Analysis of multi-seasonal meteorological storylines leading to reduced forest greenness in Europe in 2000-2020

Mauro Hermann\(^1\), Matthias Röthlisberger\(^1\), Arthur Gessler\(^2,3\), Andreas Rigling\(^2,3\), Cornelius Senf\(^4\), Thomas Wohlgemuth\(^2\), and Heini Wernli\(^1\)

\(^1\)Institute for Atmospheric and Climate Science (IAC), ETH Zürich, Zurich, Switzerland
\(^2\)Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Zurich, Switzerland
\(^3\)Institute of Terrestrial Ecosystems (ITES), ETH Zürich, Zurich, Switzerland
\(^4\)Ecosystem Dynamics and Forest Management Group, Technical University of Munich, Freising, Germany

Correspondence: Mauro Hermann (mauro.hermann@env.ethz.ch)

Abstract.

Recent forest decline in Europe is strongly influenced by meteorological conditions imposed by seasonal variations of temperature, \(T_{2m}\), and precipitation, \(P\), and can be monitored with forest greenness. This study quantitatively investigates anomalous characteristics of the three-year meteorological storyline preceding events of reduced forest greenness in Europe’s temperate and Mediterranean biome in the phase space of seasonal-mean anomalies of \((T_{2m}, P)\). A specific focus is on the amplitude, persistence, and co-variability of these anomalies. A pragmatic approach based on remote sensing observations of normalized difference vegetation index \(NDVI\) serves to identify low forest \(NDVI\) events at the 50 km scale in Europe in June to August 2000—2020. An independent forest disturbance data set is used to qualitatively validate the identified more than 1'500 low \(NDVI\) events. These events occur in summers with particularly dry and hot conditions but their meteorological storylines feature significant anomalies during multiple seasons preceding the events, with clear differences between the two biomes. In the Mediterranean biome, the anomalously dry conditions persist over more than 1.5 y prior to the events, whereas \(T_{2m}\) is anomalously warm only during the last 0.4 y. In contrast, in the temperate biome, \(T_{2m}\) is anomalously large during most of the 2.5 y prior to the events and, most interestingly, the autumn/winter preceding the events is characterized by anomalously wet and warm conditions. These anomalies potentially induce a negative legacy on the following summer drought.

The seasonal-mean anomalies of \(P\) are strongly determined by synoptic-scale weather systems, such that long dry periods are characterized by a deficit of cyclones and an excess of anticyclones. A final analysis investigates the peculiarities of low \(NDVI\) events that occur in two consecutive summers and the potential role of drought legacy effects. In the temperate biome, the second event summer of an event sequence has less hot and less dry anomalies than the first one and than during a single event. In summary, detailed investigations of the multi-annual meteorological storyline of low forest \(NDVI\) events provided clear evidence that anomalies of surface weather and synoptic-scale weather systems over time periods of up to 2.5 y can negatively impact European forest activity, with important differences between the temperate and Mediterranean biomes.
1 Introduction

European forest ecosystems have typically been in balance with their bio-physical environment and are, thus, largely adapted and acclimated to the local mean climate and its variability. This balance is increasingly disturbed by anthropogenic climate change (Seidl et al., 2017; McDowell et al., 2020). As a consequence, the exposure and vulnerability of European forests to climatic disturbances has increased in the recent two decades (Seidl et al., 2014; Bastos et al., 2020b). On the continental scale, drought has been a key driver of excess forest mortality in Europe (Senf et al., 2020). In addition to meteorological drought, i.e., reduced precipitation, high temperatures enhance the atmospheric water demand and, thus, water loss from the vegetated surface by evapotranspiration (Yuan et al., 2019; Grossiord et al., 2020). Heatwaves (Meehl and Tebaldi, 2004; IPCC, 2021), meteorological droughts (Trenberth et al., 2014; IPCC, 2021), and compound hot droughts (Allen et al., 2015) are expected to intensify future forest decline (Brodribb et al., 2020). Such meteorological conditions, for example during 2003 and 2018, led to stem dehydration (Salomón et al., 2022), reduced forest growth (Ciais et al., 2005; Trotsiuk et al., 2020), complete hydraulic failure (Schuldt et al., 2020), and ultimately increased tree mortality in Europe (Allen et al., 2010; Hansen et al., 2013). At the same time, recent hot-dry conditions fostered bark beetle outbreaks (Rouault et al., 2006; Jakoby et al., 2019) and forest fires (Seidl et al., 2017). Although some forest ecosystems are adapted to low summer water availability, e.g., in the summer-dry Mediterranean or in dry inner-Alpine regions, extended and more frequent hot droughts will strongly affect their dynamics including growth and survival (Rigling et al., 2013; Tague et al., 2019; Ogaya et al., 2020). In this context, forest greenness is an effective measure related to forest performance to monitor the recent forest decline in Europe (Orth et al., 2016; Buras et al., 2020, 2021).

While forests can endure short-term weather extremes (e.g., an individual multi-day heatwave), they are more susceptible to longer-term extreme conditions. Particularly harmful long-term extremes include persisting or sequentially occurring droughts that disturb forests beyond the summed individual impacts, as the initial drought exerts a negative legacy effect on the following one by reducing forest resilience (Anderegg et al., 2015, 2020; Bose et al., 2020). However, drought legacy effects have also been suggested to provide acclimation to following droughts (Gessler et al., 2020). Moreover, compound hot and dry periods cause an impact exceeding that of the summed univariate extremes (Zscheischler and Seneviratne, 2017). Furthermore, a stormy winter followed by a hot-dry growing season allows bark beetles to spread out and attack damaged and dying trees (Temperli et al., 2013; Biedermann et al., 2019; Jakoby et al., 2019). Additionally, persisting heat, and/or precipitation deficits can – given fuel availability – trigger forest fires year-round, occurring most intensely in the Mediterranean (Turco et al., 2017). Lastly, beneficial conditions in the early growing season can exert a negative legacy on forest activity in dry summers, as early growth fosters canopy buildup and soil moisture depletion (Bastos et al., 2020a). These examples highlight that in addition to the (co-)occurrence, magnitude, and duration of heat and drought, their position in a longer-term meteorological storyline modulates their interplay with other disturbances and, thus, also their impact on forest performance.
In the context of such interactions and legacy effects the time-scale over which the meteorological storyline is relevant to forest performance is intensively discussed. The meteorological impact has often been investigated by considering the mean over the current growing season (e.g., Hlášny et al., 2017; Seftigen et al., 2018; Seidl et al., 2020). A recent comprehensive analysis of drought-induced canopy mortality revealed that hot-dry conditions are particularly harmful in March to July (Senf et al., 2020). Neumann et al. (2017) stressed warm summer temperatures and high variability in seasonal precipitation as meteorological drivers of tree mortality. In a drought-prone region, forest drought stress was about equally determined by temperature in summer and the previous autumn, and precipitation during the cold season (Williams et al., 2013). More generally, drought-induced partial or complete tree mortality shows a threshold behaviour (Brodribb et al., 2020; Senf et al., 2020); however, water stress is not equally harmful at all times. Especially outside or at the margins of the growing season there are complicating factors such as growth compensation, soil moisture coupling, and snow melt (e.g., Harpold et al., 2015; Bastos et al., 2020a). Despite increasing understanding, a systematic analysis of the meteorological storyline of reductions in forest performance in Europe is still missing.

The meteorological processes that are relevant for low forest performance cover a wide range of timescales. Although longer-term meteorological extremes are of particular interest, they can also be composed of multiple shorter-term anomalies, which are not necessarily extreme at that timescale (Röthlisberger et al., 2020). On the timescale of about 3–10 days, atmospheric blocking can induce surface heatwaves by large-scale subsidence and high solar insulation (Pfahl and Wernli, 2012a; Zschenk-derlein et al., 2019). At the same time, the involved quasi-stationary upper-tropospheric ridge leads to reduced precipitation. In central Europe, summer heatwaves can also arise from weak synoptic forcing, which, in combination with a Scandinavian blocking, allowed for widespread hot-dry conditions in 2018 (Spensberger et al., 2020). During heatwaves with no or reduced precipitation, the soil moisture-atmosphere coupling exacerbates the near-surface warming over drying soils (Fischer et al., 2007; Seneviratne et al., 2010). Accordingly, and especially in the Mediterranean, an extremely hot summer is more likely in years of a winter/spring precipitation deficit (Russo et al., 2019). As Europe hardly experiences drought over a longer (multiannual) time scale (Schubert et al., 2016), seasonal meteorology, which is strongly linked to weather system dynamics, is of particular interest for forest performance, and, therefore, in the focus of the present study.

Despite great progress in understanding the eco-hydraulic mechanisms linking drought to events of reduced forest performance and, thus, greennis (Brodribb et al., 2020), a systematic analysis of meteorological storylines accompanying such events is still lacking. The purpose of this study is to systematically document and characterize these meteorological storylines in Europe’s temperate and Mediterranean forests. We focus on the evolution of 2-m temperature ($T_{2m90d}$) and precipitation ($P_{90d}$) as key characteristics of the three-year meteorological storylines and quantitatively address the following research questions in Sect. 3.2: (1) What is the average meteorological storyline preceding low forest greenness events in Europe’s temperate and Mediterranean forests? (2) How anomalous is the magnitude and persistence of $T_{2m90d}$ and $P_{90d}$ anomalies at different time lags prior to the events? (3) Which changes in weather system frequency go along with the meteorological storyline identified in (1)? (4) How does the meteorological storyline differ for sequences of low forest greenness events as compared
to single events? To do so, we use low values of the normalized difference vegetation index (NDVI) in summer 2000–2020 as a measure of reduced forest greenness (Sects. 2.2 & 2.3). We also aim to answer whether the identified low NDVI events feature more forest disturbances than non-events by evaluating them with a more comprehensive European forest disturbance data set by Senf and Seidl (2021a, Sect. 3.1). Finally, we critically discuss our results and the limitations of our analyses in Sect. 4.

The term “meteorological storyline” is central to this study. It encompasses the evolution of $P_{90d}$ and $T_{2m90d}$ over the three years prior to low NDVI events, and, thereby, emphasizes the two fundamental meteorological variables at the seasonal time scale. Throughout the story we focus on different periods within, and use various measures of the meteorological storyline, respectively, e.g., anomalies of $P_{90d}$ and $T_{2m90d}$ or their bi-variate evolution, to capture all its facets.
2 Data and methods

We differentiate between the temperate and Mediterranean biome according to Schultz (2005) in the domain extending from 10° W to 45° E and 35° N to 60° N, excluding all boreal forests (Fig. 1a).

![Forest grid cells](image)

**Figure 1.** (a) forest grid cells in the study domain, separated into temperate and Mediterranean forests by the black dashed line. The boreal biome is cropped by the second dashed line in the Northeast of the domain. Hatching indicates where events cannot be evaluated with the disturbance data set (see Sect. 2.5). (b) the identification of low NDVI events in an example forest grid cell (details provided in the text).

2.1 Forest cover data

Forest cover observations are available from the global tree canopy cover data version 1.7 derived from Landsat at one arc-second (~30 m) horizontal resolution from the year 2000 (Hansen et al., 2013). For comparison with the other data sets introduced below, we coarse grain fractional forest area, $FA$, to 0.05° (~5 km) and to 0.5° (~50 km), hereafter denoted as $FA^{0.05}$ and $FA^{0.5}$, respectively. For our analysis we only consider sufficiently forest-covered 0.05° × 0.05° pixels, i.e., pixels with $FA^{0.05} \geq 50\%$, termed forest pixels. We additionally request that they lie within 0.5° × 0.5° grid cells with a forest coverage $FA^{0.5} \geq 20\%$, termed forest grid cells. According to this definition, there are 1'387 and 343 forest grid cells in the temperate and Mediterranean biome, respectively, as shown in Fig. 1a.

2.2 Normalized difference vegetation index

We use the monthly normalized difference vegetation index ($NDVI$) at 0.05° horizontal resolution from February 2000 to August 2020 from NASA MODIS Terra provided on the Giovanni data system (Acker and Leptoukh, 2007). It is based on the red ($RED$) and near-infrared ($NIR$) spectral irradiance:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$
The greener a vegetated surface is, the closer its \( \text{NDVI} \) is to +1 (Tucker, 1979). Note that at forest pixels, the \( \text{NDVI} \) additionally depends on the remaining vegetation cover, i.e., on within-pixel heterogeneity in surface reflectance properties. The \( \text{NDVI} \) serves as a measure of vegetation performance and has previously been used to assess drought impact on ecosystems (Anyamba and Tucker, 2012; Orth et al., 2016; Buras et al., 2020). We linearly interpolate missing values from the previous and subsequent month, and perform a linear detrending of the entire time series as in Buras et al. (2020) due to a detected greening trend (Bastos et al., 2017). Gap interpolation is mainly necessary over water and hardly affects our analysis.

After this post-processing, we only consider normalized \( \text{NDVI} \) anomalies \( (\text{NDVI}') \) at forest pixels and in June–August (JJA). At every forest pixel \( j \) in month \( m \) and year \( n \), \( \text{NDVI}'_{j,mn} \) results from dividing the absolute anomaly by the local standard deviation \( \sigma_j(\text{NDVI}) \), which is calculated from all 63 monthly \( \text{NDVI} \) anomalies in JJA 2000–2020. The respective climatology \( \overline{\text{NDVI}}_{j,m} \) is calculated as the mean in 2000–2020.

\[
\text{NDVI}'_{j,mn} = \frac{\text{NDVI}_{j,mn} - \overline{\text{NDVI}}_{j,m}}{\sigma_j(\text{NDVI})} \quad (2)
\]

Note that \( \sigma_j(\text{NDVI}) \) is merely used as a measure of inter-annual \( \text{NDVI} \) variability in JJA, and not for quantifying the likelihood or the return period of a given value. The normalization is used here because the short data record precludes more sophisticated statistical modelling, and, at least to some extent, enables comparisons of \( \text{NDVI} \) values across Europe and over the study period. Also, according to a Shapiro-Wilk test, the assumption of normally distributed \( \text{NDVI} \) anomalies in JJA 2000–2020, which is implicit to normalization, cannot be rejected at 72% of all forest pixels (p-value > 0.05).

### 2.3 Identification of low \( \text{NDVI} \) events

The aim of the scheme presented here is to identify low \( \text{NDVI} \) events at \( 0.5^\circ \times 0.5^\circ \) forest grid cells, as this is the resolution of the meteorological data used in this study (see below). In essence, it (1) considers all forest pixels \( j \) within a forest grid cell \( J \) and flags them if their \( \text{NDVI}' \) is below a specified threshold, and (2) identifies an event in \( J \) if the majority of forest pixels inside \( J \) is flagged (Fig. 1b).

In the first step (1 in Fig. 1b), we calculate the annual minimum of the three \( \text{NDVI}'_{j,mn} \) values in JJA of year \( n \), termed \( \text{NDVI}'_{j,n}^{\text{min}} \). At every forest pixel \( j \), an event flag \( ev_{j,n} \) is determined according to the following criterion:

\[
ev_{j,n} = \begin{cases} 
1 & \text{if} \quad \text{NDVI}'_{j,n}^{\text{min}} \leq -2, \\
0 & \text{otherwise}.
\end{cases} \quad (3)
\]

In the second step (2 in Fig. 1b), we check whether more than half of the forest pixels in \( J \) are flagged as \( ev_{j,n} = 1 \). Therefore, the total area of forest pixels with \( ev_{j,n} = 1 \) in \( J \) \( (A_{j,n}^{\text{ev}}) \) has to be at least 50% of the total forest pixel area in \( J \) \( (A_{j}^{\text{for}}) \):
For the identified low $NDVI$ events ($EV_{J,n} = 1$), we calculate a measure of event intensity. It is equal to the spatial average of summer minimum $NDVI_{J,n}^{min}$ at all flagged forest pixels ($ev_{j,n} = 1$), and, hereafter, termed $NDVI_{J,n}^{min}$.

2.4 ERA5 reanalysis data

Atmospheric fields are used from ERA5 reanalyses from the European Centre for Medium-Range Weather Forecasts (ECMWF, Hersbach et al., 2020) available hourly on 137 vertical levels and interpolated to a regular grid with $0.5^\circ$ horizontal resolution.

2.4.1 Seasonal 2-m temperature and precipitation

Our analyses focuses on seasonal averages of 2-m temperature ($T_{90d}^{2m}$) and precipitation ($P_{90d}^{'}$) for the periods Sep–Nov (SON), Dec–Feb (DJF), Mar–May (MAM), and Jun–Aug (JJA). For consistency - as we also use moving seasonal (90-day) averages - we calculate seasonal mean values always over 90 days, right-aligned at the last day of the season. Each 90-day mean value, therefore, is labelled by the time step of the last value that contributes to the average. Leap days are discarded from the analysis. The respective climatologies cover SON 1999 to JJA 2020. Throughout the study, we calculate at every grid cell normalized anomalies denoted as $T_{90d}^{2m}$ and $P_{90d}^{'}$, e.g.,

$$P_{90d}^{'} = \frac{P_{90d} - \overline{P}_{90d}}{\sigma_{P_{90d}}}$$

where $\overline{P}_{90d}$ and $\sigma_{P_{90d}}$ denote the climatological seasonal mean and standard deviation in the considered 21 years, respectively. Note that as for $NDVI$ (Sect. 2.3), normalization of $P_{90d}^{'}$ and $T_{90d}^{2m}$ is used merely for scaling and spatiotemporal comparison, not to estimate the values’ return period.

2.4.2 ($T_{90d}^{2m}, P_{90d}^{'}$) phase space

Moreover, we characterize the meteorological storylines in the ($T_{90d}^{2m}, P_{90d}^{'}$) phase space, which we divide into five compartments (see dashed lines in Fig. 2). Q0 denotes the central area within a circle of radius 1 and indicates weak seasonal mean anomalies. Outside this circle, the remaining four compartments correspond to the quadrants of the phase space. We label them clockwise as Q1 (warm-wet), Q2 (hot-dry), Q3 (cold-dry), and Q4 (cold-wet). The frequency of a phase space compartment $Q$ in a given set of events is termed $f(Q)$. The climatological distributions of ($T_{90d}^{2m}, P_{90d}^{'}$) over all forest grid cells in both biomes and all seasons are shown in Fig. 2. In both biomes and all seasons, $f(Q0)$ is about 40%. $T_{90d}^{2m}$ and $P_{90d}^{'}$ are anti-correlated in MAM, JJA, and SON (Fig. 2a,c,d,e,g,h), meaning that hot-dry and cold-wet conditions occur more frequently than warm-wet and cold-dry. In the temperate (Mediterranean) biome in JJA, $f(Q2 \lor Q4) = 44.0\% (40.8\%)$, while Q1 and Q3
only make up for 14.9% (18.4%) of the season. In contrast, DJF seasonal anomalies are weakly positively correlated in both biomes (Fig. 2b,f), and are, therefore, more often in Q1 or Q3 compared to the other seasons.

Figure 2. 2D hexagonal binning of the climatological frequency distribution of $(T_{2m}^{90d}, P_{90d}^\prime)$ in (a,e) SON, (b,f) DJF, (c,g) MAM, and (d,h) JJA for the (a-d) temperate ($n = 29127$), and (e-h) Mediterranean biome ($n = 7203$), respectively. Dashed lines mark the division into the phase space compartments Q0–Q4, and numbers indicate their climatological frequency.

2.4.3 Significance assessment

We conduct two similar types of bootstrapping tests to assess significant differences in the meteorological storylines. Both tests are described in greater detail in Appendix A. Broadly, the bootstrapping produces 10’000 synthetic samples of low $NDVI$ events with a sample size equal to the number of low $NDVI$ events under consideration, and we use a significance level of $\alpha = 5\%$. In Sects. 3.2.1, 3.2.2 & 3.2.4 we test the null hypothesis $H_{0, EV}$ at different time lags $\Delta t$:

$H_{0, EV}$: The meteorological storyline at $t_0 - \Delta t$ is unrelated to the occurrence of low $NDVI$ events at $t_0$.

We use different test statistics to investigate several measures of the meteorological storyline. These measures are the event mean $T_{2m}^{90d}$ and $P_{90d}^\prime$, respectively, the fraction of $\Delta t$ that is on average covered by warm and dry periods, respectively, and the occurrence frequency of each $(T_{2m}^{90d}, P_{90d}^\prime)$ phase space compartment $Q$ per event set. Moreover, the re-sampling is
designed to retain the spatial correlation of the original $T2m$ and $P$ fields (for details see Appendix A). The median over all 10’000 synthetic measures of the meteorological storyline is referred to as the reference climatology $T_{90d}^m, P_{90d}^r$, or $f(Q^r)$. The reference confidence interval is computed from the 2.5th and 97.5th percentile of all 10’000 values. At time lags $\Delta t$ when the meteorological storyline of the events lies outside the confidence interval, $H_{0,EV}$ is rejected for time lag $\Delta t$.

In Sect. 3.2.3 we employ an analogous technique to identify statistically significant differences in the meteorological storylines of single low NDVI events (EV10) and sequences of events (EV11). Specifically, we draw 10’000 times a sample of $n_{EV11}$ from the meteorological storylines of EV10, where $n_{EV11}$ is the number of EV11. We again receive a 95% confidence interval - this time for the EV10 reference - and thereby test the null hypothesis $H_{0,EV11}$:

$H_{0,EV11}$: The meteorological storyline of EV11 at $\Delta t$ is the same as that of EV10.

2.4.4 Weather systems

Cyclones and anticyclones are identified from the outermost closed sea level pressure (SLP) contour around local SLP minima and maxima, respectively, according to Wernli and Schwierz (2006) and Sprenger et al. (2017). We calculate seasonal cyclone and anticyclone frequencies, $f_{90d}(C)$ and $f_{90d}(A)$, respectively, as right-aligned averages over 90 days. Moreover, we evaluate $T_{90d}^m, P_{90d}^r, f_{90d}(C)$, and $f_{90d}(A)$ continuously over time, i.e., moving 90-day averages.

2.5 Forest disturbances

For evaluating our set of low NDVI events we use the forest disturbance data set by Senf and Seidl (2021a) with an original resolution of 30 m. It is based on a time-series segmentation approach called LandTrendr (Kennedy et al., 2010) and identifies tree canopy mortality. The approach uses two spectral bands (shortwave infrared I and II) and two spectral indices (tasseled cap wetness and normalized burn ratio) from Tier 1 Landsat 4, 5, 7, and 8 images in Jun–Sep. For more details see Senf and Seidl (2021a). From this data set we use the annual disturbance area $D_{J,n}$, which is aggregated for every forest grid cell. $D$ is shown for some exemplary years in Fig. B1. More specifically, we use two measures: the disturbance anomaly $D'$, and the rank of $D$ among the 21 annual values $DR_{J,n}$:

$$D'_{J,n} = \frac{D_{J,n}}{\overline{D}_J} - 1$$

$$DR_{J,n} = \text{rank}(D_{J,n})$$

at forest grid cell $J$ in year $n$, with $\overline{D}_J$ denoting the climatological mean disturbance area in 2000–2020. The disturbance data overlaps with 73% of the forest grid cells (Fig. 1a), which include 81% of all low NDVI events.
3 Results

3.1 Reduced forest performance

3.1.1 Low \( NDVI \) events in 2000–2020

Low \( NDVI \) events cover substantial parts of both biomes in 2000–2020 (Fig. 3). In the temperate biome, the years with the most wide-spread event coverage - in descending order - were 2018, 2020, 2019, 2003, and 2010 (Fig. 3a,d). In 2018 close to 20% of temperate forests were affected by reduced greenness, while the two subsequent years covered 11–14% of which large parts were already affected in 2018 (Fig. 3d). Additionally, parts of France and eastern Europe that were unaffected by any event in 2018 experienced some events in the succeeding summers. The top five years in terms of affected forest area in the Mediterranean biome were 2017, 2012, 2000, 2001, and 2006, again sorted by decreasing area affected. In the most extensive event year 2017, low greenness affected 16% of the Mediterranean forest area.

Figure 3. (a) event year, (b) event intensity as measured by \( NDVI^{\text{min}} \), (c) the number of events per forest grid cell, and (d) the event-affected forest area relative to the total temperate and Mediterranean biome (in %). Years with less than 15 events are shown in white in (a). In (a) and (b) each grid cell is split into four quads, therefore showing the variables for up to four events in different years, ordered by intensity from top left to bottom right.
Furthermore, the intensity and frequency of the low NDVI events varied over Europe. Hot-spot regions with two to three events during the 21 years were Germany, France, and southeastern Europe (Fig. 3c). The latter was affected mainly in 2000, 2007, and in 2012–2013. Central Europe was largely affected by the 2018–2019 events, while further to the Southwest, 2003, 2005–2006, 2015, and 2019–2020 were the most dominant event years (Fig. 3a). Event intensity as measured by $\text{NDVI}^\text{min}$ (see Sect. 2.3) averaged at $-2.7$ in both biomes, and was strongest in events in southeastern Europe and in the 2010 event in Russia (Fig. 3b). Two forest grid cells in the French Massif Central and one in the UK were struck by four or five events, which, however, were of only intermediate intensity. Note that one third of the forest regions in the study domain never experienced an event in JJA 2000–2020 (Fig. 3c). These grid cells are in northeastern Europe or in mountainous regions including the Alps, the Carpathians, the Dinaric Alps, and the Eastern Black Sea Mountains. After a brief comparison of low NDVI events to independent forest disturbance data in the following section, we use these events in both biomes separately as the basis for specific meteorological analyses in Sect. 3.2.

### 3.1.2 Event evaluation

To evaluate the low NDVI events identified in the previous sections, we qualitatively compare them to the independent disturbance data set of Senf and Seidl (2021a). In 67% of all low NDVI events the disturbance area $D$ is larger than on average in 2000–2020 - more often in the temperate (69%) than in the Mediterranean biome (56%; Fig. 4a,c). The median disturbed area increases by +23% and +13% during events in the temperate and Mediterranean biome, respectively. Furthermore, non-events typically go along with negative $D'$ in the temperate (64% of non-events) and the Mediterranean biome (66%). Figure 4b,d additionally shows the disturbance area rank, $D_R$, from 1 (smallest $D$ in 2000–2020) to 21 (largest $D$). With 1–5 events per affected forest grid cell (Fig. 3) an event would go along with $D_R$ 17–21 if the event years were equal to the years of largest disturbed area. Most events indeed cover ranks 17–21 and 16–21 for the temperate and Mediterranean biome, respectively. We conclude that low NDVI events tend to go along with more forest disturbances, i.e., enhanced canopy mortality, and rank among the largest forest disturbances at forest grid cells.

While the two data sets correspond particularly well in years of widespread events, e.g., 2003, 2018, 2019, and 2020 (see Fig. B1), there are substantial discrepancies between the two data sets in other years. For example, 36% of non-event grid cells experience a positive $D'$. We show in Appendix B that several severe winter storms caused large $D$ but did not manifest in very low NDVI in the following summer. Examples for storms that were not followed by a low NDVI event are Lothar in western Europe in winter 1999/2000, Gudrun in Scandinavia in 2004/05, Kyrill in Germany in 2006/07, and Vaia in northern Italy in 2018/19. Furthermore, $D$ of several low NDVI events ranks among the lowest during the study period, which can be explained by a number of conceptual, technical, and physical reasons (see discussion in Sect. 4.1). Nevertheless, Fig. 4 reveals that low NDVI events are also associated with increased canopy mortality and we, thus, proceed to analyzing their meteorological storyline.
3.2 Meteorological storyline of low NDVI events

We investigate four characteristics of the meteorological storyline prior to low NDVI events separately for the two biomes. Thereby, we first consider the anomalies $T2m'_{90d}$ and $P'_{90d}$ and their link to anomalous weather system frequencies continuously (i.e., a right-aligned 90-day running mean) over the three years prior to the events (Sect. 3.2.1). Second, we assess the persistence of these anomalies (Sect. 3.2.2). Based on these results, we focus on the shorter meteorological storyline of the six previous seasons to, thirdly, compare the meteorological storyline of event sequences (i.e., two consecutive summers with low NDVI events) with those of single events (Sect. 3.2.3). Lastly in Sect. 3.2.4, we analyze anomalies in the frequency of seasonal ($T2m'_{90d}, P'_{90d}$) phase space compartments, again addressing the link to weather systems.
Figure 5. Average three-year evolution of (a,c) $P_{90d}'$ and (b,d) $T_{2m_{90d}}'$ leading up to low $NDV I$ events (red lines). The grey shading displays the confidence interval of the tested null hypothesis $H_{0, EV}$, i.e., the 2.5th – 97.5th percentile of the reference climatology (see Sect. 2.4.3). Solid and dashed lines show the 90-day (e,f) cyclone and (g,h) anticyclone frequency in climatology and during events, respectively. Colored shading indicates the respective anomalies $f_{90d}'(C)$ and $f_{90d}'(A)$. Plots apply to events in the (a,c,e,g) temperate, and (b,d,f,h) Mediterranean biome.

3.2.1 Uni-variate $P_{90d}'$ and $T_{2m_{90d}}'$ and weather system frequencies

Figure 5 displays the evolution of $P_{90d}'$ and $T_{2m_{90d}}'$ and collocated weather system frequencies for events in both biomes. For instance, the red line in Fig. 5a shows the time evolution of the 90-day averaged precipitation anomaly during the three years prior to low $NDV I$ events, averaged for all 1'228 low $NDV I$ events in the temperate biome. The events occur in JJA-0m on the very right of the diagram and are characterized by a precipitation deficit of about 1 standard deviation. This is clearly outside of the confidence interval of the hypothesis test that evaluates $H_{0, EV}$: “The meteorological storyline is not related to..."
First of all, event anomalies more than two years back (before JJA-24m) are mostly within the reference confidence interval, i.e., are not significantly different from the climatological reference. After that, periods of negative $P'_{90d}$ are highly prominent in both biomes (Fig. 5a,b). In the temperate biome, a first period of negative $P'_{90d}$ stretches over almost the entire year prior to events, and is significant at DJF-18m and MAM-15m (Fig. 5a). Interestingly, DJF-6m is wetter than usual with a peak anomaly of +0.4 that is significantly different from the reference climatology. Afterwards, $P'_{90d}$ drops abruptly and is significantly reduced in four months prior to events, reaching $-1.0$ during JJA-0m. In the Mediterranean biome during twenty months prior to events, $P'_{90d}$ is continuously negative and exerts distinct significant peak values (Fig. 5b). The $H_{0, EV}$ can further be rejected for $P'_{90d}$ in JJA-0m, when $P'_{90d}$ drops to about $-1.0$. All in all, distinct features of the events’ previous $P'_{90d}$ evolution are a dry JJA-0m, negative $P'_{90d}$ already in the previous year, and, in the temperate biome, a remarkable wet period in the cold season prior to the events.

In some contrast to $P'_{90d}$ after JJA-24m the meteorological storyline in the Mediterranean biome features positive $T2m'_{90d}$ in JJA and negative $T2m'_{90d}$ in DJF, although these anomalies are largely not statistically significant (Fig. 5d). In the temperate biome, $T2m'_{90d}$ is almost exclusively positive after JJA-24m and is unusually increased during SON-21m and SON-9m (Fig. 5c). The most prominent (and highly significant) warm anomaly of +1.2 and +0.8 arises during the three and four months prior to events in the temperate and Mediterranean biome, respectively (Fig. 5c,d). So a hot JJA-0m, and, in the temperate biome, warm periods in SON-21m and SON-9m emerge as key constituents of the meteorological storyline of low NDVI events.

Most of these distinct anomalies relate to a specific collocated weather system frequency anomaly. In Mediterranean forest regions, the climatological peak in 90-day averaged cyclone frequency, $f_{90d}(C)$, of $\sim10\%$ occurs from DJF to MAM (Fig. 5f). In particular in the 20 month-long dry period, $f_{90d}(C)$ is reduced twice during that peak phase by up to $-3\%$. The onset of that dry period in SON-21m is related to the strongest increase in anticyclone frequency during the entire meteorological storyline of $f'_{90d}(A) = +5\%$ (Fig. 5b,h). Further, $f_{90d}(A)$ is increased at the beginning and during the most intense phase of the significantly drier period from DJF-6m to JJA-0m. Similarly in the temperate biome, negative $f'_{90d}(C)$ co-occurs with the pronounced dry periods from SON-21m to JJA-12m and from MAM-3m to JJA-0m (Fig. 5a,e). Positive $f'_{90d}(C)$ only occurs during the intermittent wet period when the anticyclone frequency is reduced by $\sim-2\%$ (note the climatological value of 20–30%; Fig. 5g). As for Mediterranean forests, the three significantly drier periods in temperate forests in DJF-18m, JJA-12m, and JJA-0m correspond to peaks in $f'_{90d}(A)$ of +2 to +5%. The effect of the two weather systems on $T2m'_{90d}$ depends on the season. Anticyclones cause locally warmer conditions in JJA (e.g., in JJA-0m) and colder conditions in DJF (e.g., in DJF-18m), respectively (Fig. 5c,d,g,h). The effect of cyclones on $T2m'_{90d}$ is less systematic at the biome scale (cf. Fig. D1) and will be addressed amongst others in Sect. 3.2.4. All in all, collocated weather systems seem to have a clear imprint on the most anomalous periods of the meteorological storyline.
3.2.2 Persistence of dry and warm periods

The previous section also pointed to not necessarily intense but unusually persistent dry and warm periods, which we investigate in more detail in Figure 6. To do so we take the data displayed in Fig. 5 and compute the fraction of positive and negative anomalies of $T2m_{90d}'$ and $P'_{90d}$ from a single seasonal average to three years prior to the low $NDVI$ events (red line in Fig. 6). Moreover, analogously to Fig. 5, we contrast the respective fractions to values expected under the null hypothesis $H_{0, EV}$ that the fraction of dry and warm periods preceding the low $NDVI$ events (i.e., during $\Delta t$) are unrelated to these events (grey shading in Fig. 6; Sect. 2.4.3).

In the temperate biome, the 5-month persistence of the most recent dry period up to JJA-0m is not unusual (Fig. 6a), contrary to its magnitude (Fig. 5a). However, when considering two years before events, dry periods make up almost 80% of that time period, which is significantly different from the climatological expectation. Thus, the fact that 80% of the preceding two years are anomalously dry emerges as a further key aspect of the meteorological storyline of temperate biome low $NDVI$ events. Warm periods accumulate to a highly unusual degree when integrating over more than ten preceding months, and are interrupted by hardly any cold period (Fig. 6c). This characteristic is most distinct - i.e., most clearly rejects $H_{0, EV}$ - at a time scale of 18 months before low $NDVI$ events. So persistent and recurrent warm anomalies and an accumulation of dry periods

Figure 6. The percentage of the integration period with a biome-wide average (a,c) dry ($P'_{90d} \leq 0$) and (b,d) warm periods ($T2m'_{90d} \geq 0$) for decreasing integration period $\Delta t$ prior to low $NDVI$ events. The grey shading displays the 95% confidence interval (CI) of the reference climatology (see Sect. 2.4.3).
over the about two previous years are peculiarities of events in the temperate biome.

In the Mediterranean biome, warm periods are more frequent but not significantly so compared to the reference climatology (Fig. 6d), while dry periods accumulate to a highly unusual degree. When going back more than eight months, the frequency of dry periods is clearly unique to low NDVI events (Fig. 6b). The lowest p-value is reached 20 months prior to events, when the uninterrupted dry period starts to arise. So in the Mediterranean, prior to low NDVI events dry periods persist over 18 months and are a distinct feature of the events’ meteorological storyline considering time periods up to 31 months prior to the events. In contrast, the persistence of warm conditions appears not to be crucial for Mediterranean low NDVI events.

### 3.2.3 Event sequences

In addition to the biome-wide analyses of the previous sections, Figure 7 shows the meteorological storyline of two subgroups: two-year event sequences (EV11) and single events including their recovery (EV10). The two-digit binary code indicates whether there was an event in JJA-0m, and in JJA+12m, respectively. This time, we test the null hypothesis $H_{0, EV11}$ that the meteorological storyline of EV11 is the same as that of EV10 (blue shaded confidence interval in Fig. 7; Sect. 2.4.3). Note that the meteorological storyline of the two subgroups are fairly similar especially for the temperate biome as forest regions belonging to EV10 and EV11 can be neighbouring, and as EV11 mostly occur when EV10 are wide-spread too. The signal of EV10 stems to 42% from 2018/19 and 2019/20 when also 82% of EV11 occur. Nevertheless, the differences between the two can be instructive about why a second event occurs in EV11 but not in EV10. Moreover, the meteorological storyline of EV10 up to JJA-0m is very similar to that of all events (Fig. 5), as - due to the grouping - only sequential events ($n = 92$) and those in 2020 ($n = 186$) are excluded.

Several aspects point to a contrasting meteorological storyline during event sequences as compared to single events in temperate forests. Firstly, $P_{90d}'$ up to MAM-0m are similar for EV10 and EV11 (Fig. 7a), while $T2m_{90d}'$ is slightly increased for EV11 (Fig. 7c). In JJA-0m, when both groups experience an event, the hot-dry anomaly is more pronounced for EV11 events. Afterwards in EV10 hot and dry anomalies gradually weaken and DJF+6m to MAM+9m features a wet period of +0.4. The following JJA+12m exerts conditions close to climatology. In contrast, anomalies during EV11 are more pronounced and significantly different from EV10 from SON+3m to JJA+12m. While the wet period during the cold season peaks at +1.3, $P_{90d}'$ decreases again towards JJA+12m (Fig. 7a). $T2m_{90d}'$ is by about 0.5 larger in EV11 as compared to EV10 from SON+3m until JJA+12m (Fig. 7c). Interestingly, the meteorological storylines also contrast between the one event of EV10 (in JJA-0m), and in the second event of EV11 (in JJA+12m). Despite showing a larger wet anomaly in the previous cold season, the second event of EV11 goes along with less anomalous dry conditions in JJA+12 as compared to EV10 in JJA-0m. The same applies for the warm anomaly during JJA, and when comparing the meteorological anomalies of the second with the first event summer of EV11. These contrasts between the second event of EV11 and EV10 or the first event of EV11 in JJA are an indication of legacy effects, which are not related to the prevailing seasonal meteorology. What fosters the occurrence of another low NDVI event are, therefore, either the warm anomalies more than one year back, the hot-dry anomalies in the previous JJA, or
the warm-wet anomalies in the previous cold season, which were all significantly stronger for the second event of EV11 than for the first event or for EV10.

Figure 7. Two and a half year-evolution of (a,c) $P'_{90d}$ and (b,d) $T2m'_{90d}$ averaged over all low NDVI events in the subgroups EV10 (blue) and EV11 (brown) in (a,b) the temperate and (c,d) the Mediterranean biome. The blue shading displays the 95% confidence interval (CI) of the tested null hypothesis $H_0, EV_{11}$ (see Sect. 2.4.3). EV10 exhibits an event in JJA-0m, and EV11 does so in JJA-0m and JJA+12m, as indicated by the colored horizontal bars. Vertical grey lines show the end of each season, i.e., when the seasonal average value occurs.

Mediterranean forests show a clear sign of recovery in EV10 and fewer differences between the second event of EV11 and single events. The first clearly significant difference between the meteorological storylines of EV11 and EV10 arises in MAM-3m, when EV11 exerts much lower $T2m'_{90d}$ (Fig. 7d). Then in JJA-0m, i.e., during the first event, meteorological anomalies are not significantly different in EV11 with $(T2m'_{90d}, P'_{90d}) = (+0.6, -0.7)$ and in EV10 with $(T2m'_{90d}, P'_{90d}) = (+0.8, -0.8)$ (Fig. 7b,d). The meteorological storyline of EV10 after JJA-0m is characterized by more and more wet conditions, especially in MAM+9m, and only small $T2m'_{90d}$ (Fig. 7b,d). Contrarily, negative $P'_{90d}$ continues over all seasons prior to the second event of EV11 and is significantly different during MAM+9m and JJA+12m. In JJA+12m, both hot and dry anomalies are similar to those in the year before (Fig. 7b,d). So it appears that event recovery is aided by positive $P'_{90d}$ in the Mediterranean biome. The meteorological storyline of EV11 events significantly differs from that of EV10 events in terms of lower $P'_{90d}$ after the first event. Anomalies in that time period, i.e., prior to the second event of EV11 are, to some degree, similar to the meteorological anomalies of the initial event, which does not indicate strong legacy effects. In summary, for the temperate biomes we identify
signatures of potential legacy effects in both $T2m'_{90d}$ and $P'_{90d}$, while for the Mediterranean biome Fig. 7 suggests that legacy effects are less pronounced.

### 3.2.4 Bi-variate ($P'_{90d}, T2m'_{90d}$) and weather system frequencies

To conclude Sect. 3.2, we examine the bi-variate evolution of the meteorological storylines preceding low NDVI events and relate it to weather system frequencies. To do so we display histograms of the five ($T2m'_{90d}, P'_{90d}$) phase space compartments for the six seasons prior to low NDVI events in both biomes (Fig. 8). Analogously to the analyses presented in Figs. 5 & 6 we moreover test $H_{0,EV}$ that the occurrence of low NDVI events and the meteorological storyline (characterized here with the five compartments) are unrelated. Details about the spatial distribution of the seasonal phase space compartments are shown in Fig. C1. The meteorological storyline of most events is characterized by hot-dry conditions (Q2) in MAM-3m and JJA-0m (Figs. 8e,f,k,l). In both biomes, around +30% (MAM-3m) to +45% (JJA-0m) more forest regions experience Q2 than expected from the reference climatology. This mostly comes at the expense of cold-wet (Q4) and Q0 conditions, meaning that there is a diagonal shift in the ($T2m'_{90d}, P'_{90d}$) phase space from Q4 to Q2. In the Mediterranean biome in DJF-6m, dry conditions go along with colder than usual $T2m_{90d}$, and, thus, +20% more forest regions experience cold-dry conditions instead of Q1, manifesting a diagonal shift from Q1 to Q3 (Fig. 8j). For some regions, e.g., Greece or parts of Italy, Q2 conditions only occur in DJF-6m or MAM-3m and not, as during most events, in JJA-0m (see Fig. C1d-f).

As we have seen in earlier sections, many low NDVI events in temperate forests also experience wet periods in their meteorological storyline. For 20–30% of the events, mostly occurring in central to western Europe and the Baltic, DJF-6m and/or SON-9m are warm-wet (Figs. 8c,d & C1). At the biome scale, the increase of around +10% more Q1 in SON-9m is significantly different from the climatological reference. Note that this signal does not stem from a single event, but includes, for example, event years 2001, 2003, 2016, and 2018–2020 (Figs 3a & C1d). Further in the temperate biome, there are a few exceptional forest regions, which do not show hot-dry conditions in JJA-0m (Fig. C1f). This mostly relates to Q0 conditions in JJA 2020 affecting parts of western Europe, the Ukraine, and Russia. Also, a cold-wet JJA-0m occurred during the 2017 event in Russia. Lastly, in both biomes, cold-wet conditions are reduced over the entire meteorological storyline.

While the link of weather system frequency anomalies to $T2m'_{90d}$ and $P'_{90d}$ was rather straightforward at the biome scale (Sect. 3.2.1), the surface impact of weather systems is locally more nuanced. Details are provided in Figs. C1, D1 & E1, which show maps of the seasonal phase space compartment, the seasonal cyclone frequency anomaly $f'_{90d}(C)$, and the seasonal anticyclone frequency anomaly $f'_{90d}(A)$, respectively. Some important conclusions from these spatial considerations are:

- In the temperate biome, the prominent (warm-)wet seasons between SON-9m and MAM-3m are accompanied by large positive $f'_{90d}(C)$ in some regions (northern Europe, the Balkans), and by weak or negative $f'_{90d}(C)$ in other regions (France, eastern Europe).
Figure 8. Histograms of the seasonal phase space compartments Q0–Q4 at event grid cells in (a-f) the temperate and (g-l) the Mediterranean biome, respectively, during the six seasons prior to low NDVI events from MAM-15m to JJA-0m. Median and CI of the reference climatology are shown in black solid and dashed lines, respectively. Significantly anomalous frequencies are highlighted with a pink arrow indicating the direction of change.

- While hot-dry conditions in JJA-0m and JJA-12m mostly go along with positive $f_{90d}(A)$, $f_{90d}(A)$ in western Europe is reduced by up to $-10\%$.

- In continental eastern Europe and the Mediterranean, cold-dry winters and hot-dry summers - both peculiarities of the meteorological storyline - are caused by more frequent anticyclones.

All in all, there is large spatiotemporal variability in the coupling of synoptic-scale weather systems and surface meteorology, consistent with findings from previous studies (e.g., Hawcroft et al., 2012; Pfahl and Wernli, 2012a; Chan et al., 2019; Röthlisberger and Martius, 2019).
4 Discussion

4.1 Low NDVI events

Events of low NDVI as identified in this study typically represent summers with a drought-induced reduction in forest performance, as shown in earlier studies (Anyamba and Tucker, 2012; Orth et al., 2016; Buras et al., 2020). However, we also identify low NDVI events related to large-scale logging in 2009 (Senf and Seidl, 2021a), an ice storm in 2014 (Senf and Seidl, 2021b), or cold-wet conditions in Russia in 2017. Late frost is another potential source of low NDVI, which is not related to drought (Bascietto et al., 2018; Vitasse et al., 2019). Nevertheless, large events of drought-induced decline in forest performance are well-captured by our method. Exemplary, low NDVI events co-occurred with reduced forest growth and carbon uptake in 2003, and with premature leaf senescence, reduced productivity, and canopy mortality in 2018 (Ciais et al., 2005; Bastos et al., 2020b; Schuldt et al., 2020; Trotsiuk et al., 2020).

Contrasting our low NDVI events with forest disturbance data of Senf and Seidl (2021a), more specifically canopy mortality, helps to further characterize the nature of the identified events. Large-scale windthrow, such as that during the destructive storms Lothar, Gudrun, and Kyrill, do not appear in our event set because reflectance properties of the still green understory likely dominate the NDVI signal (McDowell et al., 2015). Nevertheless, fresh deadwood and partial tree damage from wind disturbances promote multiplication of bark beetles and, thus, may have a non-negligible effect on some of the events (Temperli et al., 2013; Jakoby et al., 2019). There is also conceptual discrepancy that arises from the fact that we only consider information at sufficiently forest-covered pixels with an area of \(\sim 5 \times 5 \text{ km}^2\), while \(D\) is originally assessed at every \(30 \times 30 \text{ m}^2\) pixel. Also, our aggregation method potentially ignores localized disturbances that affect less than half of the forest pixels within a \(\sim 50 \times 50 \text{ km}^2\) grid cell, and includes NDVI signals of only partly forest-covered pixels (Sect 2.3). So we acknowledge that the greenness-based event set identified here misses some forest disturbances that reduce forest performance. However, given the above-mentioned previous assessments and the affirming event evaluation, we are confident that the event set covers large-scale reductions in forest performance in Europe in 2000–2020.

4.2 Meteorological storyline

We first mention that the complex challenge of identifying periods when ecosystems are most sensitive to meteorology has been recognized for crops too, and was successfully tackled with storyline approaches (e.g., Goulart et al., 2021). For the forest-related low NDVI events, statistically more prevalent dry periods in the two preceding years, and the magnitude of concurrent hot-dry conditions are common features of the meteorological storyline in the two biomes. Both signals, especially the prevalence of dry conditions \((P_{90d} < 0)\) in the Mediterranean biome, are highly implausible when assuming no event-meteorology relationship at these lags. Continuously dry conditions are crucial and typical in the Mediterranean biome with potentially year-round growth (Camarero et al., 2021). Also, they play an important role for forest fires, which likely aggravate the meteorological impact on NDVI indirectly (Nagel et al., 2017; Turco et al., 2017). In both biomes, \(P_{90d}\) is reduced before \(T2m_{90d}\) is elevated during the event year. This might follow from the fact that, especially in southern Europe, soil moisture drought
in DJF and MAM increases the likelihood of a hot subsequent JJA through an enhanced soil moisture-atmosphere feedback (Seneviratne et al., 2010; Russo et al., 2019). Regardless, our results emphasize the damaging consequences of drought in the early growing season as highlighted by, e.g., Senf et al. (2020), Bigler and Vitasse (2021), and Bose et al. (2021). Despite the generally subordinate role of heat compared to drought, the identified extremely unusual accumulation of warm periods over up to three years prior to events in the temperate biome points to its indirect effects on insect populations and fire, as well as to the joint amplification of drought impacts (Seidl et al., 2017; Sommerfeld et al., 2018; Seidl et al., 2020; Forzieri et al., 2021).

In the temperate biome, meteorological drought is interrupted by a wetter than usual period in the cold season. Warm and wet conditions in SON-9m and DJF-6m, respectively, could extend the growing season of the previous year. Favorable growing conditions at the margins of the growing season can exacerbate the impact on following droughts, which was shown for an early start into the growing season (Buermann et al., 2013; Bastos et al., 2020a). Furthermore, a wet-warm cold season can relate to more rain, lower snow cover, and earlier snow disappearance, which favors spring drought in forests (Buermann et al., 2013; Blankinship et al., 2014; Harpold et al., 2015) and forest fire activity (Westerling, 2016). Given the prevalent dry periods in the previous year and the extended growing season, soil moisture drought likely penetrated to deeper levels (Barnard et al., 2021). Alternatively, the surplus in winter precipitation could locally foster growth in the dormant season or spring and, thus, promote larger canopy development which can aggravate the impact of the following summer drought (Bastos et al., 2020a). Both mechanisms suggest that the warm-wet interruption of the meteorological drought fosters soil moisture drought in the subsequent growing season, which is itself again characterized by low precipitation. Additionally, increased cyclone frequencies alongside warm-wet conditions possibly enhance windthrow and could indirectly foster bark beetle multiplication (Temperli et al., 2013; Biedermann et al., 2019). These results, therefore, suggest a negative legacy effect of warm-wet DJF-6m, in addition to drought legacy discussed in Sect. 4.4. In the context of the drier JJA-12m and JJA-0m, these wet conditions also relate to large seasonal variability in precipitation, which has been identified as driver of tree mortality in Europe (Neumann et al., 2017).

### 4.3 Weather systems perspective

Our results highlight, amongst others, that the timing and positioning of weather systems is crucially determining their impact on surface meteorology relevant for low \(NDVI\) events. The 20-month long dry period in the Mediterranean biome is accompanied by reduced seasonal cyclone frequency, especially in DJF and MAM. Cyclones are the main contributor to cold season precipitation in these forest regions (Rüdisühli et al., 2020), and also to extreme precipitation (Pfahl and Wernli, 2012b). On the other hand, more frequent anticyclones in JJA-0m likely relate to an upper-level subtropical ridge extending into the Mediterranean, which is known to be important for heat extremes (Sousa et al., 2018; Zschenderlein et al., 2019). The combination of absent cyclones and more frequent anticyclones enables hot-dry conditions as during the events’ meteorological storyline.

The distinct warm-wet anomaly in the temperate biome during SON-9m to MAM-3m is related to more frequent cyclones. In particular, regions such as Scandinavia experience a wet anomaly with a collocated increase in cyclones corresponding to a
northward shift in the cyclone steering jet stream (Messori et al., 2022). Some other regions in central Europe show weak or no increase in cyclone frequency, as they typically receive precipitation from fronts located outside the cyclone center (Rüdisühli et al., 2020). Hot-dry conditions during JJA are partly related to more frequent anticyclones. The reduced anticyclone frequency in JJA-0m in western Europe could relate to the fact that summer precipitation frequently occurs in high-pressure systems (Rüdisühli et al., 2020), i.e., a more northerly displaced jet stream favors reduced greenness in that region (Messori et al., 2022). An anticyclone centered over Europe often goes along with convective precipitation in its moist and unstable inflow over southwestern Europe (Mohr et al., 2020). Contrarily, continental to northern Europe is a region where stable European anticyclones and upper-level blocking typically cause heat waves and precipitation suppression (Pfahl and Wernli, 2012a; Zschenderlein et al., 2019). These considerations highlight the importance of weather systems and the necessity of considering their spatiotemporally varying impact on surface meteorology, also when interested in events of substantial forest impact.

### 4.4 Drought legacy effects and event recovery

In Sect. 3.2.3 we compared single events to sequences of events and found indications of drought legacy effects in the temperate biome. Drought legacy effects are known to reduce tree-level growth (Anderegg et al., 2015), and, more broadly, resilience (Bose et al., 2020). Thus, the previous drought-induced low $\text{NDVI}$ event could be a sign of reduced tree vigor and higher susceptibility for a following drought, as detected, e.g., in 2019 (Schuldt et al., 2020; Bastos et al., 2021). Moreover, forests potentially acclimate to increased drought stress, which might also lead to reduced leaf area and productivity (Gessler et al., 2020). However, drought legacy is not always reflected in $\text{NDVI}$ (Kannenberg et al., 2019). Furthermore, it is important to remember that $P_{\text{90d}}' \leq 0$ does not necessarily indicate soil moisture drought, nor critically low precipitation sums due to normalization (Zang et al., 2020, see also Sect. 4.5). Nevertheless, our meteorological findings support that increased vulnerability to drought sequences can out-compete acclimation (Anderegg et al., 2020), and also that drought legacy effects - at least partly - reflect in $\text{NDVI}$. They further suggest that temperate forests are more sensitive to droughts in consecutive years, while Mediterranean forests show a decrease in sensitivity to the second drought in agreement with Anderegg et al. (2020).

During event recovery in both biomes, hot-dry conditions as prior to low $\text{NDVI}$ events are clearly absent. In the Mediterranean biome, seasonal precipitation is around normal in DJF and increased in MAM, which are periods of increased drought-sensitivity (Bose et al., 2021; Camarero et al., 2021). In the temperate biome, it seems that average seasonal precipitation - also in the more drought-sensitive MAM (Bose et al., 2021) - allows for forest recovery after a low $\text{NDVI}$ event. This is also supported by the fact that low $\text{NDVI}$, potentially by defoliation, can be followed by re-greening instead of mortality, depending on subsequent growing conditions (Dobbertin and Brang, 2001; Kannenberg et al., 2019; Rohner et al., 2021).

### 4.5 Caveats

The main caveat of this study is the short data record, which implies that we have to (i) aggregate over large areas to get robust results, and (2) frequently perform normalization as opposed to superior statistical modelling (Sects. 2.3 & 2.4.1). The former implies that our analyses can neither account for species-specific drought responses (Scherrer et al., 2011; Vanoni et al., 2022).
2016), nor for the multi-dimensional nature of tree mortality (Allen et al., 2015; Etzold et al., 2016; Schuldt et al., 2020). The link between drought, drought response, and tree mortality is mediated by site, stand, and tree properties (Etzold et al., 2019; Vitasse et al., 2019; Wohlgemuth et al., 2020), and can further be shaped by tree species diversity within a forest (Grossiord et al., 2014) and its micro-climate (Buras et al., 2018). Also, using NDVI signals at pixels with up to 50% non-forest surface cover introduces some bias. Increasing this threshold, however, does not change the identified events hot spots, and unfavorably decreases the sample size of our analyses (not shown). The large-scale approach of the study has proven useful to connect impaired forest performance with atmospheric dynamics acting at a scale much larger than the stand level. Nevertheless, the results of this study should be confronted with more specific and local impact assessments.

We rely on normalization to compare NDVI anomalies and the meteorological storylines of the relatively scarce events over space and time. The normalization of NDVI is less problematic as it serves a pragmatic, site-specific identification of low performance, and was additionally compared to an independent data set. $P'_{90d}$ and $T'_{2m90d}$, however, do not necessarily represent the actual site-level temperature and precipitation values and their interpretation requires care (Zang et al., 2020). This specifically applies when comparing the meteorological storyline of the temperate with the Mediterranean biome, respectively, as the latter climatologically receives little precipitation during summer (Schultz, 2005). The use of basic meteorological variables instead of a more complex drought index is motivated by the overarching research question of this study, which is to characterize the meteorological storyline of these events.
5 Conclusions

This study identified specific aspects of the meteorological storyline (the three-year evolution of seasonal temperature, $T^{2m}_{90d} \; ^\prime$, and precipitation anomalies, $P^\prime_{90d}$), which is characteristic for events of low summer forest greenness at the 50 km scale in Europe in 2000–2020. Forest greenness as measured by the $NDVI$ is used to detect reduced forest performance - an early warning mechanism of forest decline (Buras et al., 2021). In temperate forests, low $NDVI$ events are preluded by extraordinarily persistent warm periods, and an unusual accumulation of dry periods over two years. Most interestingly, the preceding cold season specific to such events is warm-wet, and - as we hypothesize - potentially exerts a negative legacy effect on the following summer drought. Dry and hot anomalies are both statistically significant during spring and summer leading up to the identified events, respectively. While the warm-wet cold season anomalies are more pronounced prior to the second event of a two-year event sequence, the following growing season is less hot and dry as compared to single events, which suggests a potential role of legacy effects from the previous event. In the Mediterranean biome, a quasi-uninterrupted 1.5 year-long seasonal precipitation deficit is observed prior to low $NDVI$ events. These dry conditions in the Mediterranean go along with colder $T^{2m}_{90d}$ in winter and warmer $T^{2m}_{90d}$ in spring and summer. Closer to the event summer, negative $P^\prime_{90d}$ and positive $T^{2m}_{90d}$ are characteristically pronounced over the preceding nine and four months, respectively. In two-year event sequences, the seasonal precipitation deficit persists over an additional year while variations in $T^{2m}_{90d}$ seem to play a subordinate role.

Moreover, we provide clear evidence on the spatially varying impact of synoptic-scale weather systems over time periods of up to 2.5 years on the key characteristics of the meteorological storylines. At the biome scale, persistent drought is caused by a continuously reduced cyclone frequency, and only the intermittent wet period in temperate forests relates to more frequent cyclones. In contrast, increased anticyclone frequency relates to the onset and intensification of negative $P^\prime_{90d}$. This effect can, however, differ at a local scale, depending on which weather system is locally relevant for precipitation. For instance, western Europe often receives summer precipitation from convective cells in anticyclones and, thus, hot-dry conditions in the event summer go along with reduced anticyclone frequency.

The important differences between the meteorological storyline impacting temperate and Mediterranean forests identified in this study provide a better understanding of European forests’ response to multi-seasonal meteorology. We could, for example, bring forward meteorological indications of drought legacy effects in the temperate forests. Finally, the presented systematic investigations bridge the gap between forest dynamics and atmospheric dynamics, and, thereby, constitute progress in how expected forest decline can be linked to changing meteorological and climatic conditions under global warming.
Data availability. The data sets used in this study are freely available, namely monthly MODIS-Terra downloaded from NASA Giovanni database (https://giovanni.gsfc.nasa.gov/giovanni/), global forest cover area by Hansen et al. (2013, https://glad.earthengine.app/view/global-550-forest-change), and atmospheric fields of ERA5 from the ECMWF (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=form). We use the updated version 1.1 of forest disturbance data from Senf and Seidl (2021a) and aggregate the data to the ERA5 grid. Version 1.0 is available at https://doi.org/10.5281/zenodo.3925446.

Code and data availability. At https://github.com/corneliussenf/AggregateDisturbancesERA5 the code used to aggregate forest disturbance data can be accessed. All other code is available upon request.

Appendix A: Bootstrapping tests

For null hypothesis \( H_{0, EV} \), we generate 10'000 synthetic event sets \( EV_{J,n}^{r} \) by randomly shuffling the 21 annual chunks of \( EV_{J,n} \) (Sect. 2.3), i.e., by assigning a random year \( n^{r} \) without replacement to all events in year \( n \). The synthetic meteorological storylines (i.e., time series of normalized meteorological anomalies \( T2m_{90d}^{r} \) and \( P_{90d}^{r} \)) are generated by extracting ERA5 fields for \( EV_{J,n}^{r} \). They are then used to identify the corresponding phase space compartment \( Q^{r} \). As for the actual low NDVI events, we compute for each of the 10'000 synthetic event sets the event mean \( T2m_{90d}^{r} \), the event mean \( P_{90d}^{r} \), and the frequency \( f(Q) \) among all events for every phase space compartment. The distribution spanned up by these 10'000 synthetic values gives us the reference climatologies \( T2m_{90d}^{r} \), \( P_{90d}^{r} \), and \( Q^{r} \). Note that the shuffling of years is done prior to extracting the spatial fields of \( T2m \) and \( P \) from the ERA5 data set. This has the convenient effect that spatial correlation in these two meteorological variables is retained and synthetic meteorological storylines are, thus, constructed from a data set with exactly the same spatial correlation of \( T2m \) and \( P \) as the original data set. We then compare \( T2m_{90d}^{r}, P_{90d}^{r}, \) and \( f(Q) \) of low NDVI events to the 10'000 \( T2m_{90d}^{r}, P_{90d}^{r}, \) and \( Q^{r} \), respectively. Values outside the range of \( T2m_{90d}^{r}, P_{90d}^{r}, \) and \( Q^{r} \) receive a p-value of 0 (Röthlisberger et al., 2016). The remaining p-values are estimated from the percentiles of the synthetic data set. At the significance level of \( \alpha = 5\% \), \( H_{0, EV} \) is rejected at time lags \( \Delta t \) if the event value is outside the 95% confidence interval of the reference climatologies.

For null hypothesis \( H_{0, EV11} \), we generate 10'000 synthetic event sets \( EV10_{J,n}^{r} \) by drawing samples with \( n_{EV11} = 73 \) (19) from the original event set \( EV10_{J,n} \) with \( n_{EV10} = 971 \) (282) in the temperate (Mediterranean) biome. \( T2m \) and \( P \) data is again extracted 10'000 times for \( EV10_{J,n}^{r} \) yielding a reference of EV10-type meteorological storylines. The \( H_{0, EV11} \) (that there is no difference in \( T2m_{90d}^{r} / P_{90d}^{r} \) between EV10 and EV11) is rejected for all time lags \( \Delta t \) at which the events \( T2m_{90d}^{r} / P_{90d}^{r} \) lies outside the 2.5th – 97.5th percentile of the 10’000 reference values.
Appendix B: Low NDVI events and forest disturbances

Figure B1. Disturbance anomaly $D'$ at all overlapping forest grid cells in selected years (shading). The stippling indicates low NDVI events during JJA. Tracks of cyclones mentioned in the text, which occurred in the previous autumn or winter, are shown in black, with 3-hourly core pressure in [hPa] in colored dots. The location and extent of the cyclone at time of lowest core pressure is shown in cyan.
Figure B1 presents forest disturbance anomalies $D'$ (Sect. 2.5) for selected years together with low NDVI events (Sect. 2.3). In years of widespread low NDVI events, namely 2018, 2020, 2019, and 2003, the event coverage largely agrees with regions of positive $D'$ (Fig. B1b,f-h). As noted in Sect. 3.1.2, an important source of discrepancy between the two data sets is windthrow. For example, storm Lothar in December 1999 caused catastrophic damage reflecting in extremely high $D'$ over parts of France, Germany, and Switzerland in the following year (Fig. B1a; Usbeck et al., 2010). Other examples were the devastating storms Gudrun affecting southern Sweden in January 2005, Kyrill in Central Europe in January 2007, and Vaia in northern Italy in October 2018 (Valinger and Fridman, 2011; Senf and Seidl, 2021b). The latter manifests in low $D'$ in the following year 2019 (Fig. B1g). Similar to our data set, the acquisition period of $D$ was in summer and increased canopy mortality following autumn or winter storms is reflected in the subsequent year (Senf and Seidl, 2021a). The tracks of the four cyclones in their month of occurrence are, thus, shown for the following summer (Fig. B1a,c,d,g). Regions of large $D'$ are typically located at the southern or eastern flank of the cyclones, where typically their fronts propagate, at around the time when they reach their lowest core pressure. Lastly, human influence is another source of disagreement between $D$ and low NDVI, as, for example, salvage logging following wind disturbances in the Gascony in southwestern France in 2009 (Fig. B1e; Senf and Seidl, 2021a).
Appendix C: Seasonal phase space compartment

Figure C1. As Fig. 3a for the seasonal \((T_{2m}^{90d}, P_{90d}^{90d})\) phase space compartment in the six seasons preceding low NDVI events. See Fig. 8 for biome-wide averages and statistical evaluation.
Appendix D: Seasonal cyclone frequency anomalies

Figure D1. As Fig. 3a for seasonal cyclone frequency anomalies $f_{\text{mod}}(C)$ in the ten seasons preceding low $NDVI$ events. For biome-wide averages see Fig. 5.
Appendix E: Seasonal anticyclone frequency anomalies

Figure E1. As Fig D1 for seasonal anticyclone frequency anomalies $f'_{90d}(A)$. For biome-wide averages see Fig. 5.
Author contributions. MH performed most of the analyses and wrote a first version of the manuscript. CS contributed data to the study. All authors contributed to the design of the study, the interpretation of the results, and the writing.

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


