excelsior*) did not show negative legacies. Therefore, the lower resilience of *Fagus sylvatica* compared to other species may have partly resulted in stronger negative legacies at the pure European beech forest at DE-Lnf than at DE-Hai. In addition, contrasting legacy effects of these two sites could also be associated with different age classes and the absolute stand age since the effects of stand age modulating the heat and drought impact on carbon exchange (Arain et al., 2022) and ecosystem-level photosynthetic capacity (Musavi et al., 2017) have been recognized. However, the evidence of species diversity and age structure effects on legacy effects needs to be further explored using more sites in the future.

Stronger negative legacy effects on GPP in 2020 than those in other legacy years were found at DE-Hai in the seasonal and annual scales. This might be associated with significant tree mortality in the whole forest including the main footprint in the period 2018-2020 (about 6% year^{-1} between 2017 and 2020 compared to less than 1% year^{-1} between 2005 and 2017) mainly caused by the storm *Friedrike* in January 2018 and the heat and/or drought in summer 2018 and 2019 (unpublished data). RI of *Fagus sylvatica* in 2020 showed slightly positive legacy effects in growth, since only living trees were sampled. This might be explained by the favorable weather conditions in winter/spring 2019/2020 associated with high mineralization rates and reduced competition for nutrients, light and water of the surviving trees (Grossiord, 2020). The RI data reflected mean growth signals from individual surviving trees, while the GPP data reflected mean carbon assimilation at stand level, including positive, negative or absent legacy effects at individual tree level as well as the reduction of assimilating individuals due to higher tree mortality.

Overall, we found that the lagged impacts of drought on GPP are significant compared with concurrent drought impacts at the two sites studied here. The lagged reduction in GPP resulting from drought is usually not quantified (Ciais et al., 2005; Reichstein et al., 2007), perhaps because separating legacy effects on ecosystem carbon fluxes from observations is challenging (Kannenberg et al., 2019) and process-based models have been shown to miss such legacy effects (Bastos et al., 2021). This implies that the impact of droughts on ecosystem carbon cycling in most studies might be underestimated.

5.3 Importance of deep root-zone soil moisture data

Deep root-zone soil moisture has been recognized as an important water source for vegetation, especially during droughts (Miguez-Macho and Fan, 2021; Werner et al., 2021). Although soil moisture measurements across three soil layers are available at both sites, the deepest depth (ca. 30cm) cannot capture the entire soil water reservoir available for European beech which has been observed to have non-negligible amounts of fine roots below 30cm across different sites (Leuschner et al., 2004, Gessler et al., 2021).

We tested an initial model using anomalies of soil moisture at three layers as predictors (RF_{SM}), and found strong positive legacy effects in the late growing season in 2019 at DE-Hai (Fig. S9), which however could not be reproduced by any of the models using soil moisture information from deeper layers (Fig. S9) including the local water balance (WAI, CWD) and the reanalysis data (ERA5). Comparing the predicted time series of GPP_{anom} of the RF_{SM} model with observations, we found the predicted GPP_{anom} became much more negative in the late growing season while observed GPP_{anom} were close to zero (Fig. S10). Therefore, although soil moisture anomalies in the third layer (30cm) were largely negative when the positive residuals appeared (Fig. S11), soil moisture from layers deeper than 30 cm may maintain the water supply for photosynthesis. Also, we found the evapotranspiration from the shallow layers (0~30cm) estimated by soil moisture decrease was less than the observed evapotranspiration during dry-down periods (Fig. S11), which indicated plant water uptake from layers deeper than 30 cm during dry-down periods, in line with our hypothesis. In summary, these positive patterns are likely due to model errors from incomplete information on the soil-moisture profile rather than actual positive legacy effects.

These results highlight the importance of soil moisture measurements that capture the entire root zone for more reliable understanding of ecosystem functioning, particularly in the case of drought legacy effects.

6. Conclusions

The frequency, intensity, duration, and spatial extent of droughts are expected to increase in the next decades due to anthropogenically caused global warming in many regions (IPCC, 2022). Drought not only impacts ecosystems concurrently, but also can have legacy effects on ecosystem carbon fluxes. We developed a residual-based approach using a random forest regression model to detect drought legacies on gross primary productivity (GPP) using eddy-covariance data. The methodology proposed here allows quantifying significant drought legacy effects on GPP at the sub-seasonal and annual scales. The GPP reduction due to drought legacy effects is of comparable magnitude to the concurrent drought effects at the studied sites, which confirms the importance of legacy effects. We found contrasting legacy effects at two neighbouring forests with different species and age structures, yet the importance of these factors could not be evaluated. Future studies across a larger range of sites will be needed to understand whether the crucial role of legacy effects is general and on which mediating factors they depend.

Acknowledgements

Xin Yu and Markus Reichstein acknowledge funding from the European Research Council (ERC) Synergy Grant “Understanding and Modelling the Earth System with Machine Learning (USMILE)” under the Horizon 2020 research and innovation programme (Grant agreement No. 855187). Xin Yu acknowledges support from the International Max Planck Research School for Global Biogeochemical Cycles. René Orth acknowledges support by the German Research Foundation

(Emmy Noether grant 391059971). Franziska Koebsch and Alexander Knohl acknowledge support by Niedersächsisches Vorab (ZN 3679), Ministry of Lower-Saxony for Science and Culture (MWK). Martina Mund acknowledges support by the Integrated project CarboEurope-IP (European Commission, Directorate-General Research, Sixth Framework Programme, Priority 1.1.6.3: Global Change and Ecosystem, Contract no. GOCECT-2003-505572), the Max Planck Institute for Biogeochemistry, Germany, the German Research Foundation (DFG) (INST 186/1118-1 FUGG), the German Federal Ministry of Education and Research (BMBF; research infrastructure ICOS) and the Georg-August-University Göttingen, Germany. Sophia Walther acknowledges funding within the ESA Living Planet Fellowship on the project Vad3e mecum. Benjamin D. Stocker was funded by the Swiss National Science Foundation grant no. PCEFP2_181115. We thank the Warm Winter 2020 initiative of the Integrated Carbon Observation System (ICOS), specifically Nicola Arriga and Dario Papale, and the administration of the Hainich National Park for the opportunity for research within the National Park.

Author contributions

The study was conceived by X. Yu, A. Bastos, R. Orth, M. Reichstein, and M. Bahn. X. Yu implemented the method and performed the data analyses. A. Knohl, A. Klosterhalfen, and F. Koebsch provided the eddy-covariance data. M. Mund provided the data of radial increment and net primary productivity of fruits and leaves. J. A. Nelson helped X. Yu to process the transpiration estimation. S. Walther provided and processed the Enhanced Vegetation Index data. B. D. Stocker suggested quantitatively separating structural and physiological effects. M. Migliavacca helped to interpret the results. X. Yu, A. Bastos, R. Orth, M. Reichstein, and M. Bahn prepared the first draft and all authors contributed to discussion of results and the revisions of the manuscript.

Competing interests

At least one of the (co-)authors is a member of the editorial board of Biogeosciences. The peer-review process should be guided by an independent editor, and the authors also have no other competing interests to declare.

Code and data availability

Eddy-covariance and enhanced vegetation index data used are freely accessible. Tree ring width and net primary productivity of fruits and leaves data are available on request to Martina Mund. Our code is available on request.

References

Allen, C. D., Breshears, D. D., and McDowell, N. G.: On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene, 6, art129, https://doi.org/10.1890/ES15-00203.1, 2015.

Anderegg, W. R. L., Plavcová, L., Anderegg, L. D. L., Hacke, U. G., Berry, J. A., and Field, C. B.: Drought's legacy: multiyear hydraulic deterioration underlies widespread aspen forest die-off and portends increased future risk, 19, 1188–1196, https://doi.org/10.1111/gcb.12100, 2013.

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, 349, 528–532, https://doi.org/10.1126/science.aab1833, 2015.

Anthoni, P. M., Knohl, A., Rebmann, C., Freibauer, A., Mund, M., Ziegler, W., Kolle, O., and Schulze, E.-D.: Forest and agricultural land-use-dependent CO2 exchange in Thuringia, Germany, 10, 2005–2019, https://doi.org/10.1111/j.1365-2486.2004.00863.x, 2004.

Arain, M. A., Xu, B., Brodeur, J. J., Khomik, M., Peichl, M., Beamesderfer, E., Restrepo-Couple, N., and Thorne, R.: Heat and drought impact on carbon exchange in an age-sequence of temperate pine forests, Ecological Processes, 11, 7, https://doi.org/10.1186/s13717-021-00349-7, 2022.

Assal, T. J., Anderson, P. J., and Sibold, J.: Spatial and temporal trends of drought effects in a heterogeneous semi-arid forest ecosystem, Forest Ecology and Management, 365, 137–151, https://doi.org/10.1016/j.foreco.2016.01.017, 2016.

Aubinet, M., Grelle, A., Ibrom, A., Rannik, U., Moncrieff, J., Foken, T., Kowalski, A. S., Martin, P. H., Berbigier, P., Grunwald, T., Morgenstern, K., Pilegaard, K., Rebmann, C., Bernhofer, C., Clement, R., Elbers, J., Granier, A., Snijders, W., Valentini, R., and Vesala, T.: Estimates of the Annual Net Carbon and Water Exchange, 30, 113–174, 2000.

Bastos, A., Orth, R., Reichstein, M., Ciais, P., Viovy, N., Zaehle, S., Anthoni, P., Arneth, A., Gentine, P., Joetzjer, E., Lienert, S., Loughran, T., McGuire, P. C., O, S., Pongratz, J., and Sitch, S.: Vulnerability of European ecosystems to two compound dry and hot summers in 2018 and 2019, 12, 1015–1035, https://doi.org/10.5194/esd-12-1015-2021, 2021.

Beringer, J., Hutley, L. B., Tapper, N. J., and Cernusak, L. A.: Savanna fires and their impact on net ecosystem productivity in North Australia, Global Change Biol, 13, 990–1004, https://doi.org/10.1111/j.1365-2486.2007.01334.x, 2007.

Bose, A. K., Gessler, A., Bolte, A., Bottero, A., Buras, A., Cailleret, M., Camarero, J. J., Haeni, M., Hereş, A.-M., Hevia, A., Lévesque, M., Linares, J. C., Martinez-Vilalta, J., Matías, L., Menzel, A., Sánchez-Salguero, R., Saurer, M., Vennetier, M., Ziche, D., and Rigling, A.: Growth and resilience responses of Scots pine to extreme droughts across Europe depend on predrought growth conditions, 26, 4521–4537, https://doi.org/10.1111/gcb.15153, 2020.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Chen, J.-W., Zhang, Q., Li, X.-S., and Cao, K.-F.: Gas exchange and hydraulics in seedlings of Hevea brasiliensis during water stress and recovery, Tree Physiology, 30, 876–885, https://doi.org/10.1093/treephys/tpq043, 2010.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., Chevallier, F., De Noblet, N., Friend, A. D., Friedlingstein, P., Grünwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Papale, D., Pilegaard, K., Rambal, S., Seufert, G., Soussana, J. F., Sanz, M. J., Schulze, E. D., Vesala, T., and Valentini, R.: Europe-wide reduction in primary productivity caused by the heat and drought in 2003, 437, 529–533, https://doi.org/10.1038/nature03972, 2005.

Erbilgin, N., Zanganeh, L., Klutsch, J. G., Chen, S., Zhao, S., Ishangulyyeva, G., Burr, S. J., Gaylord, M., Hofstetter, R., Keefover-Ring, K., Raffa, K. F., and Kolb, T.: Combined drought and bark beetle attacks deplete non-structural carbohydrates and promote death of mature pine trees, 44, 3866–3881, https://doi.org/10.1111/pce.14197, 2021.

Fatichi, S., Leuzinger, S., and Körner, C.: Moving beyond photosynthesis: from carbon source to sink-driven vegetation modeling, 201, 1086–1095, https://doi.org/10.1111/nph.12614, 2014.

Frank, D., Reichstein, M., Bahn, M., Thonicke, K., Frank, D., Mahecha, M. D., Smith, P., Velde, M. van der, Vicca, S., Babst, F., Beer, C., Buchmann, N., Canadell, J. G., Ciais, P., Cramer, W., Ibrom, A., Miglietta, F., Poulter, B., Rammig, A., Seneviratne, S. I., Walz, A., Wattenbach, M., Zavala, M. A., and Zscheischler, J.: Effects of climate extremes on the terrestrial carbon cycle: concepts, processes and potential future impacts, 21, 2861–2880, https://doi.org/10.1111/gcb.12916, 2015.

Fu, Z., Ciais, P., Bastos, A., Stoy, P. C., Yang, H., Green, J. K., Wang, B., Yu, K., Huang, Y., Knohl, A., Šigut, L., Gharun, M., Cuntz, M., Arriga, N., Roland, M., Peichl, M., Migliavacca, M., Cremonese, E., Varlagin, A., Brümmer, C., Gourlez de la Motte, L., Fares, S., Buchmann, N., El-Madany, T. S., Pitacco, A., Vendrame, N., Li, Z., Vincke, C., Magliulo, E., and Koebsch, F.: Sensitivity of gross primary productivity to climatic drivers during the summer drought of 2018 in Europe, 375, 20190747, https://doi.org/10.1098/rstb.2019.0747, 2020.

Galvagno, M., Wohlfahrt, G., Cremonese, E., Rossini, M., Colombo, R., Filippa, G., Julitta, T., Manca, G., Siniscalco, C., Cella, U. M. di, and Migliavacca, M.: Phenology and carbon dioxide source/sink strength of a subalpine grassland in response to an exceptionally short snow season, Environ. Res. Lett., 8, 025008, https://doi.org/10.1088/1748-9326/8/2/025008, 2013.

Gessler, A., Bächli, L., Rouholahnejad Freund, E., Treydte, K., Schaub, M., Haeni, M., Weiler, M., Seeger, S., Marshall, J., and Hug, C.: Drought reduces water uptake in beech from the drying topsoil, but no compensatory uptake occurs from deeper soil layers, New Phytologist, 2021.

González de Andrés, E., Rosas, T., Camarero, J. J., and Martínez-Vilalta, J.: The intraspecific variation of functional traits modulates drought resilience of European beech and pubescent oak, 109, 3652–3669, https://doi.org/10.1111/1365-2745.13743, 2021.

Grossiord, C.: Having the right neighbors: how tree species diversity modulates drought impacts on forests, 228, 42–49, https://doi.org/10.1111/nph.15667, 2020.

Haberstroh, S. and Werner, C.: The role of species interactions for forest resilience to drought, n/a, https://doi.org/10.1111/plb.13415, 2022.

Hacket-Pain, A. J., Friend, A. D., Lageard, J. G. A., and Thomas, P. A.: The influence of masting phenomenon on growth–climate relationships in trees: explaining the influence of previous summers' climate on ring width, Tree Physiology, 35, 319–330, https://doi.org/10.1093/treephys/tpv007, 2015.

Hartmann, H. and Trumbore, S.: Understanding the roles of nonstructural carbohydrates in forest trees – from what we can measure to what we want to know, 211, 386–403, https://doi.org/10.1111/nph.13955, 2016.

Hartmann, H., Bahn, M., Carbone, M., and Richardson, A. D.: Plant carbon allocation in a changing world – challenges and progress: introduction to a Virtual Issue on carbon allocation, 227, 981–988, https://doi.org/10.1111/nph.16757, 2020.

Herbst, M., Mund, M., Tamrakar, R., and Knohl, A.: Differences in carbon uptake and water use between a managed and an unmanaged beech forest in central Germany, Forest Ecology and Management, 355, 101–108, https://doi.org/10.1016/j.foreco.2015.05.034, 2015.

Huang, J., Hammerbacher, A., Gershenzon, J., Dam, N. M. van, Sala, A., McDowell, N. G., Chowdhury, S., Gleixner, G., Trumbore, S., and Hartmann, H.: Storage of carbon reserves in spruce trees is prioritized over growth in the face of carbon limitation, PNAS, 118, https://doi.org/10.1073/pnas.2023297118, 2021.

Huang, M., Wang, X., Keenan, T. F., and Piao, S.: Drought timing influences the legacy of tree growth recovery, 24, 3546–3559, https://doi.org/10.1111/gcb.14294, 2018.

IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. In Press.

Kannenberg, S. A., Novick, K. A., Alexander, M. R., Maxwell, J. T., Moore, D. J. P., Phillips, R. P., and Anderegg, W. R. L.: Linking drought legacy effects across scales: From leaves to tree rings to ecosystems, 25, 2978–2992, https://doi.org/10.1111/gcb.14710, 2019.

Kannenberg, S. A., Schwalm, C. R., and Anderegg, W. R. L.: Ghosts of the past: how drought legacy effects shape forest functioning and carbon cycling, 23, 891–901, https://doi.org/10.1111/ele.13485, 2020.

Kendall, M. G.: Rank correlation methods., 1948.

Knohl, A., Schulze, E.-D., Kolle, O., and Buchmann, N.: Large carbon uptake by an unmanaged 250-year-old deciduous forest in Central Germany, Agricultural and Forest Meteorology, 118, 151–167, https://doi.org/10.1016/S0168-1923(03)00115-1, 2003.

Körner, C.: Paradigm shift in plant growth control, Current Opinion in Plant Biology, 25, 107–114, https://doi.org/10.1016/j.pbi.2015.05.003, 2015.

Krishnan, P., Black, T. A., Grant, N. J., Barr, A. G., Hogg, E. (Ted) H., Jassal, R. S., and Morgenstern, K.: Impact of changing soil moisture distribution on net ecosystem productivity of a boreal aspen forest during and following drought, Agricultural and Forest Meteorology, 139, 208–223, https://doi.org/10.1016/j.agrformet.2006.07.002, 2006.

Leuschner, C., Hertel, D., Schmid, I., Koch, O., Muhs, A., and Hölscher, D.: Stand fine root biomass and fine root morphology in old-growth beech forests as a function of precipitation and soil fertility, Plant and Soil, 258, 43–56, https://doi.org/10.1023/B:PLSO.0000016508.20173.80, 2004.

Lewis, S. L., Brando, P. M., Phillips, O. L., van der Heijden, G. M. F., and Nepstad, D.: The 2010 Amazon Drought, 331, 554–554, https://doi.org/10.1126/science.1200807, 2011.

Ma, X., Huete, A., Moran, S., Ponce-Campos, G., and Eamus, D.: Abrupt shifts in phenology and vegetation productivity under climate extremes, 120, 2036–2052, https://doi.org/10.1002/2015JG003144, 2015.

Meinshausen, N., & Ridgeway, G. (2006). Quantile regression forests. Journal of Machine Learning Research, 7(6).

Migliavacca, M., Meroni, M., Manca, G., Matteucci, G., Montagnani, L., Grassi, G., Zenone, T., Teobaldelli, M., Goded, I., Colombo, R., and Seufert, G.: Seasonal and interannual patterns of carbon and water fluxes of a poplar plantation under peculiar eco-climatic conditions, Agricultural and Forest Meteorology, 149, 1460–1476, https://doi.org/10.1016/j.agrformet.2009.04.003, 2009.

Miguez-Macho, G. and Fan, Y.: Spatiotemporal origin of soil water taken up by vegetation, 598, 624–628, https://doi.org/10.1038/s41586-021-03958-6, 2021.

Mund, M., Herbst, M., Knohl, A., Matthäus, B., Schumacher, J., Schall, P., Siebicke, L., Tamrakar, R., and Ammer, C.: It is not just a ‘trade-off’: indications for sink- and source-limitation to vegetative and regenerative growth in an old-growth beech forest, 226, 111–125, https://doi.org/10.1111/nph.16408, 2020.

Musavi, T., Migliavacca, M., Reichstein, M., Kattge, J., Wirth, C., Black, T. A., Janssens, I., Knohl, A., Loustau, D., Roupsard, O., Varlagin, A., Rambal, S., Cescatti, A., Gianelle, D., Kondo, H., Tamrakar, R., and Mahecha, M. D.: Stand age and species richness dampen interannual variation of ecosystem-level photosynthetic capacity, Nat Ecol Evol, 1, 1–7, https://doi.org/10.1038/s41559-016-0048, 2017.

Nehemy, M. F., Benettin, P., Asadollahi, M., Pratt, D., Rinaldo, A., and McDonnell, J. J.: Tree water deficit and dynamic source water partitioning, 35, e14004, https://doi.org/10.1002/hyp.14004, 2021.

Nelson, J. A., Carvalhais, N., Cuntz, M., Delpierre, N., Knauer, J., Ogée, J., Migliavacca, M., Reichstein, M., and Jung, M.: Coupling Water and Carbon Fluxes to Constrain Estimates of Transpiration: The TEA Algorithm, 123, 3617–3632, https://doi.org/10.1029/2018JG004727, 2018.

Orth, R., Destouni, G., Jung, M., and Reichstein, M.: Large-scale biospheric drought response intensifies linearly with drought duration in arid regions, 17, 2647–2656, https://doi.org/10.5194/bg-17-2647-2020, 2020.

Pardos, M., del Río, M., Pretzsch, H., Jactel, H., Bielak, K., Bravo, F., Brazaitis, G., Defossez, E., Engel, M., Godvod, K., Jacobs, K., Jansone, L., Jansons, A., Morin, X., Nothdurft, A., Oreti, L., Ponette, Q., Pach, M., Riofrío, J., Ruíz-Peinado, R., Tomao, A., Uhl, E., and Calama, R.: The greater resilience of mixed forests to drought mainly depends on their composition:

Analysis along a climate gradient across Europe, Forest Ecology and Management, 481, 118687, https://doi.org/10.1016/j.foreco.2020.118687, 2021.

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardö, J., Arkebauer, T., Arndt, S. K., Arriga, N., Aubinet, M., Aurela, M., Baldocchi, D., Barr, A., Beamesderfer, E., Marchesini, L. B., Bergeron, O., Beringer, J., Bernhofer, C., Berveiller, D., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Boike, J., Bolstad, P. V., Bonal, D., Bonnefond, J.-M., Bowling, D. R., Bracho, R., Brodeur, J., Brümmer, C., Buchmann, N., Burban, B., Burns, S. P., Buysse, P., Cale, P., Cavagna, M., Cellier, P., Chen, S., Chini, I., Christensen, T. R., Cleverly, J., Collalti, A., Consalvo, C., Cook, B. D., Cook, D., Coursolle, C., Cremonese, E., Curtis, P. S., D’Andrea, E., da Rocha, H., Dai, X., Davis, K. J., Cinti, B. D., Grandcourt, A. de, Ligne, A. D., De Oliveira, R. C., Delpierre, N., Desai, A. R., Di Bella, C. M., Tommasi, P. di, Dolman, H., Domingo, F., Dong, G., Dore, S., Duce, P., Dufrêne, E., Dunn, A., Dušek, J., Eamus, D., Eichelmann, U., ElKhidir, H. A. M., Eugster, W., Ewenz, C. M., Ewers, B., Famulari, D., Fares, S., Feigenwinter, I., Feitz, A., Fensholt, R., Filippa, G., Fischer, M., Frank, J., Galvagno, M., et al.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, Sci Data, 7, 225, https://doi.org/10.1038/s41597-020-0534-3, 2020.

Peltier, D. M. P., Guo, J., Nguyen, P., Bangs, M., Wilson, M., Samuels-Crow, K., Yocom, L. L., Liu, Y., Fell, M. K., Shaw, J. D., Auty, D., Schwalm, C., Anderegg, W. R. L., Koch, G. W., Litvak, M. E., and Ogle, K.: Temperature memory and non-structural carbohydrates mediate legacies of a hot drought in trees across the southwestern US, Tree Physiology, https://doi.org/10.1093/treephys/tpab091, 2021.

Piao, S., Zhang, X., Chen, A., Liu, Q., Lian, X., Wang, X., Peng, S., and Wu, X.: The impacts of climate extremes on the terrestrial carbon cycle: A review, Sci. China Earth Sci., 62, 1551–1563, https://doi.org/10.1007/s11430-018-9363-5, 2019.

Poyatos, R., Granda, V., Flo, V., Adams, M. A., Adorján, B., Aguadé, D., Aidar, M. P. M., Allen, S., Alvarado-Barrientos, M. S., Anderson-Teixeira, K. J., Aparecido, L. M., Arain, M. A., Aranda, I., Asbjornsen, H., Baxter, R., Beamesderfer, E., Berry, Z. C., Berveiller, D., Blakely, B., Boggs, J., Bohrer, G., Bolstad, P. V., Bonal, D., Bracho, R., Brito, P., Brodeur, J., Casanoves, F., Chave, J., Chen, H., Cisneros, C., Clark, K., Cremonese, E., Dang, H., David, J. S., David, T. S., Delpierre, N., Desai, A. R., Do, F. C., Dohnal, M., Domec, J.-C., Dzikiti, S., Edgar, C., Eichstaedt, R., El-Madany, T. S., Elbers, J., Eller, C. B., Euskirchen, E. S., Ewers, B., Fonti, P., Forner, A., Forrester, D. I., Freitas, H. C., Galvagno, M., Garcia-Tejera, O., Ghimire, C. P., Gimeno, T. E., Grace, J., Granier, A., Griebel, A., Guangyu, Y., Gush, M. B., Hanson, P. J., Hasselquist, N. J., Heinrich, I., Hernandez-Santana, V., Herrmann, V., Hölttä, T., Holwerda, F., Irvine, J., Isarangkool Na Ayutthaya, S., Jarvis, P. G., Jochheim, H., Joly, C. A., Kaplick, J., Kim, H. S., Klemedtsson, L., Kropp, H., Lagergren, F., Lane, P., Lang, P., Lapenas, A., Lechuga, V., Lee, M., Leuschner, C., Limousin, J.-M., Linares, J. C., Linderson, M.-L., Lindroth, A., Llorens, P., López-Bernal, Á., Loranty, M. M., Lüttschwager, D., Macinnis-Ng, C., Maréchaux, I., Martin, T. A., Matheny, A., McDowell, N., McMahon, S., Meir, P., et al.: Global transpiration data from sap flow measurements: the SAPFLUXNET database, 13, 2607–2649, https://doi.org/10.5194/essd-13-2607-2021, 2021.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, 11, 1424–1439, https://doi.org/10.1111/j.1365-2486.2005.001002.x, 2005.

Reichstein, M., Ciais, P., Papale, D., Valentini, R., Running, S., Viovy, N., Cramer, W., Granier, A., Ogée, J., Allard, V., Aubinet, M., Bernhofer, C., Buchmann, N., Carrara, A., Grünwald, T., Heimann, M., Heinesch, B., Knohl, A., Kutsch, W., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Pilegaard, K., Pumpanen, J., Rambal, S., Schaphoff, S., Seufert, G., Soussana, J.-F., Sanz, M.-J., Vesala, T., and Zhao, M.: Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modelling analysis, 13, 634–651, https://doi.org/10.1111/j.1365-2486.2006.01224.x, 2007.

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, 500, 287–295, https://doi.org/10.1038/nature12350, 2013.

Rocha, A. V. and Goulden, M. L.: Drought legacies influence the long-term carbon balance of a freshwater marsh, 115, https://doi.org/10.1029/2009JG001215, 2010.

Ryo, M. and Rillig, M. C.: Statistically reinforced machine learning for nonlinear patterns and variable interactions, 8, e01976, https://doi.org/10.1002/ecs2.1976, 2017.

Schwalm, C. R., Williams, C. A., Schaefer, K., Arneth, A., Bonal, D., Buchmann, N., Chen, J., Law, B. E., Lindroth, A., Luyssaert, S., Reichstein, M., and Richardson, A. D.: Assimilation exceeds respiration sensitivity to drought: A FLUXNET synthesis, 16, 657–670, https://doi.org/10.1111/j.1365-2486.2009.01991.x, 2010.

Sun, S., Sun, G., Caldwell, P., McNulty, S., Cohen, E., Xiao, J., and Zhang, Y.: Drought impacts on ecosystem functions of the U.S. National Forests and Grasslands: Part II assessment results and management implications, Forest Ecology and Management, 353, 269–279, https://doi.org/10.1016/j.foreco.2015.04.002, 2015.

Tamrakar, R., Rayment, M. B., Moyano, F., Mund, M., and Knohl, A.: Implications of structural diversity for seasonal and annual carbon dioxide fluxes in two temperate deciduous forests, Agricultural and Forest Meteorology, 263, 465–476, https://doi.org/10.1016/j.agrformet.2018.08.027, 2018.

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

Walther, S., Besnard, S., Nelson, J. A., El-Madany, T. S., Migliavacca, M., Weber, U., Ermida, S. L., Brümmer, C., Schrader, F., Prokushkin, A. S., Panov, A. V., and Jung, M.: Technical note: A view from space on global flux towers by MODIS and Landsat: The FluxnetEO dataset, Biogeochemistry: Land, https://doi.org/10.5194/bg-2021-314, 2021.

Wang, B., Chen, T., Li, C., Xu, G., Wu, G., and Liu, G.: Discrepancy in growth resilience to drought among different stand-aged forests declines going from a semi-humid region to an arid region, Forest Ecology and Management, 511, 120135, https://doi.org/10.1016/j.foreco.2022.120135, 2022.

Warm Winter 2020 Team, & ICOS Ecosystem Thematic Centre. (2022). Warm Winter 2020 ecosystem eddy covariance flux product for 73 stations in FLUXNET-Archive format—release 2022-1 (Version 1.0). ICOS Carbon Portal. https://doi.org/10.18160/2G60-ZHAK.

Werner, C., Meredith, L. K., Ladd, S. N., Ingrisch, J., Kübert, A., van Haren, J., Bahn, M., Bailey, K., Bamberger, I., Beyer, M., Blomdahl, D., Byron, J., Daber, E., Deleeuw, J., Dippold, M. A., Fudyma, J., Gil-Loaiza, J., Honeker, L. K., Hu, J., Huang, J., Klüpfel, T., Krechmer, J., Kreuzwieser, J., Kühnhammer, K., Lehmann, M. M., Meeran, K., Misztal, P. K., Ng, W.-R., Pfannerstill, E., Pugliese, G., Purser, G., Roscioli, J., Shi, L., Tfaily, M., and Williams, J.: Ecosystem fluxes during drought and recovery in an experimental forest, 374, 1514–1518, https://doi.org/10.1126/science.abj6789, 2021.

Wu, X., Liu, H., Li, X., Ciais, P., Babst, F., Guo, W., Zhang, C., Magliulo, V., Pavelka, M., Liu, S., Huang, Y., Wang, P., Shi, C., and Ma, Y.: Differentiating drought legacy effects on vegetation growth over the temperate Northern Hemisphere, 24, 504–516, https://doi.org/10.1111/gcb.13920, 2018.

Zhang, S., Yang, Y., Wu, X., Li, X., and Shi, F.: Postdrought Recovery Time Across Global Terrestrial Ecosystems, Journal of Geophysical Research: Biogeosciences, 126, e2020JG005699, 2021.

Zscheischler, J., Mahecha, M. D., Buttlar, J. von, 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, Environ. Res. Lett., 9, 035001, https://doi.org/10.1088/1748-9326/9/3/035001, 2014.