ASCMO

Open Access

Spatial patterns and indices for heat waves and droughts over Europe using a decomposition of extremal dependency

Svenja Szemkus and Petra Friederichs

Institute of Geosciences, University of Bonn, Bonn, Germany

Correspondence: Svenja Szemkus (sszemkus@uni-bonn.de)

Abstract. We present a method for the analysis and compact description of large-scale multivariate weather extremes. Spatial patterns of extreme events are identified using the tail pairwise dependence matrix (TPDM) proposed by Cooley and Thibaud (2019). We also introduce the cross-TPDM to identify patterns of common extremes in two variables. An extremal pattern index (EPI) is developed to provide a pattern-based aggregation of temperature. A heat wave definition based on EPI is able to detect the most important heat waves over Europe. As an extension for considering simultaneous extremes in two variables, we propose the threshold-based EPI (TEPI) that captures the compound character of spatial extremes. We investigate daily temperature maxima and precipitation deficits at different accumulation times and find evidence that preceding precipitation deficits have a significant influence on the development of heat waves and that heat waves often co-occur with short-term drought conditions. We exemplarily show for the European heat waves of 2003 and 2010 that TEPI is suitable for describing the large-scale compound character of heat waves.

1 Introduction

Extreme weather events over Europe, such as heat waves and droughts, but also heavy rainfall events have repeatedly attracted attention in recent years due to their dramatically high impact on socioeconomic systems. The most prominent European heat waves in the 21st century occurred in 2003, 2010, and 2018. The high temperatures, which lasted for a period of several weeks to months, in combination with compound drought periods, led to disastrous effects on socioeconomic and local natural systems, such as an increased number of wildfires, crop failure, and increased health-related deaths (see, e.g., Liu et al., 2020; Fischer and Schär, 2010). For instance, according to WHO, more than 70 000 additional deaths in 12 European countries can be attributed to the European heat wave in 2003 (Anderson et al., 2004), which was very likely the hottest summer in 500 years (Luterbacher et al., 2004).

Several studies indicate that an increase in the global mean temperature and its variability leads to more frequent heat wave and drought events that are characterized by longer du-

rations and greater severity (Gershunov and Douville, 2009; Dai et al., 2018; Guerreiro et al., 2018; Della-Marta et al., 2007). This is also confirmed in the recent IPCC report (IPCC, 2021), saying that European hot extremes (including heat wave) have significantly increased in intensity and number since 1950. Particularly severe heat waves in Europe occurred in recent years in 2018, 2019, and 2022, often in connection with drought conditions. The 2022 heat wave resulted in record-breaking temperatures in France, Spain, the United Kingdom, and Germany (Rousi et al., 2023). Our understanding of the underlying dynamics of heat waves and, consequently, the changes that can be expected under climate change have improved considerably in recent years. In contrast, the dynamics and underlying structure of co-occurring temperature and drought events are much less understood. However, studies show that concurrent drought events can strongly influence heat extremes (Zscheischler and Seneviratne, 2017). For example, for the heat wave of 2003, it has been demonstrated that the co-occurring drought was a key contributor to the extreme temperature anomalies (e.g., Fischer et al., 2007; Black et al., 2004; Ferranti and Viterbo,

2006). Therefore, when considering potential changes under climate change, joint consideration of heat and drought extremes is necessary.

Several definitions and indices for drought and heat waves can be found in the literature. Heat wave indices usually include daily maximum temperature exceeding a certain threshold for several consecutive days (Robinson, 2001). Standard indices for heat waves are, e.g., the heat and cold wave index (HCWI) (Lavaysse et al., 2018), which is implemented in the Copernicus European Drought Observatory or the heat wave duration index (HWDI) and has been recommended by the IPCC (Griggs and Noguer, 2002). Standard meteorological drought indices are, e.g., the Palmer drought severity index (PDSI) (Palmer, 1965) and the standardized precipitation index (PSI) (McKee et al., 1993). A recent, promising approach to identifying heat waves and droughts is presented by Schädler and Breil (2021). They base their evaluation on so-called regional climate networks (RCNs) and identify extremes over Germany. Indices for compound heat waves and drought are less common in the literature. Recent approaches to index definition include the dry-hot magnitude index (Wu et al., 2019) or the standardized compound event indicator (Hao et al., 2019).

Detection and attribution (D&A) provides a powerful statistical tool for analyzing climate change's complex causes and effects. Reviews of the application of D&A to meteorological data are given in Slater et al. (2021) and Ummenhofer and Meehl (2017). In the high-dimensional space of the climate system, both the natural variability of the system and some internal variability overshadow the signals of climate change (Fischer and Knutti, 2016). Therefore, local changes, especially on smaller scales, only become visible if the data are analyzed in a compressed form. Data compression is therefore of great importance for a successful and reliable D&A study.

There exists a variety of statistical methods of different complexity for information compression. For example, a relatively simple approach is to aggregate the data spatially, but this usually neglects the spatial dependence structure of the underlying extreme values. Other techniques may include filtering using Fourier or wavelet transforms or adaptive data methods such as principal component analysis (PCA). However, most data compression methods focus on the description of the bulk of the underlying distribution, as described by first- and second-order moments, and provide little to no information about the tail of the distribution. Adaptive data methods have been widely used to study the spatial structure of heat waves. For example, Stefanon et al. (2012) use a clustering approach to identify six common heat wave patterns over Europe. Schubert et al. (2014) identify and analyze spatial patterns related to precipitation and temperature based on rotated empirical orthogonal functions (EOFs) over Russia. The studies agree that the patterns identified are consistent with large-scale blocking anticyclones associated with Rossby waves.

For the D&A of extreme events, mathematical methods are needed that achieve a targeted description of extremes. The framework in which these methods operate is the multivariate extreme value theory (MEVT) with broad applications in mathematical and meteorological research (e.g., Beirlant et al., 2006; Coles et al., 2001). New approaches combine MEVT with existing information compression methods. The rich experience with these methods in atmospheric research allows their targeted application to extreme events. For example, Morris et al. (2019) use an approach based on empirical basis functions to model the spatial dependence of low-rank max-stable processes. Janßen and Wan (2020) and Saunders et al. (2021) propose clustering approaches combined with an extremal framework, and Goix et al. (2017) also use a clustering approach to design a statistical learning procedure for anomaly detection based on a dataset with standard Pareto margins. Boulaguiem et al. (2022) combine generative adversarial networks with extreme value theory to model spatial dependencies in temperature and precipitation extremes. This shows that some of these methods are suitable for the description of composite extrema in two variables. For example, Vignotto et al. (2021) use an extreme value dependence clustering approach to identify spatial patterns of situations where precipitation and wind extremes are very likely to occur simultaneously.

Another promising approach comes from Cooley and Thibaud (2019). They propose the tail pairwise dependence matrix (TPDM), which summarizes extremal dependence using a regular varying framework and a threshold-based estimator. They build on the work of Larsson and Resnick (2012), who first defined this matrix for the bivariate case and show that the TPDM has some favorable properties that it shares with the covariance matrix so that the TPDM can be used as an analog to the covariance matrix for extremes. Jiang et al. (2020) and Rohrbeck and Cooley (2023) propose a PCA study using the TPDM with applications to daily precipitation over the United States (Jiang et al., 2020) as well as historical river flow levels across northern England and southern Scotland (Rohrbeck and Cooley, 2023). Drees and Sabourin (2019) also investigate a PCA based on the TPDM but focus more on describing the mathematical background. Finally, Fix et al. (2021) use the TPDM to propose a new model for spatial dependencies inspired by the simultaneous autoregressive (SAR) model with application to a gridded dataset of daily precipitation over Colorado.

In this paper, we introduce the extremal pattern index (EPI) for spatial weather extremes. The EPI is constructed based on the leading patterns derived from a PCA for extremes, as proposed by Jiang et al. (2020). Climate indices based on PCA, e.g., for the North Atlantic Oscillation (Barnston and Livezey, 1987), are already well-established tools in climate monitoring and analysis. By replacing the covariance matrix with the TPDM, we shift our focus to identifying extreme weather patterns, such as heat waves and droughts, while leaving the PCA procedure basically unchanged.

We use daily maximum 2 m temperature (maxT2m) from COSMO-REA6 reanalysis to derive the TPDM and identify spatial patterns associated with temperature extremes. Its EPI describes the intensity and spatial extent of heat waves in the area under consideration. To test the suitability of the EPI for describing heat waves, we focus on the two European heat waves in 2003 and 2010, about which many publications already exist. We compare our findings with existing literature.

We further explore the potential use and benefits of the TPDM in describing co-occurring heat waves and meteorological droughts. We again use maxT2m and define precipitation deficit (PD) as the additive reciprocal of daily precipitation accumulated over periods ranging from 11 to 47 d. To identify the pairwise extremal dependence between both variables, we introduce the cross-TPDM and provide an estimator thereof. We decompose the estimator of the cross-TPDM using a singular value decomposition (SVD), providing insight into the co-occurrence of extreme spatial patterns in two variables. Finally, we apply an analogue definition of the EPI using the SVD of the cross-TPDM. In order to target the index to co-occurring events, a modulated threshold-based estimator of EPI (TEPI) is defined.

The remainder of the paper is organized as follows. First, the theoretical framework of regular variation underlying the definition of the TPDM is presented in Sect. 2. Here we also discuss the estimation of the TPDM and its decomposition. Accordingly, we introduce the cross-TPDM in Sect. 3. We then discuss their evaluation (Sect. 4) and the definition of the various extremal pattern indices (Sect. 5). Section 6 discusses the data and some pre-processing steps. The results are presented in Sect. 7 with a discussion first on the temperature extremes and the EPI to define heat waves and then the compound heat wave and drought events. The paper concludes with a summary and discussion in Sect. 8.

2 The tail pairwise dependence matrix (TPDM)

2.1 Definition of the TPDM

The definition of the TPDM is based on regularly varying random variables. Regularly varying variables are defined by the tail behavior of their distribution function $F(x)$ in the limit $x \to \infty$. A summary on regular variation is given by Resnick (2007) saying "regularly varying functions are those functions which behave asymptotically like power functions". Let X be a non-negative random variable with positive and measurable distribution function $F(x)$ on $\mathbb{R}_+$. We call F to be a regularly varying function with tail index $\rho \in \mathbb{R}$ (denoted by $F(X) \in RV_\rho$), if for all $t > 0$

$$\lim_{x \to \infty} \frac{F(xt)}{F(x)} = t^\rho. \tag{1}$$

If the tail function, $\overline{F}(X) = 1 - F(X)$, is a regularly varying function, i.e. $\overline{F}(X) \in RV_{-\rho}$ ($\rho > 0$), then X is a regularly varying random variable with tail index ρ (Resnick,

2007, Theorem 3.6), and we write $X \in RV_\rho$. If X is Fréchet-distributed with $F(x) = \exp(-x^{-\alpha})$ and $x \geq 0$, then X is a regularly varying random variable with tail index α (Resnick, 2007).

Now let X be a positive multivariate random variable of dimension p with Fréchet-distributed margins and tail index α. In the following, we denote $X \in RV_\alpha^+$. For multivariate regularly varying random variables, it is convenient to transform the random vector to generalized polar coordinates with the angular component ω and the radial component r. With any norm $|| \cdot ||$ we define the unit sphere $\mathbb{S}_+^{p-1}$ as

$$\mathbb{S}_+^{p-1} = \left\{ x \in \mathbb{R}_+^p : ||x|| = 1 \right\} \tag{2}$$

and the polar coordinate transform as

$$x \to \left(||x||, \frac{x}{||x||} \right) =: (r, \omega) \quad \text{for } ||x|| > 0, \tag{3}$$

with $r = ||x||$ and $\omega = x/||x||$. If X is regularly varying, then its radial and angular components become conditionally independent for $r \to \infty$.

To understand the derivation and definition of the TPDM, we explore the properties of a regularly varying random variable. This exploration is done through the lens of weak convergence (Resnick, 2007), restricted to the special case of vague convergence, denoting convergence of Radon measures. Given a multivariate, Fréchet-distributed random vector $X \in RV_\alpha^+$, the definition of regular variation says that there is a sequence b_n with $n \to \infty$ and a limit measure $\nu(\cdot)$, such that we can write $n\mathbb{P}\left[X/b_n \in \cdot \right] \xrightarrow{\nu} \nu(\cdot)$ as $n \to \infty$. Here, $\xrightarrow{\nu}$ denotes vague convergence in $M_+([0, \infty]^p \setminus \{0\})$, the space of Radon measures on $(0, \infty] \setminus \{0\}$. After the polar transform, one can specify $\nu(\mathrm{d}r \times \mathrm{d}\omega) = \rho r^{-\rho-1} \mathrm{d}r \mathrm{d}H(\omega)$ where ρ is the tail index of X, and H is some Radon measure, referred to as the angular measure on $\mathbb{S}_+^{p-1}$. It can be interpreted as a probability measure for ω (for a detailed description see, e.g., Larsson and Resnick, 2012; Cooley and Thibaud, 2019; Resnick, 2007).

Larsson and Resnick (2012) found that the measure H, on the unit sphere $\mathbb{S}_+^{p-1}$, holds information about extremal dependencies. They define the extremal dependence measure of any regularly varying, bivariate random vector based on its angular components after polar transformation.

Cooley and Thibaud (2019) expand the original approach to higher dimensions. They concentrate on Fréchet-distributed variables with tail index $\alpha = 2$ and choose the L_2 norm when transforming into polar coordinates. They define the tailed pairwise dependence between the ith and jth element of X, using its angular component ω as

$$\sigma_{ij} = \int_{\mathbb{S}_+^{p-1}} \omega_i \omega_j \mathrm{d}H(\omega). \tag{4}$$

The TPDM (Σ_X) is then a $p \times p$ matrix, which summarizes the pairwise extremal dependencies of X for all mar-

gins $i, j = 1, \ldots, p$: $\Sigma_X = (\sigma_{ij})_{i,j=1\ldots p}$. Cooley and Thibaud (2019) show that by their choice of parameters, the properties of the TPDM are very close to those of the covariance matrix. The most important property for our application is that the TPDM is symmetric and positive definite with real and positive eigenvalues. For rigorous derivations and proofs as well as additional literature, the reader is referred to Cooley and Thibaud (2019).

2.2   Estimation of the TPDM

Let x_t, $t = 1, \ldots, n$ be a sample of a p-dimensional random vector X with Fréchet margins and sample components $x_{t,i}, i = 1, \ldots, p$. The estimator of the individual elements of the TPDM is given by Jiang et al. (2020) as

$$\hat{\sigma}_{ij} = 2 n_{ij,\text{exc}}^{-1} \sum_{t=1}^{n} w_{t,i} w_{t,j} \mathbb{I}_{(r_{t,ij} > r_{0,ij})}. \tag{5}$$

with $r_{t,ij} = \sqrt{x_{t,i}^2 + x_{t,j}^2}$ as well as $w_{t,i} = x_{t,i}/r_{t,ij}$ and $w_{t,j} = x_{t,j}/r_{t,ij}$. $r_{0,ij}$ is a threshold that can be set individually, with the index function $\mathbb{I}$ ensuring that only observations larger than $r_{0,ij}$ are considered for the estimation, and $n_{ij,\text{exc}} = \sum_{t=1}^{n} \mathbb{I}_{(r_{t,ij} > r_{0,ij})}$ is the number of elements which are greater than $r_{0,ij}$. The positive definiteness of the TPDM is not guaranteed by the proposed estimator in Eq. (5). If the estimator is not positive definite, we follow Cooley and Thibaud (2019) by computing the nearest correlation matrix (Higham, 2002).

Pre-processing is required when using meteorological data. Again we follow Jiang et al. (2020) and transform the marginal distribution of the original data to Fréchet margins with tail index $\alpha = 2$. Let X^{orig} be the original random vector with marginal distribution $F_i^{\text{orig}}(X), i = 1, \ldots, p$ and X the transformed random vector with Fréchet margins $F_i(X) = \exp(-x_i^{-\alpha})$. Due to the conservation of probability, we have to ensure that $F_i(X) = F_i^{\text{orig}}(X^{\text{orig}})$, which leads us to the transformation

$$X_i = \left(-\log F_i^{\text{orig}}(X^{\text{orig}})\right)^{-1/\alpha}. \tag{6}$$

As an estimator for F_i^{orig}, we may either use the empirical distribution function or fit a parametric distribution (e.g. normal distribution).

2.3   Decomposition of TPDM

The eigenvalue decomposition (EVD) of a covariance matrix is a standard tool in climate data analysis to reduce the dimensionality of high-dimensional weather and climate data (see, e.g., Wallace et al., 1992). The eigenvectors (EVs) form an orthogonal, complete base system. Here we use an estimator based on extreme dependencies instead of covariances, which leads to a targeted consideration of extremal dependencies. For an introduction to EVD and its applications,

see, for example, Wilks (2011), Jolliffe (1986), and Björnsson and Venegas (1997). The decomposition of the TPDM, with $\Sigma_{XX} = U \Lambda U^T$, provides the eigenvector matrix U with eigenvectors $u^{(k)}, k = 1, \ldots, p$ and components $u_1^{(k)}, \ldots u_p^{(k)}$, as well as the diagonal-matrix Λ, containing the eigenvalues $\lambda_1, \ldots \lambda_p$.

The TPDM is defined on the basis of positive real numbers. The transformation to Fréchet margins leads to the input variable X being consistent with this assumption. However, the EVs of the TPDM live by construction in $\mathbb{R}^p$. For the EVs to form an orthogonal basis in $\mathbb{R}_+^p$, we have to transform the EVs from $\mathbb{R}^p$ to $\mathbb{R}_+^p$. We again follow Cooley and Thibaud (2019) and use $\tau(y) = \log(1 + \exp(y))$ with its back-transform $\tau^{-1}(y) = \log(\exp(y) - 1)$. The Fréchet distribution, as well as the orthogonality of the EVs, is preserved under this transformation. We transform the EVs component-wise with $e_j^{(k)} = \tau(u_j^{(k)})$, for each of the $j = 1, \ldots, p$ components and $k = 1, \ldots, p$ vectors. The time series of principal components (PCs), in contrast, are calculated by a projection of the back-transformed data X on the respective EV as

$$\boldsymbol{\eta} = U^T \tau^{-1}(X). \tag{7}$$

We can reconstruct a filtered data set for each sample x_t ($t = 1, \ldots, n$) of the p-dimensional input variable X using $\tilde{x}_t = \tau(\eta_t U)$.

3   The cross-TPDM

3.1   Definition of the cross-TPDM

The TPDM presented above provides information on the pairwise dependencies between the margins of a random vector. Here we introduce the cross-TPDM as a measure of extremal dependency between two random vectors, such as spatial fields of maxT2m and PD.

Let X and Y be Fréchet random vectors with tail index α and dimensions p and q, respectively. For the following definitions, we consider a joint random vector $Z = (X_1, \ldots, X_p, Y_1, \ldots, Y_q)^T$ of dimension $(p+q)$. As in Sect. 2, we use the L_2 norm $||z||$ to define the unit sphere:

$$\mathbb{S}_+^{(p+q)-1} = \{z \in \mathbb{R}_+^{p+q} : ||z|| = 1\}. \tag{8}$$

Along with the transformation of z into polar coordinates,

$$z \to \left(||z||, \frac{z}{||z||}\right) =: (r, \omega), \tag{9}$$

with $r = ||z||$ and $\omega = z/||z||$. Following Eq. (4) we define the joint TPDM, which has the dimensions $(p+q) \times (p+q)$. It separates into Σ_{XX}, Σ_{YY}, and the cross-TPDM Σ_{XY}, which is a $p \times q$ matrix. If we break down ω into its x and y components $\omega^x = (x_1/||z||, \ldots, x_p/||z||)^T$ and $\omega^y = (y_1/||z||, \ldots, y_q/||z||)^T$, the cross-TPDM Σ_{XY} has the en-

tries

$$\sigma_{ij} = \int_{\mathbb{S}_+^{(p+q)-1}} \omega_i^x \omega_j^y \, dH(\omega). \tag{10}$$

The cross-TPDM is the counterpart of the cross-covariance matrix and represents an extension of the TPDM to two variables. By definition, it is a component of the joint TPDM, obtained from the joint random vector Z. While it is conceivable that the joint TPDM could also be used to extract information about the extremal dependence between X and Y, this approach is relatively rare and involves higher computational costs.

3.2 Estimation of the cross-TPDM

Let x_t, y_t, $t = 1, \ldots, n$ be samples of two Fréchet random vectors X and Y, with components $x_{t,i}$, $i = 1, \ldots, p$ and $y_{t,j}$, $j = 1, \ldots, q$. To estimate the cross-TPDM, we compute, for the radial component,

$$r_{t,ij} = \sqrt{x_{t,i}^2 + y_{t,j}^2} \tag{11}$$

and, for the angular components,

$$w_{t,i}^x = \frac{x_{t,i}}{r_{t,ij}}; \; w_{t,j}^y = \frac{y_{t,j}}{r_{t,ij}}. \tag{12}$$

We then estimate each element of the cross-TPDM Σ_{XY} as

$$\hat{\sigma}_{ij} = 2n_{ij,\text{exc}}^{-1} \sum_{t=1}^n w_{t,i}^x w_{t,j}^y \mathbb{I}_{(r_{t,ij} > r_{0,ij})}. \tag{13}$$

with threshold $r_{0,ij}$ and $n_{ij,\text{exc}} = \sum_{t=1}^n \mathbb{I}_{(r_{t,ij} > r_{0,ij})}$, the number of exceedances.

3.3 Decomposition of the cross-TPDM

The SVD of a cross-covariance matrix is a standard tool in climate data analysis to reduce the dimensionality of high-dimensional weather and climate data (see, e.g., Wallace et al., 1992). The singular vectors (SVs) form an orthogonal basis, which can be used for information compression. Here we use an estimator based on extreme dependencies instead of cross-covariances, which leads to a targeted consideration of extremal dependencies. For an introduction to SVD and its applications, see, for example, Wilks (2011), Jolliffe (1986), and Björnsson and Venegas (1997).

The decomposition of the cross-TPDM with $\Sigma_{XY} = U \Lambda V^T$ provides left and right singular vectors (SVs), as well as the matrix Λ, which contains the singular values on its diagonal. The SVs can be interpreted as pairs of spatial patterns that explain the maximum possible proportion of extremal dependence, analogous to SVD in Newman and Sardeshmukh (1995). As they point out, one should be cautious in the physical interpretation of the patterns, since they are first of all statistical patterns. However, data reduction based on SVs does not require the physical interpretability of the patterns. The singular values provide information about the extent to which the dependencies are explained by the respective SV patterns.

The cross-TPDM is defined on the basis of positive real numbers. The transformation to Fréchet margins leads to the input variables X and Y being consistent with this assumption. However, the SVs live by construction in $\mathbb{R}^p$. For the SVs to form an orthogonal basis in $\mathbb{R}_+^p$, we have to transform the SVs from $\mathbb{R}^p$ to $\mathbb{R}_+^p$. We again follow Cooley and Thibaud (2019) and use $\tau(y) = \log(1 + \exp(y))$ with its back-transform $\tau^{-1}(y) = \log(\exp(y) - 1)$. The Fréchet distribution, as well as the orthogonality of the SVs, is preserved under this transformation. Again, we transform the SVs component-wise. The time series of expansion coefficients (ECs) are calculated by a projection of the back-transformed data X and Y on the respective SVs as

$$\boldsymbol{\eta} = U^T \tau^{-1}(X) \tag{14}$$

$$\boldsymbol{\nu} = V^T \tau^{-1}(Y). \tag{15}$$

Reconstructions of each sample x_t and y_t ($t = 1, \ldots, n$) are obtained using $\tilde{x}_t = \tau(\eta_t U)$ and $\tilde{y}_t = \tau(\nu_t V)$.

4 Evaluation

To assess how many EVs or SVs are needed to represent the main features of the data, we use the fraction of explained covariance, which is

$$\text{ECF}_k = \frac{\lambda_k}{\sum_{i=1}^p \lambda_i}, \tag{16}$$

with λ_k the kth largest eigenvalue or singular value in Λ. The term ECF (explained covariance function) in the context of (cross-)TPDM may be misleading because, strictly speaking, SVs no longer provide information about (cross-)covariance but about extremal dependence. Nevertheless, we use the term here because it is a fixed term for measuring the explained dependencies in PCA or SVD.

To examine the extent to which each pattern is dominated by individual events that occurred during the training period and find an appropriate number of modes that yields a representative subset of patterns, we perform 2-fold cross-validation. There are several measurements that aim to find the "optimal" number of modes after PCA (see, e.g., Abdi and Williams, 2010). We use Q_m^2 as a rather simple but effective measure to assess overfitting (i.e., the point at which the quality of the prediction decreases with an increasing number of modes). To derive Q_m^2 we divide the time series into two periods of equal length and calculate a (cross-)TPDM for each period, respectively.

The reconstructions follow Sect. 2.3 and 3.3. The within-sample reconstruction $\tilde{x}_t^m = \tau(\tilde{u}^T \tau^{-1}(x_t)\tilde{u})$ uses the first

m EVs/SVs $\tilde{u}$ derived from the period including t, while the out-of-sample reconstruction $\hat{x}_t^m$ uses the corresponding EVs/SVs $\hat{u}$ of the period not including t. The reconstructions of the two periods are then merged to calculate Q_m^2 defined as

$$Q_m^2 = 1 - \frac{\sum_{t=1}^n ||x_t - \hat{x}_t^m||_2}{\sum_{t=1}^n ||x_t - \tilde{x}_t^{m-1}||_2}, \tag{17}$$

with $||\cdot||_2$ the L_2 norm. Q_m^2 is a measure of information gain relative to the increase of noise. A positive Q_m^2 means that the mth EV/SV mode significantly improves an out-of-sample reconstruction relative to the in-sample reconstruction using $m-1$ EV/SV modes. The number of modes m at which Q_m^2 becomes negative indicates the point at which overfitting starts to occur. The standard error of Q_m^2 is assessed by re-sampling using a block bootstrap method of annual data with 100 samples.

5   Extremal pattern index

We define an extremal pattern index (EPI) based on the PCs using the L_2 norm. The L_2 norm is chosen since both PCA and SVD may be defined as an L_2-minimization problem (see, e.g., Abdi and Williams, 2010). Let be given m PCs η_t at time t as derived by Eq. (7) using the m leading EV of a TPDM. We further incorporate the leading eigenvalues/singular values $\lambda_{1,...,m}$ to normalize the EPI, which we define at time t as

$$\text{EPI}_t^\eta = \frac{\sqrt{\sum_{k=1}^m \eta_{t,k}^2}}{\sqrt{\sum_{j=1}^m \lambda_j}}. \tag{18}$$

The EPI represents a spatial aggregation and shows when individual patterns or a linear combination of patterns are particularly pronounced. The first EV/SVs of the (cross-)TPDM show little spatial variation (see, e.g., Cooley and Thibaud, 2019), so $\eta_{t,1}$ and $\nu_{t,1}$ essentially represent the spatial mean of the Fréchet-distributed variables. The higher modes are those representing the spatial variation patterns. Thus, the EPI can become extreme even if the first principal component is not outstanding.

The SVD of the cross-TPDM provides two ECs, for the right and the left singular vectors. We define the EPI for the cross-TPDM analogously as

$$\text{EPI}_t^{\eta,\nu} = \frac{\sqrt{\sum_{k=1}^m (\eta_{t,k}^2 + \nu_{t,k}^2)}}{\sqrt{\sum_{j=1}^m \lambda_j}}. \tag{19}$$

This definition of EPI describes scenarios where either one or both of the corresponding ECs are high. To account for the simultaneous occurrence of extremes, we introduce a threshold-based approach that considers both the left and right ECs of each mode k to be prominently expressed. We

refer to this as the threshold-based extremal pattern index (TEPI).

$$\text{TEPI}_t^{\eta,\nu} = \frac{\sqrt{\sum_{k=1}^m (\eta_{t,k}^2 + \nu_{t,k}^2) \mathbb{I}_{(|\eta_{t,k}|>q_\eta, |\nu_{t,k}|>q_\nu)}}}{\sqrt{\sum_{j=1}^m \lambda_j}}, \tag{20}$$

with threshold values q_η and q_ν. The threshold ensures that the associated patterns are simultaneously highly pronounced.

6   Application

6.1   Data description

This study relies on the COSMO-REA6 regional reanalysis (Bollmeyer et al., 2015). COSMO-REA6 was developed within the Hans Ertel Centre for Weather Research (HErZ) at the University of Bonn. COSMO-REA6 is currently available for the period from 1995 to 2019 at the opendata FTP server[1] of the German meteorological service (DWD).

We use accumulated PD, as well as daily maxT2m. PD is defined as the additive inverse of precipitation accumulated over various periods between 11 and 47 d using a centered moving-average approach (e.g., for an accumulation period of 11 d, PD is estimated at the center including the 5 d preceding and following the center date, respectively). The use of the moving average for precipitation is less common in the literature, but for our application, we consider it appropriate. Otherwise, we would rather focus on the precipitation deficit before a heat wave. (e.g., PD from 1 to 30 June in conjunction with maxT2m on 30 June). The spatial domain of COSMO-REA6 is according to the CORDEX EUR-11 specifications[2] but with a higher horizontal resolution of about 6 km horizontal grid space. We limit our analysis to land areas and also exclude countries that have little to no precipitation during northern hemispheric summer months. We further restrict our analysis to the northern hemispheric summer months from June to August, when heat waves have particularly strong impacts on social and environmental systems through associated effects such as forest fires and heat stress.

6.2   Data pre-processing

Our pre-processing requires three steps: (1) in order to reduce spatial non-stationarity we choose to standardize the data at grid point level. Along with standardization, we also remove seasonal trends, which we identify in maxT2m and PD. To this end, we estimate mean and standard deviation at each grid point and day of the year, which are then

---

[1]https://opendata.dwd.de/climate_environment/REA/COSMO_REA6/ (last access: 20 December 2023)

[2]https://euro-cordex.net/060374/ (last access: 20 December 2023)

used for standardization. (2) The spatial fields in COSMO-REA6 contain 824×848 grid points, resulting in a matrix of $698\,752 \times 698\,752$ pairwise dependence measures. To limit the computational cost, the number of grid points is reduced to a 121×117 grid. For consistency with extreme value theory, we reduce the grid size using spatial block maxima instead of standard interpolation techniques. To this end, we compute the maximum in a 7×7 neighborhood of grid points for spatial reduction. (3) To transform maxT2m and PD to a Fréchet distribution with tail index α we use the empirical distribution function with

$$\hat{F}_{Y_i}^{\text{orig}}(y) = \frac{1}{n} \sum_{t=1}^{n} \mathbf{I}(y_{t,i}^{\text{orig}} \le y). \tag{21}$$

6.2.1 Choice of threshold

The choice of the threshold value for estimating the TPDM is always critical. If the threshold is too low, there is a risk that the asymptotic limit of the extreme value model has not yet been reached, which leads to biases in the estimation. A threshold that leaves too few data points to which the model can be fitted leads to large uncertainties in the estimates. To determine a suitable threshold, we consider the rank of the TPDM and determine its stability as a function of the threshold value in Fig. 1. A stable linear downward trend is given for thresholds below the 98 % quantile (red line), but a rapid decline is seen for higher thresholds. The 98 % quantile, therefore, seems suitable as a threshold for further calculations.

Figure 2 shows maxT2m and the 15d PD at two grid points in COSMO-REA6 near Paris and Bonn. Only the data outside the gray shaded area (i.e., above the 98 % quantile) are used to estimate the cross-TPDM. The dependence between maxT2m at grid points with a distance between Paris and Bonn is strong even for the extreme values since the temperature variations are usually dominated by large-scale patterns and therefore still show dependencies, even at large distances. In contrast, the dependencies in the PD and between PD and maxT2m are much lower, which is due to the small-scale structures of the extreme values in the PD.

Figure 2 may be used to illustrate the asymptotic dependence/independence between Fréchet-distributed time series. The measure $H(\omega)$ of the angular component ω describes the tail dependence on the unit sphere $\mathbb{S}_{+}^{p-1}$ (Sect. 2). With asymptotic independence, we expect the mass of $H(\omega)$ to be at the edges, so we do not expect an extreme event to occur in both variables simultaneously. With asymptotic dependence, corresponding probability mass exists on the unit circle, and we observe events that are extreme in both components.

As Jiang et al. (2020) point out, the assumption of asymptotic dependence, which is the basis of the concept of regular variations, may not be fulfilled. Jiang et al. (2020) focus on precipitation data, which are clearly not asymptotically dependent at the continental scale. This is less clear
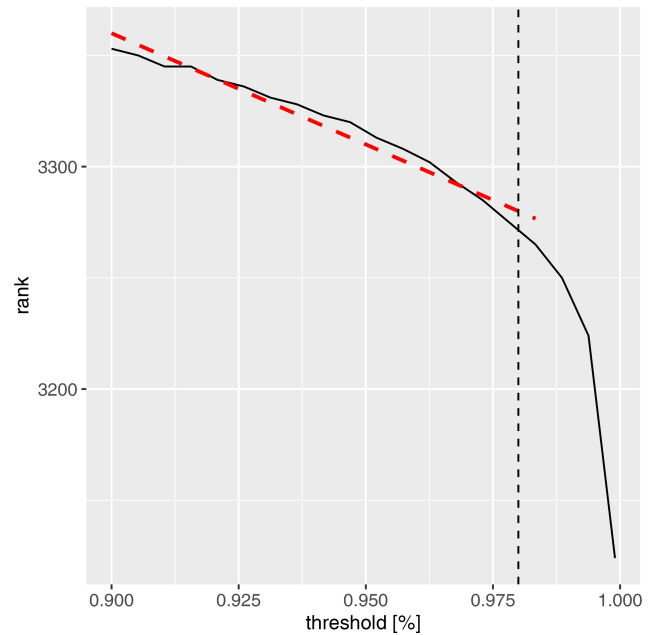

Figure 1. Rank of the TPDM, calculated for maxT2m and varying thresholds r_0 (Eq. 5). Vertical line: 98 % quantile. Red line: linear least-squares fit of the rank using a threshold between 90 % and 98 % quantile.

for regional-scale maxT2m and PD and cross-dependence. The TPDM method is nevertheless well suited to decompose (cross-)dependencies in our case since we typically do not reach the asymptotic limit for real data. For real data, we are more likely to be in the pre-asymptotic limit, which is characterized by significant dependencies, as seen in Fig. 2.

7 Results

7.1 Spatial patterns of extreme daily maximum 2m temperature

We start our analysis with the estimation of the TPDM for maxT2m. The estimation is based on maxT2m values exceeding the local 98 % quantile on the reduced grid (see Sect. 6.1). Figure 3 shows the first 12 EVs of the TPDM. The first mode represents the nonzero spatial mean and consists of nearly constant values with only small variations of order 10^{-2}. Since the data are defined in $\mathbb{R}_{+}$ (see Sect. 2.3), the mean value is not removed unlike in standard PCA (cf. Jiang et al., 2020; Cooley and Thibaud, 2019).

Higher-order EVs show large-scale patterns associated with the typical dipole and multipole structures also known from PCA. Since the EVs are transformed on $\mathbb{R}_{+}$, values $> \log(2)$ refer to positive and values $< \log(2)$ to negative anomalies (see Sect. 2.3). The wave-like structures are strongly reminiscent of westward-propagating Rossby waves, which are the main driving force for large-scale circu-
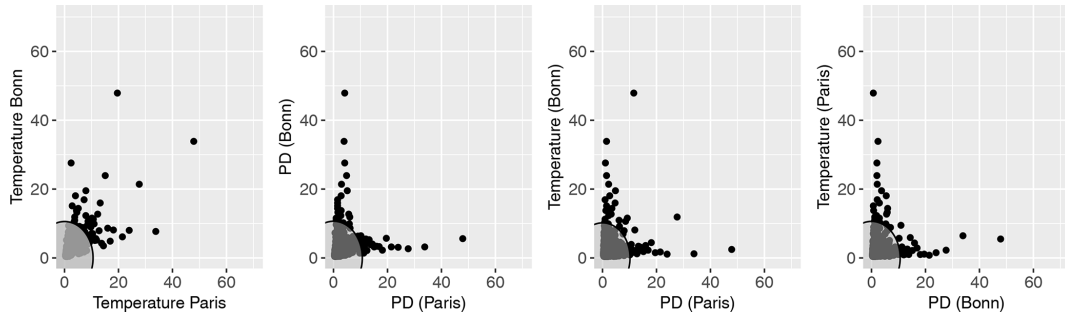
Figure 2. Scatter plots of Fréchet-transformed maxT2m and PD (black dots) at two grid points in COSMO-REA6 at Paris and Bonn. PD has been calculated from 15 d accumulated precipitation. The circle indicates the respective 98 % quantile, and the values inside the circle (gray shading) are not used for the estimation of the (cross-)TPDM.
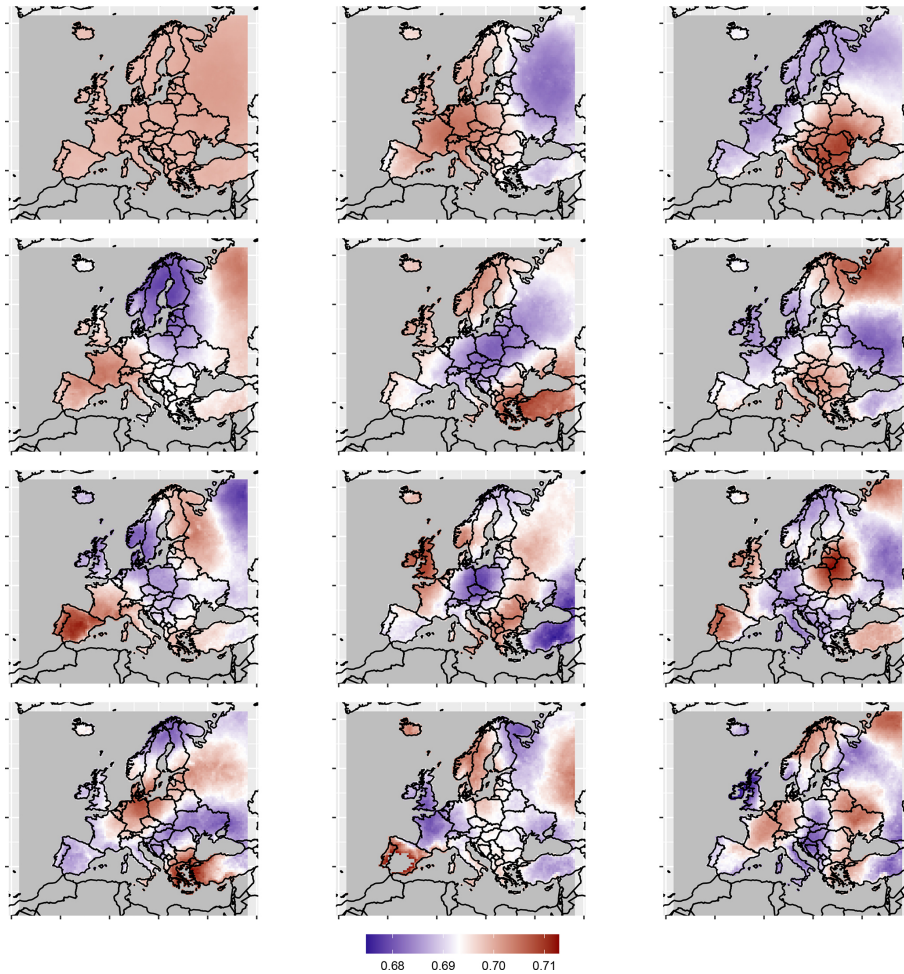

Figure 3. First 12 (from top to bottom, left to right) EVs of the TPDM for maxT2m in northern hemispheric summer months (JJA). Each vector is transformed to the positive real numbers $\mathbb{R}^p \to \mathbb{R}^p_+$ such that $\tau(0) = \log(2)$ corresponds to zero.

lation patterns in northern mid-latitudes (see, e.g., Liu et al., 2020; Schubert et al., 2011; Gershunov and Douville, 2009).

An important question is how representative the EVs are and at what number of modes significant information is no longer added by further modes but overfitting prevails. In Figs. 4 and 5 we show explained variance and Q^2 as introduced in Sect. 4. The first EV contributes to about 25 % of the total variance, followed by less than 8 % for the second EV. As in Jiang et al. (2020) and Cooley and Thibaud (2019), the first EV (i.e. changes in the spatial mean maxT2m) dom-
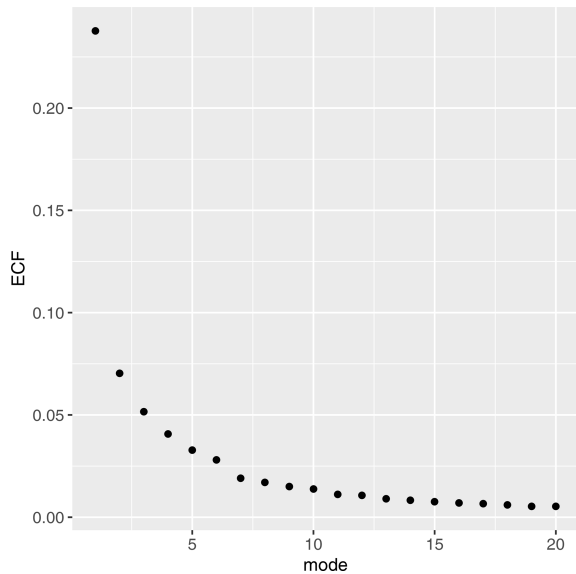
Figure 4. Explained variance ECF_k for each mode k (Eq. 16). Calculated from TPDM using maxT2m in northern hemispheric summer months (JJA).
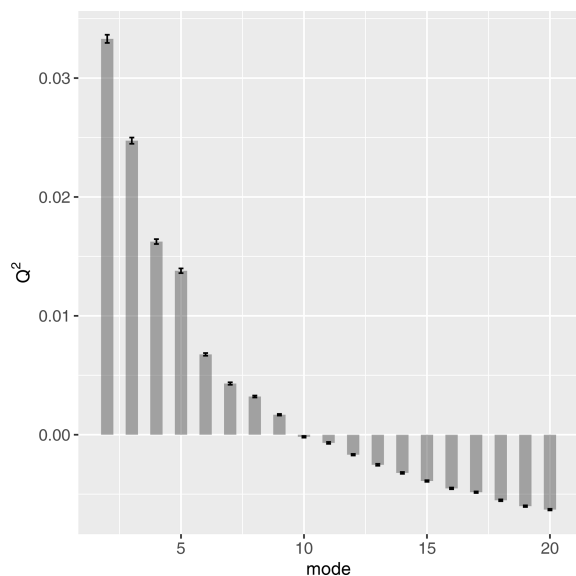

Figure 5. Q_m^2 for the first $m = 20$ modes (Eq. 17). Calculated from TPDM for maxT2m in northern hemispheric summer months (JJA). The error bars indicate the standard error determined using 100 bootstrap samples.

inates the variance. About 52 % (75 %) of the variance is explained by the first 10 (50) EVs. The out-of-sample information gain as measured by Q_m^2 is positive for the leading 10 EVs. Both measures indicate that the first 10 EVs provide a representative subset for the spatial variation patterns of extreme maxT2m.

7.2 Extreme pattern index for temperature extremes

Our calculation of the EPI^η (Eq. 18) uses the first 10 PCs η_k with $k = 1, \ldots, 10$. The time series of the PCs for the first six EVs are displayed in Fig. A1 in the Appendix. Figure 6 shows EPI^η for the June to September months from 1995 to 2019 with values varying between 0.5 and 13.5. Peaks in EPI^η indicate intense heat waves within the studied region. Although the 2010 summer heat wave prominently stands out in EPI^η, we also identify other peaks associated with typical heat waves. The large outliers in the PCs as well as in EPI^η indicate a heavy tail behavior, as already shown by Cooley and Thibaud (2019) for the PCs.

We will now examine the two heat waves of 2003 and 2010 in more detail using our framework. Both heat waves have been extensively studied and analyzed in the existing literature. The 2010 heat wave is characterized by the highest peak EPI in the period considered. This indicates that the 2010 heat wave was the most extreme event in terms of its combination of intensity and spatial extent over the period considered. The 2003 heat wave, on the other hand, did not reach the highest peak but is characterized by its long, highly pronounced duration. Finally, both heat waves had significant impacts on socio-economic systems, underlining their importance for understanding extreme heat events (see Beniston, 2004; Barriopedro et al., 2011; Kueh and Lin, 2020; Schubert et al., 2011; Gershunov and Douville, 2009). Figure 7 (bottom panel) shows EPI^η for the June to September months in 2003. The respective averaged maxT2m anomalies are shown in the upper panels for the periods during periods as marked with (a) to (c). We observe a first heat wave in mid-July (a) in western Europe, which is most pronounced over the United Kingdom, Iceland, and Norway. This short-term event is followed by a few non-extreme days (b), after which we observe a very severe heat wave over central and southern Europe from late July to mid-August (c) (see also García-Herrera et al., 2010; Liu et al., 2020). In Fig. 7 we also display the reconstruction of the maxT2m fields using only the leading 10 EVs. The spatial structure of the heat wave pattern in maxT2m is well represented within the 10-dimensional subspace. The reconstructed patterns are somewhat smoother and thus more clearly highlight the areas affected by the heat wave. Thus, the subspace spanned by the 10 eigenvectors of the TPDM is sufficient to describe the 2003 heat wave and fulfills the goal of a compact description of large-scale heat waves.

In Fig. 8, we analyze the Russian heat wave in 2010. High values of EPI^η consistently capture their evolution from the second week of July to mid-August, as analyzed, e.g., in Barriopedro et al. (2011) or Liu et al. (2020). Since we observe the strongest expressions of EPI^η in the first half of August, we focus on this period in the following. The beginning of August is characterized by the most intensive phase of the heat wave according to EPI^η (a), with its center located over Moscow, western Russia. In the second week, we observe a
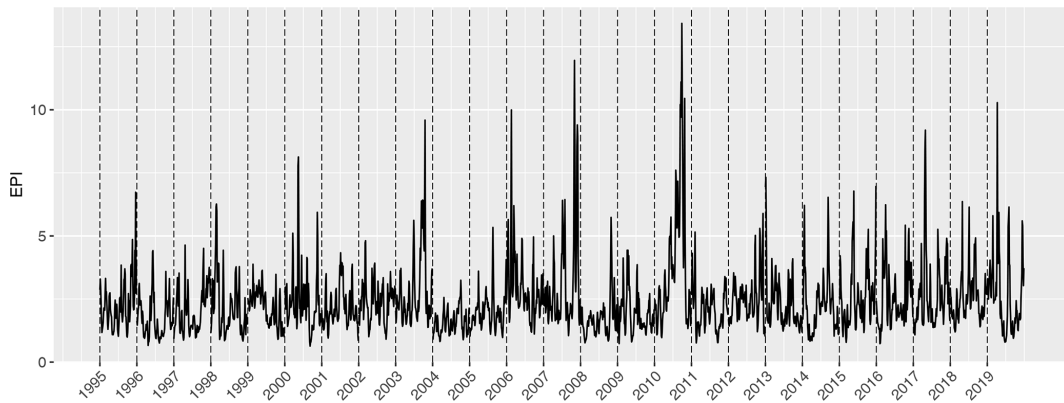
Figure 6. EPI^η as calculated from first 10 PCs after a PCA of the TPDM for maxT2m. Shown are the northern hemispheric summer months (JJA) from 1995 to 2019.

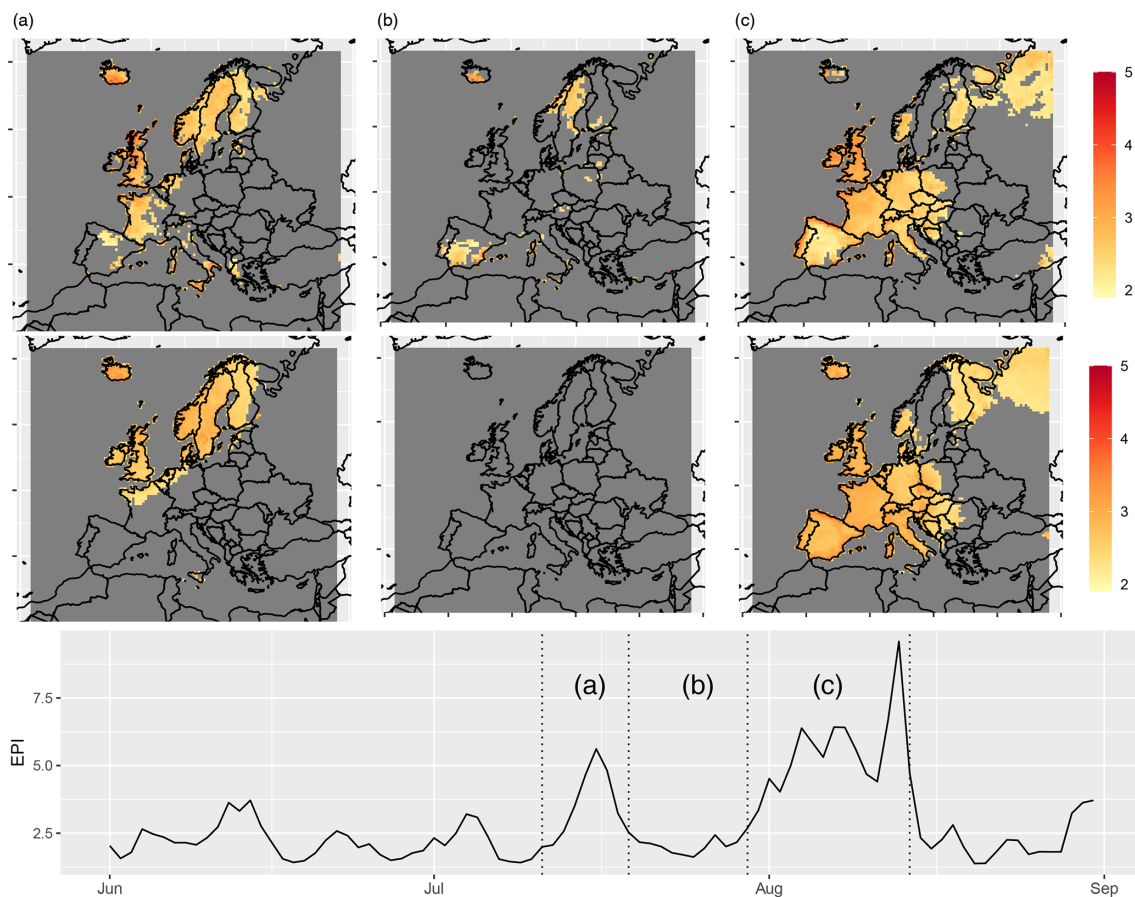**Figure 7.** Overview of the temporal evolution of the 2003 heat wave. Top: pattern of mean anomalies within the periods **(a)** 11–19 July 2003, **(b)** 19–30 July 2003, and **(c)** 30 July–14 August 2003. We show the maxT2m anomalies (upper pattern) and the reconstruction, using the first 10 modes of PCs (lower pattern). All patterns use the same color scale and only show values exceeding the 98 % quantile, respectively. Bottom: EPI^η as in Fig. 6 but for JJA months in 2003.

westward shift of the maxT2m anomalies, accompanied by a short drop in EPI^η (b). Then the heat wave intensifies again for 2 d, with a new center over Belarus and Lithuania (c), before it collapses in mid-August. This particular feature of the

2010 heat wave has received little attention in the literature, although it is clearly visible in station measurements in, e.g., Moscow (Grumm, 2011). Consequently, the Russian heat wave can also be well described by our method.
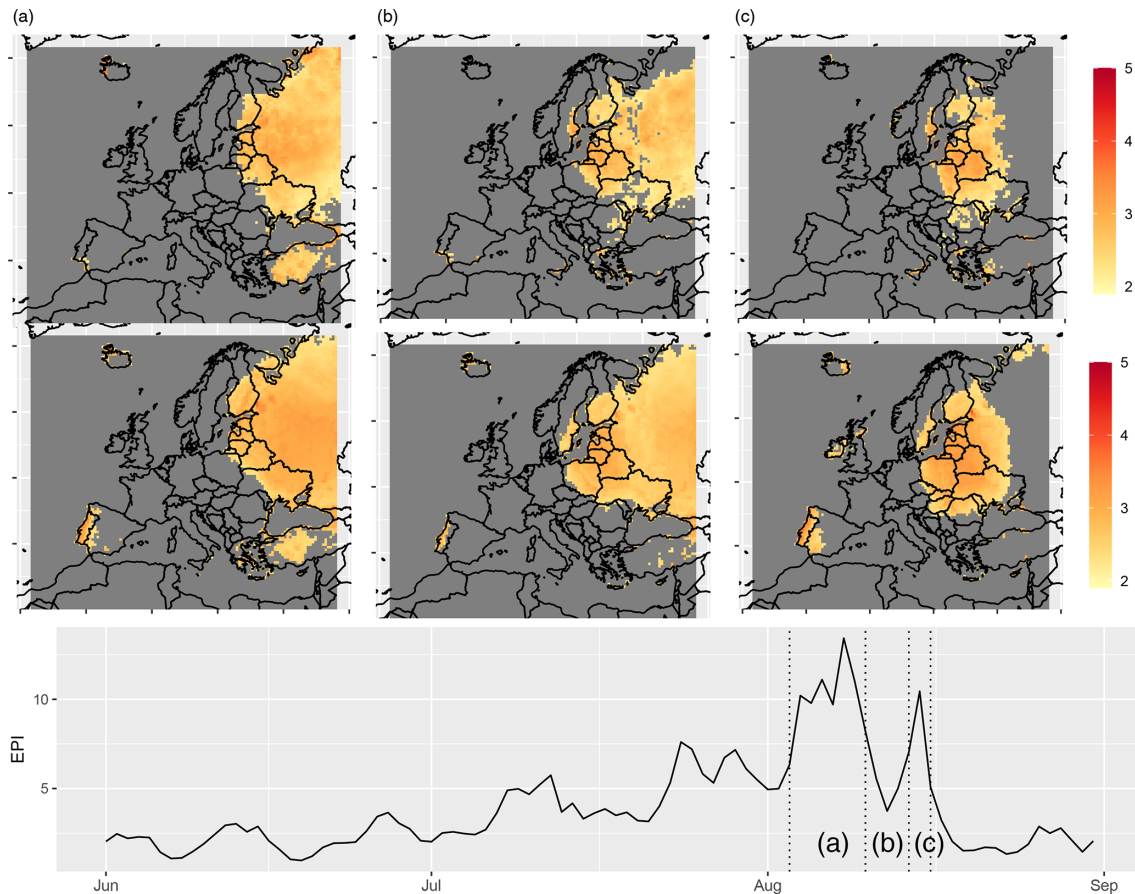
Figure 8. Overview of the temporal evolution of the 2010 heat wave. Top: pattern of mean anomalies within the periods **(a)** 3–10 August 2010, **(b)** 10–14 August 2010, and **(c)** 14–16 August 2010. We show the maxT2m anomalies (upper pattern) and the reconstruction, using the first 10 modes of PCs (lower pattern). All patterns use the same color scale and only show values exceeding the 98 % quantile, respectively. Bottom: EPI^{η} as in Fig. 6 but for JJA months in 2010.

To further evaluate the suitability of EPI^{η} for heat wave detection and representation, we compare event definitions using EPI^{η} with other heat wave definitions. A heat wave is usually defined on the basis of grid points (hereafter GRD event), e.g., a 5 d continuous exceedance of the 98 % quantile of maxT2m, with seasonal trends removed. A GRD event exists when a 5 d continuous exceedance occurs in at least 3 % of the grid points. Using maxT2m of COSMO-REA6, we obtain a number of 32 GRD events during the summer months in 1995–2019 in Europe. To rank heat wave events in terms of their "extremeness", DWD compares predefined heat waves (i.e. GRD events) based on the properties of duration, maximum number of grid points affected (GRDsize), and average magnitude of maxT2m anomalies (GRDmean)[3]. The time period of the 15 strongest GRD events with respect to GRDsize and GRDmean is displayed in black in Table 1.

We define an EPI event as a period when EPI^{η} exceeds the 80 % quantile for at least 5 consecutive days. The thresh-

[3]https://www.dwd.de/EN/ourservices/rcccm/int/rcccm_int_hwkltr.html (last access: 7 July 2022)

old is chosen so that we get a similar number of heat waves for the EPI events as for the GRD events, yielding a total of 28 heat waves. In Table 1, we compare the top 15 EPI events according to their average EPI^{η} (EPImean), with the top 15 GRD events according to their GRDsize and GRDmean. Of the top 15 GRD events, 3 (June 2002, August 2008, and August 2012) are not classified as EPI events. For each of these events, the maxT2m fields exceeding the 98 % quantile and the EPI^{η} are shown in the Appendix Fig. A2. We find EPI^{η} exceeds the 80 % quantile in all three GRD events, but the duration is less than 5 d in each case. Of the top 15 EPI events, 13 are identified as being among the top 15 GRD events in terms of GRDsize as well, suggesting that EPI^{η} is particularly successful in the identification of large-scale extremes. This is largely due to the property of PCA, which detects large-scale patterns mainly in the first few modes. Large-scale extremes are thus weighted more strongly than small-scale events.

The Cooley and Thibaud framework requires the assumption of independent and identically distributed samples.

Table 1. Periods of strongest heat waves defined by EPI and GRD events. Periods indicated in black are within the top 15 strongest events using the respective index (EPImean, GRDsize, and GRDmean). Periods in bold are heat waves with a lower rank than 15, and underlined periods show no correspondence in the respective counterpart (EPI or GRD).

EPImean	GRDsize	GRDmean
1995-08-17–1995-08-22	**1995-08-17–1995-08-22**	1995-08-17 1995-08-22
1998-06-11–1998-06-21	1998-06-11–1998-06-21	**1998-06-11–1998-06-21**
	2002-06-08–2002-06-12	<u>2002-06-08 2002-06-12</u>
2003-07-14–2003-07-18	2003-07-14–2003-07-19	2003-07-14 2003-07-19
2003-07-31–2003-08-14	2003-07-31–2003-08-14	2003-07-31 2003-08-14
2006-06-02–2006-06-06	2006-06-02–2006-06-06	2006-06-02 2006-06-06
2006-06-18–2006-06-23	**2006-06-18–2006-06-23**	**2006-06-18–2006-06-23**
2006-07-06–2006-07-11	**2006-07-06 2006-07-10**	2006-07-06 2006-07-10
2007-07-15–2007-07-25	2007-07-15–2007-07-25	2007-07-15 2007-07-25
2007-08-14–2007-08-19	2007-08-13–2007-08-20	**2007-08-13–2007-08-20**
2007-08-21–2007-08-27	2007-08-22–2007-08-27	2007-08-22 2007-08-27
	<u>2008-08-15 2008-08-19</u>	**2008-08-15 2008-08-19**
2009-06-25–2009-06-29	**2009-06-25–2009-06-29**	2009-06-25 2009-06-29
2010-07-07–2010-08-17	2010-07-07–2010-07-14	**2010-07-07–2010-07-14**
2010-07-07–2010-08-17	2010-07-16–2010-08-17	**2010-07-16–2010-08-17**
	2012-08-17 –2012-08-21	<u>2012-08-17 2012-08-21</u>
2014-08-03–2014-08-07	**2014-08-03–2014-08-07**	**2014-08-03–2014-08-07**
2015-06-30–2015-07-08	2015-06-30–2015-07-08	2015-06-30 2015-07-08
2016-06-17–2016-06-27	**2016-06-17–2016-06-27**	**2016-06-17–2016-06-27**
2017-06-29–2017-07-04	2017-06-29–2017-07-04	2017-06-29 2017-07-04
2017-08-01–2017-08-06	**2017-08-01–2017-08-06**	2017-08-01 2017-08-06
2019-06-12–2019-06-16	2019-06-12–2019-06-16	**2017-08-01–2017-08-06**
2019-06-25–2019-07-02	2019-06-25–2019-07-02	2019-06-25 2019-07-02
2019-07-23–2019-07-27	2019-07-23–2019-07-27	2019-07-23 2019-07-27

This assumption is generally not given for daily values of maxT2m and PD. However, it turns out that our method is quite robust in this respect. We evaluate this in Appendix Fig. A3 using the EPI calculated on the one hand from the patterns of daily maxT2m data and on the other hand from the patterns of weekly maxima of maxT2m. It can be seen that the EPI is robust to small deviations in individual patterns. The range of EPI values changes slightly as the first 10 modes explain a higher proportion of the variance.

7.3 Decomposition of the cross-TPDM for maxT2m and PD

We have shown that the decomposition of the TPDM provides a set of spatial patterns to describe extremal dependencies in maxT2m and leads to a suitable definition of a heat wave index. We now turn to the cross-TPDM, which encodes the extremal covariability of maxT2m and PD. Commonly used indices, e.g., the standard precipitation index (SPI), are calculated on relatively long accumulation times from 3 to 48 months (McKee et al., 1993). However, so-called "flash droughts", e.g., occur instantaneously at sub-seasonal scales. They are characterized by rapid development and intensification of drought, usually accompanied by exceptionally high temperatures (see, e.g., Peters et al., 2002; Hunt et al.,

2009; Cook et al., 2018). To capture this structure, Mo and Lettenmaier (2016), e.g., calculate PD using 5 d precipitation averages. Here, we want to capture both perspectives on the dynamics and the evolution of compound heat wave and drought events and therefore assess extremal covariability for different accumulation times from 11 to 47 d.

To allow for a computation of the cross-TPDM for many different accumulations times, we need to lower the computational costs by further reducing the grid size. For this purpose, we increase the neighborhood in the COSMO-REA6 grid to 21×21 grid points and take the respective maximum of PD and maxT2m as the value (in Sect. 7.2 we used a neighborhood of 7×7 grid points). With this approach, we obtain a grid of size 40×39. To ensure that no important information gets lost in this new reduction, in Appendix Fig. A4 we exemplarily show the EPI$^\eta$ as calculated in Sect. 7.2 against the EPI$^\eta$ on the highly reduced grid. The differences in the temporal behavior are small, and the EPI$^\eta$ is robust to this further data reduction. Although not shown here, we have produced similar diagrams for the other indices we have defined, with similar results.

Not only the duration of a dry spell but also when it occurs in relation to the heat wave can be important for dynamic interpretation. For example, there is evidence that a dry phase can promote the development of an intense heat period (see,

e.g., Sousa et al., 2020). To account for different feedback mechanisms between PD and maxT2m, we introduce a time shift between PD and maxT2m, in the sense that PD leads maxT2m in time when there is a negative shift. The cross-TPDM and its decomposition using SVD are calculated for each time shift and accumulation time, respectively.

We assess statistical significance based on the singular values. To this end, we generate a 100-head bootstrap sample where we draw maxT2m and PD of a season independently so that the temporal relationship is destroyed while seasonality is preserved in the data. In Fig. 9, we show the first and the sum of the first and second singular values of the cross-TPDM, normalized by the respective average over the bootstrap samples. This is shown as a function of time shift and accumulation level. For values > 1 the measure of the extremal covariability between maxT2m and PD is larger than on average in the bootstrap sample. White dots indicate values that exceeded the 95 % confidence interval of the values in the bootstrap sample with "no statistical association".

For short accumulation times, maximum association is obtained for instantaneous anomalies. For higher accumulation times, the association is maximal for negative shifts when anomalies in PD precede the maxT2m anomalies. This indicates that a heat wave is preferentially formed when it is preceded by a drought period. Considering the sum of more singular values (see Appendix Fig. A5 for the first 10 singular values), this time-shifted dependence structure weakens but remains significant. We conclude that the statistical association between the large-scale patterns in matT2m and PD is significant and that the accumulation time of PD must be evaluated together with a temporal shift.

We now turn to the large-scale patterns of extremes in maxT2m and PD. In Fig. 10, we show the patterns for an accumulation time of 11 d and a time shift of 0 d. The large-scale patterns again show the characteristic wave-like structure previously discussed in Sect. 7.1. The corresponding anomalies in maxT2m and PD have a very similar structure, with the regions of the strongest gradients in PD slightly shifted compared to the gradients in maxT2m (second pattern – eastwards, third pattern – southwards). Although a physical interpretation of the SVs should always be approached with caution (see Sect. 3.3), it is possible that eastward-propagating Rossby waves (Sect. 7.1) could explain this pattern if, for example, strong gradients in the temperature fields are associated with frontal systems typically associated with precipitation. The leading patterns are also very similar for other accumulation times and time shifts.

7.4 Extremal pattern indices for compound maxT2m and PD extremes

As in Sect. 5 we base our calculation of an extremal pattern index on the expansion coefficients, where η_k represents the respective expansion coefficient of maxT2m and ν_k of PD. We calculate the EPI$^{\eta,\nu}$ (Eq. 19) and the TEPI$^{\eta,\nu}$ (Eq. 20) us-

ing η_k and ν_k of the leading 10 SVD modes of maxT2m and PC, respectively. As a threshold for the TEPI$^{\eta,\nu}$, we use the 90 % quantile, respectively. This gives us a sufficient number of events without losing focus on the extremes. In this way, we obtain a flexible tool for the analysis of different event characteristics of compound extremes in maxT2m and PD.

Our investigation focuses on three cases with varying accumulation periods and lead times for PD anomalies: one with 11 d accumulations and immediate anomalies, another with 35 d accumulations and a PD anomaly lead time of 8 d, and a third case with 93 d accumulations and a PD anomaly lead time of 38 d. The first case represents short-term, co-incident events, while the second and third cases represent monthly and seasonal PD driving a heat wave. In all cases, we find significant statistical dependence between maxT2m and PD (Fig. A5). Although not shown, we obtain a significant value of 1.13 for the third case 93 d accumulation and 38 d shift. Figure 11 displays the TEPI$^{\eta,\nu}$ for all cases. The heat waves in 2003 and 2010 are indicated with colors.

Both the 2003 and 2010 heat waves show a strong signal in instantaneous TEPI. TEPI$^{\eta,\nu}$ events coincide with heat wave events as indicated by the EPI$^{\eta}$, suggesting that both heat waves were accompanied by short-term PD. This indicates typical weather conditions during heat waves with clear skies, strong solar radiation, and little to no precipitation (Fischer et al., 2007).

Furthermore, the 2003 heat wave was preceded by a PD that was mainly evident on a seasonal timescale, whereas the 2010 heat wave was preceded by a more pronounced PD on a monthly timescale. In both cases, the high agreement of TEPI$^{\eta,\nu}$ and EPI$^{\eta}$ suggests that PD contributed to the extreme maxT2m during 2003 and 2010 (see also Fischer et al., 2007; Fink et al., 2004; Zaitchik et al., 2006; Christian et al., 2020).

Thus, the heat waves in 2003 and 2010 display different dynamics concerning common extremes of PD and maxT2m. Our TEPI definition effectively captures these characteristics, demonstrating that both events meet the IPCC definition of compound events as "two or more extreme events occurring simultaneously or sequentially" (Seneviratne et al., 2012).

We thus demonstrate that our definition of TEPI for different accumulation times and temporal shifts summarizes many characteristics of compound extremes and represents a suitable index.

8 Discussion and conclusion

In this study, we apply the promising approach of Cooley and Thibaud (2019) to describe extremal dependencies using the TPDM. The TPDM has comparable properties to the covariance matrix and can be decomposed similar to PCA while focusing on the tail behavior of the variables. To describe the dynamics of large-scale temperature extremes, we
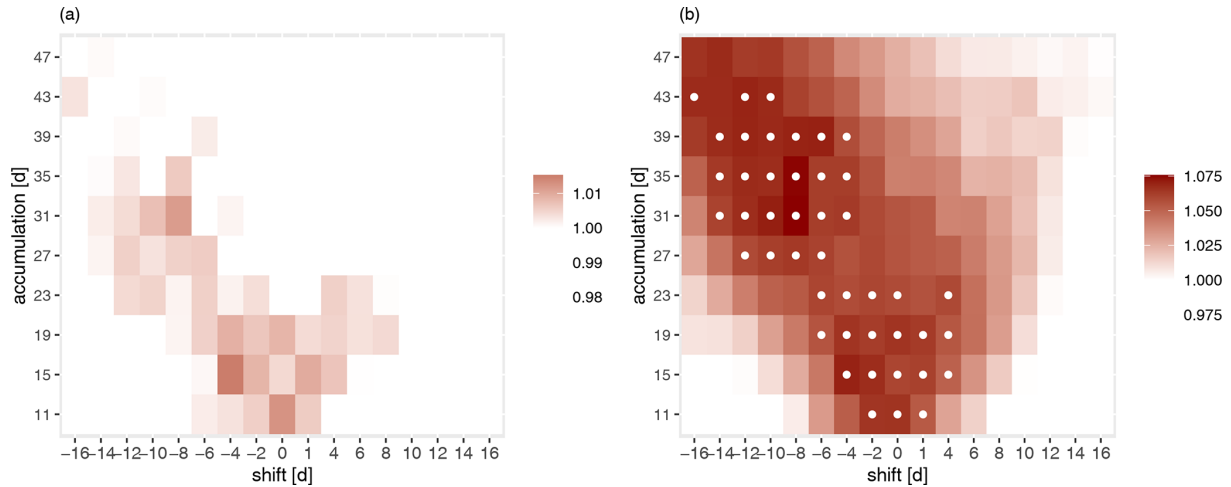
Figure 9. First and first + second singular values of the cross-TPDM calculated from maxT2m and PD in northern hemispheric summer months for varying shift and accumulation times. Each value has been normalized by the respective mean of 100 bootstrap samples where the temporal relation is destroyed. White dots indicate values exceeding the 95 % confidence interval of "no statistical relation".
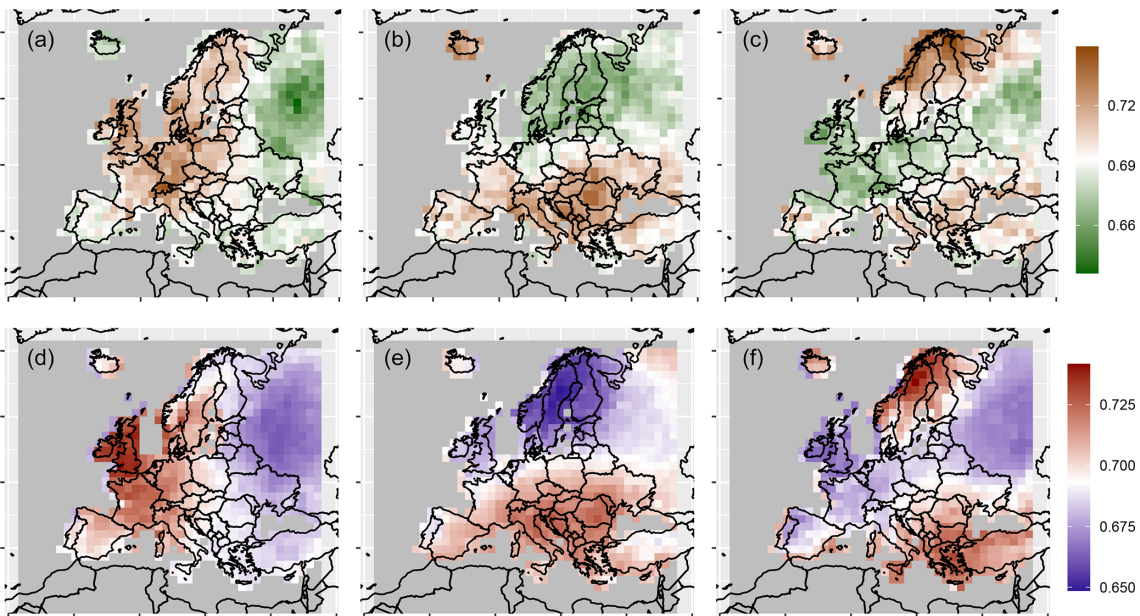

Figure 10. Second, third, and fourth left and right SVs of the cross-TPDM, associated with anomalies in maxT2m **(d, e, f)** and 11 d PD **(a, b, c)**. As in Fig. 3, we transformed each vector to the positive real numbers $\mathbb{R}^p \to \mathbb{R}^p_+$ such that $\tau(0) = \log(2)$ corresponds to zero.

calculated the TPDM based on maxT2m and identified typical temperature patterns of European heat waves. Our results are in line with previous studies that identify large-scale blocking events as the primary mechanism behind European heat waves (e.g. Schubert et al., 2014; Stefanon et al., 2012). We determine the optimal number of patterns to describe the main features of European heat waves and show exemplarily for the 2003 and 2010 heat waves that the description of the dynamics of large-scale temperature extremes is suitable with only the 10 leading eigenmodes.

We define the extremal pattern index (EPI) based on the time series of PCs and obtain a pattern-based spatial aggregation index. Thus, we can specifically identify spatially related extreme events with minimal reliance on pre-defined thresholds and regions of influence in grid point space. Although we cannot completely eliminate the use of a threshold, as it is a fundamental factor in estimating extreme dependencies based on the TPDM, all data, even those below the threshold, are included in the EPI estimator. We can show that this results in little information being lost with respect to the heat wave description.
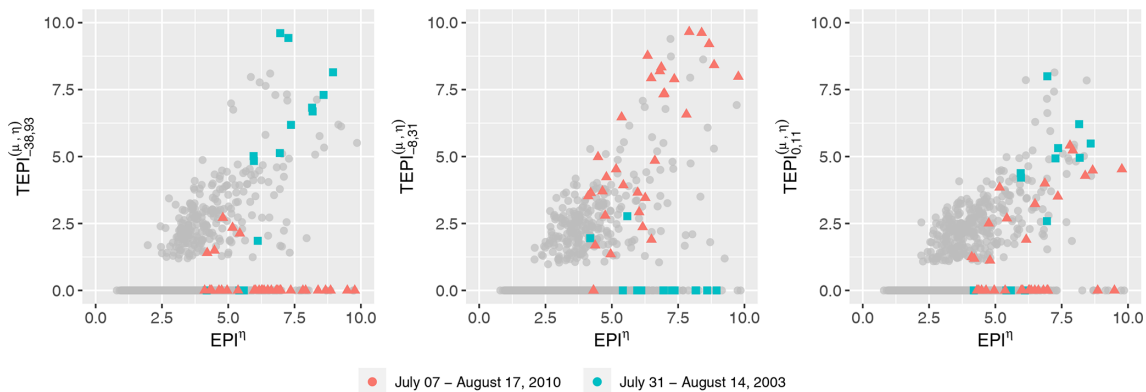
Figure 11. Scatter plot of EPI^{η} and $TEPI^{\mu,\eta}$. $TEPI^{\mu,\eta}$ is shown for calculation from PD using three different combinations of shift and accumulation (11 d accumulation, 0 d shift; 35 d accumulations, 8 d shift; and 93 d accumulation, 46 d shift). All indices were calculated using the first 10 PCs of the cross-TPDM from maxT2m and PD for northern hemispheric summer months. The 2010 and 2003 heat waves (31 July–14 August 2003; 7 July–17 August 2010), as identified in Table 1, are highlighted in blue and red, respectively.

Using the 2003 and 2010 European heat waves as examples, we examine how well the EPI captures the essential characteristics of heat waves. The temporal progression of both heat waves can be effectively illustrated using the EPI and aligns with the findings of previous analyses. Some characteristics, such as the collapse of the 2010 heat wave between 10 and 14 August, can be clearly identified using the EPI yet have received little exploration in the existing literature. An additional condition of heat waves besides spatial extent and intensity is usually temporal persistence. This condition can also be applied to EPI by defining heat waves. In Sect. 7, e.g., we define a heat wave (i.e., an EPI event) such that the EPI exceeds a threshold for at least 5 consecutive days. We propose an EPI-based definition of heat waves that, compared to a conventional definition, gives very consistent results. This shows that the EPI is suitable as an alternative to conventional heat wave indices with minimum reliance on user-defined thresholds. The EPI is thus a powerful tool for, e.g., attribution studies, climate monitoring, and characterization of extreme weather conditions.

Since pronounced heat waves over Europe cover a large part of the area studied, the spatial mean describes the course of the heat waves relatively well. Nevertheless, the higher modes of the PCs contribute significantly to the description of heat waves, as we discuss in Sect. 7.1. It can be assumed that the dominant influence of the first PC on the EPI would decrease when analyzing a larger geographical area.

To include the temporal persistence of heat waves as another condition, one could include only those patterns that exhibit a certain persistence and then aggregate them over time. However, too hard a condition could be detrimental.

After initially focusing on the description of heat waves using maxT2m, we extend our approach and go beyond the examination of individual variables. We deal with common extremes in maxT2m and PD. To this end, we propose a cross-TPDM model as an analog to the cross-covariance ma-

trix that describes pairwise extremal dependencies between the two variables. Previous studies for compound events have usually referred to monthly or even seasonal variables (Wu et al., 2019; Hao et al., 2019). We calculate PD from daily precipitation using different duration scales. To generalize our approach and apply it to compound events as well, we shift the time series against each other. This allows us to investigate the dynamics of compound extremes in maxT2m and PD, even at relatively short duration levels. Using the singular values following an SVD, we analyze the dynamics of compound extremes in PD and maxT2m and show that preceding extremes in PD exert a significant influence on the evolution of heat waves (see, e.g., Christian et al., 2020). The investigation of additional variables such as hydrological drought indices and boundary layer heights would undoubtedly be interesting and may provide new insights into physical relationships.

Last but not least, we extend the EPI from considering a single variable to describing simultaneous extremes in two variables. We introduce a threshold-based estimator, the TEPI, which ensures simultaneous extremes of the underlying patterns. To demonstrate the utility of the TEPI, we again focus on the 2003 and 2010 heat waves and analyze their different dynamics in terms of compound extremes in maxT2m and PD using the TEPI. We analyze both heat waves using two event types: short-term coincident events and long-term precipitation deficit triggering a heat wave. In both cases, we can use TEPI to reconstruct the known characteristics of these events and demonstrate its effectiveness as a valuable tool for analyzing compound extremes. TEPI is a way to focus on specific events. A disadvantage of our formulation of TEPI is the hard condition of exceeding a threshold (see Fig. 11). We need to investigate whether a less strict selection or weighting would be beneficial.

There are several indicators for assessing climate indices (see, e.g., Keyantash and Dracup, 2002; Redmond, 1991).

Both EPI and TEPI are robust to application to different data sources, although we believe they are most successful in identifying large-scale extremes. In principle, they are applicable to point (e.g., measured data) and gridded data sets. The mathematical theory and derivation are quite complex, but the application is comparatively simple since existing functions for PCA and SVD, etc., can be used in various programming languages. Jiang et al. (2020) also make their R code for calculating TPDM, data pre-processing, and performing PCA available.

The cost-intensiveness of calculating TPDM based on high-resolution climate models is a universal problem that affects methods for dealing with spatial dependencies. This can possibly be solved by clever mathematical reformulation or extension of the working memory. As shown in Fig. A4, the data reduction approach we propose, using the spatial block maximum, provides a robust estimate of the EPI.

The TPDM assumes asymptotic dependence. However, we recognize that this assumption may not hold for maxT2m, PD, and their interdependencies. As Jiang et al. (2020), we nevertheless choose to decompose using the TPDM, as it opens up a new way to examine extreme patterns in maxT2m and PD. The resulting patterns are consistent with previous findings in the field and provide a meaningful dynamic context.

Appendix A
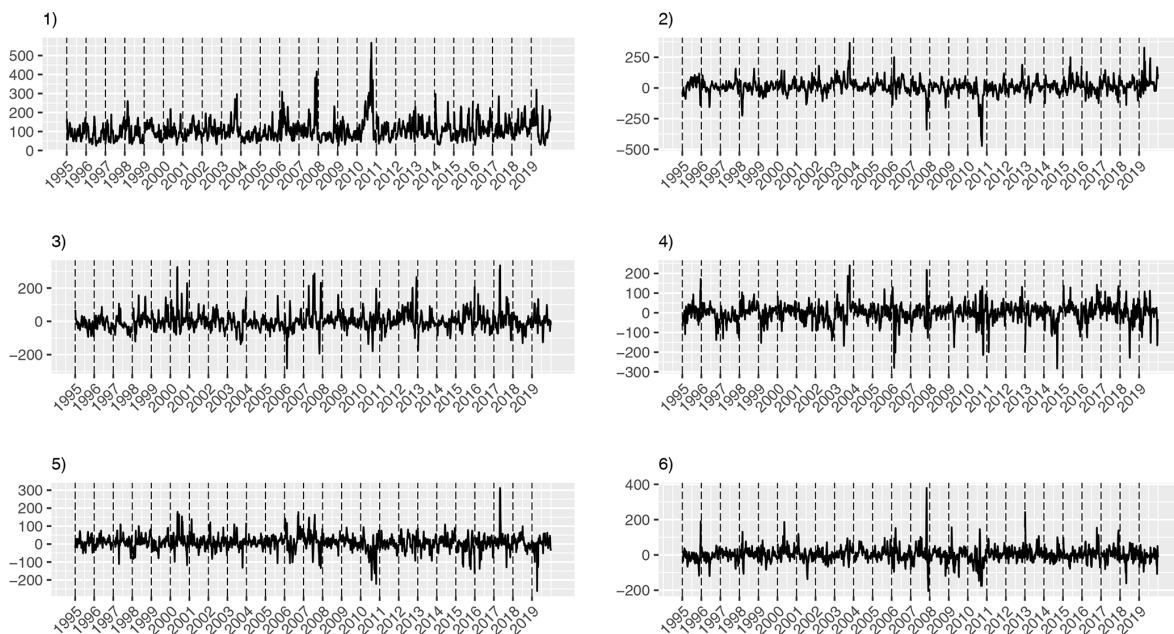
Figure A1. First six (from left to right, top to bottom) PCs of the TPDM as calculated from maxT2m in northern hemispheric summer months (JJA).
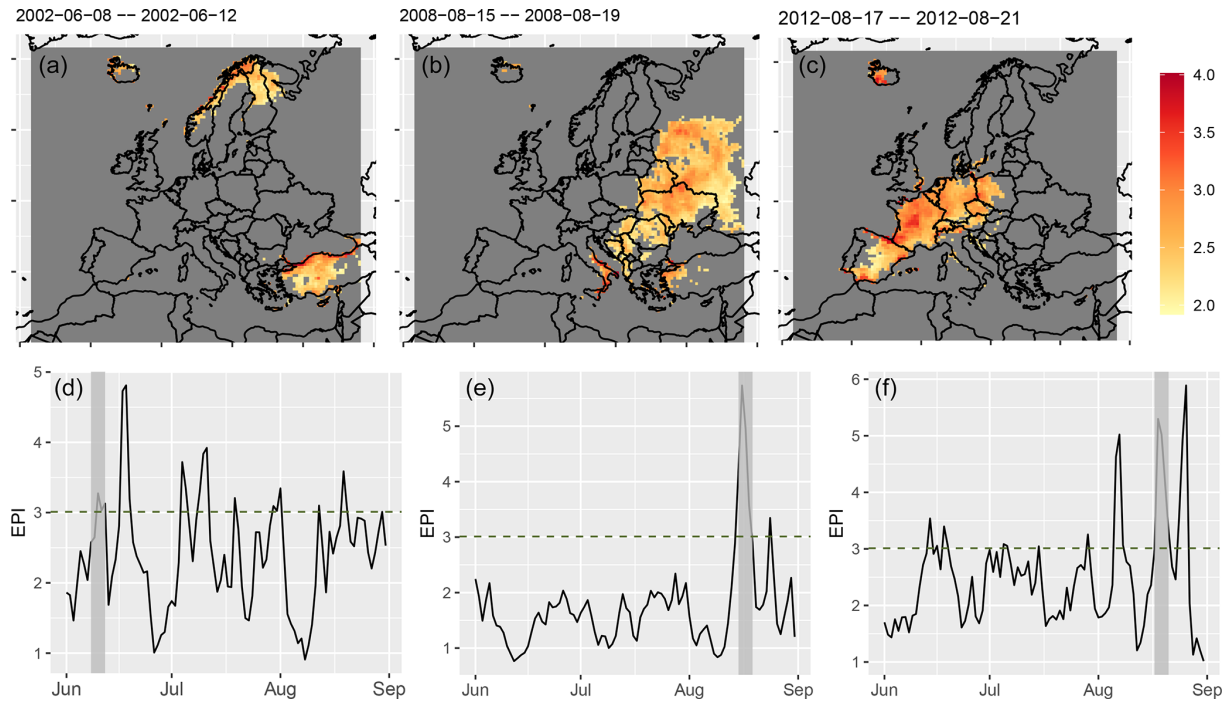
Figure A2. EPI$^{\eta}$ and maximum maxT2m anomaly of top 15 EPI events (according to EPImean) which are not identified as GRD events (Sect. 7.2). **(a, b, c)** Maximum maxT2m anomaly. Shown are only grid points exceeding the 98 % quantile. **(d, e, f)** EPI$^{\eta}$ as shown in Fig. 6 but for single years 2002, 2008, and 2012. Time periods of corresponding EPI events are marked in gray.
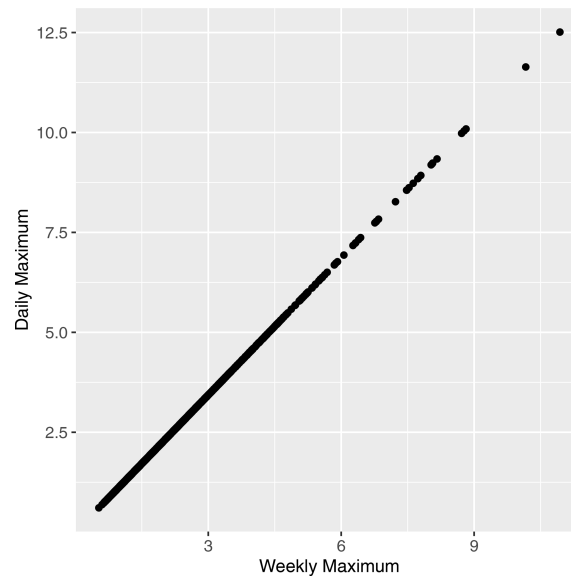

Figure A3. Scatter plot representation of EPI$^{\eta}$ calculated on EVs and eigenvalues of daily maxT2m and weekly maximum of maxT2m.
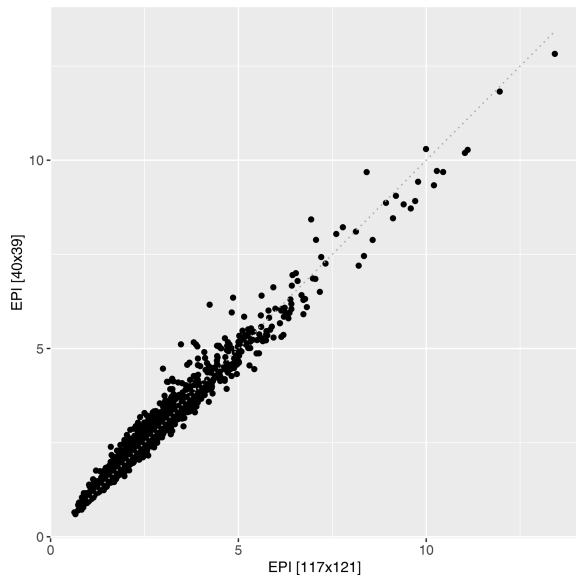
Figure A4. Scatter plot representation of EPI^η calculated on $[117 \times 121]$ grid (see Sect. 7.2) vs. EPI^η calculated on $[40 \times 39]$ grid (Sect. 3).
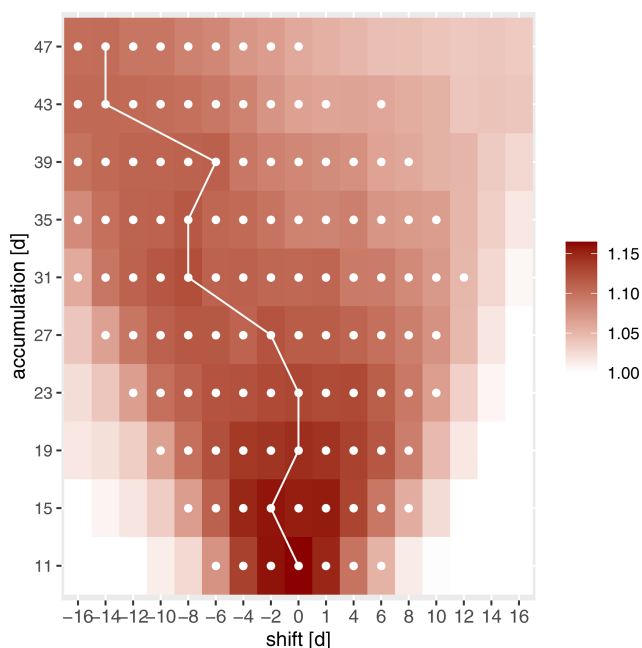
Figure A5. As in Fig. 9 but for the sum of first 10 singular values. The white line connects the maximum values over each accumulation stage.

Code availability. Code for data pre-processing and estimation of (cross-)TPDM and (T)EPI is provided as the R package "ExtrPatt". It is available at https://CRAN.R-project.org/package=ExtrPatt (Szemkus et al., 2023).

Competing interests. The contact author has declared that neither of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Special issue statement. This article is part of the special issue "Past and future European atmospheric extreme events under climate change". It is not associated with a conference.

Review statement. This paper was edited by Dan Cooley and reviewed by three anonymous referees.

References

Abdi, H. and Williams, L. J.: Principal component analysis, Wires Comput. Stat., 2, 433–459, https://doi.org/10.1002/wics.101, 2010.

Anderson, H. R., Atkinson, R. W., Peacock, J., Marston, L., and Konstantinou, K.: Meta-analysis of time-series studies and panel studies of particulate matter (PM) and ozone (O3): report of a WHO task group, Tech. rep., Copenhagen: WHO Regional Office for Europe, https://iris.who.int/handle/10665/107557, 2004.

Barnston, A. G. and Livezey, R. E.: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns, Mon. Weather Rev., 115, 1083–1126, 1987.

Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., and García-Herrera, R.: The hot summer of 2010: redrawing the temperature record map of Europe, Science, 332, 220–224, 2011.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J. L., De Waal, D. and Ferro, C.: Statistics of extremes: theory and applications, John Wiley & Sons, 2006.

Beniston, M.: The 2003 heat wave in Europe: A shape of things to come? An analysis based on Swiss climatological data and model simulations, Geophys. Res. Lett., 31, L02202, https://doi.org/10.1029/2003GL018857, 2004.

Björnsson, H. and Venegas, S.: A manual for EOF and SVD analyses of climatic data, CCGCR Report, 97, 112–134, 1997.

Black, E., Blackburn, M., Harrison, G., Hoskins, B., and Methven, J.: Factors contributing to the summer 2003 European heatwave, Weather, 59, 217–223, 2004.

Bollmeyer, C., Keller, J., Ohlwein, C., Wahl, S., Crewell, S., Friederichs, P., Hense, A., Keune, J., Kneifel, S., Pscheidt, I., Redl, S., and Steinke, S.: Towards a high-resolution regional reanalysis for the European CORDEX domain, Q. J. Roy. Meteor. Soc., 141, 1–15, 2015.

Boulaguiem, Y., Zscheischler, J., Vignotto, E., van der Wiel, K., and Engelke, S.: Modeling and simulating spatial extremes by combining extreme value theory with generative adversarial networks, Environmental Data Science, 1, 4, https://doi.org/10.1017/eds.2022.4, 2022.

Christian, J. I., Basara, J. B., Hunt, E. D., Otkin, J. A., and Xiao, X.: Flash drought development and cascading impacts associated with the 2010 Russian heatwave, Environ. Res. Lett., 15, 094078, https://doi.org/10.1088/1748-9326/ab9faf, 2020.

Coles, S., Bawa, J., Trenner, L., and Dorazio, P.: An introduction to statistical modeling of extreme values, vol. 208, Springer, https://doi.org/10.1007/978-1-4471-3675-0, 2001.

Cook, B. I., Mankin, J. S., and Anchukaitis, K. J.: Climate change and drought: From past to future, Current Climate Change Reports, 4, 164–179, 2018.

Cooley, D. and Thibaud, E.: Decompositions of dependence for high-dimensional extremes, Biometrika, 106, 587–604, 2019.

Dai, A., Zhao, T., and Chen, J.: Climate change and drought: a precipitation and evaporation perspective, Current Climate Change Reports, 4, 301–312, 2018.

Della-Marta, P. M., Haylock, M. R., Luterbacher, J., and Wanner, H.: Doubled length of western European summer heat waves since 1880, J. Geophys. Res.-Atmos., 112, D15103, https://doi.org/10.1029/2007JD008510, 2007.

Drees, H. and Sabourin, A.: Principal component analysis for multivariate extremes, arXiv [preprint], https://doi.org/10.48550/arXiv.1906.11043, 26 June 2019.

DWD/HErZ – Deutscher Wetterdienst – Climate Data Center/Hans Ertel Centre for Weather Research: COSMO-REA6 regional reanalysis, https://opendata.dwd.de/climate_environment/REA/COSMO_REA6/, last access: 20 December 2023.

Ferranti, L. and Viterbo, P.: The European summer of 2003: Sensitivity to soil water initial conditions, J. Climate, 19, 3659–3680, 2006.

Fink, A. H., Brücher, T., Krüger, A., Leckebusch, G. C., Pinto, J. G., and Ulbrich, U.: The 2003 European summer heatwaves and drought-synoptic diagnosis and impacts, Weather, 59, 209–216, 2004.

Fischer, E. M. and Knutti, R.: Observed heavy precipitation increase confirms theory and early models, Nat. Clim. Change, 6, 986–991, 2016.

Fischer, E. M. and Schär, C.: Consistent geographical patterns of changes in high-impact European heatwaves, Nat. Geosci., 3, 398–403, 2010.

Fischer, E. M., Seneviratne, S. I., Vidale, P. L., Lüthi, D., and Schär, C.: Soil moisture–atmosphere interactions during the 2003 European summer heat wave, J. Climate, 20, 5081–5099, 2007.

Fix, M. J., Cooley, D. S., and Thibaud, E.: Simultaneous autoregressive models for spatial extremes, Environmetrics, 32, e2656, https://doi.org/10.1002/env.2656, 2021.

García-Herrera, R., Díaz, J., Trigo, R. M., Luterbacher, J., and Fischer, E. M.: A review of the European summer heat wave of 2003, Crit. Rev. Env. Sci. Tec., 40, 267–306, 2010.

Gershunov, A. and Douville, H.: Extensive summer hot and cold spells under current and possible future climatic conditions: Europe and North America, Climate Extremes and Society, 74–98, https://doi.org/10.1017/Cbo9780511535840.008, 2009.

Goix, N., Sabourin, A., and Clémençon, S.: Sparse representation of multivariate extremes with applications to anomaly detection, J. Multivariate Anal., 161, 12–31, https://doi.org/10.1016/j.jmva.2017.06.010, 2017.

Griggs, D. J. and Noguer, M.: Climate change 2001: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change, Weather, 57, 267–269, 2002.

Grumm, R. H.: The central European and Russian heat event of July–August 2010, B. Am. Meteor. Soc., 92, 1285–1296, 2011.

Guerreiro, S. B., Dawson, R. J., Kilsby, C., Lewis, E., and Ford, A.: Future heat-waves, droughts and floods in 571 European cities, Environ. Res. Lett., 13, 034009, https://doi.org/10.1088/1748-9326/aaaad3, 2018.

Hao, Z., Hao, F., Singh, V. P., and Zhang, X.: Statistical prediction of the severity of compound dry-hot events based on El Niño-Southern Oscillation, J. Hydrol., 572, 243–250, 2019.

Higham, N. J.: Computing the nearest correlation matrix – a problem from finance, IMA J. Numer. Anal., 22, 329–343, 2002.

Hunt, E. D., Hubbard, K. G., Wilhite, D. A., Arkebauer, T. J., and Dutcher, A. L.: The development and evaluation of a soil moisture index, Int. J. Climatol., 29, 747–759, 2009.

IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896, 2021.

Janßen, A. and Wan, P.: k-means clustering of extremes, Electron. J. Stat., 14, 1211–1233, 2020.

Jiang, Y., Cooley, D., and Wehner, M. F.: Principal Component Analysis for Extremes and Application to US Precipitation, J. Climate, 33, 6441–6451, 2020.

Jolliffe, I. T.: Principal components in regression analysis, in: Principal component analysis, 129–155, Springer, https://doi.org/10.1007/978-1-4757-1904-8_8, 1986.

Keyantash, J. and Dracup, J. A.: The quantification of drought: an evaluation of drought indices, B. Am. Meteor. Soc., 83, 1167–1180, 2002.

Kueh, M.-T. and Lin, C.-Y.: The 2018 summer heatwaves over northwestern Europe and its extended-range prediction, Sci. Rep.-UK, 10, 1–18, 2020.

Larsson, M. and Resnick, S. I.: Extremal dependence measure and extremogram: the regularly varying case, Extremes, 15, 231–256, 2012.

Lavaysse, C., Cammalleri, C., Dosio, A., van der Schrier, G., Toreti, A., and Vogt, J.: Towards a monitoring system of temperature extremes in Europe, Nat. Hazards Earth Syst. Sci., 18, 91–104, https://doi.org/10.5194/nhess-18-91-2018, 2018.

Liu, X., He, B., Guo, L., Huang, L., and Chen, D.: Similarities and differences in the mechanisms causing the European summer heatwaves in 2003, 2010, and 2018, Earth's Future, 8, e2019EF001386, https://doi.org/10.1029/2019EF001386, 2020.

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., and Wanner, H.: European seasonal and annual temperature variability, trends, and extremes since 1500, Science, 303, 1499–1503, 2004.

McKee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration to time scales, in: 8th Conf. Appl. Climatol., Anaheim, CA, https://climate.colostate.edu/pdfs/relationshipofdroughtfrequency.pdf (last access: 23 December 2023), 17–22 January 1993.

Mo, K. C. and Lettenmaier, D. P.: Precipitation deficit flash droughts over the United States, J. Hydrometeorol., 17, 1169–1184, 2016.

Morris, S. A., Reich, B. J., and Thibaud, E.: Exploration and inference in spatial extremes using empirical basis functions, J. Agr. Biol. Envir. St., 24, 555–572, 2019.

Newman, M. and Sardeshmukh, P. D.: A caveat concerning singular value decomposition, J. Climate, 8, 352–360, 1995.

Palmer, W. C.: Meteorological drought, vol. 30, US Department of Commerce, Weather Bureau, https://books.google.de/books?id=kyYZgnEk-L8C (last access: 23 December 2023), 1965.

Peters, A. J., Walter-Shea, E. A., Ji, L., Vina, A., Hayes, M., and Svoboda, M. D.: Drought monitoring with NDVI-based standardized vegetation index, Photogramm. Eng. Rem. S., 68, 71–75, 2002.

Redmond, K.: Climate monitoring and indices, Drought management and planning: international drought information center, Department of Agricultural Meteorology, Am. Metorol. Soc., 9, 29–33, 1991.

Resnick, S. I.: Heavy-tail phenomena: probabilistic and statistical modeling, Springer Science & Business Media, ISBN 0387450246, 9780387450247, 2007.

Robinson, P. J.: On the definition of a heat wave, J. Appl. Meteorol. Climatol., 40, 762–775, 2001.

Rohrbeck, C. and Cooley, D.: Simulating flood event sets using extremal principal components, Ann. Appl. Stat., 17, 1333–1352, 2023.

Rousi, E., Fink, A. H., Andersen, L. S., Becker, F. N., Beobide-Arsuaga, G., Breil, M., Cozzi, G., Heinke, J., Jach, L., Niermann, D., Petrovic, D., Richling, A., Riebold, J., Steidl, S., Suarez-Gutierrez, L., Tradowsky, J. S., Coumou, D., Düsterhus, A., Ellsäßer, F., Fragkoulidis, G., Gliksman, D., Handorf, D., Haustein, K., Kornhuber, K., Kunstmann, H., Pinto, J. G., Warrach-Sagi, K., and Xoplaki, E.: The extremely hot and dry 2018 summer in central and northern Europe from a multi-faceted weather and climate perspective, Nat. Hazards Earth Syst. Sci., 23, 1699–1718, https://doi.org/10.5194/nhess-23-1699-2023, 2023.

Saunders, K., Stephenson, A., and Karoly, D.: A regionalisation approach for rainfall based on extremal dependence, Extremes, 24, 215–240, 2021.

Schädler, G. and Breil, M.: Identification of droughts and heatwaves in Germany with regional climate networks, Nonlin. Processes Geophys., 28, 231–245, https://doi.org/10.5194/npg-28-231-2021, 2021.

Schubert, S., Wang, H., and Suarez, M.: Warm season subseasonal variability and climate extremes in the Northern Hemisphere: The role of stationary Rossby waves, J. Climate, 24, 4773–4792, 2011.

Schubert, S. D., Wang, H., Koster, R. D., Suarez, M. J., and Groisman, P. Y.: Northern Eurasian heat waves and droughts, J. Climate, 27, 3169–3207, 2014.

Seneviratne, S., Nicholls, N., Easterling, D., Goodess, C., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., and Zhang, X.: Changes in climate extremes and their impacts on the natural physical environment, in: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, edited by: Field, C. B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L., Mastrandrea, M. D., Mach, K. J., Plattner, G.-K., Allen, S. K., Tignor, M., and Midgley, P. M., A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC), Cambridge University Press, Cambridge, UK, and New York, NY, USA, 109–230, 2012.

Slater, L. J., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., Kelder, T., Kowal, K., Lees, T., Matthews, T., Murphy, C., and Wilby, R. L.: Nonstationary weather and water extremes: a review of methods for their detection, attribution, and management, Hydrol. Earth Syst. Sci., 25, 3897–3935, https://doi.org/10.5194/hess-25-3897-2021, 2021.

Szemkus, S., Cooley, D., and Jiang, Y.: ExtrPatt: Spatial Dependencies and Indices for Extremes, https://CRAN.R-project.org/package=ExtrPatt (last access: 20 December 2023), 2023.

Sousa, P. M., Barriopedro, D., García-Herrera, R., Ordóñez, C., Soares, P. M., and Trigo, R. M.: Distinct influences of large-scale circulation and regional feedbacks in two exceptional 2019 European heatwaves, Commun. Earth & Environ., 1, 1–13, 2020.

Stefanon, M., D'Andrea, F., and Drobinski, P.: Heatwave classification over Europe and the Mediterranean region, Environ. Res. Lett., 7, 014023, https://doi.org/10.1088/1748-9326/7/1/014023, 2012.

Ummenhofer, C. C. and Meehl, G. A.: Extreme weather and climate events with ecological relevance: a review, Philos. T. Roy. Soc. B, 372, 20160135, https://doi.org/10.1098/rstb.2016.0135, 2017.

Vignotto, E., Engelke, S., and Zscheischler, J.: Clustering bivariate dependencies of compound precipitation and wind extremes over Great Britain and Ireland, Weather and Climate Extremes, 32, 100318, https://doi.org/10.1016/j.wace.2021.100318, 2021.

Wallace, J. M., Smith, C., and Bretherton, C. S.: Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies, J. Climate, 5, 561–576, 1992.

Wilks, D. S.: Statistical methods in the atmospheric sciences, vol. 100, Academic press, ISBN 0123850223, 9780123850225, 2011.

Wu, X., Hao, Z., Hao, F., Singh, V. P., and Zhang, X.: Dry-hot magnitude index: A joint indicator for compound event analysis,

Environ. Res. Lett., 14, 064017, https://doi.org/10.1088/1748-9326/ab1ec7, 2019.

Zaitchik, B. F., Macalady, A. K., Bonneau, L. R., and Smith, R. B.: Europe's 2003 heat wave: a satellite view of impacts and land–atmosphere feedbacks, Int. J. Climatol., 26, 743–769, 2006.

Zscheischler, J. and Seneviratne, S. I.: Dependence of drivers affects risks associated with compound events, Sci. Adv., 3, e1700263, https://doi.org/10.1126/sciadv.1700263, 2017.