Increased vulnerability of European ecosystems to two compound dry and hot summers in 2018 and 2019

Ana Bastos¹, René Orth¹, Markus Reichstein¹, Philippe Ciais², Nicolas Viovy², Sönke Zaehle¹, Peter Anthoni³, Almut Arneth³, Pierre Gentine⁴,⁵, Emilie Joetzjer⁶, Sebastian Lienert⁷, Tammas Loughran⁸, Patrick C. McGuire⁹, Sungmin O¹, Julia Pongratz⁷, and Stephen Sitch¹⁰

¹Max-Planck Institute for Biogeochemistry, Hans-Knöll Str. 10, 07745, Jena, Germany
²Laboratoire des Sciences du Climat et de l’Environnement, Gif-sur-Yvette, France
³KIT, Atmospheric Environmental Research, Garmisch-Partenkirchen, Germany
⁴Dept. of Earth and Environmental Engineering, Columbia University, NY 10027, USA
⁵Earth Institute and Data Science Institute, Columbia University, NY 10027, USA
⁶CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
⁷Climate and Environmental Physics, Physics Institute and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
⁸Ludwig-Maximilian University, Geography Dept., Luisenstr. 37, 80333, Munich, Germany
⁹Department of Meteorology, Department of Geography Environmental Science, and National Centre for Atmospheric Science, University of Reading, Earley Gate, RG66BB Reading, UK
¹⁰College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4RJ, UK

Correspondence: Ana Bastos (abastos@bgc-jena.mpg.de)

Abstract. In 2018 and 2019, central Europe was struck affected by two consecutive extreme dry and hot summers (DH2018 and DH2019). The DH2018 had severe impacts on ecosystems and likely affected vegetation activity in the subsequent year, for example through depletion of carbon reserves or damage from drought. Such legacies from drought and heat stress can further increase vegetation susceptibility to additional hazards. Temporally compound extremes such as DH2018 and DH2019, DH18 and DH19 can, therefore, result in an amplification of impacts by preconditioning effects of past disturbance legacies.

Here, we evaluate how these two consecutive extreme summers impacted ecosystems in central Europe and how the vegetation responses to the first compound event (DH2018DH18) modulated the impacts of the second (DH2019DH19). To quantify the modulating role of vegetation responses to the impacts of changes in vegetation vulnerability to each compound event, we first train a set of statistical models for the period 2001-2017, which are then used to predict the impacts of DH2018 and DH2019 on Enhanced Vegetation Index (EVI) anomalies from MODIS. These estimates can be seen as the correspond to expected EVI anomalies, had the impacts of DH2018 and DH2019 been consistent with in DH18 and DH19 based on past sensitivity to climate. These can then be Large departures from the predicted values can indicate changes in vulnerability to dry and hot conditions, and used to identify modulating effects by vegetation activity and composition or other environmental factors such as elevated CO₂ or warming trends on observed impacts.

We find two regions in which the impacts of the two DH events were significantly stronger than those expected based on previous climate–vegetation relationships. One region, largely dominated by grasslands and crops, showed much stronger
impacts than expected in both DH events due to an amplification of their sensitivity to heat and drought, possibly linked to changing background CO$_2$ and temperature conditions. A second region, dominated by forests and grasslands, showed browning from DH2018 to DH2019. DH18 to DH19, even though dry and hot conditions were partly alleviated in 2019. This browning trajectory was mainly explained by the preconditioning role of DH2018-DH18 to the observed response to DH2019 DH19 through legacy effects, and possibly by increased susceptibility to biotic disturbances, which are also promoted by warm conditions.

Dry and hot summers are expected to become more frequent in the coming decades posing a major threat to the stability of European forests. We show that state-of-the-art process based models miss these legacy effects. These gaps may result in an overestimation of the resilience and stability of temperate ecosystems in future model projections.

Copyright statement. No copyrights

1 Introduction

Extreme dry and hot summers in western and central Europe have become more frequent over the past decades (Coumou and Rahmstorf, 2012; Seneviratne et al., 2014), a trend that is expected to continue as global mean temperatures rise (Barriopedro et al., 2011). Hot extremes in Europe are promoted by changes in atmospheric circulation (Coumou et al., 2015; Drouard et al., 2019) and amplified by strong feedbacks between the land-surface and the atmosphere, being therefore also associated with severe droughts (Miralles et al., 2014; Samaniego et al., 2018), i.e. compound hot and dry and hot events (DH).

In Europe, DH events have usually strong negative impacts on ecosystems, such as reduced ecosystem productivity (Ciais et al., 2005; Bastos et al., 2020b). After severe drought and heat stress, plant recovery can be lagged, for example due to leaf shedding reduced growth, or non-reversible losses in hydraulic conductance or carbon reserve depletion (Ruehr et al., 2019). This, in turn may increase vulnerability if another DH to another DH, if it occurs before complete recovery. A second DH event. Repeated droughts have been linked to increased forest vulnerability in the northern mid-latitudes, although with variable responses (Anderegg et al., 2020). Impaired functioning during the recovery period can additionally increase the hazard of other disturbances such as fires or insect outbreaks (Rouault et al., 2006; ?, ?). More frequent extremes may, therefore, threaten ecosystem stability by compounding multiple hazards and concurrent and lagged effects from highly impactful DH events subsequent disturbances, e.g. insect outbreaks (Rouault et al., 2006). However, reductions in leaf area, increases in root allocation (McDowell et al., 2008) or reduced growth, by reducing evaporative tissue and enhancing water uptake capacity, could also confer an advantage to subsequent droughts (Gessler et al., 2020). It remains unclear whether the increased vulnerability to a subsequent drought can be explained by compounding hazards (e.g. accumulated water-deficits or compound heat) or modulating effects due to vegetation responses to the first event.

At European scale
In Europe, the summer in-of 2018 was the hottest since 1500 (Sousa et al., 2020) while at the same time leading to and associated with an unprecedented area affected by drought (Albergel et al., 2019; Bastos et al., 2020a). In 2019, central Europe was stricken by This DH event resulted in decreases in ecosystem productivity by up to 50% in central Europe (Bastos et al., 2020a; Buras et al., 2019) and crop yield losses (Beillouin et al., 2020). Part of the central European region affected by DH18 registered another extremely hot and dry summer (Boergens et al., 2020; Sousa et al., 2020). In 2019 (Boergens et al., 2020; Sousa et al., 2020). In 2019, central Europe was stricken by this DH event resulted in decreases in ecosystem productivity by up to 50% in central Europe (Bastos et al., 2020a; Buras et al., 2019) and crop yield losses (Beillouin et al., 2020). Part of the central European region affected by DH18 registered another extremely hot and dry summer (Boergens et al., 2020; Sousa et al., 2020).

From a hydrometeorological perspective, the dry and hot summers in 2018 and 2019 (DH2018 and DH2019 and DH18, respectively) could be considered individually as two compound events in that both high temperatures and strong drought conditions were observed (Zscheischler and Fischer, 2020). Taken together, they constitute a temporally compound event (Zscheischler et al., 2020). For example, Boergens et al. (2020) have shown that while soil-moisture deficits in summer 2019 were not as pronounced as in 2018, total water storage was lower in 2019 due to the water storage deficit resulting from the 2018 event.

From an ecological perspective, these events are more complex as they constitute a combination of temporally (two consecutive extreme summers) and preconditioning (changes in ecosystem functioning) compound events. The DH2018 was one of the strongest in the past decades, leading to decreases in ecosystem productivity by up to 50% in central Europe (Bastos et al., 2020a; Buras et al., 2019). Compared to previous extreme summers, DH2018 was associated with increased sensitivity to temperature (Bastos et al., 2020b). This can be a sign of increased plant vulnerability to “hotter” droughts (Allen et al., 2015), or of detrimental effects from increased growth in response to the previous sunny and warm spring (Bastos et al., 2020a). Given the unprecedented magnitude of DH2018 and its severe impacts, it likely imposed such legacies throughout the subsequent year, for example through reduced growth or carbon reserve depletion that may have increased vulnerability to yet another dry and hot summer—DH18, it is likely that at least some ecosystems had not yet fully recovered in 2019.

Repeated droughts have been linked to increased forest vulnerability in the northern mid-latitudes, although with variable responses (Anderegg et al., 2020). It remains unclear whether the increased vulnerability to a subsequent drought can be explained by physical drivers (e.g., accumulated water deficits or compound heat) or by modulating effects by vegetation responses during and following the first drought. Moreover, responses to drought are expected to be modulated by long-term increases in CO₂, but the direction of this effect is not clear.

Environmental changes; water-savings from reduced stomatal conductance should attenuate drought stress (Peters et al., 2018), but concurrent decrease in evapotranspiration cooling along with “hotter droughts” may amplify heat stress (Obermeier et al., 2018) (Allen et al., 2015; Obermeier et al., 2018) (Fig. 1).

Separating the modulating effects controlled by vegetation responses to global change or by legacies from past disturbances (Kannenberg et al., 2020) and seasonal legacy effects (Buermann et al., 2018) in observations is problematic as it requires considering the compounding effects of multiple drivers (e.g., compound heatwave and drought) and separating the role of seasonal and inter-annual legacies both in physical variables (e.g., soil-moisture depletion) and in vegetation vulnerability to those drivers. Such effects have been separated for seasonal legacies using model experiments and...
designing counter-factual scenarios ((Lian et al., 2020; Bastos et al., 2020a)) to force process-based models, as recently done to evaluate seasonal legacy effects of hot and dry springs (Lian et al., 2020; Bastos et al., 2020a). However, it has been argued that Earth System models have been reported to fail at modelling woody biomass trajectories following droughts (Anderegg et al., 2015), but so that they might miss inter-annual legacy effects from DH events, although no simulations designed to isolate the individual impact of drought over subsequent years have been conducted. The simulations by land-surface models (LSMs) in Bastos et al. (2020a) separated the individual impact of DH2018 on carbon and water fluxes by using an additional factorial simulation. When extended to 2019, these simulations allow evaluating how models simulate inter-annual legacy effects of DH2018 and vulnerability to consecutive droughts (DH2018 and DH2019) performed. Alternatively, statistical models can be used to separate such effects based on observational data (Chan et al., 2021).

The occurrence of two consecutive hot and dry summers is uncommon in central Europe but may become more likely in the coming decades Barriopedro et al. (2011); Boergens et al. (2020). Therefore, the DH2018 and DH2019 can provide insights on how resilient might European ecosystems be.

Here, we focus on DH18 and DH19 to gain insights on the resilience of European ecosystems to repeated hot and dry summers in the coming decades. Here, we attempt to answer this question by using both observations and models to: (i) evaluate the vulnerability of ecosystems to DH2018 and DH2019; (ii) detect fingerprints of global change and disturbance legacy effects in ecosystem vulnerability to DH2018 and DH2019; (iii) assess the ability of state-of-the-art LSMs to simulate the impacts on ecosystem of these two events.

Our results show an increasingly important contribution of vegetation condition, evaluated here by greenness and productivity, in preconditioning the response to both DH2018 and DH2019, suggesting that feedbacks between climate extremes and ecosystem functioning may increased vulnerability to climate change. Using both remote-sensing data and an update of the simulations by Bastos et al. (2020a), we first evaluate whether there are signs of increased vegetation vulnerability to repeated dry and hot summers of 2018 and 2019, and try to identify possible modulating effects. We then compare observation-based
results to updated simulations by state-of-the-art land-surface models and dynamic global vegetation models (for simplicity referred to as LSMs) designed to isolate the impacts of DH18 and legacy effects (Bastos et al., 2020a).

2 Data

2.1 Climate variables

For In ecological studies, drought is better characterized by soil-moisture anomalies i.e. agricultural drought (Sherriff et al., 2011; Seneviratne et al., 2012; Samaniego et al., 2018) than atmospheric drought indices. We therefore base our drought assessment on two complementary soil-moisture datasets. The first is the observation-based soil moisture data obtained from SoMo.ml (Sungmin and Orth, 2021), used as reference in this study, and the second, for comparison with SoMo.ml, is give given by ERA5 volumetric soil-water content (Hersbach et al., 2020).

The SoMo.ml data are generated using a Long Short-Term Memory neural network model trained with meteorological forcing from ERA5 and land surface characteristics as inputs and global in-situ soil moisture measurements (Dorigo et al., 2011; Zeri, 2020) as target variables. The data cover the period 2000-2019 soil-moisture in the first 50cm of the soil and are available at 0.25°lat/lon resolution and daily time-steps for the period 2000–2019. We remapped the fields to the finer resolution of the MODIS grid and aggregated the data to monthly means. We then subtracted the mean seasonal cycle and long-term linear trend, and divided by the corresponding standard deviation to obtain standardized soil-moisture anomalies ($SM_{anom}$).

We use monthly temperature and volumetric soil-water content (layers 1 and 2, top 28cm) from the ECMWF ERA5 Reanalysis. ERA5 uses an improved land surface data assimilation system that makes use of remotely sensed and in situ observations, and shows improved skill compared to previous reanalyses (Hersbach et al., 2020) and good temporal agreement with a range of global soil moisture networks (?). Data were obtained from the Copernicus Climate Change Service at 0.25°lat/lon resolution (Hersbach et al., 2020) at monthly time-steps and selected for the period 2000-2019 (common with SoMo.ml) and remapped to the finer resolution of the MODIS grid using conservative regridding/remapping. Standardized anomalies were calculated by removing the mean seasonal cycle and long-term linear trend and then dividing by the corresponding standard deviation of as described for $SM_{anom}$ for ERA5 temperature and soil-moisture fields ($T_{anom}, SM_{anom}^{ERA5}$). Soil-moisture anomalies from ERA5 in layers 1–2 (top 28cm) are used for comparison of drought conditions with those estimated by SoMo.ml ($SM_{anom}$ for simplicity), although it is worth noting that, although the two datasets are not fully independent.

2.2 Vegetation and soil data

We used the 16-day Enhanced vegetation Index (EVI) from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor from the MOD13C1 CMG product. The MOD13C1 CMG follows a strict quality control and uses a gap-filling scheme to provide continuous cloud-free spatial composites from 1km data (Didan et al., 2015) projected on a 0.05°lat/lon grid covering (Didan et al., 2015), and were selected for the period 2001–2019. MOD13C1 CMG is, therefore, a higher quality product.
especially suitable for spatiotemporal analysis and for comparison with LSMs as intended here. Standardized EVI anomalies ($EVI_{anom}$) were calculated by removing the mean seasonal cycle and long-term linear trend, and were then scaled by the corresponding pixel-level temporal standard deviation following the same approach as for climate variables. The standardization allows comparing the relative magnitude of anomalies for pixels with distinct temporal variability patterns and with vegetation productivity simulated by LSMs, which have different physical units.

For land-cover distribution we used in 2018 from the ESA Climate Change Initiative land-cover map (Kirches et al., 2014) (CCI-LC) for the year 2018. The data are originally provided in land-cover classes at 300m spatial resolution and were converted to fractional cover at 0.05°lat/lon resolution for forest, grassland, crop classes using the LC-CCI user–tool.

Isohydricity—We used isohydricity fields from global satellite measurements from Konings et al. (Konings et al., 2017) are available at 1°lat/lon resolution from the ESA Climate Change Initiative land-cover product (Konings et al., 2017). Anisohydric plants (low isohydricity) show weak regulation of stomatal opening, and prioritize carbon assimilation over water savings during droughts. High isohydric plants show strong stomatal regulation of productivity and thereby preserve water at the cost of carbon assimilation during drought.

We use soil Available Water Capacity (AWC) from Ballabio et al. (2016) and Panagos et al. (2012), which used the Land Use and Cover Area frame Statistical survey (LUCAS) topsoil database to map soil properties at continental scale.

The data are provided by the European Soil Data Centre (ESDAC) (esdac.jrc.ec.europa.eu).

2.3 Land-surface Outputs from land-surface and global dynamic vegetation models

Standardized anomalies of gross primary productivity ($GPP_{anom}$) and soil-moisture ($SM_{anom}$) were estimated by the mean of seven land-surface models and dynamic global vegetation models (for simplicity referred to as LSMs) between 1979–2019 from an extension of Bastos et al. (2020a) simulations: a baseline simulation for comparison with observations and a factorial simulation to quantify the individual impact of summer 2018 and its legacy effects, when compared to the reference simulation.

For the reference simulation, the models were forced with observed CO$_2$ concentration from NOAA/ESRL and changing climate between 1979—A detailed description of the models used and 2019 from ERA5 and fixed land-cover map from 2010 from LULCv2 (Hurtt et al., 2011). An additional simulation was run where the models were forced with changing climate, except June–August 2018, where climatological summer climate conditions were used to force the models as described in Bastos et al. (2020a). This simulation allows evaluating the direct impact of DH2018 and its legacy effects. The model simulations were run for most models at 0.25°spatial resolution for the European domain (32–75°N and 11–65°E), following a spin up to equilibrate carbon pools. For more details on the simulation protocol, we refer to (Bastos et al., 2020a). These simulations were extended here to 2019, using climate fields from ERA5. Here we analyse gross primary productivity (GPP) and simulated soil moisture, the simulation protocol is provided in the Appendix A.

The seven LSMs followed the protocol and extended the simulations in Bastos et al. (2020a) up to 2019. These models are: ISBA-CTRIP (Joetzjer et al., 2015), JSBACH (Mauritsen et al., 2018), LPJ-GUESS (Smith et al., 2014), LPX-Bern (Lienert and Joos, 2018), OCN (Zaeble et al., 2010), ORCHIDEE (Krinner et al., 2005) and SDGVM (Walker et al., 2017).
First, all model outputs were remapped to a common 0.25 degree grid, and the multi-model ensemble mean was calculated for the common period with MODIS (2001–2019). The variables were then deseasonalized, detrended and standardized as done for the other variables in the study.

3 Methods

3.1 Drought characterization

We use the observation-based SoMo.ml as a reference dataset to define agricultural drought conditions. Regions with average \( SM_{anom} \) below \(-1\sigma\) (Seneviratne et al., 2012) during summer (JJA) are considered drought-affected areas during the DH events. Then, a regional domain affected by both DH2018 and DH2019 events is selected to evaluate the impacts of two consecutive DH events. Within this region most pixels had negative \( SM_{anom} \) and the majority registered \( SM_{anom} < -1.5\sigma \), but they differ in the magnitude of agricultural drought in 2019. This allows comparing responses across pixels for different combinations of stress between DH2018 and DH2019. Since we are interested in evaluating how recovery from DH18 affected impacts of DH19, we limit our analysis to pixels with negative \( EVI_{anom} \) in DH18.

3.2 Compound DH2018-DH18 and DH2019-DH19 events

3.2.1 DH2018-DH18 and DH2019-DH19 impact characterization

To characterize different response “types” to DH2018 and DH2019, we group pixels in an unsupervised way (K-means clustering) based on the EVI impacts during the two extreme summers. Using an unsupervised method allows avoiding making assumptions about the magnitude of impacts or the trajectory between DH2018 and DH2019 when grouping pixels. For this, we applied a K-means cluster analysis (Hamerly and Elkan, 2003) using two features, corresponding to the \( EVI_{anom} \) fields in DH2018 and DH2019, for pixels with negative \( EVI_{anom} \) in DH2018 and DH19. Four clusters captured the most significant differences in the impacts of DH2018 and corresponding DH2018 and DH19 responses: moderate/strong DH2018-DH18 impacts and moderate/strong impacts by DH2019-DH19. These clusters were then used to evaluate how LSMs simulate the summer \( GPP_{anom} \) and \( SM_{anom} \).

3.2.2 Detecting increased vulnerability to drought and heat stress

We propose that the two events can be considered a combination of temporally and preconditioning compound events (Fig. 1): a sequence of two DH events, whose impacts may be preconditioned by ecosystem vulnerability to DH, especially in the case of DH2019. The DH impacts and ecosystem vulnerability, i.e. the propensity to be negatively impacted by a given event, are DH19. Vulnerability to DH is defined as the
impact of the physical hazard (hot and dry conditions) on vegetation and assessed by remotely-sensed EVI and modelled GPP anomalies.

Conceptual description of the compound DH2018 and DH2019 events. Dry and hot conditions in both summers were a result of compounding atmospheric drivers (synoptic patterns, preceding climate anomalies, land-atmosphere interactions). The DH2018 impacts were modulated by seasonal legacy effects in ecosystem functioning from a sunny and warm spring. We hypothesise that legacies from the DH2018 event also contributed to modulate the response to DH2019. These impacts can be further modulated by long-term changes in ecosystem vulnerability to drought and heat stress due to anthropogenic climate change and increasing CO₂.

The difference between the reference and factorial simulations by LSMS allows separating the modulating effects of DH2018 DH18 legacies to the DH2019 DH19 impacts (dashed arrow in Fig. 1). Separating the legacies in observations is more challenging, because the EVI signal at any time-step includes both signals. We hypothesize that preconditioning effects from legacies from past disturbance (modulating DH2019) but also from due to past disturbance legacies (modulating DH19) and global change (modulating DH2018 and DH2019 DH18 and DH19) should be detectable by changes in ecosystem sensitivity to similar hazards. Increased vulnerability can thus be detected if corresponds thus to EVI\textsubscript{anom} values are lower (more negative or less positive) than those expected for a given drought or temperature anomaly based on past sensitivities. Inversely, acclimation to drought could increased resistance would result in EVI\textsubscript{anom} being less negative or more positive than expected for a given SM\textsubscript{anom}.

We test whether increased assess whether changes in the sensitivity to climate anomalies is detected for DH2018 and DH2019 in DH18 and DH19 using a statistical modelling approach to predict EVI\textsubscript{anom} in DH18 and DH19 based on 2001–2017 climate–vegetation relationships. We do this in two steps: a linear case, and another including non-linear and first by fitting a linear regression model for mean EVI\textsubscript{anom} in each cluster, and then, for more detailed insights, by fitting a random forest model at pixel scale, in which we include potential seasonal legacy effects. In both steps, we characterize EVI climate relationships for the period 2001–2017 and predict the EVI anomalies in DH2018 and DH2019. In the cases, the training period includes other DH events (Ciais et al., 2005; Orth et al., 2016), with similar climate anomalies, particularly 2003, thereby reducing the risk of attempting to predict EVI\textsubscript{anom} based on “unseen” climatic conditions.

On a first step, we estimate the EVI climate relationships for each cluster by univariate linear regression model of using for the spatially-averaged variables within each cluster, we fit the following models:

\[
\text{EVI}_{\text{anom}}^{C_i} = b_0 + b_1 \times \text{VAR}_{\text{anom}}^{C_i}
\]  

(1)

Where EVI\textsubscript{anom} and VAR\textsubscript{anom} correspond to the spatial average values of EVI\textsubscript{anom} and climate variable (growing-season SM\textsubscript{anom} or T\textsubscript{anom} as predictors. Because impacts on EVI could), respectively, b₀, b₁ are the coefficients of each linear regression trained on 2001–2017 values. Each model is then used to estimate DH18 and DH19 EVI\textsubscript{anom}. Negative
model residuals (observations minus predictions) can indicate increased vulnerability, while positive residuals can be a sign of increased resistance.

However, departures from a linear model could also result from non-linear interactions between soil-moisture and temperature or from legacy effects from spring (Bastos et al., 2020a; Lian et al., 2020). To account for such effects and evaluate potential spatial asymmetries in the departures from long-term climate–vegetation relationships, we fit a random-forest (RF) regression model using as target variable $EVI_{anom}$ in each pixel ($i$) from 2001–2017, and the corresponding $SM_{anom}$ and $T_{anom}$ in spring (MAM) and summer $SM_{anom}$ or $T_{anom}$. We in summer (JJA) as predictors:

$$EVI_{anom-i} = RF(T^{spr}_{anom-i}, SM^{spr}_{anom-i}, T^{sm}_{anom-i}, SM^{sm}_{anom-i})$$  \hspace{1cm} (2)

To reduce the risk of over-fitting due to the small sample size (17 years) and large number of predictors (4), we fit the RF model on a per-pixel basis but on 3x3 moving windows, in order to increase the sample size centered around each pixel (i.e. $17 \times 9^2$) for each regression and reduce overfitting. To further control for possible over-fitting and poor predictive skill, we assess the model performance outside of the training samples. We calculate the RF model by calculating the out of bag scores in addition to the training sample scores. The importance of each predictor is then estimated by the Shapley additive explanation values (Lundberg and Lee, 2017). We then predict $EVI_{anom}$ in DH18 and DH19 using the respective anomalies in $T^{spr}_{anom}$, $SM^{spr}_{anom}$, $T^{sm}_{anom}$, $SM^{sm}_{anom}$.

The $EVI_{anom}$ predicted by the RF model for DH2018 and DH2019 DH18 and DH19 correspond to the expected DH impacts if no changes ecosystem vulnerability to drought and heat were present, i.e. considering only links between atmospheric drivers, from past relationships between the hazards and impacts in Fig. 1. It should be noted that the training period includes other DH events, particularly 2003 and 2015 (Ciais et al., 2005; Orth et al., 2016) thereby reducing the risk of attempting to predict anomalies out of the training sample. The As for the linear case, the difference between the RF model predictions and the actual $EVI_{anom}$ (model residuals) provides thus an estimate of the contribution of an indication of changes in ecosystem vulnerability to the DH2018 and DH2019-DH18 and DH19 impacts.

For comparison with LSM simulations, the $EVI_{anom}$, clusters were remapped to 0.25 degree by largest area fraction calculation, and subsequently $GPP_{anom}$ and $SM_{anom}$ model ensemble means for each cluster were compared with corresponding $EVI_{anom}$ and ERA5 $SM_{anom}$. We first evaluate the linear relationships between the averaged $GPP_{anom}$ for each cluster and the corresponding climate anomalies. for comparison with $EVI_{anom}$. Then, we estimate the legacy effects from DH2018 to 2019-DH18 on $GPP_{anom}$ during 2019 based on the difference between the reference and factorial LSM simulations.

### 3.2.3 Modulating effects

To understand how land-cover can contribute to modulate the impacts of DH2018 and DH2019-DH18 and DH19 we analyse the land-cover composition of each cluster. Given that central Europe is characterized mostly by mixed pixels, we do this by calculating the a very heterogeneous landscape, we calculate land-cover selectivity in each cluster for forests, natural grasslands and croplands. Selectivity is defined as the difference between the probability a given land-cover class being present within
a cluster compared to its overall presence in the whole region. The probabilities are calculated by fitting a kernel-distribution function to the fractional cover fields for the whole region and for separate clusters. Positive (negative) selectivity means that a given land-cover type is more (less) common likely to be found in a given cluster than compared to its overall presence in the region.

Finally, we try to explain the changes in ecosystem vulnerability, which are given by departures of $EVI_{anom}$ residuals ($EVI_{anom}$, observed minus predicted) from the range of residuals in the training period. To do this, we For other modulating effects we evaluate how the spatial distribution of $EVI_{anom}$ residuals for DH2018 and DH2019–DH18 and DH19 relates to climatic and ecological variables: $SM_{anom}$ and $T_{anom}$ in spring and summer, number of dry months in the year of the DH event and the preceding year (i.e. 2017–2018 for DH2018–DH18, and 2018–2019 for DH2019–DH19), $EVI_{anom}$ in the preceding summer ($EVI_{anom}^{yr-1}$), the number of dry months in a given year and its preceding year (DM), forest, cropland and grassland cover fractions from CCI-LC, isohydricity (IsoH) and AWC, available water capacity (AWC, related to the maximum amount of water available for plants).

We include some of the drivers used to train the temporal climate-driven RF model to diagnose possible changes in the vulnerability to climate, i.e. the impact is still driven by climate conditions, but vegetation responds more strongly explained by stronger vegetation sensitivity to climate anomalies than in the training period. $EVI_{anom}^{yr-1}$ is used to evaluate the preconditioning role of legacy effects from past extreme summers or disturbances (summer is the peak of the growing season in this region). The number of dry months and AWC (related to the maximum amount of water available for plants) are also included as they may explain non-linear relationships between $SM_{anom}$ and vegetation stress. Isohydricity provides a measure of the degree of stomatal regulation by plants. Since many of these variables have strong spatial covariability (e.g. $T_{anom}$ and $SM_{anom}$, or to some extent tree/grassland cover and IsoH), we evaluate their relationships with $EVI_{anom}$ residuals by calculating the partial rank correlation (Spearman’s $\rho$ ) between each variable, controlling for the others separately. Since these effects might depend on land-cover type, we analyse separately pixels with high and low forest cover.

To further evaluate long-term importance of how inter-annual legacy effects in vegetation activity affect long-term vegetation dynamics, we apply a second temporal RF model to pixel-level $EVI_{anom}$ (Section 3.2.2) where we additionally include with $EVI_{anom}^{yr-1}$ as a predictor of the regression. additional predictor. The model is trained for the period 2002–2017 also on 3×3 moving windows and spring and is then used to predict $EVI_{anom}$ in DH2018 and DH2019–DH18 and DH19. The resulting model residuals were then compared to those of the climate–driven RF model.

4 Results

4.1 DH2018–DH18 and DH2019–DH19 impacts

Following the extreme summer in central Europe in 2018, mild temperatures and strong soil-moisture deficits remained until January 2019, when $SM_{anom}$ returned to normal conditions (Fig. B1, Fig. B2). In central Europe, June 2019 was extremely hot, but July and August 2019 were milder (Fig. B1, (Sousa et al., 2020)), and soil-moisture deficits were became very pronounced in July (Fig. B2). In this region, excepting except April 2019, the months preceding summer were not particularly dry and
were even slightly wetter than average in February, March and May, the latter also colder than average. Therefore, the DH2018 and DH2019, DH18 and DH19 constitute more a sequence of two compound events than a long-single drought. The areas experiencing severe dry and hot conditions in both summers correspond to a region covering central and eastern Europe and southern Sweden. This region is our study domain and indicated by the rectangle in Fig.2).

Spatial patterns of temperature ($T_{\text{anom}}$), soil moisture ($SM_{\text{anom}}$) and EVI ($EVI_{\text{anom}}$) anomalies during summer 2018 (top panel) and summer 2019 (bottom panel) for the study region. The study region corresponds to a domain with dry and hot conditions in both 2018 and 2019 summers (DH2018 and DH2019).

Both DH events led to vegetation browning, though negative $EVI_{\text{anom}}$ were more widespread in DH2018 than DH2019. Within the study region, 79% of the area showing negative $EVI_{\text{anom}}$ in DH2018 ($EVI_{\text{anom}}^{\text{DH2018}}$DH18 $EVI_{\text{anom}}^{\text{DH19}}$) also registered negative $EVI_{\text{anom}}$ in DH2019 ($EVI_{\text{anom}}^{\text{DH2019}}$), although greening can be found in some areas, DH19 ($EVI_{\text{anom}}^{\text{DH19}}$).

In this study, we limit our analysis to pixels negatively impacted by DH2018 and evaluate subsequent responses to DH2019 by grouping pixels based on ($EVI_{\text{anom}}^{\text{DH2018}}$, $EVI_{\text{anom}}^{\text{DH2019}}$) values using unsupervised clustering.

The spatial distribution of the resulting clusters is shown in Fig. 3 (left panel) and the corresponding ($EVI_{\text{anom}}^{\text{DH2018}}$, $EVI_{\text{anom}}^{\text{DH2019}}$) pairs clusters resulting from the unsupervised classification based on ($EVI_{\text{anom}}^{\text{DH2018}}$, $EVI_{\text{anom}}^{\text{DH2019}}$) are shown in the top right panel. For comparison, ($SM_{\text{anom}}^{\text{DH18}}$, $SM_{\text{anom}}^{\text{DH19}}$) pairs are also shown.

Classification of impact groups within the study region in central Europe. The left panel shows the spatial distribution of the four clusters from unsupervised classification of ($EVI_{\text{anom}}^{\text{DH2018}}$, $EVI_{\text{anom}}^{\text{DH2019}}$). The corresponding ($EVI_{\text{anom}}^{\text{DH2018}}$, $EVI_{\text{anom}}^{\text{DH2019}}$) distribution in each cluster are indicated in the top right panel (circles indicate the spatial mean and the lines spatial standard deviation within each cluster). The corresponding distribution of $SM_{\text{anom}}$ and $T_{\text{anom}}$ pairs are shown in the centre right and bottom right panels respectively. The grey line, indicates similar anomalies in the two DH events.

The $T_{\text{anom}}^{\text{DH18}}$, $T_{\text{anom}}^{\text{DH19}}$ (center and bottom right panels). The four clusters aggregate pixels according to different impacts in DH2018 and DH2019. Cluster C1 covers DH18 and DH19. One cluster, covering 20% of the area, includes pixels with moderate impacts in DH2018, DH18 and further browning in DH2019 ($EVI_{\text{anom}}^{\text{DH2019}}$DH19, being therefore referred to as (CDecline) (dark brown, $EVI_{\text{anom}}^{\text{DH19}}$ below the 1:1 line in Fig.3, top right panel). This cluster is associated with mixed cover of forests (10-40%, dominated by needle-leaved) and grasslands (15-60%), (Fig.B3). Cluster C2 (CHighly (high vulnerability, covering 15% of the area) corresponds to pixels experiencing strong impacts in both events and is associated with high grassland and cropland fractions and low forest cover. Pixels with strong impacts in DH2018 and DH18 and weakly negative $EVI_{\text{anom}}^{\text{DH19}}$, i.e. partial recovery in DH2019 (C3 (CRecov, 21% of the area), are mainly dominated by croplands. Finally, a group of pixels shows moderate $EVI_{\text{anom}}^{\text{DH18}}$ and positive $EVI_{\text{anom}}^{\text{DH19}}$ (CGreening), while pixels showing positive $EVI_{\text{anom}}^{\text{DH19}}$ (C4, 44%) correspond, corresponding mostly to mixed forest-grassland pixels (30-65% of forest, dominated by needle-leaved).

All clusters have align along proportional DH18:DH19 values of $SM_{\text{anom}}$ and $T_{\text{anom}}$, with predominantly negative $SM_{\text{anom}}$ and positive $T_{\text{anom}}$ in both DH events, but show but alleviation of soil-moisture deficits and heat stress in DH2019 compared to DH2018, DH19 compared to DH18 (Fig. 3). Clusters align along proportional $SM_{\text{anom}}$ and $T_{\text{anom}}$ in DH2018 vs DH2019.
Figure 2. Spatial patterns of temperature ($T_{anom}$), soil-moisture ($SM_{anom}$) and EVI ($EVI_{anom}$) anomalies during summer 2018 (top panel) and summer 2019 (bottom panel) for the study region. The study region corresponds to a domain with dry and hot conditions in both 2018 and 2019 summers (DH18 and DH19), with overlapping distributions. The two recovery clusters (C3 and C4) correspond to pixels with less severe drought conditions and milder temperatures in DH2019, and C4 (greening) corresponds to pixels where dry and hot conditions in DH2018 were also more moderate. C2 (HighV) corresponds to pixels experiencing drier and hotter anomalies in both summers and shows accordingly stronger impacts. Cluster C4 (Decline), however, shows increasing browning in DH2019 in spite of drought and heat stress alleviation (Fig.3), which suggests. The distributions of climate anomalies for each cluster overlap each other and, in some cases, the 1:1 line, indicating that the intensity of the hazard hazards (temperature, drought) alone cannot account for the resulting impacts.

4.2 Ecosystem vulnerability to DH2018 DH18 and DH2019 DH19
Figure 3. Classification of impact groups within the study region in central Europe. The left panel shows the spatial distribution of the four clusters from unsupervised classification of \((EVI_{DH18}^{DH19}, EVI_{DH19}^{DH19})\) values. The corresponding \((EVI_{DH18}^{DH19}, EVI_{DH19}^{DH19})\) distribution in each cluster are indicated in the top right panel (circles indicate the spatial mean and the lines spatial standard deviation within each cluster). The corresponding distribution of \(SM_{anom}\) and \(T_{anom}\) pairs are shown in the center right and bottom right panels respectively. The grey line indicates similar anomalies in the two DH events. Only pixels with negative \(EVI_{anom}^{DH18}\) are considered.

We evaluate ecosystem vulnerability to the two compound events by comparing \(EVI_{anom}\) in DH2018 and DH2019 with past \(EVI_{anom} - SM_{anom}\) and \(EVI_{anom} - T_{anom}\) relationships (Fig. 4) for each cluster separately.

All clusters show significant positive linear relationships between summer \(EVI_{anom}\) and \(SM_{anom}\) and negative linear relationships with \(T_{anom}\) in 2001–2017 consistent with a general summer water limited regime. (Fig. 4). The relationships include the two extreme summers of 2003 and 2015 which had comparable \(T_{anom}\) and \(SM_{anom}\) to DH2018 and DH2019 DH18 and DH19 in most clusters. The long-term sensitivities estimated are, though, robust even if these summers are excluded.

All three clusters with strong impacts in one event (C1 DH2019, C3 DH2018) or both (C2) show...
The results correspond to a general summer water-limited regime, especially in clusters C_{Decline}, C_{HighV} and C_{PRecov}, which show stronger sensitivities to T_{anom} and S_{M_{anom}} (slopes in Fig. 4) and higher variance explained by both models ($R^2\approx 0.58-0.68$ for $S_{M_{anom}}$ and 0.49-0.55 for $T_{anom}$). For these clusters, $EVI_{anom}$ is below the 95% confidence interval of the long-term linear relationships with both relationships for DH18 (C_{PRecov} and C_{HighV}) and DH19 (C_{Decline} and C_{HighV}). $S_{M_{anom}}$ and $T_{anom}$ for that event. These departures are not likely explained simply by compounding heat and drought impacts, as $S_{M_{anom}}$ and $T_{anom}$ in 2003 are similar, or even stronger, than in DH2018 and DH2019. DH18 and DH19 are generally similar to those of 2003, but DH18 was drier than 2003 in C_{PRecov} and C_{HighV}.

These departures may be related with seasonal legacy effects from the warm spring in DH2018 and or the onset of non-linear responses to heat and drought under long-term changing environmental conditions. To account for these modulating effects, we model long-term (2001–2017) $EVI_{anom}$–climate relationships using spring and summer $S_{M_{anom}}$ and $T_{anom}$ as predictors using random forest regression (see Section 3.2.2). For most pixels, the model is able to explain predict 48–90% (median and maximum out of bag score across pixels) of the pixel-level temporal variability of summer $EVI_{anom}$ in 2001–2017 (Fig.B4) and consistently. Analysis of the variable importance shows that the model estimates summer water limitation and negative legacy effects from spring warming (Fig. B5), consistent with a summer water-limited regime and process-based modeling studies (Bastos et al., 2020a; Lian et al., 2020).

In DH2018 As in the linear case, the RF model estimates less negative or more positive $EVI_{anom}$ in DH18 and DH2019, the model has predominantly negative residuals, i.e. observed $EVI_{anom}$ is more negative or less positive than predicted from past vegetation–climate relationships, as found in the linear case DH19 than observations (Fig. 5). Consistent with the results by the linear models, the residuals are below the range of the training period for the high impact clusters: C1 and C3 in DH2018 and DH2018C_{Decline} and C_{PRecov} in DH19 and DH18, respectively, and C2C_{HighV} in both (Fig. 5, bottom panel). In the DH2019 “greening cluster” (C4)C_{Greening}, residuals are predominantly positive (i.e. observed $EVI_{anom}$ more positive than predicted), but still partly overlap with the range of residuals in the training period (Fig. 5).

We evaluate the role of diverse environmental variables in explaining the spatial distribution of residuals:

(i) the same climatic variables as used to train the RF model, indicating increased ecosystem sensitivity to climate;
(ii) $EVI_{anom}$ in the previous summer to account for inter annual legacy effects;
(iii) isohydricity and land cover composition to evaluate the modulating role of vegetation functioning differences;
(iv) soil available water capacity and number of dry months, which can impose thresholds in water limitation.

To do this, we calculate the Pixels with high tree cover tend to show less negative or more positive residuals than pixels with low tree cover in both DH events (Fig. 6), but in DH19 the range of residuals is larger and includes pixels with strongly negative values. The partial rank correlation of the spatial distribution of $EVI_{anom}$ residuals with respect to the explanatory variables selected (different explanatory variables is shown for pixels with high and low forest cover in Fig. 6). Given the large number of pixels, all correlations are significant except those for croplands. In DH2018, vegetation condition in the previous year’s growing season (i.e. summer 2017, + positive effect), tree cover (+),

In DH18, $T_{anom}$ in spring ($T_{spr_{anom}}$, for high and low tree cover) and summer $S_{M_{anom}}$ ($S_{M_{anom}}$, ) show stronger for high tree cover and + for low tree cover) show the strongest relationships with $EVI_{anom}$ residuals. In DH2019 DH19, $EVI_{anom}$
Figure 4. Departure of $EVI_{anom}$ in DH2018–DH18 and DH2019–DH19 from long-term climate-driven variability. Relationship between $EVI_{anom}$ and $SM_{anom}$ (top panel) and between $EVI_{anom}$ and $SM_{anom}^{-1}$ $T_{anom}$ (bottom panel) for each individual summer between 2001 and 2019 over the study region. The results are shown separately for the four clusters defined in Fig. 3. The black line and shaded areas show the relationship and respective 95% confidence intervals obtained by ordinary least-squares linear regression between $EVI_{anom}$ and the respective climate variable for all years between 2001–2017. Values of $(EVI_{anom}, SM_{anom})$ that deviate from the long-term relationships show increased sensitivity to climate anomalies, which can be a sign of increased vulnerability or degradation trajectories decline. The colors indicate individual years, ranging from 2001 (red) to 2019 (purple) and square markers indicate 2018 and 2019.

$(+)_T^{spr}_{anom}$ are also the most relevant variables, but $EVI_{anom}^{-1}$ shows stronger correlation with $EVI_{anom}$ residuals than in DH2018, and $T^{spr}_{anom}$ has an opposite sign and $T^{sm}_{anom}$ (-) show strong correlations, with consistent sign for both high and low tree cover pixels. DH19 residuals of pixels with high tree cover show strong correlation with $SM_{anom}$ with opposite signs.
Figure 5. Spatial distribution of $EVI_{anom}$ residuals in DH2018-DH18 (top panel) and DH2019-DH19 (central panel) estimated by the temporal RF model trained for 2001–2017 with spring and summer $SM_{anom}$ and $T_{anom}$ as predictors. The corresponding distribution per cluster for each DH event is shown by the boxplots in the bottom panel. The shaded grey envelope indicates the range of residuals in the training period.
Spatial partial correlation (spearman) of \( EVI_{anom} \) residuals with environmental variables in DH2018 and DH2019. The variables considered are: spring and summer \( T_{anom} \) and \( SM_{anom} \) (indicated by superscripts \( spr \) and \( sm \), respectively), \( EVI_{anom} \) in the previous growing season (\( EVI_{yr-1} \)), tree, crop and grassland cover, number of dry months, soil available water capacity (AWC) and plant isohydricity (IsoH). Because of the large number of pixels considered, all correlations are significant \((p-val << 0.01)\), except for cropland cover.

**Figure 6.** Spatial partial correlation (spearman) between \( EVI_{anom} \) residuals and environmental variables in DH18 (top panels) and DH19 (bottom panels), for pixels with high (dark green, top 5% cover fraction) and low (light green, lower 5% cover fraction) tree cover (left panels). The variables considered are: spring and summer \( T_{anom} \) and \( SM_{anom} \) (indicated by superscripts \( spr \) and \( sm \), respectively), \( EVI_{anom} \) in the previous growing season (\( EVI_{yr-1} \)), plant isohydricity (IsoH) and the number of dry months (DM). Because of the large number of pixels considered, all correlations are significant \((p-val << 0.01)\). The right panels show the distribution of residuals for pixels with high and low tree cover.
in spring (+) and summer (−). In DH2019, $T^\text{sm}_{\text{anom}}$ and $SM^\text{sm}_{\text{anom}}$ show stronger relationships with and with AWC (−). In DH19, pixels with low tree cover show negative correlation between IsoH and $EV\text{I}_{\text{anom}}$ residuals as do the variables relating to water availability (AWC, dry months and IsoH), all except $T^\text{sm}_{\text{anom}}$ with opposite sign as in DH2018 residuals.

To test whether the importance of $EV\text{I}^\text{yr−1}_{\text{anom}}$ is particular to the two DH events DH19, or if it may reflect long-term interannual legacy effects of anomalies in vegetation activity, we fit a second temporal RF model where $EV\text{I}^\text{yr−1}_{\text{anom}}$ is used as an additional predictor (Figs. B4 and B6). Including vegetation condition in the previous summer improves the predictive power of the long-term RF model (72–97% out of bag score, compared to 48–90% for the model trained with climate drivers only). However, even though the residuals for DH2018 and DH2019 are comparable to those of the training period are considerably reduced relative to the climate–driven model, while the residuals for the training period are considerably reduced DH18 and DH19 are comparable.

4.3 DH2018–DH18 and DH2019–DH19 impacts simulated by LSMs

The GPP from the LSM multi-model ensemble mean matches well the relative differences in the impact of DH18 between clusters compared to $EV\text{I}_{\text{anom}}$ differences in impacts between clusters in DH18 (Fig. 7, top and middle panels). The temporal evolution of monthly GPP anomalies during the 2018 growing season (April to September, Table 1) also agrees with that of $EV\text{I}_{\text{anom}}$, with correlations between with $EV\text{I}_{\text{anom}}$ of 0.74–0.90. However, 2019 trajectories from GPP simulated by LSMs indicate above average spring and early summer productivity for all clusters, and strong positive Even though the root mean squared error (RMSE) is comparable in the two growing seasons, the correlations of GPP$_{\text{anom}}$ for both C3 and C4 during DH2019. However, correlations between with growing-season $EV\text{I}_{\text{anom}}$ and GPP$_{\text{anom}}$ are much lower in 2019 DH19 (−0.09–0.43).

The disagreement in GPP$_{\text{anom}}$ by LSMs is above-average in spring and early summer 2019 cannot be solely explained by errors in simulated soil moisture anomalies, since simulated SM$_{\text{anom}}$ shows very good agreement with both observation-based SM$_{\text{anom}}$ from SoMo.ml and SM$_{\text{ERA5}}$ (correlations of 0.94–0.98, Table 1). For all clusters, and anomalies in DH19 are either more positive or less negative, compared to $EV\text{I}_{\text{anom}}$.

LSMs simulate a stronger attenuation of drought compared to the observation-based SM$_{\text{anom}}$, though with consistent relative differences in SM$_{\text{anom}}$ between clusters (compare Fig. B7 and Fig. 3). The recovery simulated by LSMs in 2019 can be partly explained by too strong recovery of modelled soil moisture (Fig. B7), but may also result from limited ability of LSMs in simulating changes in ecosystem vulnerability during the two DH events. LSMs simulate well the temporal evolution of SM$_{\text{anom}}$ in the two growing seasons, with high correlation with both SoMo.ml and SM$_{\text{ERA5}}$ (correlations of 0.81–0.98).

The RMSE for simulated SM$_{\text{anom}}$ is generally lower than that of GPP$_{\text{anom}}$.

The sensitivity of GPP$_{\text{anom}}$ from LSMs to summer (JJA) soil moisture anomalies (to simulated SM$_{\text{anom}}$ and to T$_{\text{anom}}$ (Fig. B8) is consistent with that of $EV\text{I}_{\text{anom}}$ in all clusters (Fig. 4). The sensitivity of GPP$_{\text{anom}}$ to temperature is also consistent with that of $EV\text{I}_{\text{anom}}$ for clusters C1 and C2, while for C3 and C4, although for C$_{\text{Preco}}$ and C$_{\text{Greening}}$ LSMs estimate non-significant negative relationships between GPP$_{\text{anom}}$ and T$_{\text{anom}}$. Simulated GPP agrees well with $EV\text{I}_{\text{anom}}$ during the 2018 growing season ($r=0.74–0.90$, Table 1, Fig. 7) and the deviation The deviations of GPP$_{\text{anom}}$ from the linear
Figure 7. Observed and process-based model simulations of 2018/19 impacts. Seasonal evolution of $EVI_{anom}$ (top panel) and standardized GPP anomalies ($GPP_{anom}$, central panel) over the two year period for each cluster (defined in Fig. 3 and shown for LSM grid in Fig. B7). The bottom panel shows the difference between the reference and factorial simulations, and indicates the impacts of DH2018 on $GPP_{anom}$ simulated by models during the event and in the subsequent months until December 2019.
Table 1. Correlation of growing season (April–September) $SM_{anom}$ simulated by LSMs with $SM_{anom}$ from SoMo.ml and ERA5, and of $EVI_{anom}$ with GPP simulated by LSMs.

<table>
<thead>
<tr>
<th>height</th>
<th>$C_1$</th>
<th>$C_{Decline}$</th>
<th>$C_2$</th>
<th>$C_{HighV}$</th>
<th>$C_3$</th>
<th>$C_{Preco}$</th>
<th>$C_4$</th>
<th>$C_{Greening}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SM_{anom}$ gs. 2018</td>
<td>0.98</td>
<td>0.33</td>
<td>0.98</td>
<td>0.66</td>
<td>0.97</td>
<td>0.43</td>
<td>0.97</td>
<td>0.21</td>
</tr>
<tr>
<td>$SM_{anom}$ gs. 2019</td>
<td>0.94</td>
<td>0.63</td>
<td>0.97</td>
<td>0.47</td>
<td>0.98</td>
<td>0.12</td>
<td>0.95</td>
<td>0.77</td>
</tr>
<tr>
<td>height $SM_{anom}^{ERA5}$ gs. 2018</td>
<td>0.98</td>
<td>0.87</td>
<td>0.97</td>
<td>0.56</td>
<td>0.98</td>
<td>0.85</td>
<td>0.87</td>
<td>0.64</td>
</tr>
<tr>
<td>$SM_{anom}^{ERA5}$ gs. 2019</td>
<td>0.94</td>
<td>0.71</td>
<td>0.95</td>
<td>0.72</td>
<td>0.94</td>
<td>0.90</td>
<td>0.91</td>
<td>0.70</td>
</tr>
<tr>
<td>height $EVI_{anom}$ gs. 2018</td>
<td>0.80</td>
<td>1.0</td>
<td>0.90</td>
<td>1.2</td>
<td>0.74</td>
<td>1.2</td>
<td>0.79</td>
<td>0.86</td>
</tr>
<tr>
<td>$EVI_{anom}$ gs. 2019</td>
<td>0.34</td>
<td>1.1</td>
<td>0.43</td>
<td>1.1</td>
<td>0.26</td>
<td>1.1</td>
<td>-0.09</td>
<td>1.1</td>
</tr>
</tbody>
</table>

response for $C_2$ and $C_3$ in DH2018. $C_{HighV}$ and $C_{Preco}$ in DH18 are correctly captured by LSMs. In 2019, though, simulated growing season $GPP_{anom}$ shows low or even negative correlations with $EVI_{anom}$ in all clusters ($r = -0.09$, $-0.43$). This cannot be explained by poor simulation of $SM_{anom}$, which shows high correlation to both observation based soil moisture datasets, $r$, but not that of DH19 in $C_{Decline}$.

5 Discussion

5.1 Early signs of increased vulnerability

Our results indicate that the extremely negative

For three clusters covering 56% of the pixels negatively impacted by DH18, the extremely low $EVI_{anom}$ in response to DH2018–DH18 and DH19 cannot be predicted from past $EVI$–climate relationships, even when non-linear and seasonal legacy effects are considered in 2001–2017 (Figs. 4, 5). This suggests an important role of increased ecosystem vulnerability to climate (e.g., “hotter droughts”) and of environmental factors (e.g., vegetation preconditioning effects), in explaining DH impacts.

The climatic variables used to train the temporal RF model still appear as relevant contributors to the spatial distribution of $EVI_{anom}$ residuals (Fig. 6) which supports a contribution of climate-driven increased vulnerability under the two extreme summers. The sign of the correlation between $T_{spr}$ and $SM_{spr}$ with $EVI_{anom}$ residuals indicates a positive effect of spring warming in partly offsetting vegetation growth (observed $EVI_{anom}$ more positive or less negative than modelled), but with associated water depletion in spring amplifying the impacts of DH2018 in summer (Bastos et al., 2020a). This negative seasonal legacy effect through soil moisture is also found in DH2019, but spring warming (or cooling, see Fig. B1) resulting in much stronger (weaker) impacts than those predicted by past vegetation–climate relationships (strong negative correlation). The positive correlation of $EVI_{anom}$ residuals in DH2018 with $SM_{spr}$ and $T_{spr}$ are also in line with increased sensitivity to water availability and temperature stress reported by Bastos et al. (2020b), i.e. stronger browning than predicted associated
with warmer and drier conditions. In DH2019, the negative impact of summer \( T_{anom}^{sm} \) was amplified, indicating increased vulnerability to heat stress, but the correlation of residuals with \( SM_{anom}^{sm} \) is the opposite as in DH2018. This is consistent with the results for clusters C1 and C4, where browning in DH2019 happened in spite of drought alleviation, in the case of C1, or greening happened in spite of persisting soil moisture deficits, in C4. However, these results indicate that other environmental effects may have an important role in modulating DH impacts, especially in DH2019.

The values of \( EVI_{anom} \) in the previous summer show the strongest correlations with the residuals in both years. These departures reveal increased sensitivity to dry and hot conditions, and can be a sign of increased ecosystem vulnerability to such events. It should be noted, though, that we focused on pixels which were negatively impacted by DH18, but some pixels in the regional domain selected showed greening, even in DH18 (Fig. 6). The higher correlation between \( EVI_{anom} \) residuals and \( EVI_{anom}^{yr-1} \) in DH2019 than DH2018 points to a stronger contribution of legacy effects preconditioning the impacts of DH2019, resulting from the heat/drought stress imposed by DH2018. Even though considering inter annual legacy effects mediated by vegetation condition improves the predictive skill of the RF model, this does not reduce the residuals in DH2018 and DH2019. The strong spatial relationship between \( EVI_{anom}^{yr-1} \) with \( EVI_{anom} \) residuals suggests that the preconditioning role of vegetation condition in DH2018 and DH2019 was amplified due a predominance of pixels with poorer vegetation condition before the DH events. This supports the important role of legacy effects from past stress conditions in impairing vegetation resistance to subsequent stressors, e.g. from to defoliation or damage from embolism (Ruehr et al., 2019) or higher susceptibility to diseases and pests due to reduced health (McDowell et al., 2020). The fact that \( EVI_{anom}^{DH2018} \) is the most relevant predictor for residuals in DH2019 further supports the importance of damage from past heat/drought stress in amplifying the impacts of a subsequent DH event (Anderegg et al., 2020).

These regional asymmetries result in partial regional compensation of the DH18 impacts, as shown in Bastos et al. (2020b).

In both DH2018 and DH2019, higher tree cover fraction is associated with more positive or less negative residuals (Fig. 6), indicating that trees buffered the impacts of DH conditions on ecosystem activity. This is consistent with the predominance of crops and grasslands in C2 and C3, which had strong negative residuals in DH2018 both events, and of high tree cover in C4, where residuals are mostly within the range of residuals in the training period and even slightly positive (Figs. 4 and 5). Forests in C3 also support this effect. Trees can better cope with drought with their deeper rooting depth (Fan et al., 2017) and through the use of carbon reserves to support activity under stress conditions (Wiley, 2020). Moreover, some trees and grasses with stronger stomatal regulation can buffer the drought progression and its impacts by avoiding hydraulic failure (McDowell et al., 2020; Teuling et al., 2010). Even though isohydricity is strongly species-dependent (Konings et al., 2017), this effect is reflected in the small but positive effect of isohydricity in explaining DH2018 residuals.

The negative residuals in DH2019 for C2 are consistent with C2 showing the driest and hottest anomalies and predominantly cropland cover, but in C1 the strongly negative \( EVI_{anom}^{DH2019} \) are associated with mixed pixels (up to 40% forest, relationship between isohydricity and 20–60% grassland cover) and higher isohydricity. This points to increased vulnerability and degradation occurring mainly in natural ecosystems with stronger stomatal regulation, which is consistent with the negative relationship
of isohydricity with $EV \text{T}^{DH2019}$ residuals. The large negative $EV \text{T}^{DH2019}$ residuals and browning in response to DH2019

$EV \text{T}^{anom}$ residuals in C1 may be linked to impaired growth due to pixels with high forest cover.

Increased vulnerability may be explained by modulating effects of global change on vegetation condition (e.g., “hotter droughts” (Allen et al., 2015), Fig. 1) and, in the case of DH19, it may be further linked to inter-annual legacies from the impact of DH18. The first should be expressed by relationships between $EV \text{T}^{anom}$ residuals and climatic variables. The latter are more difficult to assess without comprehensive data about different competing factors, e.g., defoliation or damage from embolism, defoliation, or depletion of carbon reserves (Ruehr et al., 2019) under longer drought conditions (negative effect of dry months). However, increased (Ruehr et al., 2019), higher susceptibility to diseases and pests due to reduced health (McDowell et al., 2020) or increased hazard of insect disturbances due to warm conditions (Rouault et al., 2006). The relationships between $EV \text{T}^{anom}$ residuals and $EV \text{T}^{yr-1}$ provide an approximation, but do not allow to identify the underlying drivers.

In DH18, we find a positive effect of spring warming in vegetation growth, leading to weaker departures from long-term vegetation–climate relationships (observed $EV \text{T}^{anom}$ more positive or less negative than modelled), but with associated water depletion amplifying the impacts of DH18 in summer in pixels with low tree cover. These results are in line with Bastos et al. (2020a) that showed contrasting seasonal legacy effects of warm springs in crop versus forest dominated regions.

On the contrary, spring and summer $T^{anom}$ in 2019 (or cooling, see Fig. B1) are negative correlated with $EV \text{T}^{anom}$ residuals in both high and low tree cover pixels. This indicates increasing damage from heat stress, for example due to reductions in evapotranspirative cooling (Obermeier et al., 2018) or cascading effects impacts of compound heat and drought, such as the increased susceptibility of stressed trees to insect attacks and diseases cannot be excluded—e.g., insect attacks (Rouault et al., 2006).

The $EV \text{T}^{yr-1}$ in the long-term RF regression model improves the predictive skill for 2001–2017, but does not reduce the residuals in DH18 and DH19. The high correlation between $EV \text{T}^{anom}$ residuals and $EV \text{T}^{yr-1}$ in DH19 can indicate either that pixels strongly impacted by DH18 were associated with amplified impacts by DH19 (negative residuals), or that pixels affected moderately by DH18 (less negative $EV \text{T}^{DH18}$) were associated with positive residuals, i.e., stronger recovery.

Damage to roots and tissues or depletion of carbon reserves from DH18 leading to higher vulnerability to DH19 could explain the positive correlation in high tree cover pixels in C_Dehaese. Conversely, the moderate DH18 impacts in C_Greening may have resulted in increased resistance to DH19. The strong correlation found in low tree cover pixels is, though, surprising, as European crop species tend to be annual plants, and annual species can also be found in many grasslands. For these pixels, it is more likely that the positive correlation is explained by management practices, e.g., through earlier harvest or active reduction of stand density in DH19 (Bodner et al., 2015).

$C_{\text{Decline}}$ stands out from the other clusters, in that browning is found in spite of drought alleviation in DH19. The strong negative correlation of residuals with $SM^{anom}$ and AWC in forest dominated pixels is counter-intuitive and suggests that other environmental effects not considered in our analysis may modulate DH19 impacts. Insect outbreaks are a potential candidate to explain such effects; the stronger correlation of residuals with $EV \text{T}^{yr-1}$ in DH2019/DH19 could reflect increased susceptibility of impaired trees, combined with favourable climatic conditions for insect growth, reflected in stronger negative effects of
$T_{anom}^{sm}$ in DH2019, (Rouault et al., 2006)), indicates that such cascading effects could also have contributed to amplify the impacts of DH2019–DH19 in high tree cover pixels.

Results from field inventories and forest plots support this hypothesis. An increase in Increased tree mortality and insect outbreaks in central Europe during 2018 has have been reported (Schuldt et al., 2020). A recent assessment by the German Federal Minister for Food and Agriculture (BMEL, 2020) reported crown damage in 36% of all tree types in summer 2019, a 7% increase compared to 2018 and predominating in trees over 60 years of age. According to this report, the mortality rate in both needle-leaved and broad-leaved trees almost tripled from 2018 to 2019. Although no large scale data on insect outbreaks is currently available, local authorities in regions where C1–C_{Decline} is prevalent report increase in tree mortality from bark-beetle infestations: the Environment Ministry of North Rhine Westphalia in western Germany reported soaring rates of spruce affected by severe bark-beetle infestations, from about 1% in 2018 to over 12% in 2019 (MULNV-NRW, 2019). In the Czech Republic, rates of spruce damaged by bark-beetles more than tripled, leading to increased mortality (Hlásny et al., 2021). In Belgium, a “bark bettle task force” was created in September 2018 by the economic office of Wallonia (OEW, 2018). Increased tree mortality and bark-beetle infestations have also been reported in eastern France (ONF, 2020).

5.2 Implications for earth system modelling

Temperate ecosystems are an important global sink of CO2 (Pan et al., 2011) and are not usually considered hot-spots of drought risk and environmental degradation under climate change (Vicente-Serrano et al., 2020). Our results show that the past two extreme summers in central Europe reveal first signs of large-scale enhanced vulnerability in response to DH events (C2–C_{HighV2–C_{Preco}}), and of potential degradation trajectories induced by consecutive events (C1–C_{Decline}). Even though limited to 20% of the study area, the patterns in C1–C_{Decline} highlight the risks associated with more frequent and intense droughts and heatwaves expected in the coming decades (Barriopedro et al., 2011; Boergens et al., 2020; Hari et al., 2020). At the same time progressive warming conditions are likely to promote can increase the likelihood of compound occurrence of multiple disturbances, such as droughts and insect outbreaks, both promoted by warm and dry conditions. Interactions between compounding disturbances can further contribute to forest C losses (Seidl et al., 2017; Kleinman et al., 2019). To anticipate such impacts, process-based modelling of ecosystem response to such events is needed.

The LSMs perform well in simulating the magnitude and evolution of productivity anomalies in 2018, but not in 2019. The recovery simulated by LSMs in DH19 can be partly explained by a strong recovery of modelled soil-moisture (Fig. B7), but may also result from limited ability of LSMs in simulating changes in ecosystem vulnerability during the two DH events. The latter is supported by the fact that simulated $SM_{anom}$ shows good agreement in the temporal evolution of soil-moisture anomalies with both observation-based datasets but not of $GPP_{anom}$ (Table 1).

The comparison of the reference and factorial simulations allows showing that the poor performance in 2019 may be related with interannual legacy effects. LSMs estimate legacies from DH2018–DH18 only in the early growing season (March to May 2019), but do not estimate any legacy effects in summer (Fig. 7 bottom panel). The poor relationships between $EVI_{anom}$ and simulated $GPP_{anom}$ in response to DH2019–DH19 indicate that processes controlling legacy effects such as damage from
embolism, carbon-starvation and resulting tree-mortality or disturbances induced by drought and heat such as insect outbreaks, currently missing in LSMs, likely explain the amplified impacts of DH2019DH19.

LSMs are known to have limited ability to simulate drought-induced stress and tree mortality (Wang et al., 2012), and lack impacts of biotic disturbances, although rudimentary approaches have been attempted (Kautz et al., 2018). These model shortcomings add to limitations in simulating soil-moisture variability and transitions between energy-limited and water-limited regimes. Attributing the LSM errors to specific climatic or non-climatic processes here is challenging since up-to-date datasets on tree mortality, tree carbon reserves or spatially-explicit information on biotic disturbances are very limited. Nevertheless, our results show that LSMs can simulate well the impacts of one strong drought (DH2018DH18) on ecosystem dynamics but have limited skill in simulating the impacts of a subsequent compound extreme event (DH2019DH19) by missing important inter-annual legacy effects.

6 Conclusions

The two consecutive extreme dry and hot summers in central Europe (DH2018 and DH2019DH18 and DH19) had stronger impacts on vegetation activity than those expected by previous vegetation–climate sensitivity. This hints at large-scale increase in the vulnerability of ecosystems to compound heat and drought events, possibly modulated by vegetation responses to the long-term warming and increasing CO\textsubscript{2} trends, environmental changes. We find signs of degradation trajectories in 20% of the study area, where EVI decreased even with drought alleviation in the following year. We attribute these trajectories to legacies from DH2018DH18 amplifying the impacts of DH2019DH19, which indicates that more frequent extreme summers may pose a major threat to the stability of temperate forests.

State-of-the-art land-surface models were able to simulate the exceptional impacts of DH2018DH18, but they underestimated the impacts of DH2019DH19. This is explained by LSMs missing the preconditioning effect of DH2018 in DH2019DH18 in DH19 impacts as they cannot simulate inter-annual legacy effects from DH events on ecosystem activity. In addition, LSMs also lack representation of biotic disturbances, which are triggered by DH conditions and further promoted by plant stress in response to DH. Because DH events may become more common in the coming decades, overlooking these effects may result in an overestimation of the resilience of the CO\textsubscript{2} sink to climate change in temperate regions.

Data availability. The MOD13C1 data are available through NASA’s data catalog at https://lpdaac.usgs.gov/products/mod13c1v006/. SoMo.ml v1.0 is publicly available via https://doi.org/10.17871/bgi_somo.ml_v1_2020. Isohydricity fields are available at https://github.com/agkonings/isohydricity. AWC data are provided by the European Soil Data Centre (ESDAC) through esdac.jrc.ec.europa.eu. The multimodel mean fields from the LSMs are provided as supplementary material. The individual LSM model outputs are available upon request to abastos@bgc-jena.mpg.de.
Author contributions. AB designed the study and methodology, conducted the data analysis and wrote the manuscript. RO, MR, PC, NV, and SZ contributed to initial development of study and to the first manuscript draft. SS and JP helped designing the LSM simulation protocol and SS coordinated the LSM modelling effort. SO provided the SoMo.ml dataset. PG contributed with expert knowledge. NV, SZ, PA, AA, PM, EJ, SL and TL ran the LSM simulations. All authors participated in the writing of the final version of the manuscript.

Competing interests. The authors have no competing interests

Acknowledgements. AB thanks Ulrich Weber for preprocessing of the MODIS data, Corinne Le Quéré for providing updated atmospheric CO2 fields for the model simulations. We thank the European Soil Data Centre (ESDAC), esdac.jrc.ec.europa.eu, European Commission, Joint Research Centre for making AWC data available and Alexandra Konings for providing the isohydricity dataset. AB received funding by the European Space Agency Climate Change Initiative ESA-CCI RECCAP2 project (ESRIN/ 4000123002/18/I-NB). SL has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 821003 (project 4C, Climate-Carbon Interactions in the Coming Century) and SNSF (grant no. 20020_172476). RO and SO acknowledge support by the German Research Foundation (Emmy Noether grant 391059971). ORCHIDEE simulations was performed using HPC resources from GENCI-TGCC (grant 2020-A0070106328).
References


Land surface and global dynamic vegetation model simulations

We have used output of gross primary productivity (GPP) and simulated soil-moisture from seven models that followed the protocol and extended the simulations in Bastos et al. (2020a) up to 2019. These models are: ISBA-CTRIP (Joetzjer et al., 2015), JSBACH (Mauritsen et al., 2018), LPJ-GUESS (Smith et al., 2014), LPX-Bern (Lienert and Joos, 2018), OCN (Zaehle et al., 2010), ORCHIDEE (Krinner et al., 2005) and SDGVM (Walker et al., 2017).

The model simulations were run for most models at 0.25 ° spatial resolution for the European domain (32–75°N and -11–65°E), following a spin-up to equilibrate carbon-pools. For the reference simulation, the models were forced with observed CO₂ concentration from NOAA/ESRL and changing climate between 1979 and 2019 from ERA5 and fixed land-cover map from 2010 from LUH2v2 (Hurtt et al., 2011). An additional simulation was ran where the models were forced with changing climate, except June–August 2018, where climatological summer climate conditions were used to force the models as described in Bastos et al. (2020a). This simulation, extended up to December 2019 allows evaluating the direct impact of DH18 and its inter-annual legacy effects.

For more details on the simulation protocol, we refer to (Bastos et al., 2020a).
Appendix B: **Supplementary Figures**
Figure B1. Monthly temperature anomalies during 2018 and 2019. The rectangle indicates the study region.
Figure B2. Monthly soil-moisture anomalies during 2018 and 2019. The rectangle indicates the study region, i.e. the areas experiencing drought conditions ($SM_{\text{anom}} < -1\sigma$) during both DH2018-DH18 and DH2019-DH19.
Figure B3. Selectivity of different land-cover composition for each cluster (Fig. 3). Selectivity is evaluated as the difference between the probability distribution of a given land-cover type (forest, left; grassland, middle; cropland, right) and the probability distribution of that land-cover type in the selected region. If selectivity is positive, the cluster is preferentially composed by the given land-cover type and the opposite for negative values. The 2018 land-cover classification maps from from ESA CCI-LC are used.
Figure B4. Performance of the temporal RF model in predicting $EVI_{anom}$, given by the out of bag scores. The top panel shows the scores for the climate-driven RF model and the bottom panel the corresponding results for the same model, but including $EVI_{anom-1}$ as an additional predictor.
Figure B5. Importance of the four predictors used in the RF model to predict \( EVI_{anom} \), spring (left) and summer (right), \( SM_{anom} \) (top) and \( T_{anom} \) (bottom), calculated from the Shapley additive explanation values (Methods).
Figure B6. As in Fig. 5 bottom panel, but for the RF model trained using spring and summer $S_{anom}$ and $T_{anom}$ as predictors, as well as $EVI_{yr-1}^{anom}$. 
Figure B7. The left panel shows the spatial distribution of the four clusters from unsupervised classification of \((EVI_{DH2018}^{anom}, EVI_{DH2018}^{anom}, EVI_{DH2019}^{anom}, EVI_{DH19}^{anom})\) values remapped to the coarser grid of LSMS. The corresponding \((GPP_{DH2018}^{anom}, GPP_{DH2018}^{anom}, GPP_{DH2019}^{anom}, GPP_{DH19}^{anom})\) values simulated by the multi-model mean in each cluster are indicated in the top right panel (circles indicate the spatial mean and the lines spatial standard deviation within each cluster). The corresponding distribution of simulated \(SM_{anom}\) pairs in each cluster are shown in the bottom right panel. The grey line, indicates similar anomalies in the two DH events.
Figure B8. Same as Fig. 4 but for GPP and soil-moisture anomalies simulated by a subset of land-surface models from (Bastos et al., 2020a) extended up to December 2019.