Weather and Climate Extremes 38 (2022) 100518

Contents lists available at ScienceDirect

Weather and Climate Extremes

journal homepage: www.elsevier.com/locate/wace

On the temporal clustering of European extreme precipitation events and its relationship to persistent and transient large-scale atmospheric drivers

Yannick Barton [a,*], Pauline Rivoire [a,c], Jonathan Koh [b], Mubashshir Ali S. [a,d], Jérôme Kopp [a], Olivia Martius [a]

[a] Oeschger Centre for Climate Change Research and Institute of Geography, University of Bern, Bern, Switzerland
[b] Oeschger Centre for Climate Change Research and Institute of Mathematical Statistics and Actuarial Science, University of Bern, Bern, Switzerland
[c] Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland
[d] Risk Management Solutions, London, United Kingdom

ARTICLE INFO

Dataset link: https://apps.ecmwf.int/datasets/

Keywords:
Extreme precipitation
Temporal clustering
Serial clustering
General additive model

ABSTRACT

Extreme precipitation events that occur in close succession can have important societal and economic repercussions. Here we use 42 years of reanalysis data (ERA-5) to investigate the link between Euro-Atlantic large-scale pattern of weather and climate variability and the temporal clustering of extreme rainfall events over Europe. We implicitly model the seasonal rate of extreme occurrences as part of a Poisson General Additive Model (GAM) using cyclic regression cubic splines. The smoothed seasonal rate of extreme rainfall occurrences is used to (i) infer the frequency of significant temporal clustering and (ii) implicitly serves as the baseline rate when modeling the effects of atmospheric drivers on extreme rainfall clustering. We use GAMs to model the association between the temporal clustering of extreme rainfall events and seven predominant year-round weather regimes in the Euro-Atlantic sector as well as a measure of synoptic-scale transient recurrent Rossby wave packets. Sub-seasonal clustering of precipitation events is significant at all grid-points over Europe; the proportion of extreme rainfall events that cluster in time ranges between 2% to 27%. The most relevant weather regime is the Atlantic Trough (corresponding to NAO+ with a southward shift of the jet) explaining most of the significant increase in clustering probability over Europe. The Greenland Blocking regime explains most of the clustering over the Iberian Peninsula. The Scandinavian Blocking regime is associated with a significant increase in clustering probability over the western Mediterranean, with a northwards shift in the signal to central Europe in summer.

1. Introduction

Sub-seasonal temporal clustering of extreme precipitation events (EPEs) can have profound societal and economic repercussions. First, the close succession of rainfall events can increase the time for soil moisture, rivers and lakes to return to climatological values in catchments with high-retention capacity and cause devastating floods (e.g., Barton et al., 2016; Galarneau et al., 2012; Grams et al., 2014; Guo et al., 2020; Huntingford et al., 2014; Lau and Kim, 2012; Martius et al., 2013; Priestley et al., 2017; Tuel and Martius, 2021a). Second, the short recovery time between events can overburden emergency services and prevent proper clean-up of damages and efficient repairing of damaged protective structures (Raymond et al., 2020). Third the temporal clustering may generate unexpected financial implications for the insurance/reinsurance industry (Vitolo et al., 2009) Hence, a sequence of extreme rainfall events can amplify impacts when compared with a single hazard. It is therefore a temporally compounding event (Zscheischler et al., 2020; Tuel and Martius, 2021a). Process understanding of the temporal clustering of EPEs is needed to assess and mitigate associated risks and to improve predictability on sub-seasonal time-scales.

Extratropical modes of climate variability such as the Artic Oscillation (AO), the North Atlantic Oscillation (NAO), the East Atlantic pattern (EAP) and the Scandinavian pattern (SCA) capture predominant atmospheric circulation patterns on sub-seasonal time-scales. They have been shown to govern dry/wet conditions and heavy precipitation frequency across wide parts of Europe (e.g. Comas-Bru and McDermott, 2014; Kenyon and Hegerl, 2010; Krichak et al., 2014; Casanueva et al., 2014; Scaife et al., 2008). These modes of climate variability are further related to the position and strength of the jet stream and thereby the temporal clustering of cyclones (e.g. Mailier et al., 2006; Priestley et al., 2017). Modes of variability have been used to investigate if

* Corresponding author.
E-mail address: yannick.barton@giub.unibe.ch (Y. Barton).

https://doi.org/10.1016/j.wace.2022.100518
Received 9 February 2022; Received in revised form 7 October 2022; Accepted 14 October 2022
Available online 27 October 2022

reanalyzes (Yang and Villarini, 2019) and global climate models (Yang and Villarini, 2021) capture temporal clustering of heavy precipitation events across Europe. Building on the modes of variability, weather regimes (WRs) summarize regional climate variability on timescales of several weeks (e.g. Cassou, 2010; Ferranti et al., 2015; Grams et al., 2017). Recently, Grams et al. (2017) presented seven WRs in the Euro-Atlantic region designed to capture year-round, preferred states of the large-scale atmospheric circulation.

The recurrent nature of weather situations is potentially of relevance for the understanding of the temporal clustering of EPEs. For example, synoptic-scale transient recurrent Rossby wave packets, termed as RRWPs (Röthlisberger et al., 2019), affect persistent wet spells over the European sector, increasing wet spell duration in western Europe and western Russia during summer, and over eastern Europe and the Mediterranean in winter (e.g. Ali et al., 2021). How recurrent Rossby wave packets modulate the temporal clustering of individual extreme events remains to be investigated. Besides recurrent weather situations, stationary blocking anticyclones can also lead to temporal clustering of heavy precipitation as demonstrated by Barton et al. (2016), Lenggenhager and Martius (2019) on the Alpine south-side and by Tuel and Martius (2022) for selected areas around the Northern Hemisphere. Currently systematic quantification of the link between blocks and temporal clustering of heavy precipitation in Europe is still missing.

Temporal clustering is typically analyzed using methods that assume point processes, such as the dispersion statistic (e.g. Mailier et al., 2006; Vitolo et al., 2009; Pinto et al., 2013; Villarini et al., 2011, 2012), Ripley's K function (Barton et al., 2016; Tuel and Martius, 2021a,b) or count-based approaches (Bevacqua et al., 2020; Kopp et al., 2021). In agreement with this assumption, EPEs are here defined as events lasting from one to several days consisting of independent point events along the time axis. Independence is assured by grouping consecutive extreme precipitation days into single events (Barton et al., 2016; Tuel and Martius, 2021a,b; Kopp et al., 2021).

Previous studies have investigated the association between sub-seasonal temporal clustering of EPEs and the time-varying effect of large-scale atmospheric drivers in the Euro-Atlantic sector (e.g. Yang and Villarini, 2019, 2021). In these studies, this association is modeled using a Cox regression model, i.e. a Poisson process with a randomly varying rate of occurrence (e.g. Cox and Isham, 1980; Smith and Karr, 1986). Such a model has also been used to relate the temporal clustering of flood events (Villarini et al., 2013) and EPEs across the central United States (Mallakpour et al., 2017) to modes of climate variability. Several studies have relied on Poisson regression to model sub-seasonal temporal clustering of extratropical cyclones over Europe (Mailier et al., 2006; Vitolo et al., 2009), flood events in Austria (Villarini et al., 2012) and heavy precipitation in the North American Midwest (Villarini et al., 2011). However, these modeling approaches assume a linear combination of parameters and cannot parametrically account for the seasonal cycle of EPEs, which is omnipresent across Europe (Fig. 5). In a Cox regression model, seasonality is implicitly accounted for in the model through the baseline hazard, though it is not parametrically specified (Villarini et al., 2013). In Poisson regression, seasonality can either be accounted for by including binary indicator variables for each month (Mailier et al., 2006; Vitolo et al., 2009, e.g.) or by fitting a separate model for each season. A general additive model (GAM) lifts the linearity constraints; for example, Villarini et al. (2012) assesses non-linear effects of the NAO on the rate of occurrence of yearly flood counts using a Poisson GAM.

A GAM is essentially a Generalized Linear Model (GLM; McCullagh and Nelder, 2019) in which the relationships between the predictor variables and the response is specified by flexible smooth functions (refer to Hastie, 2017; Wood, 2017, for technical details). The flexibility that such a model offers has many advantages over classical GLMs. A GAM does not impose any parametric form of the link between the response and the predictor variables. This flexibility allows non-linear relationships between the predictor variables, i.e. atmospheric drivers, and the response, i.e. EPE counts, to be modeled. In practical terms, it is possible to implicitly account for the seasonal cycle in EPE counts by modeling them as a smoothly varying cyclic function, e.g. a cyclic regression cubic spline, of time, i.e. the day of the year. As such, it is possible to use one parsimonious model instead of separate models for each season; parameters are better estimated and temporal discontinuities between seasons are avoided.

In this study, we propose a novel approach to investigate sub-seasonal temporal clustering of EPEs across Europe using 42 years of daily precipitation data from ERA5 (Hersbach et al., 2020). This sub-seasonal temporal clustering is defined as random spurts of activity beyond the seasonal cycle of EPE counts, which we capture through Poisson GAMs fitted using data from all seasons. Our model results provide insight into the clustering frequency over Europe, i.e. to quantify the significance of temporal clustering in EPEs and to identify unique clustering episodes. Moreover, we investigate potential links between atmospheric dynamics and temporal clustering of precipitation extremes by inspecting the effects of seven predominant year-round WRs (Grams et al., 2017), atmospheric blocks being implicitly captured by the WRs, and RRWPs (Röthlisberger et al., 2019; Ali et al., 2021) in the Euro-Atlantic sector on the temporal clustering of EPEs on sub-seasonal time-scales. More specifically, we address the following questions: (1) where do we find significant sub-seasonal temporal clustering of EPEs in Europe beyond the seasonal cycle and how is it spatially distributed? (2) If temporal clustering is detected, where and when do the atmospheric predictors modulate temporal clustering of EPEs, i.e. explain positive deviations from the seasonal cycle in EPE counts?

The paper is organized as follows: the data is introduced in Section 2. Section 3 contains methodological aspects on EPE definition, modeling the temporal clustering using a GAM, WRs definition and quantification of the clustering. The results and discussion are presented in Sections 4 and 5 respectively. Finally, general conclusions and future research avenues are presented in Section 6.

2. ERA-5 data

We use European Centre for Medium-Range Weather Forecast's (ECMWF) latest reanalysis product, ERA-5. It is the fifth generation atmospheric reanalysis of the global climate from ECMWF and can be accessed publicly. We use precipitation fields from the ERA-5 reanalysis data set at 0.5° horizontal resolution from 1 January 1979 to 31 December 2020 (Hersbach et al., 2020). ERA-5 precipitation data are available at an hourly resolution that we aggregate to daily precipitation. ERA-5 precipitation stems from short-term numerical model forecasts and has been shown to be in good agreement with observational data (Rivoire et al., 2021) and capture temporal clustering of precipitation well (Tuel and Martius, 2021b). In this study, we model the temporal clustering of EPEs at 6140 gridpoint locations across Europe in the domain 30N-72N and 25W-40E. We use the ERA-5 land-sea mask to restrict the analysis only on land locations for computational efficiency and relevance in terms of impact. For convenience, we omit February 29 from all leap years contained in the 42 years of data.

3. Methods

3.1. Modeling the temporal clustering of extreme precipitation events

3.1.1. Definition of extreme precipitation events

We apply a Peak Over Threshold approach to select extreme rainfall days (Coles et al., 2001). The threshold for extreme rainfall days is the all-day 98th percentile of the local daily precipitation accumulation at each gridpoint (Fig. 1a). This threshold is chosen based on a trade-off between a high enough precipitation level while retaining enough

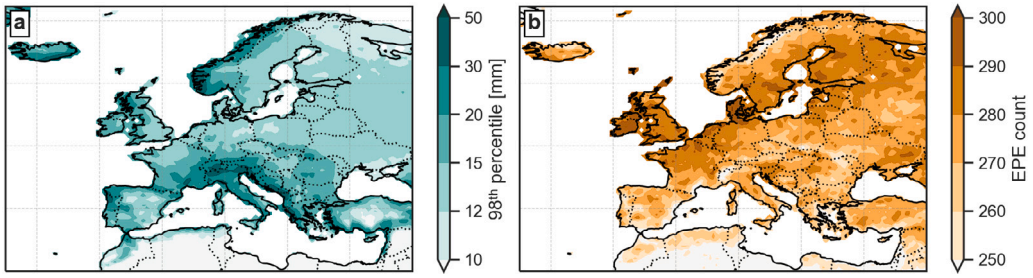

Fig. 1. Gridpoint-based (a) all-day 98th percentile of the daily precipitation distribution and (b) total number of EPEs.

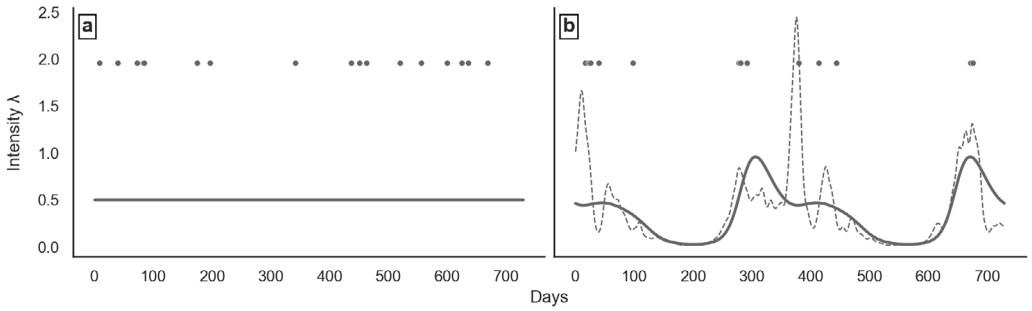

Fig. 2. Illustration of (a) a homogeneous Poisson process (dots) with constant intensity (solid line) and (b) a non-homogeneous Poisson process (dots) driven by seasonal variability (solid line) and predictor variables (dashed line) over 730 days along the time axis.

data. An extreme precipitation event is defined as an independent point or occurrence along the time axis and therefore runs of consecutive exceedances, .i.e. extreme rainfall days, are declustered. We apply runs declustering with a run length of 2 days, i.e. exceedances separated by no more than one day belong to the same event. A run length of 2 days has been determined based on objective statistical methods in relation to the extremal index (Ferro and Segers, 2003; Barton et al., 2016; Fukutome et al., 2015) and corresponds to the average time that a same mid-latitude weather system would generate extreme precipitation a given location (Lackmann, 2011). For events lasting longer than one day, the date of occurrence of the event corresponds to the day of the precipitation maximum within the event. The total number of independent EPEs ranges from 202 to 301 across the study domain (Fig. 1b).

3.1.2. EPEs as a non-homogeneous Poisson process

The one-dimensional Poisson process (Cox and Isham, 1980) is a stochastic point process that is used to model the occurrence of random points in time, i.e. events. The rate (or intensity) λ of the process is the only parameter of this model. It indicates the average number of events per time interval and its variance. If the intensity remains constant through time, i.e., λ is constant, then the point process corresponds to the homogeneous Poisson process (Fig. 2a). The non-homogeneous Poisson process has a rate $\lambda(t)$ that varies over time (Mailier et al., 2006). The EPE occurrence can be modeled with this process, where in Fig. 2b, λ is a function of (i) the annual cycle – a cyclic component – and/or (ii) continuous predictor variables.

3.1.3. Poisson general additive model

We extract 730 counts Y_t of EPEs in successive 21-day non-overlapping time intervals (indexed by t) from the 15330 day-long binary time series of extreme rainfall exceedances. The choice of a 21-day window allows to capture enough EPEs necessary for the modeling step and is suitable for studying impact relevant persistent atmospheric processes conducive the sub-seasonal clustering of EPEs without impinging upon longer seasonal time-scales. The sensitivity of the sub-seasonal clustering to the choice of the interval length has been investigated in Kopp et al. (2021), Tuel and Martius (2021b) and it is small. We assume that each observed count Y_t follows a Poisson distribution and has a probability of occurrence given by the probability mass distribution:

$$P(Y_t = y|\lambda_t) = \frac{e^{-\lambda_t}\lambda_t^y}{y!}, y = 0, 1, 2, \ldots$$

The Poisson distribution is characterized by a single rate parameter, here λ_t as defined in Section 3.1.2 (Fig. 2b). Only λ_t needs to be estimated to model the observed counts Y_t. It is possible to determine λ_t empirically, e.g. as the averaged number of events in each time interval as in the case of the homogeneous Poisson process (Fig. 2a). Here, we are interested in estimating λ_t from eight atmospheric predictors (as described in Section 3.2 and the day of the year, using a Poisson GAM. For each gridpoint, we model nine Poisson GAMs grouped into two types (Table 1). Regardless of the GAM, the model's base structure is identical and is composed of one term that accounts for the annual cycle of EPEs – the day of the year, d_t – and external continuous atmospheric predictors $x_{j,t}$, where j is the index of the atmospheric predictor. The effect of d_t is represented as a smooth function f_{cyclic} using periodic/cyclic cubic regression splines. The cyclic implementation is motivated by the expectation that f_{cyclic} needs to be temporally continuous from one year to the next, i.e. f_{cyclic} has the same value and first two derivatives at its upper and lower boundaries at day 365 and 1 (e.g. Fig. 3a). The inclusion of seasonality of the EPE occurrence in the model ensures that the atmospheric predictors explain departure from seasonality, i.e. random spurts of activity and inactivity (see dashed line in Fig. 2b), rather than the seasonality itself. The effects of atmospheric predictor j (e.g. Fig. 3b) is represented as a thin plate regression spline in one dimension, $f_{\text{tp},j}$. Additionally, we account for the seasonal effects of each of the atmospheric predictors by modeling their interaction with f_{cyclic}. In a GAM, this can be done by using a tensor product construction of the two marginal smooth functions, which results in a smooth 3-dimensional effect surface (e.g. Fig. 3c). In practice, all smooth terms have to be approximated with a finite number of basis functions. Suppose we have chosen the set of appropriate basis functions $\{a_1, \ldots, a_{n_\alpha}\}$ and $\{b_{1,j}, \ldots, b_{n_j,j}\}_{j=1,\ldots,8}$, such that $f_{\text{cyclic}}(d_t) = \sum_{i=1}^{n_\alpha} \alpha_i a_i(d_t)$ and $f_{\text{tp},j}(x_{j,t}) = \sum_{k=1}^{n_j} \beta_{k,j} b_{k,j}(x_{j,t})$, where n_α, $n_1, \ldots, n_p$ are the number of basis functions for the smooth terms and

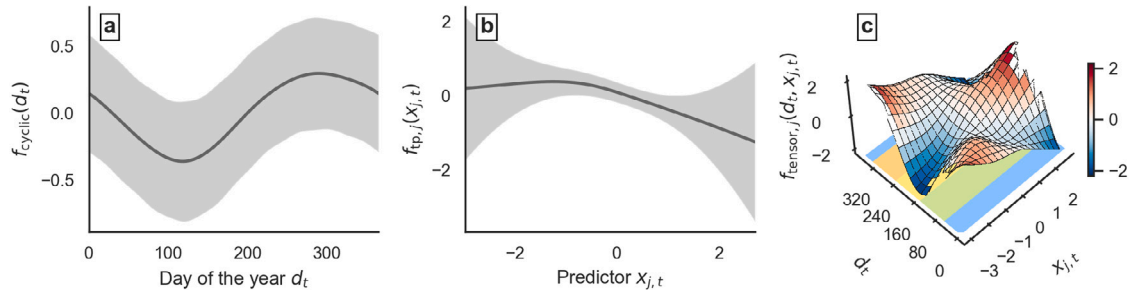

Fig. 3. Example of smooth functions of a fitted Poisson GAM on the scale of the linear predictor showing (a) the cyclic smooth function of d_t f_{cyclic}, (b) a smooth function of a continuous atmospheric predictor $f_{\text{tp},j}$ and (c) a tensor product interaction term between (b) and (c). Solid dark lines show the fitted functions, the gray shaded areas show the 95% confidence intervals, the 3D surface in (c) shows the fitted values of the tensor product and the colored sectors below the 3D surface delineate seasons over which the atmospheric predictor functions are averaged (refer to Section 3.1.4). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Summary of the nine GAM models used in this study. Brackets in the first column indicate the number of models in each type. Model 1 only includes the day of the year (d_t) as a predictor, whereas models 2-9 also include continuous atmospheric predictors, which are the 7 weather regimes (WRs) and the local R-metric (RRWPs).

	Purpose	Formula
Type 1 (1)	Quantification of clustering significance	$\log(\lambda_t) = \beta_0 + f_{\text{cyclic}}(d_t)$
Type 2 (8)	Modeling effects of 7 WRs and local RRWPs	$\log(\lambda_t) = \beta_0 + f_{\text{tensor},j}(d_t, x_{j,t})$

the α_i's and $\beta_{k,j}$'s are constant coefficients. Then, $f_{\text{tensor},j}(d_t, x_{j,t}) = \sum_{i=1}^{n_\alpha}\sum_{k=1}^{n_j}\eta_{i,k,j}a_i(d_t)b_{k,j}(x_{j,t})$, where the $\eta_{i,k,j}$'s are coefficients that need to be estimated. Without loss of generality, if the first basis functions a_1, $b_{1,1},\ldots,b_{1,8}$ are the unit functions (Chapter 5.6.3 of Wood, 2012), then we can decompose $f_{\text{tensor},j}$ into its main and interaction effects, i.e.,

$$f_{\text{tensor},j}(d_t, x_{j,t}) = \sum_{i=1}^{n_\alpha}\eta_{i,1,j}a_i(d_t) + \sum_{k=1}^{n_j}\eta_{1,k,j}b_{k,j}(x_{j,t}) + \sum_{i=2}^{n_\alpha}\sum_{k=2}^{n_j}\eta_{i,k,j}a_i(d_t)b_{k,j}(x_{j,t}), \quad (1)$$

which allows significance testing of the main and interaction effects separately.

The Poisson GAMs (Table 1) are constructed as follows: type 1 model uses $f_{\text{cyclic}}(d_t)$ as a single predictor and is used to model the baseline climatological rate of EPE occurrence based on which the clustering frequency is quantified. In addition to $f_{\text{cyclic}}(d_t)$, type 2 models uses the atmospheric predictors, i.e. weather regime indices and longitudinally averaged values of RRWPs (described in Section 3.2). The reasoning behind using separate atmospheric GAMs is that WRs and RRWPs share the same underlying data, which can lead to collinearity issues, i.e. one predictor can be linearly predicted from the others with a substantial degree of accuracy. We use one separate model for each WR, so that each index's effect can be easily interpreted. If the goal were model accuracy, we could consider incorporating all WR indices into one model instead. All predictors are standardized (mean of 0, standard deviation of 1) before the modeling step. The optimal smoothing parameters for the smooth terms are estimated via restricted maximum-likelihood; we use the penalized iterative reweighted least squares algorithm which also prevents overfitting (Wood, 2017). Note that the atmospheric predictors may implicitly explain/absorb part of the seasonal cycle but this aspect is not discussed. The p-values are calculated by Wald tests with null hypotheses that the smooth function is zero everywhere that it is defined on; low p-values indicate low likelihood that the basis functions that make up the function are jointly zero (Wood, 2012).

3.1.4. Reporting GAM effects

An effect is meaningful in a GAM if it is significantly different from a flat zero line, i.e. anything from a linear slope to a complex wiggly curve. To obtain a measure which is analogous to the regression slope of a predictor in a GLM, here we use a finite differences approach to approximate the sign and strength of a predictor's effect (Simpson, 2022). For each atmospheric predictor, we evaluate the linear slope at 365 equidistantly located points along the function to derive an acceptable global estimate $m = \Delta y/\Delta x$ of the first derivative of the function. We retrieve slope values m for each season, obtained after adding up the main effect function with the interaction effect function through the decomposition in Eq. (1), averaged by season (colored sectors in Fig. 3c). Since the slopes are computed on the log scale, we report back the exponent of the slope values, which corresponds to a multiplicative factor (or percent change) of the expected rate λ_t for a one unit change (standard deviation) in the predictor variable, all else held constant. Inference about the shape of the effect functions is made using the effective degrees of freedom (edf) estimated from the GAMs, which can be used as a proxy for the degree of non-linearity in stressor-response relationships (Zuur et al., 2009). An edf of 1 is equivalent to a linear relationship whereas an edf > 1 indicates a non-linear relationship. Only significant main or interaction effects for which the adjusted p-value is 0.05 or below are shown. We use the Benjamini and Hochberg procedure (Benjamini and Hochberg, 1995) to control for the false discovery rate (FDR). When multiple hypothesis tests are conducted, FDR is way to identify as many significant features as possible while incurring a relatively low proportion of false positives.

3.2. Low-frequency atmospheric predictors

3.2.1. Weather regimes

The WRs definition relies on standard approaches using EOF-clustering to identify the seven prevailing year-round large-scale atmospheric patterns in the Euro-Atlantic sector (for a detailed description of the methodology, refer to supplementary material of Grams et al., 2017). The WRs consist of three variants of cyclonic regimes, that is the Atlantic Trough (AT), the Zonal regime (ZO) and the Scandinavian Trough (ScTr); moreover, four blocked regimes, namely the Atlantic Ridge (AR), the European Blocking (EuBL), the Scandinavian Blocking (ScBL) and the Greenland Blocking (GL) are identified. To that end, EOF analysis is performed on 10-day low-pass-filtered (LPF) Z500 anomaly (Z500a), using a 90-day running mean at the respective calendar time as reference climatology, in the domain 30N-90N and 90W-40E. Then, K-means clustering (with $k = 7$) is applied on the seven leading EOF time series (79% of explained variance). Prior to the EOF-clustering, the seasonal cycle in the amplitude of the Z500a is removed to compensate for lower amplitudes in summer compared to winter. The seven cluster mean EOF patterns are shown in Figs. 6–13. Weather regime indices are constructed by projecting the

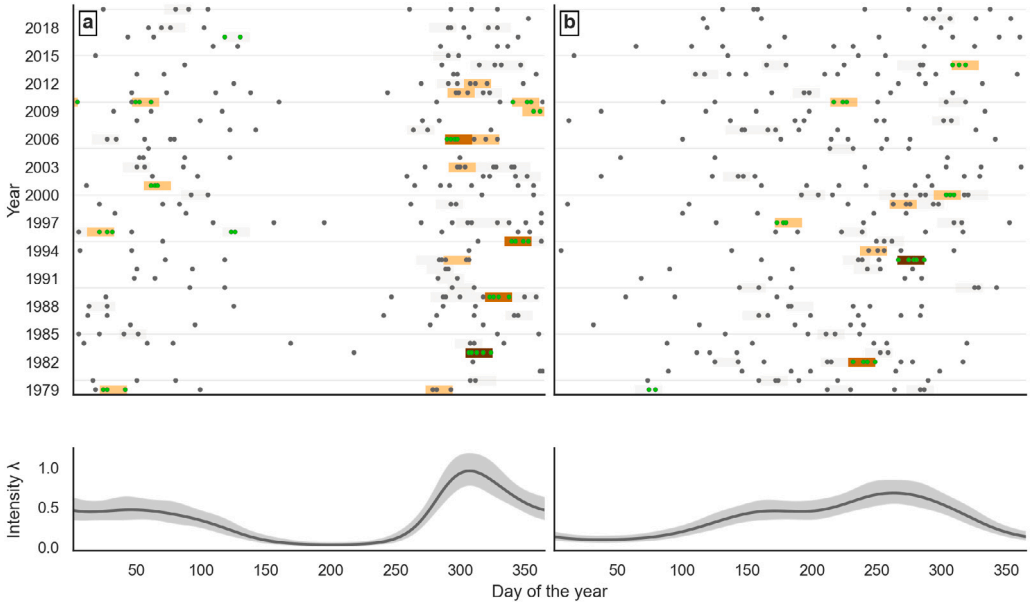

Fig. 4. Significantly clustered EPEs are shown as green points for the cities of (a) Sevilla in Spain and (b) Lugano in Switzerland. Points show the distribution of all EPEs as a function of the day of the year (x-axis) and the year (y-axis). 21-day time intervals with empirical counts of 2 and above are shown in shades of color (gray = 2, light brown = 3, medium brown = 4 and dark brown = 5). The bottom panels show the climatological expected rate for each calendar day as modeled with the simple GAM model that includes only the seasonal variability, see Section 3.1.3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

instantaneous normalized Z500a in the WR patterns following Michel and Rivière (2011); it is those indices that are used as atmospheric predictor variables for the modeling step. A positive index reveals strong resemblance to the cluster mean pattern, whereas a negative index reveals a mirrored anomaly pattern.

3.2.2. Recurrent Rossby waves packets: R-metric

Additionally, we use a predictor variable that quantifies the recurrence of phase-locked transient Rossby waves in quick succession: the *R*-metric (refer to Röthlisberger et al. (2019) for technical details). It is a simple measure of RRWPs strength derived from Hovmöller (time-longitude) diagrams of instantaneous meridionally averaged (35° N–65° N) meridional wind at 250 hPa (V250). The *R*-metric is in short the envelope of a time (15-day running mean) and wavenumber (WN) filtered (WN = 4 to WN = 15) wave given by V250 field (Ali et al., 2021). The metric is available for each longitude at each time step. For our model predictor, we use daily *R*-metric values at each longitudinal point corresponding to the one of the precipitation gridpoint. The *R*-metric values are first averaged over a 60° sector centered on each longitudinal gridpoint to account for the effects of the neighboring gridpoints (as in Röthlisberger et al., 2019; Ali et al., 2021).

3.3. Quantification of clustering significance

In this section, we describe the approach we use to determine the proportion of significantly clustered EPEs. Temporal clustering is identified by comparing the empirical distribution of counts in a 21-day disjoint time interval C_{21} to the fitted Poisson distribution with baseline climatological rate. We illustrate the approach for two gripdoint locations corresponding approximately to the cities of Sevilla in Spain and Lugano in Switzerland (refer to Fig. 4). The probability that the count C_{21} is equal to or greater than $k = 0, 1, 2, 3 \ldots$ expected events from a Poisson distribution is, assuming that the Poisson model is appropriate:

$$P(C_{21} \geq k) = 1 - \sum_{i=0}^{k-1} \frac{\lambda_t^i e^{-\lambda_t}}{i!} \tag{2}$$

where λ_t corresponds to the climatological expected rate of occurrence for each calendar day (see bottom panels in Fig. 4) and is predicted with a type one Poisson GAM for which only $f_{\text{cyclic}}(d_t)$ is used as a predictor (refer to 1). The value of λ_t used to derive the probabilities in Eq. (2) is taken on the day at the center of the respective 21-day interval. Since λ_t comes with an associated uncertainty (standard errors of the prediction), we use the upper boundary of the point-wise confidence interval of λ_t, more robust, instead of the fitted value itself. For each of the 730 counts from the 21-day disjoint intervals, we identify probabilities below the 5% threshold, thus selecting those time intervals for which the probability of observing a specific count is less than 5%, label the EPEs within those time intervals as significantly clustered (only for k ≥ 2; shown as green dots in Fig. 4) and report back the proportion of clustered EPEs with respect to all EPEs, which corresponds to 15% and 9% for Sevilla and Lugano respectively.

The advantage of the present approach is that it allows to select exact time periods when temporal clustering occurs and quantify its frequency and severity after removal of the seasonal cycle in the response. Thus, if temporal clustering is detected, then seasonality in the EPE counts is not able to explain temporal clustering by itself which is an indication that there must be other drivers at play. For example, with reference to Fig. 4, the climatological expected rate of occurrence in Lugano is highest in autumn. Nevertheless, the time interval spanning from September 23 to October 13 1993 counts 5 EPEs, which is significantly larger than expected on average, meaning that seasonality alone does not explain clustering. In contrast, the time interval spanning from September 17 to October 7 1999 counts 3 EPEs, which is a high count but within the expected range of the climatological rate of occurrence, thus leading to non-significantly clustered EPEs, likely explained by seasonality.

4. Results

4.1. Estimated seasonal cycle in EPE occurrences

Fig. 5 depicts the climatological expected rate of occurrence for each calendar day obtained with a simple GAM (model type 1) computed at each of the 6140 gridpoints in Europe and displays the range in rate of the estimated functions. Most locations reveal a pronounced seasonal cycle in the occurrence of EPEs as revealed by the positive values in the range between the minimum and maximum rate of occurrence of

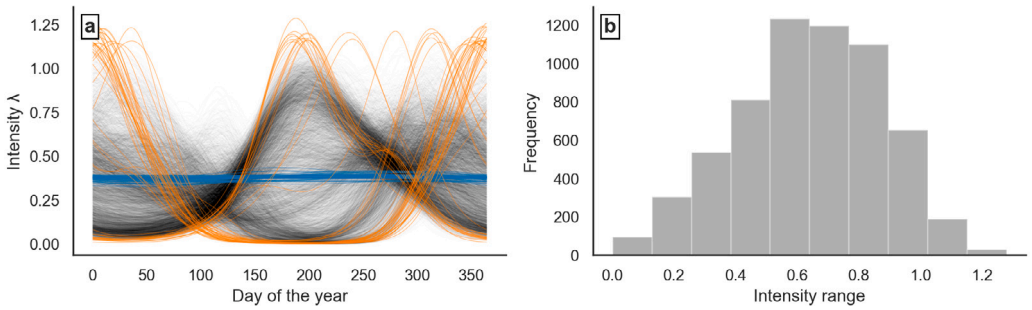

Fig. 5. Estimated seasonal cycle of EPEs at each gridpoint over Europe. The left sub-figure displays the climatological expected rate of occurrence for each calendar day (cyclic smooth functions) computed at each gridpoint in Europe. The day of the year is shown on the x-axis and the 21-day intensity λ on the y-axis. Shown in orange (blue) are the top 50 functions in terms of largest (smallest) range in rate. The right sub-figure displays the frequency distribution of all ranges of seasonal functions for all gridpoints in Europe. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

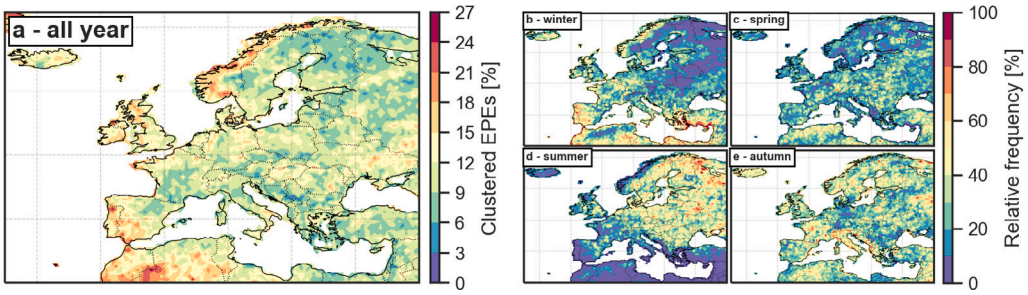

Fig. 6. Spatial distribution of the proportion of significantly clustered EPEs (in percent) at each gridpoint in Europe over the entire analysis period (a) and its relative frequency (in percent) in the winter (b), spring (c), summer (d) and autumn (e).

the estimated functions in Fig. 5b. The European average range is 0.64, which means that the average difference in the number of events in a 21-day period is 0.64 during the active season compared to the inactive season. Fig. 5a reveals two general clusters, one with a cycle peaking in summer, another one peaking during the extended winter season. Non-significant seasonal cycles (p-values > 0.05) are found in 78 out of 6140 gridpoint locations, i.e. the estimated function is not significantly different from a flat zero line. The top 50 curves in terms of amplitude (shown in orange in Fig. 5a) are characterized by pronounced seasonal differences in the climatological expected rate of occurrence, ranging from 0 to nearly 1.3 expected events in a 21-day time period at these particular locations.

4.2. Spatial distribution of EPE clustering

Fig. 6a depicts the spatial distribution of the proportion of EPEs that are found significantly clustered on sub-seasonal (21-day) timescales at each gridpoint in Europe between 1979 and 2020. Between 2% to 27% of all EPEs cluster in time. Local extremes are found over northern Ireland and northern Portugal where over 24% of the EPEs significantly cluster in time. Over 21% of the EPEs significantly cluster in time along the westernmost, northern and southern coast of Norway, westernmost tip and southwestern coast of France, northern Portugal and southern tip of the Iberian Peninsula. Gridpoints where more than 18% of EPEs are clustered are located along the western coasts of Norway, Great Britain and France, as well as over the western half of the Iberian Peninsula. Clustering frequency is generally lower inland with however some larger values in regions north of the Black Sea, over northeastern Germany and Hungary, where over 12%–15% of the EPEs significantly cluster in time. Gridpoints where less than 6% of EPEs are clustered are located over northern Sweden, Finland and Russia, over the Baltic countries, southern parts of Poland, Austria and France as well as over the northeastern Dinaric Alps with a European minima of 3% and less over southwestern Serbia. EPE clustering is characterized by seasonal variability. Most of the temporal clustering of EPEs present over Greece, southern Turkey, along the southern Mediterranean coastline, the western coast of Norway and the northwestern coast of France, as well as over the western half of the Iberian Peninsula occurs in the winter (Fig. 6b). In regions with larger values located inland, temporal clustering of EPEs tends to occur more frequently in the summer (Fig. 6d). Clustering occurs most frequently in autumn over Iceland, Ireland, central France, central and northern Italy, along the eastern coast of Spain and along the coastline of the Adriatic Sea. For Switzerland, the seasonal occurrence of EPE clustering depends on the geographic region. While 60%–90% of the clustering occurs in autumn in southern Switzerland, clustering affects northeastern Switzerland mostly in spring, while the central Alps are mostly concerned in winter.

Note that the proportion of significantly clustered EPEs can vary depending on the starting day of the counting process. Here, the largest detectable frequency is shown.

4.3. Effects of weather regimes on EPE clustering

The univariate effects of the 7 Euro-Atlantic weather regimes on modulating the rate of occurrence of EPEs for each season are displayed in Figs. 7 to 13. The effects are shown as a multiplicative factor and should be understood as the change in the expected rate of occurrence for a 1 standard deviation increment of the predictor variable. A multiplicative factor greater than 1 (in shades of red), indicates a larger than expected number of EPEs with respect to the baseline climatological rate of occurrence and therefore a tendency towards clustering. A multiplicative factor smaller than 1 (in shades of blue), indicates a smaller than expected number of EPEs with respect to the baseline climatological rate of occurrence.

With increasing AT regime (Fig. 7), the increase in the probability of EPE clustering (in red) is strongest in northwestern France where it is more than doubled with respect to the climatological expected rate in winter. An increase by 20% to 50%, regionally by 50% to 100% of the clustering probability due to AT, is seen in northwestern Iceland, over the British Isles, the northwestern part of the Iberian Peninsula, parts

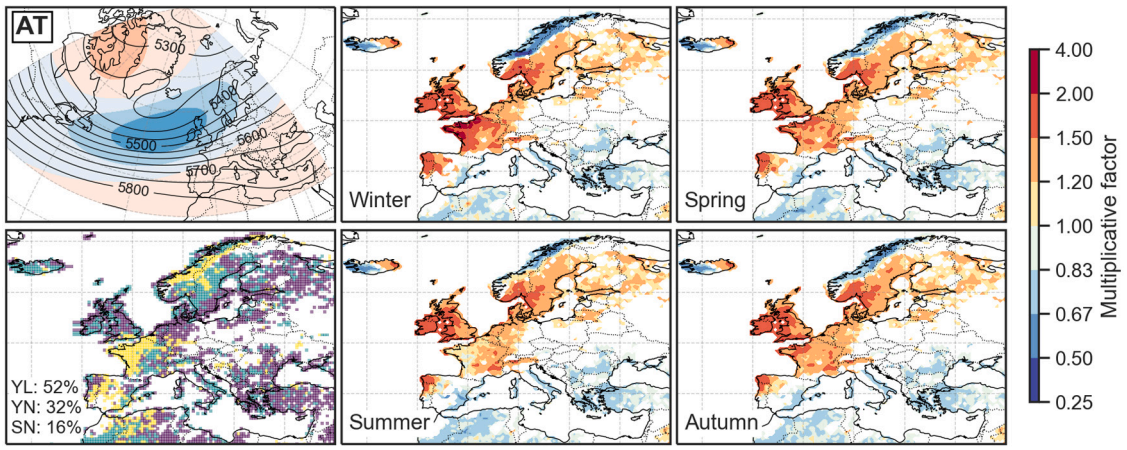

Fig. 7. Mean LPF (10 days) Z500 anomaly (shading, every 40 geopotential meters) and mean absolute Z500 (black contours, every 50 geopotential meters) of the Atlantic Trough (AT) weather regime (top left) and its effect (multiplicative factor) on EPE clustering for winter (top center), spring (top right), summer (bottom center) and autumn (bottom right). Positive (negative) effects are shown in shades of red (blue) and non-significant estimates in gray shading. The bottom left inset displays whether the predictor is year-round linear (YL; purple), year-round non-linear (YN; green) or seasonal non-linear (SN; yellow) along with proportions belonging to each class. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

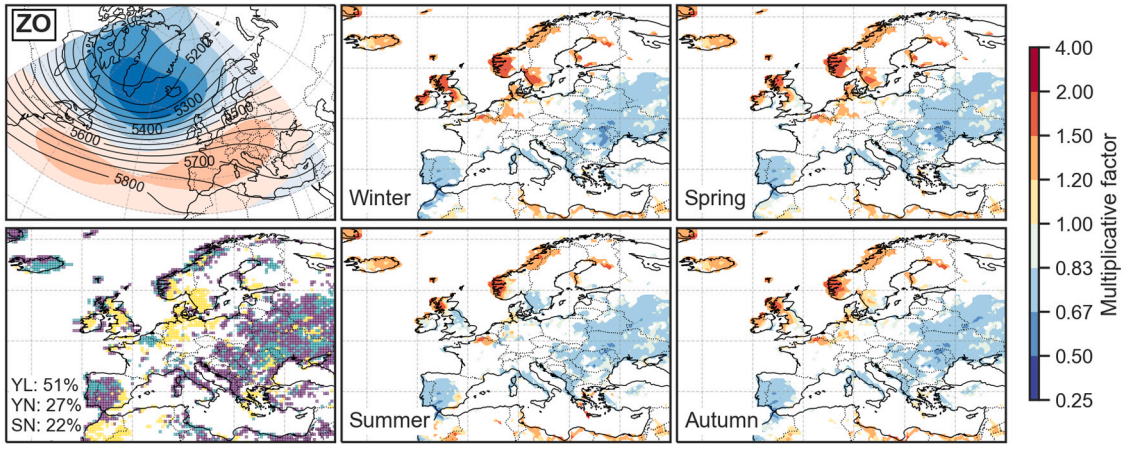

Fig. 8. Same as Fig. 7 for the ZO weather regime. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of western Europe, over southern Scandinavia and along the southern Alpine mountain range, particularly in southern Switzerland. Moreover, an increase by 20% to 50% is revealed over parts of northeastern Europe. In northwestern France, the effect of AT on clustering is characterized by seasonal variability, with a weaker increase in summer (below 20%); elsewhere in Europe, the effect is equal across all seasons. A decrease in the probability of EPE clustering (in blue) is seen over southwestern Iceland, over central and northern Norway and generally over southern Europe, particularly along the eastern flanks of the Iberian Peninsula and Italy, as well over Greece, western Turkey and regions to the west of the Black Sea. The decrease is characterized by seasonal variability in central Norway where it is largest across Europe with more than a halving of the clustering probability in winter. In southern Europe, the decrease is mostly equal across all seasons and corresponds to 20% to 50% of the climatological expected rate. The log-linear relationship of the effect of AT on clustering is year-round linear in 52%, year-round non-linear in 32% and seasonal non-linear in 16% of all significant grid-points across Europe.

With increasing ZO regime (Fig. 8), the probability of EPE clustering increases (in red) by 50% to 100% over northwestern regions of the British Isles, western Norway, southwestern Sweden, the western coast of Denmark and more locally over northern France and Finland; an increase by 20% to 50% is seen over western Iceland, central Norway and northwestern Germany. In the aforementioned locations, the effect of ZO on EPE clustering is mostly equal across all seasons with the exception of central England, Germany, Denmark and Sweden, characterized by seasonal variability and a reversed effect in summer with a decrease in clustering probability by 20% to 50%. Moreover, the positive effect of ZO on clustering extends further inland in winter over parts of the British Isles and western Norway. A decrease in the probability of EPE clustering (in blue) is seen generally over southern Europe, particularly in the southeastern half of the Iberian Peninsula and Italy, as well over Greece and regions to the north and west of the Black Sea. In these regions, the decrease is mostly equal across all seasons and corresponds to a 20% to 50% decrease (locally 50% to 100%) of the climatological rate over parts of southern Spain and western Ukraine. Southwestern parts of Spain and Greece are characterized by seasonal variability with an increase in the clustering probability in summer. The log-linear relationship of the effect of ZO on clustering is year-round linear in 51%, year-round non-linear in 27% and seasonal non-linear in 22% of all significant grid-points across Europe.

A one standard deviation increase of the ScTr regime index (Fig. 9) increases the probability of EPE clustering (in red) by 20% to 50% along the northern coastline of Iceland, central Norway, most of Finland and over some parts of central Europe; by 50% to 100% over northwestern Scotland and western Norway and more than a doubling locally over the southwestern coast of Finland in winter. The latter shows a seasonally variable effect of the ScTr on clustering, strongest

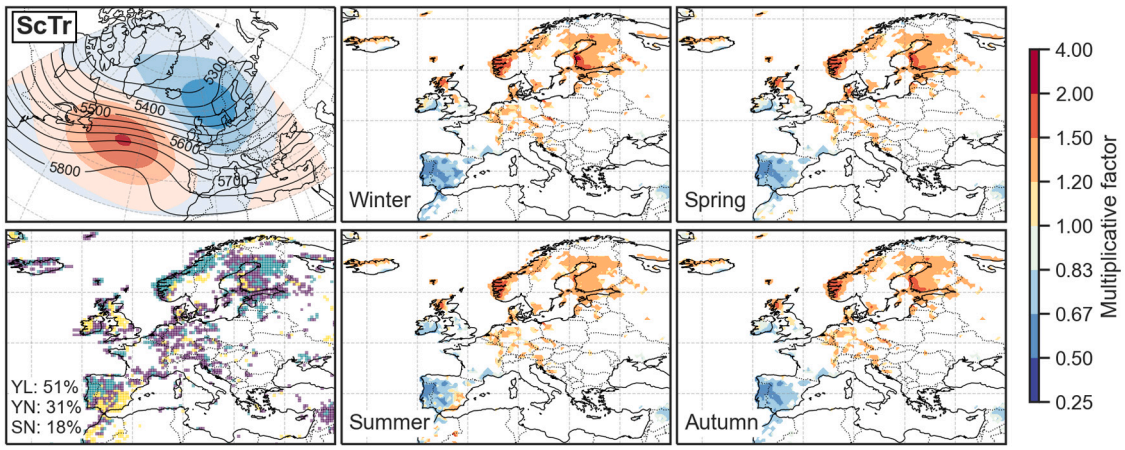

Fig. 9. Same as Fig. 7 for the ScTr weather regime. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

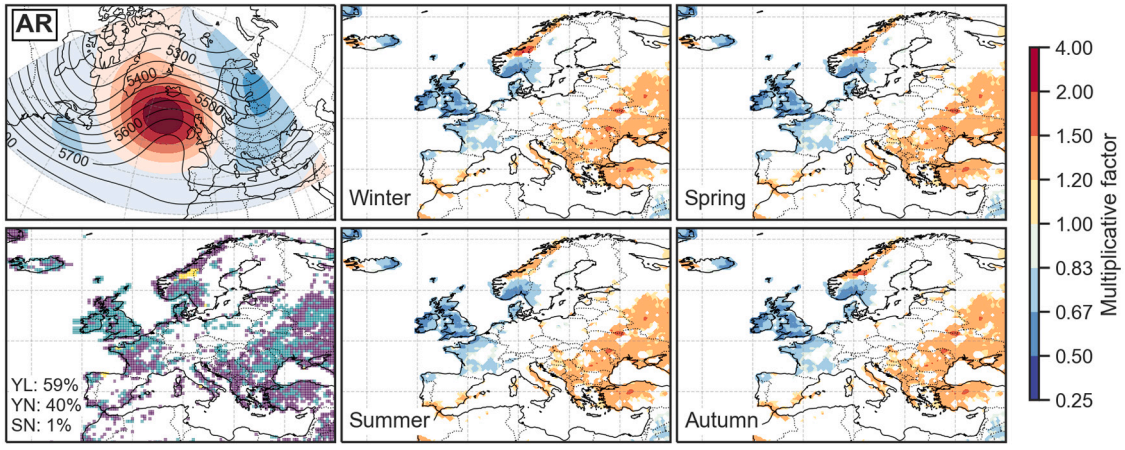

Fig. 10. Same as Fig. 7 for the AR weather regime. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in winter, weakest in summer (20% to 50%), whereas the remaining aforementioned locations reveal equal effects across all seasons. A decrease in the probability of EPE clustering (in blue) is seen generally over southwestern Europe, particularly in the southeastern half of the Iberian Peninsula, along the french Mediterranean coast of the Gulf of Lion and southern Ireland. In these regions, the decrease is mostly consistent throughout the year and corresponds to a 20% to 50% decrease (50% to 100% over the Iberian Peninsula). Southwestern Spain is characterized by seasonal variability with a reversed effect in summer, that is an increase of 20% to 50% in clustering probability. The log-linear relationship of the effect of ScTr on clustering is year-round linear in 51%, year-round non-linear in 31% and seasonal non-linear in 18% of all significant grid-points across Europe.

With increasing AR regime (Fig. 10), the increase of the probability of EPE clustering (in red) is strongest over central Norway, with more than a doubling of the clustering rate in winter; this increase is characterized by seasonal variability, weakest in summer (20% to 50%). Elsewhere in Europe, the effect of AR on clustering is equal across all seasons with an increase of the clustering probability of 20% to 50% over the westernmost part of Iceland, over northwestern Norway and parts of southern Europe, such as southeastern and southwestern Spain, along the eastern flanks of the Apennine Mountains, over the Balkans, western Turkey and in regions to the north and west of the Black Sea. An AR regime is associated with a decrease in the probability of EPE clustering (in blue) by 20% to 50% over eastern Iceland, the British Isles, the northwestern quarter of the Iberian Peninsula, parts of France, Denmark, southern Norway and Sweden as well as along the southern Alpine foothills and locally by 50% to 100% over the eastermost part of Iceland, the western parts of the British Isles and southern Norway. The log-linear relationship of the effect of AR on clustering is year-round linear in 59%, year-round non-linear in 40% and seasonal non-linear in 1% of all significant grid-points across Europe.

With increasing EuBL regime (Fig. 11), the increase of the probability of EPE clustering (in red) is most prominent over northern Norway, with locally more than a doubling of the clustering rate in winter; this increase is characterized by seasonal variability and is weaker in the remaining seasons (50% to 100%). The EuBL regime increases clustering probability consistently throughout the year by 50% to 100% over southwestern Iceland and by 20% to 50% over parts of southern Europe, such as the southeastern parts of Spain, Italy and Greece. The effect of EuBL on clustering is characterized by seasonal variability revealing contrasting effects over most of France, the southwestern Iberian Peninsula, north of the Dinaric Alps and in southern Ukraine, with an increase in clustering probability by 20% to 50% in summer (50% to 100% over the Iberian Peninsula), whereas the remaining seasons show a weaker increase or even a decrease (particularly over France and the Iberian Peninsula; 50% to 100%). An AR regime is associated with a decrease in the probability of EPE clustering (in blue) over wide parts of Europe, particularly in winter, by 20% to 50% over the British Isles, western Europe, southern Scandinavia and northeastern Europe, and regionally by 50% to 100% over the northwestern quarter of the Iberian Peninsula, western and eastern France, the Netherlands, southwestern England, western and southern Switzerland. The log-linear relationship of the effect of EuBL on clustering is year-round linear in 40%, year-round non-linear in

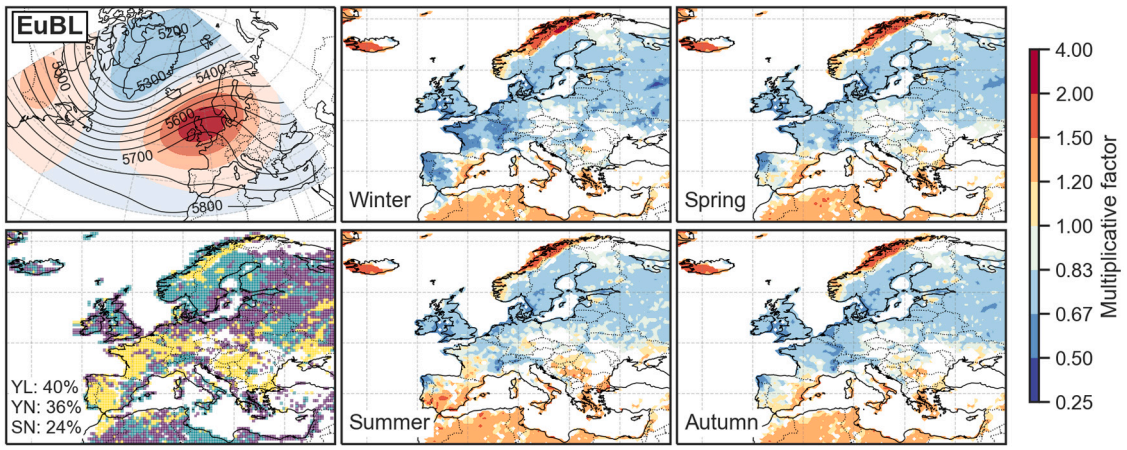

Fig. 11. Same as Fig. 7 for the EuBL weather regime. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

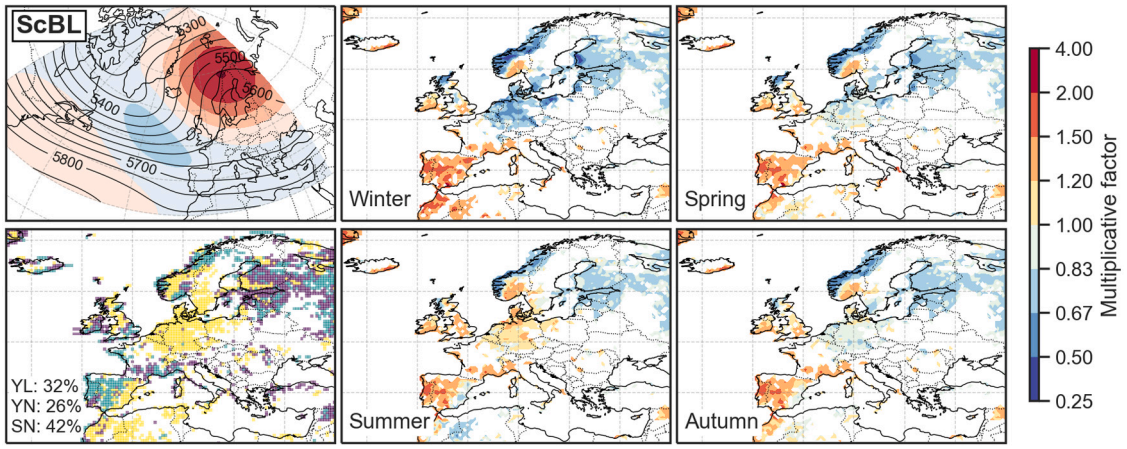

Fig. 12. Same as Fig. 7 for the ScBL weather regime. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

36% and seasonal non-linear in 24% of all significant grid-points across Europe.

An increasing ScBL regime (Fig. 12) is associated with an increase in the clustering probability (in red), equal across all seasons, of 50% to 100% over parts of the western Iberian Peninsula and along the southern coast of Iceland, and of 20% to 50% over the southern parts of British Isles, France and Norway. The effect of ScBL on clustering is characterized by seasonal variability over central Europe, revealing contrasting effects over Germany, the Benelux, Switzerland, northwestern Poland, Denmark, southern Sweden and northern Scotland, with an increase in clustering probability of up to 50% in summer, whereas the remaining seasons show a decrease, peaking in winter by up to 100%. Seasonal variability also concerns eastern Spain, where clustering probability increases by 50% to 100% in winter, but decrease by 20% to 50% in summer. While a decrease in the expected rate under a ScBL regime is seen equal across all seasons over the Baltic countries (20% to 50%), southern Finland (20% to 50%) and northwestern Norway (50% to 100%), a seasonally varying decrease is observed over parts of northern Scotland, southern Sweden, western Norway and Finland, with more than a halving of the rate in the latter two. The log-linear relationship of the effect of ScBL on clustering is year-round linear in 32%, year-round non-linear in 26% and seasonal non-linear in 42% of all significant grid-points across Europe.

With increasing GL regime (Fig. 13), the increase of the probability of EPE clustering (in red) is most prominent over the Iberian Peninsula, with a 50% to 100% increase over northwestern Portugal and a similar increase in southern Spain in autumn and winter; the latter is characterized by seasonal variability with a weaker increase in summer (up to 20%) and a reversed effect with a 20% to 50% decrease on the Mediterranean coast. Contrasting effects are observed over central parts of the British Isles and southern Sweden with a 50% to 100% increase in summer, but a reversed effect in winter revealing a decrease of 20% to 50%. A consistent throughout the year increase of 20% to 50% is seen over western parts of Ireland and England, as well as over France, western and southeastern Switzerland, the western coasts of Italy and the Balkans, southeastern Sweden and eastern European regions north of the Black Sea. The latter shows, regionally, a seasonal effect, stronger in autumn and winter. A GL regime is associated with a decrease, equal across all seasons, in the probability of EPE clustering (in blue) over southwestern Iceland (20% to 50%, locally up to 50% to 100% inland) and northwestern Norway (20% to 50%). A decrease characterized by seasonal variability is seen over western Norway and northwestern Scotland, by 50% to 100% in winter and spring, weaker in summer and autumn. The log-linear relationship of the effect of GL on clustering is year-round linear in 48%, year-round non-linear in 25% and seasonal non-linear in 28% of all significant grid-points across Europe.

4.4. Weather regime predominance on EPE clustering

Fig. 14 displays the spatial distribution of dominant Euro-Atlantic weather regimes that maximize clustering of extreme precipitation at each gridpoint per season. General patterns equal across all seasons are first described and concern over 80% of all gridpoints in the study domain.

The AT regime exhibits to most extensive effect on clustering across all seasons in Europe affecting most of western and northern regions

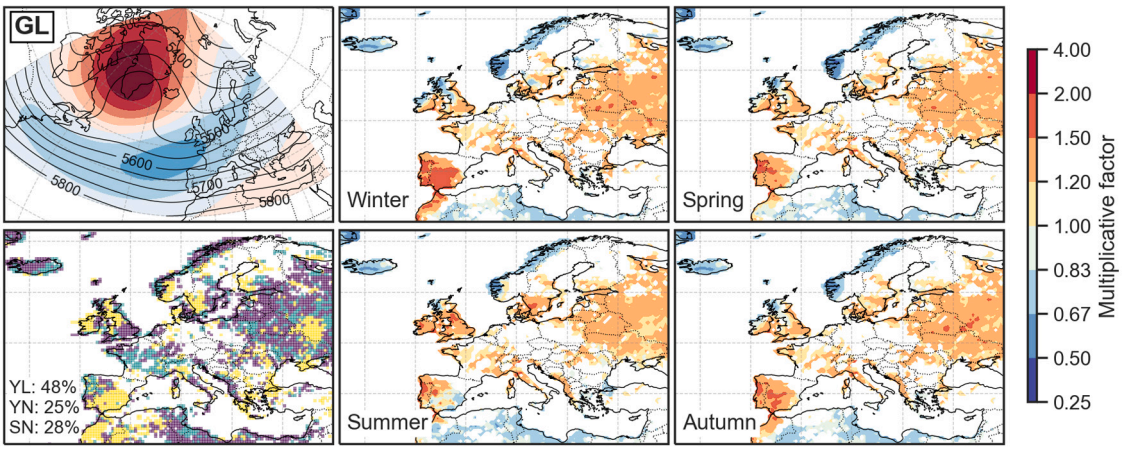

Fig. 13. Same as Fig. 7 for the GL weather regime. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

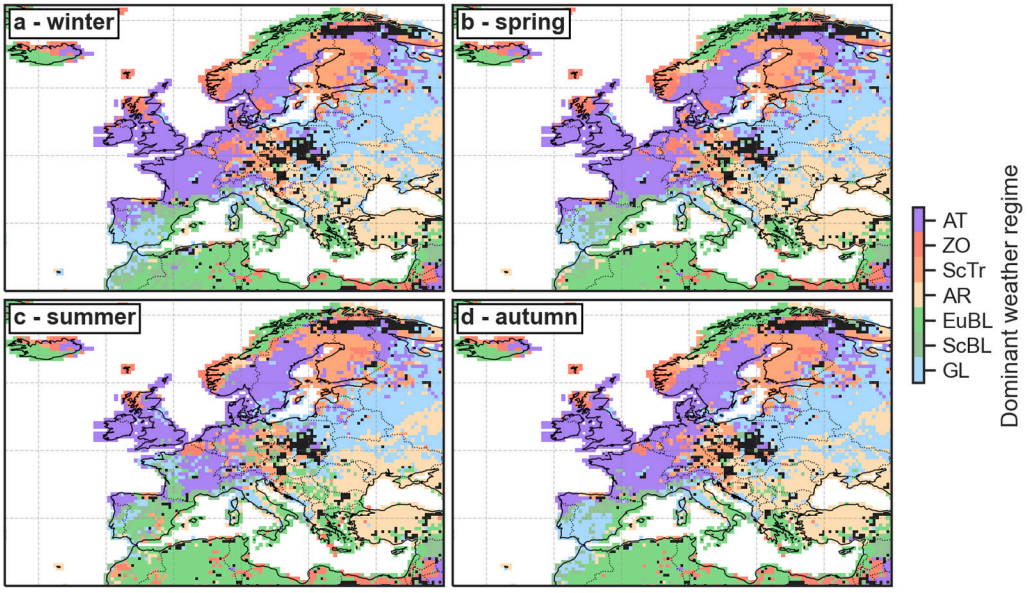

Fig. 14. Spatial distribution of dominant Euro-Atlantic weather regimes that maximize clustering of extreme precipitation at each gridpoint for winter, spring, summer and autumn. Gridpoint locations revealing no significance for all weather regimes or a reduction of the clustering probability are shown in dark gray shading. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

such as eastern Iceland, Great Britain, France, the northwestern Iberian Peninsula, western and southern Switzerland, Benelux, western Germany, southern Scandinavia and northeastern Finland. The second most extensive regime in Europe is the GL regime, affecting most of the Iberian Peninsula, eastern Europe, Corsica, part of northern Italy along the coasts of the Tyrrhenian Sea and along the southern Balkan coastline. The EuBL is the dominant regime in the southwestern part of Iceland, in the northern half of Norway, along the Mediterranean coasts of the Iberian Peninsula, the eastern Balearic Islands and Sardinia, southern Sicily, southeastern Italy and Greece. The AR regime is the dominant regime affecting clustering over central parts of Norway, Italy and Sardinia, as well as Ibiza, most of southeastern Europe such as Turkey, the western parts of Greece, the Balkan countries and regions to the wester of the Black Sea. The ScTr predominates clustering of EPEs over parts of Scandinavia, such as northern Sweden, most of Finland and the northern part of western Norway, as well as parts of central Europe and to a lesser extent the northwestern coastline of Iceland. A ZO regime affects northern Europe to a lesser extent in that it maximizes year-round clustering probability over some parts of northeastern Iceland, northwestern Scotland and the southern part of western Norway. Finally, the ScBL regime is the least predominant weather regime in terms of clustering of EPEs over Europe affecting the southeastern coastline of Iceland, northern-central parts of the Iberian Peninsula, the french Mediterranean coast of the Gulf of Lion, northeasternmost Sicily and locally over northwestern and southeastern Italy. Several regions exhibit seasonal variability in the predominance of weather regime patterns. In summer, the ScBL regime becomes the dominant flow pattern in terms of clustering over parts of Germany, the Czech Republic, western Poland, over western France and parts of the southern Iberian Peninsula. Also in summer, EuBL becomes predominant in parts of the southern Iberian Peninsula and southern France along the northern slopes of the Pyrenees. In summer, clustering tends to be predominated by the GL regime over northwestern France. For the latter, clustering is predominated by the ZO regime for the northernmost region in summer and autumn only. Finally, clustering in parts of northwestern Germany, Denmark and southwestern Sweden is predominantly driven by ZO only in winter and spring.

4.5. *Effects of recurrent transient waves on EPE clustering*

Fig. 15 reveals statistically significant effects of the R-metric, i.e. persistent and/or recurrent meridional amplification of Rossby waves

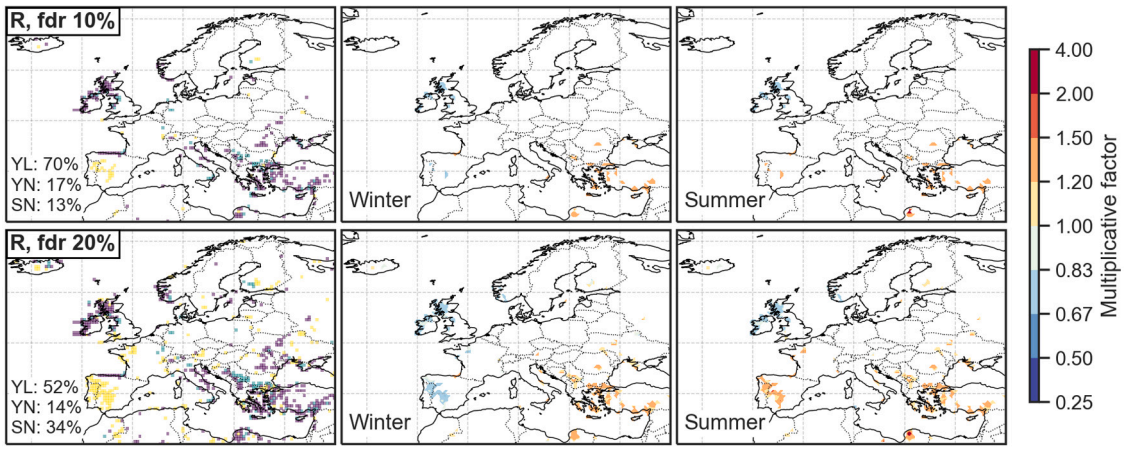

Fig. 15. Effect (multiplicative factor) of the R-metric on EPE clustering with an FDR control of 10% (top row) and 20% (bottom row) for winter (center) and summer (right). Positive (negative) effects are shown in shades of red (blue) and non-significant estimates in gray shading. The left inset displays whether the predictor is year-round linear (YL; purple), year-round non-linear (YN; green) or seasonal non-linear (SN; yellow) along with proportions belonging to each class. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

over a fixed region, on the temporal clustering of EPEs. Autumn and spring reveal very similar results than winter and summer; therefore only the latter two are shown. The effect of the *R*-metric appears to be much less pronounced than the effects of the weather regimes. Nevertheless, coherent patterns can be observed over the northern parts of the British Isles, the Iberian Peninsula and southeastern Europe. With increasing *R*-metric values, clustering probability of EPE occurrence tends to increase consistently throughout the year by 20% to 50% over parts of Greece, Turkey and the Dinaric Alps. A similar increase is observed over parts of the Iberian Peninsula in summer and autumn; there, the effect is reversed in winter and spring with a decrease of the clustering probability by 20% to 50%. A decrease of similar magnitude but equal across all seasons is seen over the northern British Isles.

5. Discussion

5.1. Advantages and limitations of using a GAM

Using GAMs to model temporal clustering of EPEs has many advantages. Its high flexibility allows the use of a single year-round model; temporal inconsistencies arising at the transition from one season to the next are avoided and uncertainty estimation is improved since more data is available and strength can be borrowed directly from neighboring data points in adjacent seasons. GAMs allow to tackle the stationarity issue naturally present in most point processes such as in EPEs by including a simple term to account for time, i.e. the day of the year, estimated by a cyclic cubic regression spline smooth, which tends to a straight line if there is no seasonal effect. A major drawback that comes with increasing complexity is the loss of interpretability. For example it is not possible to rely on a single slope parameter to report back the sign and magnitude of a predictor's effect.

5.2. Seasonal cycle and temporal clustering of EPEs

Accounting for the seasonal cycle in the EPE occurrence is justified since significant year-round non-stationarity in the EPE occurrence is identified in 99% of gridpoint locations in Europe (refer to Fig. 5). Significantly temporally clustered EPEs can be identified at every European gridpoint, ranging from 2% to 27%; this poses an inevitable risk that needs to be taken seriously. The location and spatial extent of the clustered EPEs coincide with findings presented by recent studies (e.g. Tuel and Martius, 2021b; Kopp et al., 2021); they concern the southwestern Iberian Peninsula, the northwestern coast of France, Scotland and Norway. Our approach is less conservative on the significance assessment than compared to methods that test against simulated samples of homogeneous Poisson series, however allows identification of isolated clustering episodes and therefore significant clustering is also detected elsewhere in Europe, such as over eastern Europe.

5.3. Weather regimes explain part of clustering

The sub-seasonal variability of the North Atlantic atmospheric circulation can be described as transitions between a limited number of recurrent and quasi-stationary states, i.e. weather regimes (Ferranti et al., 2015) and affect surface weather conditions on timescales of several days to weeks (Grams et al., 2017). Our results reveal the importance of weather regimes in describing sub-seasonal temporal clustering in EPEs over Europe, explaining up to 30% of the variance over the western Iberian Peninsula and Norway, and generally over 20% in regions exhibiting high clustering frequencies such as northwestern Europe, as reported by the deviance explained (Fig. 16).

5.4. Implications of weather persistence on clustering

The AT affects increase in clustering probability most extensively in Europe in all seasons, except in summer where it is equally important as the EuBL regime, followed by the GL regime (Fig. 14). This finding strongly suggests that clustering of EPEs over western and northern Europe is governed by the position and strength of the polar-front jet, which in turn controls the rate at which cyclonic low pressure systems and associated fronts elapse. Further to the point, the southward position of the jet during the AT compared to the ZO regime reveals that the positive phase of the NAO, most closely related to the ZO regime (Grams et al., 2017), explains only a small part of the European clustering as opposed to the AT regime.

The ZO mainly affects western coastal regions of northern Europe. Here, the zonal orientation of the jet is crucial for generating EPE clustering, as emphasized by the significant effect of a high R-metric (i.e. persistent synoptic-scale recurrence of upper-level meridional flow) on decreasing clustering probability (Fig. 15).

The climatological position of the polar-front jet being shifted southward during the northern hemisphere cold season is reflected in the clustering patterns. For example, the effect of the AT is more pronounced in France in winter than in summer. Moreover, the positive feedback of ZO on clustering is shifted southward in winter exposing also Germany and Sweden to enhanced clustering probability.

The GL regime is the second most important regime in explaining clustering in Europe in the northern hemisphere cold season and particularly affects the Iberian Peninsula in winter. There, clustering of

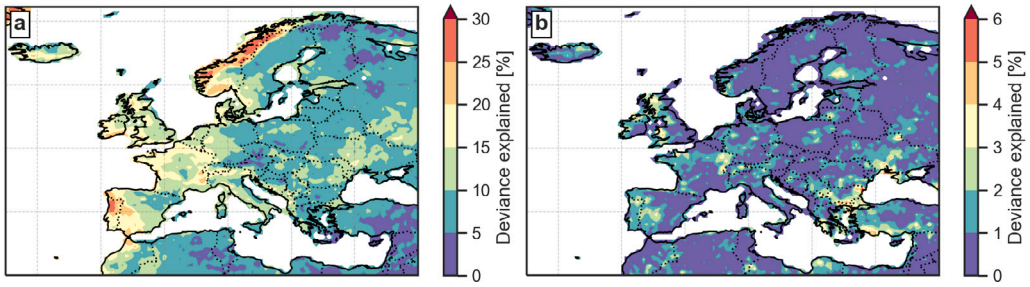

Fig. 16. Difference in deviance explained between the model using only the day of the year and (a) the WRs model (b) the R-metric model.

precipitation extremes can be attributed to a general slackening and southward shift of the zonal flow suggesting advection of moist air from the sub-tropical Atlantic towards western Europe. Our results also emphasize the importance of meridional wave amplification over the eastern Atlantic sector and the Iberian Peninsula, as revealed by large R-metric values in Fig. 15 at these locations. This is particularly the case in summer and autumn when, concomitant to a Greenland block, the phase-locked downstream Rossby wave propagation reveals preferred trough formation over the eastern Atlantic sector, exposing the Iberian Peninsula to persistent northwesterly to southwesterly upper-level flow. Ali et al. (2021) found a significant increase in wet-spell persistence under high R-metric values for the northern Iberian Peninsula in the extended summer season, which corroborates our findings. Without being too speculative, recurrent trough formation over the eastern Atlantic/western Europe is associated with warm and moist air advection from the Atlantic and Mediterranean towards the Alps. In fact as shown by Fig. 13, a GL regime significantly and considerably increases clustering of extreme precipitation along the Jura Mountains and eastern France, and along the southern Swiss Alps; this is a typical spatial pattern observed during southwesterly flow situations, often associated with orographic lifting and forced convection on the windward side of the mountainous slopes in Switzerland.

The remaining weather regimes have a more regionalized effect on European EPE clustering. The AR regime leads to a decrease (increase) in the clustering probability over some parts of northwestern Europe (over northwestern Norway). The decrease can be partly attributed to a persistent East Atlantic block forcing the polar-front jet to circumvent it over northern Europe. The deflection of the moist jet takes a westerly to northwesterly turn when it impinges on the western coasts of Norway perpendicular to the axis of the Scandinavian Mountains favoring clustering of EPEs, particularly in autumn and winter. Most of the EPEs in this area are a result of orographically enhanced precipitation from fronts embedded in North Atlantic low-pressure systems (Sandvik et al., 2018). In response to the persistent East Atlantic ridge, the increase in the clustering of EPEs over the Balkans and Turkey can be tied to the downstream meridional extension of a trough/streamer towards the eastern Mediterranean; this is supported by significant R-metric values in these regions (Fig. 15). This upper-level anomaly may evolve into the formation of cut-off lows and surface low-pressure systems, wandering over the Mediterranean Sea towards the west in direction of the southeastern flank of the Atlantic block, which could partly explain the clustering signal along the eastern coasts of Italy, Sardinia and Spain. In fact, Mastrantonas et al. (2021) found that Balkan and Black Sea Lows are highly associated with EPEs over the Balkans and Turkey, Sicilian Lows favor EPEs in locations east of the Apennines while EPEs over eastern Sardinia and the Spanish Mediterranean coasts are tied to Iberian Lows. These low-pressure systems associated with EPEs have been shown to be linked to atmospheric instability over the Mediterranean sector, which are particularly present in the south/southeast flank of blocking systems (Sousa et al., 2017).

ScBL/EuBL regimes preclude clustering of EPEs over wide parts of central and northern Europe, and affect regions directly located below the core of the anticyclonic block and regions usually exposed to moisture advection from the west now under the influence of dry continental air advection from the east, e.g. Norwegian coast for ScBL and Scotland for ScBL/EuBL. Cassou (2010) found an increase in the frequency of rainfall extremes under a Scandinavian block over the Spanish and French Mediterranean coast and along the eastern coasts of Sardinia, Sicily and southern Italy; our results reveal similar regional patterns in the increase of EPE clustering however during a EuBL. The spatial patterns of increased clustering in the Mediterranean appear to be similar as during an AR regime, but shifted to the south during a EuBL and to the west during a ScBL, and highly resemble EPE patterns in connection with Iberian and Sicilian Lows (Mastrantonas et al., 2021). As for the AR regime, the polar-front jet circumvents the western Mediterranean regions (or has a weak influence), generating a context of weak pressure gradients, in which instable conditions favor errant cut-off lows over the Mediterranean. Parts of central Europe experiencing a pronounced decrease in clustering probability in winter during a ScBL regime reveal an inverse signal, that is an increase in EPE clustering in summer. A similar shift is observed by Lenggenhager and Martius (2019) in the odds ratio of moderate precipitation extremes. The associated increase in atmospheric instability (and consequently in convective processes) to the south/southeast of the blocking center shift northward during summer, and as a consequence, positive precipitation anomalies also extend further north than during other seasons (Sousa et al., 2017). A recent example are the central European floods in July 2021 during which a persistent Scandinavian Block was present over nearly a month. The block favored the formation of a series of cut-off lows to wander over western and central Europe, in particular the near-stationary cut-off low-pressure system Bernd centered over Germany, generating a succession of intense convective rainfall events.

6. Conclusions

We use Poisson GAMs to quantify sub-seasonal temporal clustering of extreme precipitation events from 1979 to 2020 and model its dependence to seven Euro-Atlantic weather regimes and a measure synoptic-scale transient recurrent Rossby wave packets. Significant 21-day clustering episodes are found at all grid-points over Europe. The fraction of EPEs that cluster in time ranges between 2% to 27% (Fig. 6). Temporal clustering is most frequent along the western coasts of Norway, Great Britain and France as well as over the western half of the Iberian Peninsula, where it mainly occurs during autumn and winter.

The GAMs allow us to model clustering of EPEs using one model for the entire by accounting for seasonality in the EPEs with a cyclic smooth function of the day of the year. The GAMs provide seasonal estimates for each atmospheric predictor variable. Weather regimes explain up to 30% of the temporal clustering over the western Iberian Peninsula and on average over 20% in regions exhibiting high clustering frequencies (Fig. 16). The following weather regimes are linked to clustering of EPE in Europe (in decreasing order of importance): the Atlantic Trough, the Greenland Blocking, the European Blocking, the Atlantic Ridge, the Scandinavia Trough, the Zonal regime and the Scandinavian Blocking.

Future perspectives involve examining the catalogue of European clustering episodes determined in this study in more depth. For example by performing case studies of the most improbable clustering episodes to gain deeper insight on the complexity and variety of processes leading to sub-seasonal clustering of precipitation extremes. Second, our results have demonstrated the potential of weather regimes for predicting temporal clustering of EPEs, which could be further investigated with the use of teleconnection patterns on longer time-scales that drive weather regimes themselves. This could be achieved by expanding the GAMs used in this study through the addition of interaction terms. Last, our approach can easily be transferred to other fields of application and presents a flexible basis for the investigation of the temporal clustering of any desired variable on various time-scales.

Code availability

https://github.com/ywbarton/GAM_temporal_clustering.git.

CRediT authorship contribution statement

Yannick Barton: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing, Visualization. **Pauline Rivoire:** Methodology, Writing – original draft. **Jonathan Koh:** Methodology, Writing – original draft, Writing – review & editing. **Mubashshir Ali S.:** Writing – original draft. **Jérôme Kopp:** Writing – original draft. **Olivia Martius:** Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Code and data availability

ERA5 reanalysis data are available from 1979 to present and can be downloaded from https://apps.ecmwf.int/datasets/. All calculations related to GAMs and the Cox regression model are performed in R (R Core Team, 2020) using the freely available mgcv (Wood, 2011) and survival (Therneau, 2020) packages whereas the remaining tasks are performed in Python. Code availability: https://github.com/ywbarton/GAM_temporal_clustering.git.

Acknowledgments

The authors thank two anonymous reviewers for their constructive and helpful comments that lead to the improvement of this paper. Special thanks go to Christian Grams for providing weather regime data (cluster-mean spatial patterns and inidices). The authors also thank Alexandre Tuel for constructive discussions on the topic. YB and OM acknowledge support from the Swiss Science Foundation, grant number 178751.

References

Ali, S Mubashshir, Martius, Olivia, Röthlisberger, Matthias, 2021. Recurrent rossby wave packets modulate the persistence of dry and wet spells across the globe. Geophys. Res. Lett. 48 (5), e2020GL091452.

Barton, Yannick, Giannakaki, Paraskevi, Von Waldow, Harald, Chevalier, Clément, Pfahl, Stephan, Martius, Olivia, 2016. Clustering of regional-scale extreme precipitation events in southern Switzerland. Mon. Weather Rev. 144 (1), 347–369.

Benjamini, Yoav, Hochberg, Yosef, 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Stat. Methodol. 57 (1), 289–300.

Bevacqua, Emanuele, Zappa, Giuseppe, Shepherd, Theodore G., 2020. Shorter cyclone clusters modulate changes in European wintertime precipitation extremes. Environ. Res. Lett. 15 (12), 124005.

Casanueva, A, Rodríguez-Puebla, Concepción, Frías, MD, González-Reviriego, Nube, 2014. Variability of extreme precipitation over europe and its relationships with teleconnection patterns. Hydrol. Earth Syst. Sci. 18 (2), 709–725.

Cassou, Christophe, 2010. Euro-atlantic regimes and their teleconnections. In: Proceedings: ECMWF Seminar on Predictability in the European and Atlantic Regions. pp. 6–9.

Coles, Stuart, Bawa, Joanna, Trenner, Lesley, Dorazio, Pat, 2001. An Introduction to Statistical Modeling of Extreme Values, vol. 208. Springer.

Comas-Bru, Laia, McDermott, Frank, 2014. Impacts of the EA and SCA patterns on the European twentieth century NAO–winter climate relationship. Q. J. R. Meteorol. Soc. 140 (679), 354–363.

Cox, David Roxbee, Isham, Valerie, 1980. Point Processes, vol. 12. CRC Press.

Ferranti, Laura, Corti, Susanna, Janousek, Martin, 2015. Flow-dependent verification of the ECMWF ensemble over the euro-atlantic sector. Q. J. R. Meteorol. Soc. 141 (688), 916–924.

Ferro, Christopher A.T., Segers, Johan, 2003. Inference for clusters of extreme values. J. R. Stat. Soc. Ser. B Stat. Methodol. 65 (2), 545–556.

Fukutome, Sophie, Liniger, M.A., Süveges, M., 2015. Automatic threshold and run parameter selection: a climatology for extreme hourly precipitation in Switzerland. Theor. Appl. Climatol. 120 (3), 403–416.

Galarneau, Jr., Thomas J, Hamill, Thomas M, Dole, Randall M, Perlwitz, Judith, 2012. A multiscale analysis of the extreme weather events over western Russia and northern Pakistan during July 2010. Mon. Weather Rev. 140 (5), 1639–1664.

Grams, Christian M, Beerli, Remo, Pfenninger, Stefan, Staffell, Iain, Wernli, Heini, 2017. Balancing Europe's wind-power output through spatial deployment informed by weather regimes. Nature Clim. Change 7 (8), 557–562.

Grams, Christian M, Binder, Hanin, Pfahl, Stephan, Piaget, Nicolas, Wernli, Heini, 2014. Atmospheric processes triggering the central European floods in June 2013. Nat. Hazards Earth Syst. Sci. 14 (7), 1691–1702.

Guo, Yuming, Wu, Yao, Wen, Bo, Huang, Wenzhong, Ju, Ke, Gao, Yuan, Li, Shanshan, 2020. Floods in China, COVID-19, and climate change. Lancet Planet. Health 4 (10), e443–e444.

Hastie, Trevor J., 2017. Generalized Additive Models. Routledge.

Hersbach, Hans, Bell, Bill, Berrisford, Paul, Hirahara, Shoji, Horányi, András, Muñoz-Sabater, Joaquín, Nicolas, Julien, Peubey, Carole, Radu, Raluca, Schepers, Dinand, et al., 2020. The era5 global reanalysis. Q. J. R. Meteorol. Soc. 146 (730), 1999–2049.

Huntingford, Chris, Marsh, Terry, Scaife, Adam A, Kendon, Elizabeth J, Hannaford, Jamie, Kay, Alison L, Lockwood, Mike, Prudhomme, Christel, Reynard, Nick S, Parry, Simon, et al., 2014. Potential influences on the United Kingdom's floods of winter 2013/14. Nature Clim. Change 4 (9), 769–777.

Kenyon, Jesse, Hegerl, Gabriele C., 2010. Influence of modes of climate variability on global precipitation extremes. J. Clim. 23 (23), 6248–6262.

Kopp, Jérôme, Rivoire, Pauline, Ali, S Mubashshir, Barton, Yannick, Martius, Olivia, 2021. A novel method to identify sub-seasonal clustering episodes of extreme precipitation events and their contributions to large accumulation periods. Hydrol. Earth Syst. Sci. Discuss. 1–27.

Krichak, Simon O, Breitgand, Joseph S, Gualdi, Silvio, Feldstein, Steven B, 2014. Teleconnection–extreme precipitation relationships over the mediterranean region. Theor. Appl. Climatol. 117 (3), 679–692.

Lackmann, Gary, 2011. Midlatitude Synoptic Meteorology. American Meteorological Society.

Lau, William K.M., Kim, Kyu-Myong, 2012. The 2010 Pakistan flood and Russian heat wave: Teleconnection of hydrometeorological extremes. J. Hydrometeorol. 13 (1), 392–403.

Lenggenhager, Sina, Martius, Olivia, 2019. Atmospheric blocks modulate the odds of heavy precipitation events in Europe. Clim. Dynam. 53 (7), 4155–4171.

Mailier, Pascal J, Stephenson, David B, Ferro, Christopher AT, Hodges, Kevin I, 2006. Serial clustering of extratropical cyclones. Mon. Weather Rev. 134 (8), 2224–2240.

Mallakpour, Iman, Villarini, Gabriele, Jones, Michael P, Smith, James A, 2017. On the use of Cox regression to examine the temporal clustering of flooding and heavy precipitation across the central United States. Glob. Planet. Change 155, 98–108.

Martius, O, Sodemann, H, Joos, H, Pfahl, S, Winschall, A, Croci-Maspoli, M, Graf, M, Madonna, E, Mueller, B, Schemm, S, et al., 2013. The role of upper-level dynamics and surface processes for the Pakistan flood of July 2010. Q. J. R. Meteorol. Soc. 139 (676), 1780–1797.

Mastrantonas, Nikolaos, Herrera-Lormendez, Pedro, Magnusson, Linus, Pappenberger, Florian, Matschullat, Jörg, 2021. Extreme precipitation events in the Mediterranean: Spatiotemporal characteristics and connection to large-scale atmospheric flow patterns. Int. J. Climatol. 41 (4), 2710–2728.

McCullagh, Peter, Nelder, John A., 2019. Generalized Linear Models. Routledge.

Michel, Clio, Rivière, Gwendal, 2011. The link between Rossby wave breakings and weather regime transitions. J. Atmos. Sci. 68 (8), 1730–1748.

Pinto, Joaquim G, Bellenbaum, Nina, Karremann, Melanie K, Della-Marta, Paul M, 2013. Serial clustering of extratropical cyclones over the North Atlantic and Europe under recent and future climate conditions. J. Geophys. Res.: Atmos. 118 (22), 12–476.

Priestley, Matthew DK, Pinto, Joaquim G, Dacre, Helen F, Shaffrey, Len C, 2017. The role of cyclone clustering during the stormy winter of 2013/2014. Weather 72 (7), 187–192.

R Core Team, 2020. R: A language and environment for statistical computing. URL https://www.R-project.org/.

Raymond, Colin, Horton, Radley M, Zscheischler, Jakob, Martius, Olivia, AghaKouchak, Amir, Balch, Jennifer, Bowen, Steven G, Camargo, Suzana J, Hess, Jeremy, Kornhuber, Kai, et al., 2020. Understanding and managing connected extreme events. Nature Clim. Change 10 (7), 611–621.

Rivoire, Pauline, Martius, Olivia, Naveau, Philippe, 2021. A comparison of moderate and extreme ERA-5 daily precipitation with two observational data sets. Earth Space Sci. 8 (4), e2020EA001633.

Röthlisberger, Matthias, Frossard, Linda, Bosart, Lance F, Keyser, Daniel, Martius, Olivia, 2019. Recurrent synoptic-scale Rossby wave patterns and their effect on the persistence of cold and hot spells. J. Clim. 32 (11), 3207–3226.

Sandvik, Mari Ingeborg, Sorteberg, Asgeir, Rasmussen, Roy, 2018. Sensitivity of historical orographically enhanced extreme precipitation events to idealized temperature perturbations. Clim. Dynam. 50 (1), 143–157.

Scaife, Adam A, Folland, Chris K, Alexander, Lisa V, Moberg, Anders, Knight, Jeff R, 2008. European climate extremes and the North Atlantic Oscillation. J. Clim. 21 (1), 72–83.

Simpson, Gavin L., 2022. gratia: Graceful ggplot-based graphics and other functions for GAMs fitted using mgcv. URL https://gavinsimpson.github.io/gratia/ R package version 0.6.9850.

Smith, James A., Karr, Alan F., 1986. Flood frequency analysis using the Cox regression model. Water Resour. Res. 22 (6), 890–896.

Sousa, Pedro M, Trigo, Ricardo M, Barriopedro, David, Soares, Pedro MM, Ramos, Alexandre M, Liberato, Margarida LR, 2017. Responses of European precipitation distributions and regimes to different blocking locations. Clim. Dynam. 48 (3), 1141–1160.

Therneau, Terry M., 2020. A package for survival analysis in r. URL https://CRAN.R-project.org/package=survival R package version 3.2-7.

Tuel, Alexandre, Martius, Olivia, 2021a. A climatology of sub-seasonal temporal clustering of extreme precipitation in Switzerland and its impacts. Nat. Hazards Earth Syst. Sci. Discuss. 1–28.

Tuel, Alexandre, Martius, Olivia, 2021b. A global perspective on the sub-seasonal clustering of precipitation extremes. Weather Clim. Extremes 33, 100348.

Tuel, Alexandre, Martius, Olivia, 2022. Subseasonal temporal clustering of extreme precipitation in the northern hemisphere: Regionalization and physical drivers. J. Clim. 35 (11), 3537–3555.

Villarini, Gabriele, Smith, James A, Baeck, Mary Lynn, Vitolo, Renato, Stephenson, David B, Krajewski, Witold F, 2011. On the frequency of heavy rainfall for the Midwest of the United States. J. Hydrol. 400 (1–2), 103–120.

Villarini, Gabriele, Smith, James A, Serinaldi, Francesco, Ntelekos, Alexandros A, Schwarz, Ulrich, 2012. Analyses of extreme flooding in Austria over the period 1951–2006. Int. J. Climatol. 32 (8), 1178–1192.

Villarini, Gabriele, Smith, James A, Vitolo, Renato, Stephenson, David B, 2013. On the temporal clustering of US floods and its relationship to climate teleconnection patterns. Int. J. Climatol. 33 (3), 629–640.

Vitolo, Renato, Stephenson, David B, Cook, Ian M, Mitchell-Wallace, Kirsten, 2009. Serial clustering of intense European storms. Meteorol. Z 18 (4), 411–424.

Wood, Simon N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J. R. Stat. Soc. (B) 73 (1), 3–36.

Wood, Simon N., 2012. On p-values for smooth components of an extended generalized additive model. Biometrika (ISSN: 0006-3444) 100 (1), 221–228. http://dx.doi.org/10.1093/biomet/ass048, arXiv:https://academic.oup.com/biomet/article-pdf/100/1/221/478658/ass048.pdf.

Wood, Simon N., 2017. Generalized Additive Models: An Introduction with R. CRC Press.

Yang, Zhiqi, Villarini, Gabriele, 2019. Examining the capability of reanalyses in capturing the temporal clustering of heavy precipitation across Europe. Clim. Dynam. 53 (3), 1845–1857.

Yang, Zhiqi, Villarini, Gabriele, 2021. Evaluation of the capability of global climate models in reproducing the temporal clustering in heavy precipitation over Europe. Int. J. Climatol. 41 (1), 131–145.

Zscheischler, Jakob, Martius, Olivia, Westra, Seth, Bevacqua, Emanuele, Raymond, Colin, Horton, Radley M, van den Hurk, Bart, AghaKouchak, Amir, Jézéquel, Aglaé, Mahecha, Miguel D, et al., 2020. A typology of compound weather and climate events. Nat. Rev. Earth Environ. 1 (7), 333–347.

Zuur, Alain F, Ieno, Elena N, Walker, Neil J, Saveliev, Anatoly A, Smith, Graham M, et al., 2009. Mixed Effects Models and Extensions in Ecology with R, vol. 574. Springer.