Investigating the typicality of the dynamics leading to extreme temperatures in the IPSL-CM6A-LR model

Robin Noyelle, Pascal Yiou, Davide Faranda

Investigating the typicality of the dynamics leading to extreme temperatures in the IPSL-CM6A-LR model

Robin Noyelle[1*], Pascal Yiou[1] and Davide Faranda[1,2,3]

[1]Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay & IPSL, Gif-sur-Yvette, 91191, France.
[2]London Mathematical Laboratory, 8 Margravine Gardens London, W6 8RH, London, United Kingdom.
[3]Laboratoire de Météorologie Dynamique, Ecole Normale Supérieure, PSL research University & IPSL, Paris, France.

*Corresponding author(s). E-mail(s): robin.noyelle@lsce.ipsl.fr;
Contributing authors: pascal.yiou@lsce.ipsl.fr;
davide.faranda@lsce.ipsl.fr;

Abstract

Determining the underlying mechanisms leading to extreme events in dynamical systems is a challenging task. Under mild hypotheses, large deviations theory predicts that as one increases the threshold defining an extreme, dynamical trajectories which reach the extreme will look more and more like one another: they converge towards a typical, i.e. most probable, one called the instanton. In this paper, we use a 2000-year simulation of the IPSL-CM6A-LR model under a stationary pre-industrial climate to test this prediction on the case of hot extremes. We investigate whether the physical mechanisms leading to extreme temperatures at four locations in Europe are more similar with increasing extreme temperatures. Our results show that most physical variables exhibit the expected convergence towards a most probable trajectory, with some geographical and temporal variations. In particular, we observe the presence of a cut-off low in some trajectories, which suggests the existence of multiple pathways leading to extreme temperatures. These findings confirm the relevance of instanton dynamics in understanding the physical mechanisms driving extreme events in climate models.

Keywords: extreme events, large deviations theory, instanton, typicality, atmospheric dynamics

1 Introduction

Extreme weather events can have tremendous impacts on societies and ecosystems (Pörtner et al, 2022). Among them, heatwaves have been the focus of extensive attention due to their increasing frequency with global warming (Seneviratne et al, 2021). Their impacts include adverse health effects, increased power consumption, infrastructure damages, forest fires, droughts and crop failures (Koppe et al, 2004; Zuo et al, 2015; Yaghmaei, 2020). In the last decade, extensive research has been conducted to better understand the dynamics leading to heatwaves (Perkins, 2015; Horton et al, 2016; Domeisen et al, 2022a). The general scenario combines specific atmospheric synoptic conditions and anomalously low soil moisture, which can interact to further increase the intensity of the event.

Heatwaves in the midlatitudes are usually associated with a slow moving, sometimes called quasi-stationary, high-amplitude Rossby wave (Petoukhov et al, 2013). This structure is often embedded in a hemispheric pattern of wave patterns 5-7, which can trigger extreme heat and rainfalls simultaneously at different places (Coumou et al, 2014; Kornhuber et al, 2020; Di Capua et al, 2021). The mechanisms and causes of the amplification of such wave patterns are still discussed, especially their dependence on climate change (Screen and Simmonds, 2014; Petoukhov et al, 2016; Kornhuber et al, 2017; Mann et al, 2017, 2018; Kornhuber and Tamarin-Brodsky, 2021). Over the heatwave region, an anticyclone builds up — a situation called 'blocked' — at mid- and upper-level troposphere in conjunction with a change in the jet stream's climatological path, towards a large poleward meridional meander. The anticyclone sustains the poleward advection of warm air along its western flank, adiabatic warming by subsidence and clear skies at its center. Those conditions favor warming by short-wave insolation, especially at the peak of the seasonal cycle. Close to the ground, a positive feedback loop is initiated as anomalously dry soils favor the partition of incoming solar energy into sensible rather than latent heat, which enhances surface evaporation and may reinforce the anticyclonic structure (Hirschi et al, 2011; Miralles et al, 2012, 2014; Rasmijn et al, 2018; Dirmeyer et al, 2021).

These mechanisms all played a role in the record-breaking heatwaves of 2003 in western Europe (García-Herrera et al, 2010) and 2010 in Russia (Dole et al, 2011; Otto et al, 2012; Trenberth and Fasullo, 2012; Di Capua et al, 2021). The same mechanisms are also present in the exceptionally intense heatwave that occurred in Western North America in June 2021. Previous records of temperature were broken by up to 5°C (Philip et al, 2021) making it one of the most intense heatwaves ever recorded (Thompson et al, 2022). Due to its

exceptional nature, this event triggered extensive research. Dry soils (Zhang et al, 2021, 2022) likely combined with an intense omega blocking anticyclone resulting from a wave breaking event, associated with southern excursion of the polar vortex (Overland, 2021; Neal et al, 2022), and interacting with an atmospheric wave emanating from the tropical Pacific (Bartusek et al, 2022). Some authors also suggested the role of latent heat release through moisture advection by an unusual atmospheric river through the North Pacific (Qian et al, 2022; Lin et al, 2022; Mo et al, 2022). Lucarini et al (2022) however show using a long simulation in a pre-industrial climate that this event was *typical* with respect to other intense events simulated by the model at this location.

Extreme Value Theory (EVT) has been used to determine statistical models for maxima (or minima) of climate variables of interest (typically temperature or precipitation) (Ghil et al, 2011). EVT is based on a convergence principle of the probability distribution of maxima or peaks-over-threshold (Coles et al, 2001). It allows to compute return values corresponding to very large return periods (i.e. longer than the period of observations), even in non stationary contexts (Cheng et al, 2014). In practice, EVT has been an efficient framework to estimate statistical variations on short lived extremes (e.g. the highest daily temperature). In principle multi-variate EVT provides a framework to investigate events that combine several variables (e.g. temperature, precipitation and the atmospheric circulation) (Tawn, 1990), but such analyses are not designed to investigate long lasting events, for which the time persistence is a key factor in the extreme (although using the so-called extremal index partially alleviates this issue, e.g. Moloney et al (2019)).

Large Deviations Theory (LDT) is however the key statistical framework employed in statistical physics (Touchette, 2009) and has begun to gain momentum in climate sciences (Galfi et al, 2021). In contrast to EVT, it provides asymptotic laws at the leading exponential order for extremes of sums of random variables. One of its applications is the Freidlin-Wentzell theory of large deviations in random dynamical systems (Freidlin and Wentzell, 1987). This theory studies the dynamics leading to extremes by describing it as the optimal noise pushing the system in the direction leading to extremes. In particular, it predicts that, under mild hypotheses, the dynamics leading to extremes of any observable concentrates around a single most probable trajectory, usually called the *instanton* (Chetrite and Touchette, 2015; Dematteis et al, 2019a; Grafke and Vanden-Eijnden, 2019). This framework can be extended to systems without any explicit stochasticity, where it amounts to finding the "optimal" initial conditions (Dematteis et al, 2019b; Lucarini et al, 2022) which will lead the dynamics to reaching extremes of an observable under the hypothesis that the observable admits a large deviations principle at its right (or left) tail.

In this paper, we expand the work of Galfi and Lucarini (2021), Galfi et al (2021) and Lucarini et al (2022) who pioneered the use of LDT to study extreme events in geophysical systems. We also investigate the prediction of LDT concerning the concentration of trajectories leading to extremes of an

observable around a most probable one in a dynamical system. We address the question whether the dynamics leading to extreme 2m air temperature events is *typical* in a long simulation of a climate model.

As a case study, we take the 2000 years pre-industrial control run of the IPSL-CM6A-LR Earth System Model (ESM) (Boucher et al, 2020). We examine the dynamic characteristics of trajectories leading to heatwaves as their duration and strength vary. We evaluate how typical the heatwaves are by examining the key dynamic factors that contribute to their formation. Lastly, we study how the dynamics change when the location from which the heatwaves are observed is altered.

This paper is organized as follows: in Section 2, we detail the mathematical paradigm of large deviations theory in dynamical systems and make our assumptions explicit. In Section 3, we introduce the methodology employed to isolate the dynamics leading to extremes and the normalized variance metrics used to measure convergence. Section 4 presents the results of the analysis, focusing first on one observable and then on three observables at different locations. Finally, the discussion of the results and the conclusions drawn from our analysis are presented in section 5.

2 Mathematical preliminaries

The principle of convergence of trajectories around a most probable one for extreme events, the so-called instanton, can be presented either in systems with an explicit stochasticity — where it is equivalent to finding an "optimal" noise to perturb the system (Freidlin and Wentzell, 1987; Chetrite and Touchette, 2015; Dematteis et al, 2019a; Grafke and Vanden-Eijnden, 2019) — or with a deterministic evolution — where it is equivalent to finding the "optimal" initial conditions (Dematteis et al, 2019b; Lucarini et al, 2022). There is no explicit stochasticity in the climate model we study, therefore we follow here the presentation by Lucarini et al (2022).

We consider a chaotic dynamical system evolving continuously in time. Let $x(t) \in \mathbb{R}^d$ be the state vector at time t, where d is the number of dimensions needed to describe the system. We assume that x evolves according to the following ordinary differential equation

$$\frac{dx}{dt} = b(x) \tag{1}$$

where $b : x \mapsto b(x) \in \mathbb{R}^d$ defines the dynamics of the system. In the following we always assume that the transients have died out and all trajectories considered belong to the attractor of the system. We assume that there is a unique physical invariant measure μ on this attractor. We are interested in the statistics of observables computed on the attractor. Observables are smooth functions of phase-space variables $F : \mathbb{R}^d \to \mathbb{R}$ which have a physical interest.

Our observables will be of the following type:

$$F_r(x_t) = \frac{1}{r}\int_{-r/2}^{r/2} f(x_{t+t'})dt' \tag{2}$$

where f is also an observable and $r \in \mathbb{R}$ is the size of the rolling mean window. In the following, f is the daily mean temperature over a specific area. For $q \in \mathbb{R}$, we define the set Ω_q as the set of points x on the attractor such that $F_r(x) \in [q, q+dq]$. In other words, we are interested in the states of the system for which the value of the observable F_r is close to q, where q will correspond to extreme quantiles of the distribution of F_r. We now explicitly assume that F_r admits a large deviations principle, i.e. the set Ω_q is exponentially rare with respect to the measure μ with increasing q. Then, one can write:

$$\mathbb{P}(F_r(x_t) = q) = \mu(\Omega_q) \asymp \exp\left(-\min_{x\in\Omega_q} I(x))\right) \tag{3}$$

where $\asymp$ means that the ratio of the logarithms of both sides tends to unity when $q \to +\infty$. Here q plays the role of the large deviations parameter even though it does not appear explicitly as such in Eq. (3). The functional I is called the rate function and is given by the Legendre transform (Dematteis et al, 2019b):

$$I(x) = \max_p(\langle p, x\rangle - S(p)) \tag{4}$$

of $S(p) = \log \mathbb{E}_\mu(e^{\langle p,x\rangle})$, the cumulant generating function of x under the measure μ with $\langle\cdot\,,\cdot\rangle$ the inner product in $\mathbb{R}^d$.

When $q \to +\infty$ in Eq. (3), the probability is exponentially dominated by the point $\widehat{x}$ which minimizes the rate function: $\widehat{x} = \operatorname{argmin}_{x\in\Omega_q} I(x)$. The minimizer $\widehat{x}$ represents physically the state with the maximum likelihood of realization. Under the stated conditions, the probability measure accumulates near $\widehat{x}$: this *most likely* of the *least likely* states is usually called the *instanton*.

The instanton formalism gives powerful results about the paths in the phase space leading to extreme events in dynamical systems. If the observable admits a large deviations principle, then the paths leading to any extreme are dominated exponentially in probability by one unique path. One should note that in general the uniqueness is not fully guaranteed because the rate function I may have several minima. The intuition behind this result is the following: because extremes of an observable are rare (i.e. have low probability), the system has few paths in the phase space to reach those extremes. If the system had many paths to reach the extremes, then these would not be rare. This idea of uniqueness of limit behaviors in a dynamical system is reminiscent of the uniqueness of the limit distribution in EVT (Coles et al, 2001).

This mathematical derivation suggests that if we examine the extreme daily temperatures at a particular location, the paths that lead to these extremes are to converge around a single path. We investigate this prediction in the following. To do so, we employ the instanton filtering procedure introduced by Grafke et al (2013). It consists of averaging independent events x taken

from a long simulation of the dynamical system for which the value of the observable $F_r(x)$ is close to an extreme level. According to the theory presented above, this procedure leads — up to a good approximation — to the instanton reaching the extreme level of the observable F_r. In other words, we investigate the typical state $\mathbb{E}[x \mid F_r(x) = q]$ conditional on $F_r(x) = q$ where q is a given extreme level of the distribution of F_r.

Contrary to Lucarini et al (2022), we do not consider events above a certain level $\mathbb{E}[x \mid F_r(x) \geq q]$. The rationale of this choice is twofold. First we want to investigate the convergence properties for different values of q for a fixed number of independent events in each case. Second, large deviations theory predicts that for sufficiently extreme values of q the events for which $F_r(x) \sim q$ are exponentially favored compared to the ones with $F_r(x) > q$, but we cannot a priori rule out the possibility of a non linear response of the dynamics $\mathbb{E}[x \mid F_r(x) = q]$ with respect to q when reaching extremes. Therefore, the dynamical mechanisms needed to reach high temperatures may be different than the ones to reach *very* high temperatures.

3 Data and methods

We use the output of the pre-industrial control run of the IPSL-CM6A-LR (Boucher et al, 2020) model in the context of the CMIP6 intercomparison project (Eyring et al, 2016). This simulation is 2000 years long and represents a stationary climate with a CO_2 concentration corresponding to pre-industrial levels. The model has an horizontal atmospheric resolution of 2.5° in longitude and 1.3° in latitude. During the 2000 years, the global mean 2 meter air temperature of the Earth slightly drifts by 0.25K but we will assume that this drift can be neglected when it comes to studying extremes in the midlatitudes, because we are interested in deviations that can exceed several K. We use the sea level pressure (SLP), upper level soil moisture (SM), air temperature at 2m (T2M), air temperature at 850hPa (T850), geopotential height at 500hPa (Z500) and the meridional component of the wind at 250hPa (V250). These variables are used with a daily frequency. Due to a technical issue, approximately 1/4 of the data for the Z500 variable are missing. In the following, we therefore present results for Z500 only for the dates available in our data set. We have checked that when restricting to the period when we have the Z500 variable available for the other variables do not affect the results. We therefore present the results for those other variables over the whole 2000 years.

In the following, we consider four observables derived from T2M: T2M at three grid points situated in Southern (38-39°N, 5.25-3.75°W), Western (49-50°N, 1.25-3.75°E) and Northern Europe (59-60°N, 13.75-16.25°E) and T2M averaged over a region extending in Western and Central Europe (46-53.5°N, 0-25°E). These observables are named respectively S, W, N and WCE. We want to investigate the highest values reached by these observables, therefore we restrict the analysis to the three months of the meteorological summer: June, July and August (JJA). Except specified explicitly, we do not consider

detrended or deseasonalized variables. To investigate how the dynamics can change for longer events, we consider the extremes of the observable after applying a rolling mean window of $r = 1$, 5 and 15 days. The size of these windows were chosen to investigate both short and long lasting heatwaves events.

We consider the time series of one given observable F_r among S, W, N and WCE regions during the summer months for a given rolling mean window r. For a quantile q_α of a given order α of the empirical distribution of F_r, we select the pool of the $n = 50$ independent events x_i for which the values of their observable $F_r(x_i)$ are the closest from the value of the quantile q_α. In other words, we find the dates of the n-nearest neighbors of the quantile q_α of the observable. These events are searched over any of the days in the JJA months. To ensure that these events are independent one from another, we impose that for a nearest neighbor to be chosen, it must be separated by more than 15 days from any nearest neighbor already present in the pool. The choice of this timescale was made with regards to the typical chaotic timescale of the atmosphere (around 10 days). Our procedure is equivalent to defining the set Ω_q of section 2 as:

$$\Omega_{q_\alpha} := \Omega_\alpha = \{x \mid F_r(x) \in [q_\alpha - \eta, q_\alpha + \eta]\}$$

for η as small as possible to ensure that the number of elements in Ω_α is $\#\Omega_\alpha = n$. We consider the quantiles of order $\alpha = 0.75$, 0.95, 0.99 and 0.999 of the empirical distributions of observables S, W, N and WCE. These quantiles are arbitrary and are chosen to ensure homogeneity between the different observables for which the extreme values may be very different.

Figure 1 presents the histograms of the empirical distributions of the four observables (rows) for the different rolling mean windows (columns). The vertical lines in color present the quantiles for the different orders considered. The horizontal lines of the same colors present the spread η of the observables for the points found with our procedure. For the smallest orders (0.75, 0.95 and 0.99), the spread is barely visible, which means that the values of the observable for points in Ω_α are very close to the value of the quantile. For the order 0.999, the spread is bigger but still small compared to the standard deviation of the full empirical distribution.

For each value of the α-th order quantile, a given observable F_r and a rolling mean window r, we therefore consider fields (SLP, SM, T2M, T850, Z500 and V250) $\psi_{\alpha,r,F_r}(\phi, \theta, t, m)$ that have four components: the longitude ϕ and latitude θ, the time t — where the time is expressed relative to the day t when the observable is such that $F_r(x_t) \simeq q_\alpha$ —, and the number m among the points in Ω_α ($m = 1, 2, \ldots, n$). In the following we skip the α, r, F_r indices for simplicity, but note that the quantities introduced are always relative to a given value of this triplet. For any field ψ, we use the term composite to denote the average $\widehat{A}_\psi$ over the points in Ω_α and the rolling mean window r:

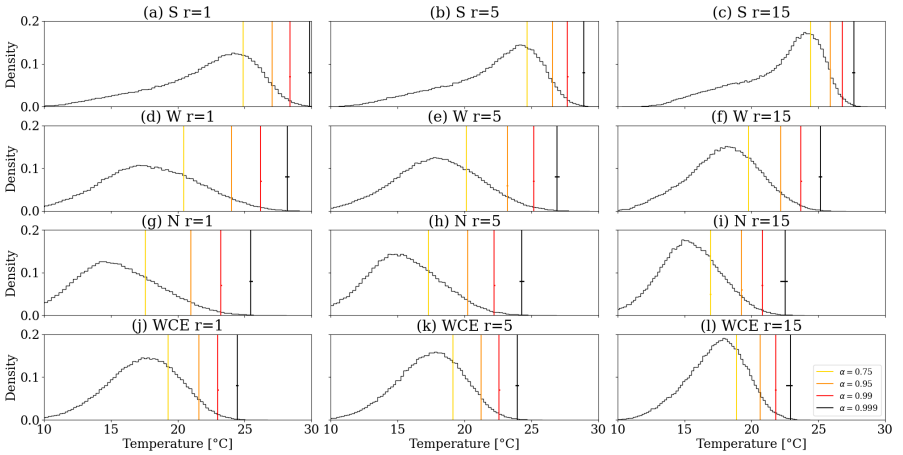

Fig. 1 Histograms of the observables and associated quantiles. In rows: observables derived from the T2M variable at three grid points situated in Southern Europe (S: 38-39°N, 5.25-3.75°W), Western Europe (W: 49-50°N, 1.25-3.75°E) and Northern Europe (N: 59-60°N, 13.75-16.25°E) and over one grid box extending in the North-West Europe (NWE, 46-53.5°N, 0-25°E). In columns: rolling mean windows for the computation of the observable of r =1, 5 and 15 days. The vertical bars show the value of the quantiles of order α of the time series for α =0.75, 0.95, 0.99 and 0.999. The horizontal lines show the spread of the observables for the points in Ω_α.

$$\widehat{A}_\psi(\phi,\theta) = \frac{1}{n}\sum_{m=1}^{n}\frac{1}{r}\sum_{t=-r/2}^{r/2}\psi(\phi,\theta,t,m) \tag{5}$$

To measure the clustering of these points, we define two normalized variances:

$$\widehat{V}_\psi(\phi,\theta) = \mathrm{Var}_m\left[\frac{1}{r}\sum_{t=-r/2}^{r/2}\psi(\phi,\theta,t,m)\right] \,/\, \mathrm{Var}_{\psi,clim}(\phi,\theta) \tag{6}$$

and:

$$\widetilde{V}_\psi(\phi,\theta,t) = \mathrm{Var}_m\left[\psi(\phi,\theta,t,m)\right] \,/\, \mathrm{Var}_{\psi,clim}(\phi,\theta) \tag{7}$$

where $\mathrm{Var}_{\psi,clim}(\phi,\theta)$ is the climatological variance of the field ψ considered over the whole summer (JJA). Both of these variances are normalized in the sense that they are compared to the variance observed at a specific location over the whole summer. In the following, they are thus expressed in percentage. The lower the value of this percentage, the smaller the variance between the points in Ω_α either over the rolling mean period ($\widehat{V}_\psi$) or at a specific time ($\widetilde{V}_\psi$). If we were to select n days randomly over the full data set, we could expect to find a variance close to $\mathrm{Var}_{\psi,clim}(\phi,\theta)$. Therefore, the closer $\widetilde{V}_\psi(\phi,\theta,t)$ is to one (or 100%), the less specific is the dynamics of the averaged field $\widetilde{\psi}(\phi,\theta,t) = \frac{1}{m}\sum_{m=1}^{n}\psi(\phi,\theta,t,m)$ at (ϕ,θ,t). The closer it is to zero, the more concentrated

are the n points in Ω_α. For $\widehat{V}_\psi$ the situation is different because when the size of the rolling window r increases, the variance $\widehat{V}_\psi$ spontaneously decreases. Nonetheless, it is still possible to compare the values of $\widehat{V}_\psi$ at specific locations when the order α of the quantile changes for a fixed r. Finally, we emphasize that $\widehat{V}_\psi$ is not equal to the average of $\widetilde{V}_\psi$ over the rolling mean window r.

In order to have simple diagnosing metrics, we also average the $\widetilde{V}_\psi$ variance spatially:

$$\langle \widetilde{V}_\psi(t) \rangle = \sum_{\phi,\theta} \widetilde{V}_\psi(\phi,\theta,t)\cos(\theta) \;/\; \sum_{\phi,\theta} \cos(\theta) \tag{8}$$

over either the North-Hemisphere ($\theta \in [22.5°\text{N}, 90°\text{N}]$ and $\phi \in [0°\text{E}, 360°\text{E}]$) or the Euro-Atlantic sector ($\theta \in [22.5°\text{N}, 70°\text{N}]$ and $\phi \in [80°\text{W}, 50°\text{E}]$). For simplicity, we drop the index ψ in the following where it is not ambiguous.

4 Results

4.1 Extreme temperatures in Western Europe

In this section we present the results for the extremes of T2M at the grid point in Western Europe (observable W). Figure 2 presents the composite maps of anomalies of T2M and Z500 over the Euro-Atlantic sector for the $n = 50$ points for which the values of their observable is the closest to the quantile and over the rolling window $r = 5$ days for the four quantiles considered. The anomalous T2M values and their spatial extension increase as the α-th order quantile increases. As α increases, the deviation in Z500 also increases, reaching a maximum of 160m at the center of the high-pressure system located above North-Western Europe. There is also a warm anomaly in Eastern North-America and a cold anomaly in Western Russia. Figure A1 in annex shows the same results over the whole North Hemisphere. Anticyclonic anomalies are present all over the Hemisphere, with a distinct wave number 4 hemispheric pattern at mid-troposphere for the highest quantiles. These structures coincide with warm anomalies at the ground.

Figure 3 shows in contours the composite T2M field and in colors the normalized variance $\widehat{V}$. For all values of α, the lowest values of the normalized variance are located in Western Europe, i.e. around the location where the observable is computed. When α increases, the normalized variance decreases: for example, the normalized variance exceeds 60% in the North Atlantic region for $\alpha = 0.75$, and it decreases to less than 30% for $\alpha = 0.999$. Figure 4 presents the same analysis for Z500. The extension of regions of highest variance also decreases when increasing the order α. The regions with the lowest variance is again centered at the location where the observable is computed. For $\alpha = 0.999$ for example, a large region of very low normalized variance ($\widehat{V} < 10\%$) embraces most of Western Europe. The decrease of variance is however not uniform, with high variance remaining downstream of the anticyclonic region and a smaller localized region west of the Iberian peninsula. We

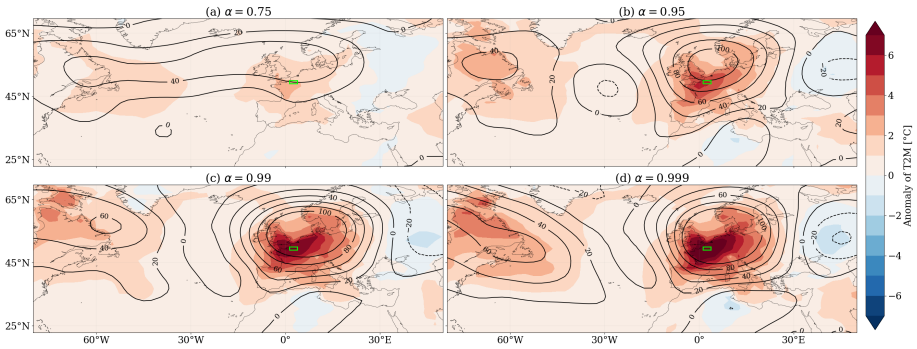

Fig. 2 Composite maps $\widehat{A}$ of anomalies of T2M and Z500 for increasing α-th order of the quantile of the W observable's empirical distribution. The figure is computed for a rolling mean window of $r = 5$ days. Colors: anomaly of air temperature at 2m (T2M, [K]). Contours: anomaly of geopotential height at 500hPa (Z500, [m]). Anomalies are computed with respect to the JJA average. The green box displays the location where the observable is computed.

come back to this latter structure below. We finally remark that on Fig. 4, for the highest quantiles the composite Z500 field shows that the large anticyclone over Western Europe is not centered just above the location of the observable, but is rather situated to its south-east.

Figures A2 and A3 in annex display the same analysis over the whole North-Hemisphere. The reduction of variance is seen over remote regions of the atmosphere, for example with a region of low variance ($\widehat{V} < 20\%$) in the North-Western Pacific for $\alpha = 0.999$ for Z500 (Fig. A3 panel (d)). We emphasize that for all those maps, the number of points in Ω_α is always the same ($n = 50$ for T2M, $n \sim 30$ for Z500 due to the missing data). The outcome shown implies that the dynamics leading to extreme temperatures concentrates within a large geographical region.

So far we have presented the results for a rolling window of $r = 5$ days. The results for $r = 1$ day and $r = 15$ days for the composite maps are similar and are thus not shown here. However, as can be seen in figures A4 and A5 in annex for $r = 1$ day and in figures A6 and A7 in annex for $r = 15$ days, the results for the variance strongly depend on r. As we mentionned earlier, this is expected in so far as averaging temporally naturally reduces the variance $\widetilde{V}$. We however still observe in those figures a reduction of the variance when α increases.

Figure 5 presents composites (colors) and zones of high variance ($\widehat{V} > 50\%$, hatches) for the other variables: anomaly of soil moisture (SM, panel(a)), anomaly of sea-level pressure (SLP, panel (b)), air temperature at 850hPa (T850, panel(c)) and meridional wind speed at 250hPa (V250, panel (d)). The threshold 50% for discriminating between "high" and "low" variance is chosen arbitrarily to highlight differences between regions. This figure is drawn for $r = 5$ days and $\alpha = 0.999$.

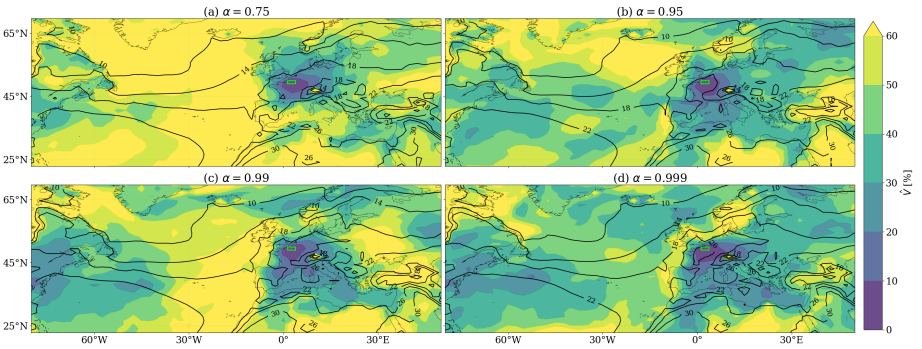

Fig. 3 Composite $\widehat{A}$ (contours) [°C] and normalized variance $\widehat{V}$ (colors) of the T2M field for increasing α-th order of the quantile of the W observable's empirical distribution. The figure is computed for a rolling mean window of $r = 5$ days. The normalized variance is computed after averaging over the rolling window and is expressed in %. The green box displays the location where the observable is computed.

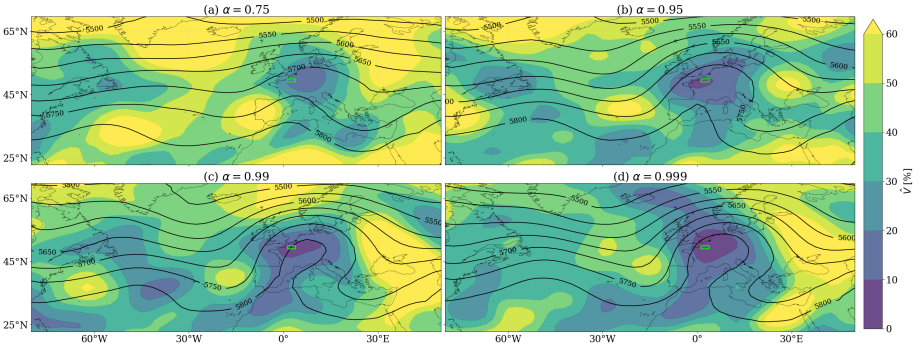

Fig. 4 Composite $\widehat{A}$ (contours) [m] and normalized variance $\widehat{V}$ (colors) of the Z500 field for increasing α-th order of the quantile of the W observable's empirical distribution. The figure is computed for a rolling mean window of $r = 5$ days. The normalized variance is computed after averaging over the rolling window and is expressed in %. The green box displays the location where the observable is computed.

Panel (a) shows the large anomalously dry soils over most of Western and Central Europe. The small variance over this region demonstrate the key importance of local dry soils for reaching very high temperatures. The convergence is stronger to the east of the location of the observable, which agrees with the results of Zschenderlein et al (2019). This pattern is similar to the one found by Faranda et al (2022) using the SPEI index (Beguería et al, 2014). We also note that when checking at $t = -15$d (i.e. 15 days before the event, not shown), the anomaly is still strong ($\simeq$4kg/m^2) and the variance still small ($\widetilde{V} < 30\%$), which supports our preceding statement for the role of dry soils. Contrary to other fields (Fig. 5 panels (b), (c) and (d)), the region of low

variance is concentrated around the observable and do not extend over large regions.

Panel (b) shows a large positive anomaly of SLP north-east of the observable location and the low-level thermal low associated with the heatwave in Western Europe. The normalized variance is low over these regions. The positive anomaly situated above the Western Atlantic ocean may not be a relevant feature of the typical dynamics because the variance is high at this place, contrary to the anomalous low pressure over Greenland. Panel (c) shows the composite map of T850. The large intrusion of warm air from the south is a key feature of the dynamics. Even though the variance is low over mainland Western Europe, we detect a region of high variance extending from the west of the Iberian peninsula to Ireland and southern Norway and Sweden, i.e. to the western flank of the anticylonic structure (Fig. 4 panel (d)) . This tongue of high variance is located at the highest gradient of T850. It is therefore difficult to decide between the two following explanations for this feature: either different dynamical mechanisms between the points in Ω_α (e.g. advection of warm air from the tropics) or a slightly displaced anticyclone which, combined with the strong gradients, would display such a strong variance tongue.

Finally, panel (d) shows the dynamics in the upper troposphere with the meridional wind speed at 250hPa. The situation is characterized by a strongly meridional circulation west of the observable, which is coherent with the anticyclonic situation presented in figure 2. The regions of high variance also coincide with the regions with the highest gradients of V250. The one situated above the Balkans and the Black Sea regions suggests the presence of an arm of the jet stream oriented to the North. We checked this explanation using the zonal wind speed at 250hPa (not shown) and found that the synoptic situation over the event corresponds indeed to a splitting of the jet caused by the large blocked anticyclonic situation over Western Europe. This explanation is coherent with the double jet dynamics associated with heatwaves (Rousi et al, 2022).

Figure 6 presents the same analysis over the entire North Hemisphere. For soil moisture (panel (a)), the only region outside the Euro-Atlantic sector where there are important anomalies and low variance is the western coast of the USA. This region is also characterized by positive T2M temperature anomalies as can be seen in figure A1. Panel (b) shows a good agreement among points in Ω_α for the low-level low over Greenland and the Ural. We also notice a positive SLP anomaly over the North-Pacific, but this region is associated with high variance. Apart from the high variance tongue in Western Europe, and high variance in the Arctic regions, panel (c) shows a good agreement between points in Ω_α for the T850 field, including over most of the Atlantic and Pacific oceans. Finally, panel (d) shows a wave pattern 6-7 extending over the entire North-Hemisphere upper-troposphere. It should be noted, however, that the anomaly is stronger and the variance is smaller in the upstream area compared to the downstream area of the observable. Specifically, there is little structure visible above the Western Pacific region. This situation may be the

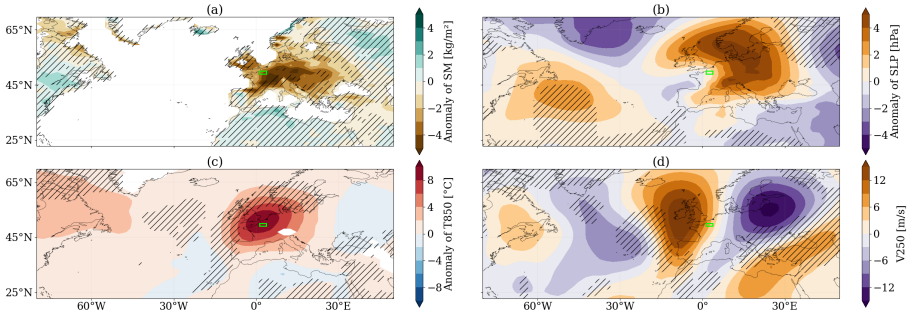

Fig. 5 Composite maps $\widehat{A}$ for SM, SLP, T850 and V250 for the quantile of order $\alpha = 0.999$ of the W observable's empirical distribution. The hatched areas correspond to $\widehat{V} > 50\%$. The figure is computed for a rolling mean window of $r = 5$ days. Anomaly of (a) soil moisture (SM), (b) sea-level pressure (SLP) and (c) temperature at 850hPa (T850) with respect to their average over the summer (JJA), and (d) meridional wind speed at 250hPa (V250). The green box displays the location where the observable is computed.

result of either an hemispheric quasi-stationary pattern (Coumou et al, 2014; Kornhuber et al, 2020) or a transient Rossby wave packet (Fragkoulidis et al, 2018).

The analyses presented so far support the instanton interpretation of extreme events: the higher the value of the quantile q_α, the stronger the concentration of trajectories reaching this quantile around a single trajectory. To check whether our visual inspection is correct, we present in figure 7 the temporal evolution of the normalized variance $\langle \widetilde{V}(t) \rangle$ averaged over both the Euro-Atlantic sector (plain lines) and the entire North-Hemisphere (dashed lines) from $t = -15$ to $t = +15$ days with respect to the event for all the variables and all the rolling windows r. The general picture drawn above is validated by the results presented in this figure: overall, the higher the value of the order α for the quantiles of the observable empirical distribution, the smaller the value of the normalized variance $\langle \widetilde{V}(t) \rangle$. This result holds when averaging over either only the Euro-Atlantic sector — where the observable is computed and the synoptic situation is the most relevant for the event — or the entire North-Hemisphere. However, when we focus on the details of the different panels, we see that this behavior is less clear for some variables. The convergence is clear for T2m (panels (a), (b) and (c)) and T850 (panels (m), (n) and (o)) for which the normalized variance $\langle \widetilde{V}(t) \rangle$ is a decreasing function of α for almost all values of t. For dynamical variables such as Z500, we see that the convergence is stronger for higher values of the rolling window (compare for example panel (f) and panel (d)). This could be explained by the fact that a long extreme needs a persistent anticylonic circulation, which is more likely to be "typical" than for a short extreme. For the SM and SLP variables the situation is the most blurred. Higher values of α broadly correspond to smaller values of the normalized variance $\langle \widetilde{V}(t) \rangle$, but the ordering of the quantiles changes with both r and t.

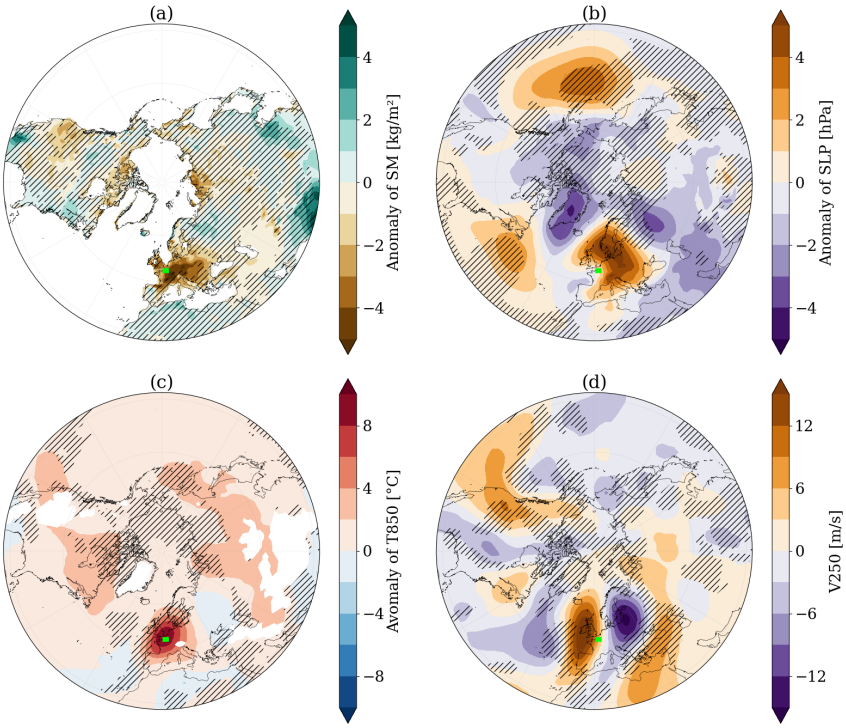

Fig. 6 Composite maps $\widehat{A}$ for SM, SLP, T850 and V250 for the quantile of order $\alpha = 0.999$ of the W observable's empirical distribution. The hatch areas correspond to $\widehat{V} > 50\%$. The figure is computed for a rolling mean window of $r = 5$ days. Anomaly of (a) soil moisture (SM), (b) sea-level pressure (SLP) and (c) temperature at 850hPa (T850) with respect to their average over the summer (JJA), and (d) meridional wind speed at 250hPa (V250). The green box displays the location where the observable is computed.

In this section we analyzed the dynamics $\mathbb{E}[x \mid F_r(x) = q]$ conditional on the value reached by an observable F_r. We demonstrated the concentration of the trajectories as the value q of the observable reaches extremes. We can now sketch the mechanisms associated with extreme temperatures at the grid point considered in Western Europe (observable W). The mechanisms summarized in the reviews of Perkins (2015), Horton et al (2016) and Domeisen et al (2022a) are present. To ensure very high temperatures, one needs dry soils, a large mid-troposphere anticyclone centered slightly to the South-East of the location of the extreme and an upper level Rossby wave train of 6-7 wave number. This situation ensures both the advection of warm air from the south at the west flank of the anticyclone, subsidence and associated adiabatic heating at the center of the anticyclone, and clear skies caused by the high-pressure system that allows for more radiative heating of the lowest layer of the atmosphere in conjunction with reduced water evaporation. We also note that the dates at which these extremes are reached are less dispersed around the peak of the

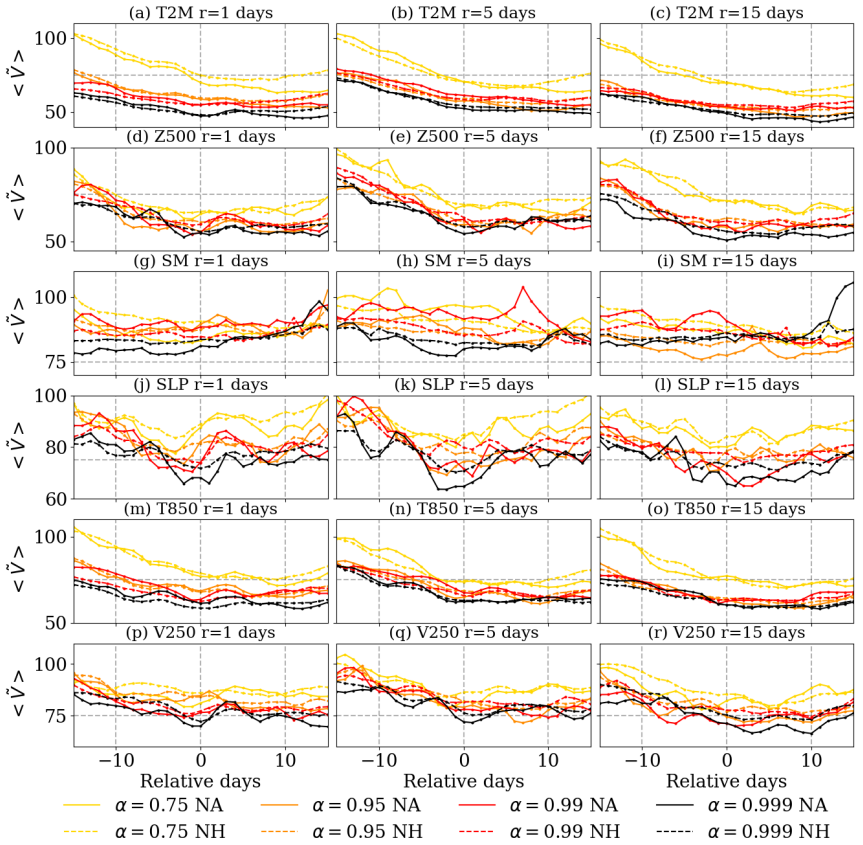

Fig. 7 Evolution of the normalized variance $\langle \widetilde{V}(t) \rangle$ averaged over the Euro-Atlantic sector (plain lines) and the North-Hemisphere (dashed lines) for the different variables (W observable). The normalized variance is expressed in %. The colors show the α-th order quantile of the observable empirical distribution. The time is expressed relative to the day when the observable is such that $F_r(x_t) \simeq q_\alpha$. The gray dashed line shows the 75% level.

T2M seasonality when α increases: the standard deviation in the calendar days of the points in Ω_α goes from 23.1 days for $\alpha = 0.75$ to 15.5 days for $\alpha = 0.999$.

In figure 4 there is a small isolated region of high variance situated west of the Iberian peninsula for all quantiles. We investigated this discrepancy by looking at individual events. It turns out that the synoptic dynamics at the mid-troposphere associated with some of the high temperature events are characterized by the presence of a cut-off low around this location. Table A1 in appendix presents the percentage of such events. The presence of a cut-off low was investigated in a semi-objective way by looking at the existence of an isolated minimum of the stream function at 500hPa located within [-30°S,+5°N] and [-40°W,+0°E] from the location where the observable W is computed and

not embedded in the upper-level jet (250hPa) (Muñoz et al, 2020). We however note that it is sometimes difficult to distinguish between a true cut-off low and a large meander of the jet caused by a deep, almost isolated, low pressure system over the Atlantic. The percentage of cut-off lows vary between 15 to 30%, with an average around 20%, depending on the quantiles and the size of the rolling mean window considered but without any clear trend (the differences may be due to the limited sample size). These cut-offs are not visible on the averaged maps (e.g. Fig. 4) not only because they represent only 20% of the events but also because their characteristic size is of the same order as the variance in the location of their center. Therefore even when considering only events with a cut-off low, they tend to be averaged out. The fact that even for very high quantiles there is a substantial amount of events with such a cut-off (around 20%) is in contradiction with the unique path hypothesis presented in section 2 only if one assumes that the rate function I has a single minimum. It therefore suggests that there may be at least a bi-modality in the typical dynamical paths to reach extremes for the W observable. We however note that if there is indeed two minima, the convergence of trajectories that we showed above using all trajectories suggests that they are close to one another in the phase space.

4.2 Results for the other locations

In this section we present the same analysis applied to the three other observables: T2M at two grid-points situated in the south (observable S) and north of Europe (observable N), and T2M averaged over a large area in Western and Central Europe (WCE observable). The results for the evolution of their normalized variance $\langle \widetilde{V}(t) \rangle$ is presented in figures A8, A9 and A10 respectively in annex. As above, the general picture of decreasing variance with increasing α is still valid but there are substantial inter-location variations. We note for example that the dynamical signal represented by the variables Z500, SLP and V250 is much clearer for locations farther to the North (Fig. A8 vs Fig. A9 panels (d), (e), (f) for example). In the contrary, soil moisture plays a bigger role for the locations situated farther to the south (Fig. A8 vs Fig. A9 panels (g), (h), (i)). This suggests that extremes of temperature at places situated to the south are more "local" in the sense that they require less large scale organized circulation to be reached, which is coherent with the results of Sousa et al (2018) who showed that a ridge situation better described the occurrence of heatwaves in southern Europe than the blocked situation as in northern Europe. This interpretation however needs to be validated at other locations. We also note that these dynamical differences between lower and higher latitudes may be related the skewness differences in the distribution of the summer temperature (Fig. 1).

The observable over a large area (observable WCE) has a similar behavior to the W and N observables, with a stronger concentration of trajectories for dynamical variables than for soil moisture. We also note that the absolute values of the normalized variances for this observable are smaller than

the one for the observable W which is situated around the same latitudes. It suggests that the typicality of the dynamics leading to anomalies of T2M over a large geographical area is stronger than for a localized observable. This seems reasonable in so far as it is less likely to have anomalies over a larger than a smaller area situated inside the larger one, hence the smaller number of synoptic conditions which can lead to an extreme for an extended observable.

Figure 8 presents the composite maps and the normalized variance maps for the S observable for the order $\alpha = 0.999$ and a rolling mean window of $r = 5$ days. The situation over the North-Hemisphere is presented in annex in figure A11. The synoptic situation is characterized by a large anticyclone centered south-east of the observable (panel(b)) and a positive SLP anomaly extending from Algeria to Northern France (panel (d)). The upper-level circulation displays a short wavelength Rossby wave (panel (f), wave number 6-7). We however note that its amplitude is smaller than for the W observable. As previously, this dynamics leads to high T2M and T850 values and it has been anticipated by dry soils over most of southern Europe and northern Africa. The variance is lowest close to the location where the observable is computed. We note a large region of high variance downstream of the observable for V250 (panel(f)), but upstream for SLP and Z500 (panels (b) and (d)). As for the W observable, we see a high variance region at the west of the maximum gradients of T850 (panel (e)). The percentage of cut-offs associated with these events (Table A1) is between 15 and 30%, similar to the W observable.

Figure 9 presents the same analysis for the N observable. The situation over the North-Hemisphere is presented in annex in figure A12. The synoptic situation is characterized by an anticyclone centered south of the observable (panel (b)), positive anomalies of SLP north of the observable (panel (d)) and a large amplitude Rossby wave at 250hPa (panel (f)). The soil moisture situation (panel (c)) is almost the opposite of the one of the observable S, with a moist southern Europe and a dry northern Europe. The dry soils over Northern Europe, as for the region of anomalous T2M (panel (a)), extends over a large region encompassing Northern Europe and the Scandinavian and Baltic areas. This is in opposition to the extension of the anomaly in Fig. 8 panel (a), which is confined to the Iberian peninsula. This support the idea of "local" extremes in the south compared to more extended ones in the north. We also note the region of low variance associated with strong negative anomaly of soil moisture east of the observable (panel(c)) (Zschenderlein et al, 2019). The percentage of cut-offs is similar to the W and S observables but when looking at individual events we remark that a majority (>50%) of these cut-offs are embedded in a so-called modon structure (Butchart et al, 1989) with a blocking high above Scandinavia and a symmetric low above the eastern Mediterranean, splitting the jet into two branches. Again, for the reasons explained above, these structures are averaged out on composite maps.

Figure 10 presents the results for the WCE observable. The situation over the North-Hemisphere is presented in annex in figure A13. Contrary to the observables S, W and N, WCE has a large spatial extension as it encompasses

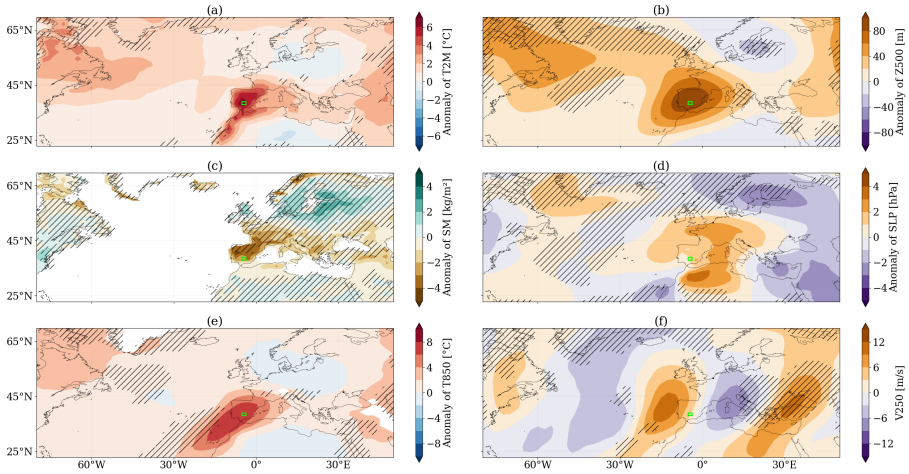

Fig. 8 Composite maps $\widehat{A}$ for T2M, Z500, SM, SLP, T850 and V250 for the quantile of order $\alpha = 0.999$ of the S observable's empirical distribution. The hatch areas correspond to $\widehat{V} > 50\%$. The figure is computed for a rolling mean window of $r = 5$ days. Anomaly of (a) air temperature at 2m (T2M), (b) geopotential height at 500hPa (Z500), (c) soil moisture (SM), (d) sea-level pressure (SLP) and (e) temperature at 850hPa (T850), and (f) meridional wind speed at 250hPa (V250). The anomalies are computed with respect to their average over the summer (JJA). The green box displays the location where the observable is computed.

most of West and Central Europe. In this case the synoptic situation is characterized by an anticyclone centered just above the observable (panel (b)), contrary to the precedent ones. The upper-level circulation looks very similar to the previous ones, with a short wavelength Rossby wave (panel (f), wave number 6-7). The soil moisture anomalies (panel (c)) extend over a large region of Northern, Central and Eastern Europe and this feature is consistent across points in Ω_α. We also note a region of high variance above southern Sweden for T850 (panel (e)), which is a feature similar to what was found for observable W (cf. Fig. 5 panel (c)). This feature could reflect different advecting dynamics of warm air above the boundary layer. The percentage of cut-offs is much higher than for the other observables (around 50%). As for the N observable most of these cut-offs are embedded in a modon structure.

5 Discussion and conclusion

The application of large deviations theory to dynamical systems (Freidlin and Wentzell, 1987; Grafke and Vanden-Eijnden, 2019; Dematteis et al, 2019b) predicts a concentration of trajectories leading to extremes for any observable: this is the so-called instanton hypothesis. In this paper we have investigated this prediction using air temperature at 2m (T2M) at four locations in Europe in the IPSL-CM6A-LR model (Boucher et al, 2020) pre-industrial control run as the observables of interest. Using the 2000-year simulation, we employed

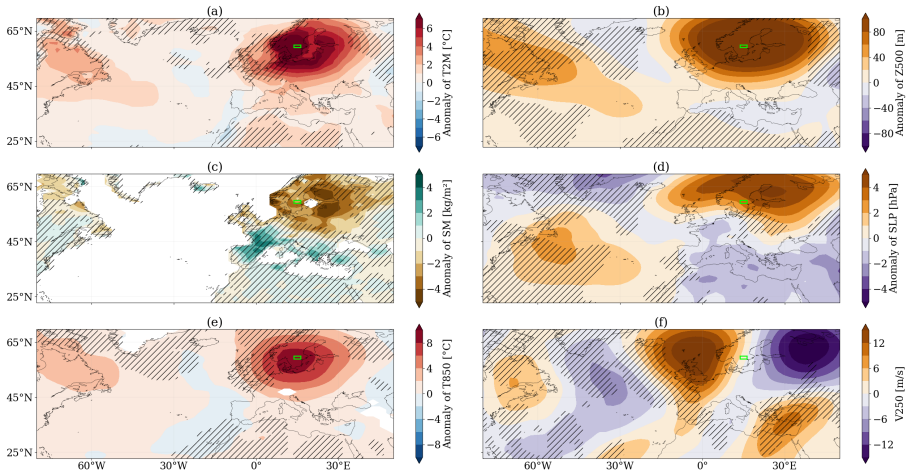

Fig. 9 Composite maps $\widehat{A}$ for T2M, Z500, SM, SLP, T850 and V250 for the quantile of order $\alpha = 0.999$ of the N observable's empirical distribution. The hatch areas correspond to $\widehat{V} > 50\%$. The figure is computed for a rolling mean window of $r = 5$ days. Anomaly of (a) air temperature at 2m (T2M), (b) geopotential height at 500hPa (Z500), (c) soil moisture (SM), (d) sea-level pressure (SLP) and (e) temperature at 850hPa (T850), and (f) meridional wind speed at 250hPa (V250). The anomalies are computed with respect to their average over the summer (JJA). The green box displays the location where the observable is computed.

an instanton filtering procedure (Grafke et al, 2013) consisting in averaging trajectories which reach a similar extreme value of these observables.

We have shown that the variance between trajectories reaching the extreme decreases as the level of extremeness of the observable increases. In other words, the more intense the extreme of T2M, the more likely that the trajectories reaching this extreme all look the same. We demonstrated this convergence on all variables considered: air temperature at 2m (T2M), geopotential height at 500hPa (Z500), upper level soil moisture (SM), sea-level pressure (SLP), temperature at 850hPa (T850) and meridional wind speed at 250hPa (V250) for a grid point observable in western Europe (observable W). Remarkably, the variance decreases even at places far from the location of the observable, which suggests a hemispheric dynamics leading to very intense heatwaves. We also showed a stronger decrease of variance for higher values of the rolling mean window r, indicating that the typicality is more easily reached when looking at longer time averages as suggested by Lucarini et al (2022).

The instanton dynamics found with our analysis is coherent with the mechanisms identified by previous literature for heatwaves dynamics in mid-latitudes (Perkins, 2015; Horton et al, 2016; Domeisen et al, 2022a). In the IPSL-CM6A-LR model with a pre-industrial CO_2 level, very high air temperature at 2m are reached by a combination of dry soils, a large mid-troposphere anticyclone and an upper level Rossby wave of 6-7 wave number. This situation ensures both the advection of warm air from the south at the west flank of

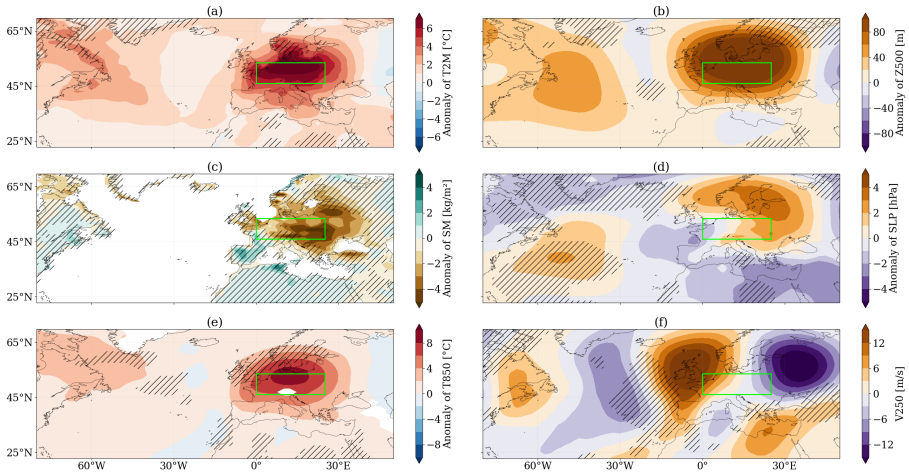

Fig. 10 Composite maps $\widehat{A}$ for T2M, Z500, SM, SLP, T850 and V250 for the quantile of order $\alpha = 0.999$ of the WCE observable's empirical distribution. The hatch areas correspond to $\widehat{V} > 50\%$. The figure is computed for a rolling mean window of $r = 5$ days. Anomaly of (a) air temperature at 2m (T2M), (b) geopotential height at 500hPa (Z500), (c) soil moisture (SM), (d) sea-level pressure (SLP) and (e) temperature at 850hPa (T850), and (f) meridional wind speed at 250hPa (V250). The anomalies are computed with respect to their average over the summer (JJA). The green box displays the location where the observable is computed.

the anticyclone over the whole troposphere, subsidence, associated adiabatic heating, and clear skies which favors the radiative heating of the lowest layers of the atmosphere in combination with reduced water evaporation.

We investigated the instanton hypothesis for three other observables: T2M at two grid points situated in the South (observable S) and North (observable N) of Europe and T2M averaged over an area covering most of Western and Central Europe (observable WCE). We found a similar convergence mechanism of different trajectories, but with some discrepancies. The convergence of variance is much stronger for observable N than observable S for dynamical variables (SLP, Z500 and V250), and also stronger for observable WCE than observable W for most variables. We showed that the global dynamics described above is similar for reaching extremes at these locations. Overall we found that the instanton hypothesis is consistent with our results, but the convergence is stronger for observable farther to the north, observables computed on extended spatial locations and for longer extremes (higher values of the rolling mean window r). We also investigated the local discrepancies on the convergence of variance and found that a substantial amount of dynamical paths (around 20% for the grid point observables) are associated to the presence of cut-off lows. This suggests that the rate function I may have several minima and therefore that there may be a multi-modality in the dynamical paths reaching very intense hot events. This is however not clear what are the

precise effects - either dynamical or thermodynamical - of the presence of these cut-offs on the intensity of the observed heatwaves.

The observed discrepancies in the decrease of the variance for some variables at certain location however do not allow to disprove the instanton hypothesis. Indeed, with an order $\alpha = 0.999$ for the quantile of the observable's empirical distribution over the JJA months and 2000 summers, it corresponds to roughly choosing $n = 50$ heatwaves with a return time around 40 years. This amount of data is greater than what has been recorded through observations since the start of the satellite era. Nonetheless, this may not be sufficient to ensure full convergence and this could explain why the variance remain high at certain locations. A much longer data set would be needed to investigate the dynamics leading to more extreme temperatures, especially daily records. For such very intense events, the bi-modality suggested here may disappear. Even if our analysis suggests that very high extremes are reached by a strengthening of the mechanisms leading to high extremes, it cannot be completely ruled out that the mechanisms can change if one wants to reach even higher values. One possible explanation could be the release of latent heat from tropical air in conjunction with an atmospheric river over the North Atlantic, which seems to be one of the reinforcing structures of the 2021 North-Western America heatwave (Qian et al, 2022; Lin et al, 2022; Mo et al, 2022). To the best of our knowledge, such a dynamics has never been observed for heatwaves in Western Europe.

In this paper we did not study long term potential precursors of heatwaves such as anomalous sea surface temperature (SST) patterns or large-scale modes of climate variability (Domeisen et al, 2022a). Anomalous SST patterns are known to be present in observed heatwaves (Black and Sutton, 2007; Duchez et al, 2016; McKinnon et al, 2016), and a convergence of the surface oceanic dynamics to reach extreme land temperature is therefore likely in climate models. It may also be the case for large-scale modes of climate variability such as El Nino-Southern Oscillation (ENSO) (Martija-Díez et al, 2021), Atlantic Multidecadal Oscillation (AMO) or Pacific Decadal Oscillation (PDO). These processes are nonetheless suggested by the large regions of low variance for T2M above the Atlantic and the Pacific oceans found with our analysis.

One may wonder how the paradigm of the typical dynamics leading to heatwaves is coherent with the presence of different clusters of heatwaves demonstrated by several studies on empirical data (e.g. Stefanon et al (2012); Gibson et al (2017); Wang et al (2018); Keellings and Moradkhani (2020); Mondal and Mishra (2021)). As we have shown here, the typical dynamics highly depends on the observable considered. It may therefore be possible that the typical dynamics leading to extremes of neighboring grid points is very similar and changes dramatically when crossing relevant physical barriers, such as mountains (Lucarini et al, 2022). Moreover, if one considers less extreme values of the distribution of grid point observables, it is possible that the typicality is not reached for this observable but it may be reached for another observable encompassing a broader region. One could therefore compare the

event observed to the typical events for the observable which maximizes its spatiotemporal rarity as proposed by Cattiaux and Ribes (2018) in the context of attribution of extreme events.

Despite the lack of data for extremely rare events, the framework of large deviations theory offers a significant simplification by predicting that studying such events is equivalent to studying a single trajectory, assuming that the rate function has only one minimum. Therefore, this theoretical framework is encouraging for gaining predictive power on the dynamics leading to extreme events and it may explain why summer heatwaves are among the most predictable meteorological extremes on subseasonal timescales (Vitart and Robertson, 2018; Vitart et al, 2019; Domeisen et al, 2022b). Here we studied the extremes of air temperature close to the surface. For other variables of interest our preceding statement of the validity of the instanton hypothesis should also be tested. More generally, this paper documents a method to study the dynamics leading to extreme events in non-equilibrium physical systems. Our framework suggests a connection between the statistical method of studying extreme events and the in-depth examination of specific events through case studies.

Acknowledgments. The authors would like to thank V. Lucarini, M. Galfi, G. Messori and A. Caubel for fruitful discussions.

Declarations

- Funding: This work has received support from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101003469 (XAIDA: PY, FP, AJ) and the grant ANR-20-CE01-0008-01 (SAMPRACE: PY).
- Conflict of interest: The authors declare no conflict of interest.
- Ethics approval : Not applicable.
- Consent to participate: Not applicable.
- Consent for publication: Not applicable.
- Availability of data and materials: The control run of the IPSL model is available upon request.
- Code availability: The main results of this work were obtained using Python. The scripts are available upon request.
- Authors' contributions: RN did the data analysis. All the authors contributed to writing and reviewing the article.

Appendix A Supplementary information

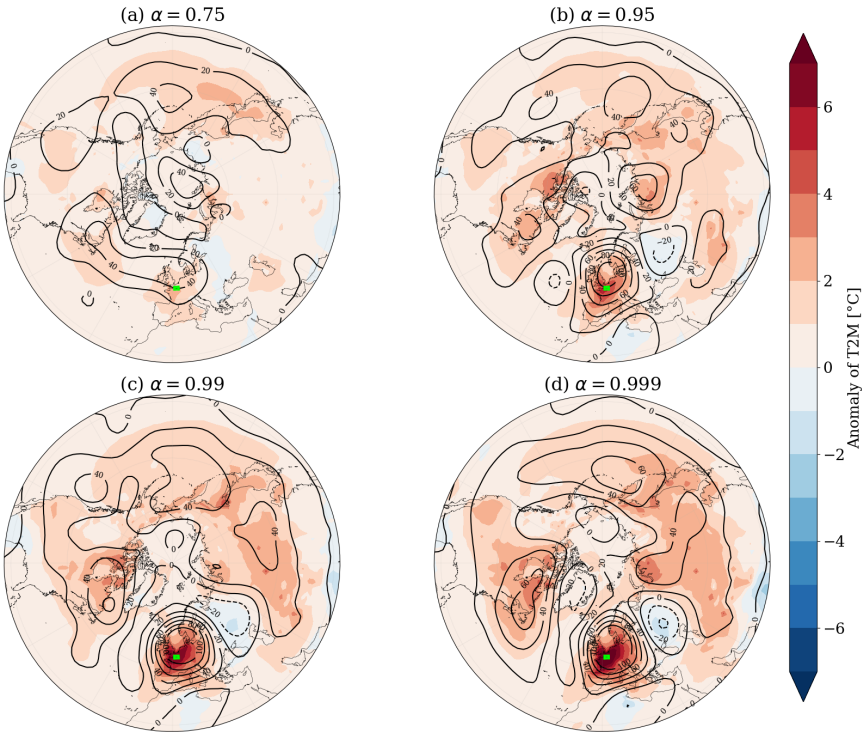

Fig. A1 Composite maps $\widehat{A}$ of anomalies of T2M and Z500 for increasing α-th order of the quantile of the W observable's empirical distribution. The figure is computed for a rolling mean window of $r = 5$ days. Colors: anomaly of T2M [K]. Contours: anomaly of Z500 [m]. Anomalies are computed with respect to the JJA average. The green box displays the location where the observable is computed.

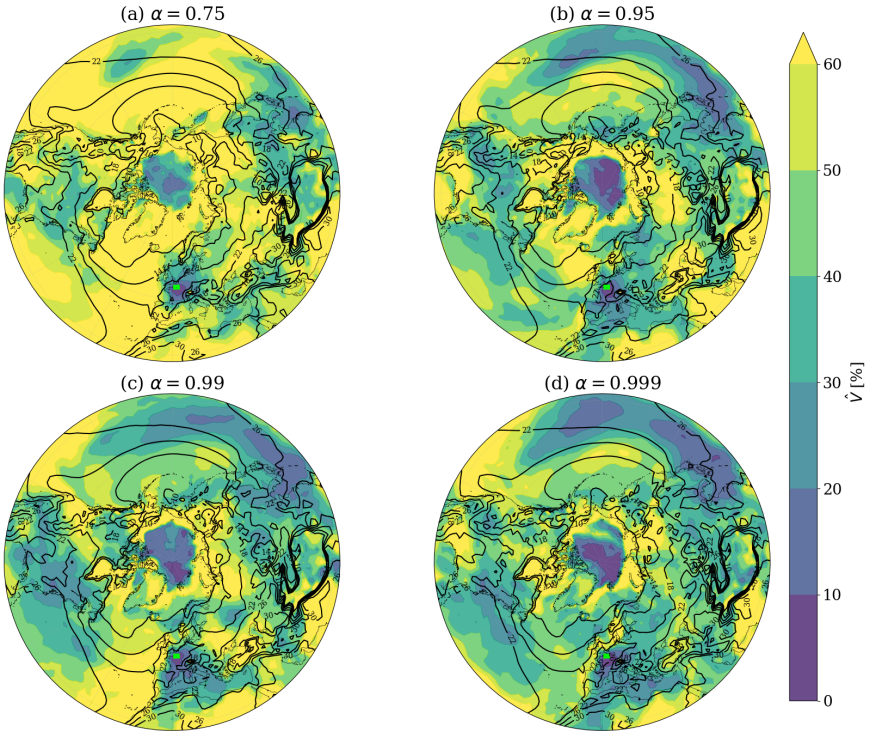

Fig. A2 Composite $\widehat{A}$ (contours) [°C] and normalized variance $\widehat{V}$ (colors) of the T2M field for increasing α-th order of the quantile of the W observable's empirical distribution. The figure is computed for a rolling mean window of $r = 5$ days. The normalized variance is computed after averaging over the rolling window and is expressed in %. The green box displays the location where the observable is computed.

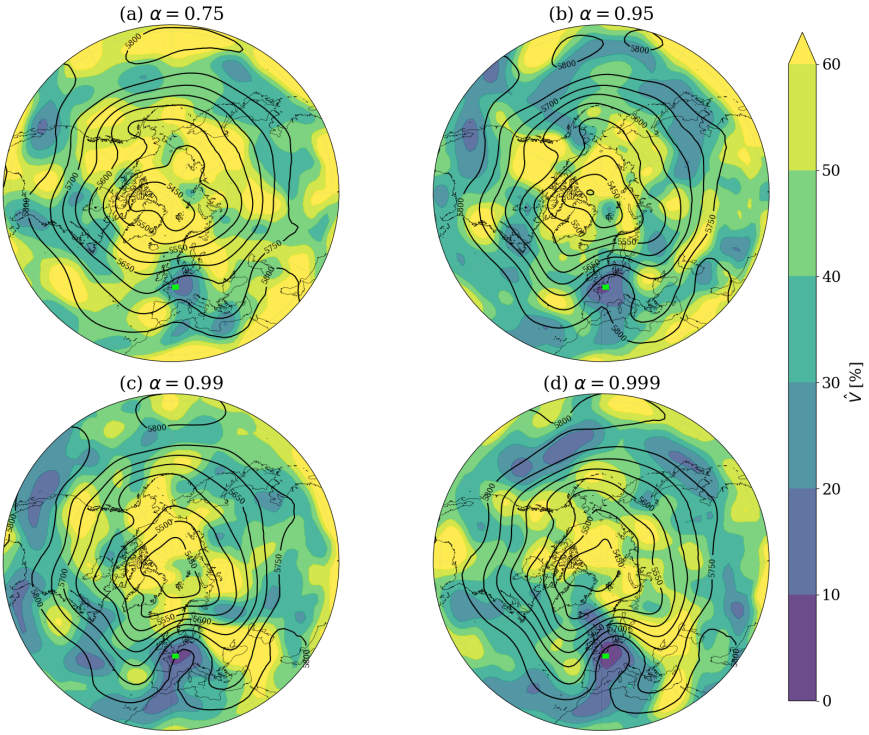

Fig. A3 Composite $\widehat{A}$ (contours) [m] and normalized variance $\widehat{V}$ (colors) of the Z500 field for increasing α-th order of the quantile of the W observable's empirical distribution. The figure is computed for a rolling mean window of $r = 5$ days. The normalized variance is computed after averaging over the rolling window and is expressed in %. The green box displays the location where the observable is computed.

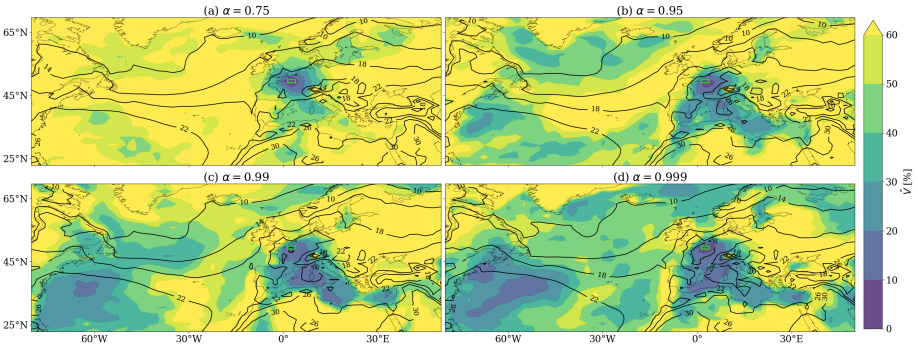

Fig. A4 Composite $\widehat{A}$ (contours) [°C] and normalized variance $\widehat{V}$ (colors) of the T2M field for increasing α-th order of the quantile of the W observable's empirical distribution. The figure is computed for a rolling mean window of $r = 1$ day. The normalized variance is computed after averaging over the rolling window and is expressed in %. The green box displays the location where the observable is computed.

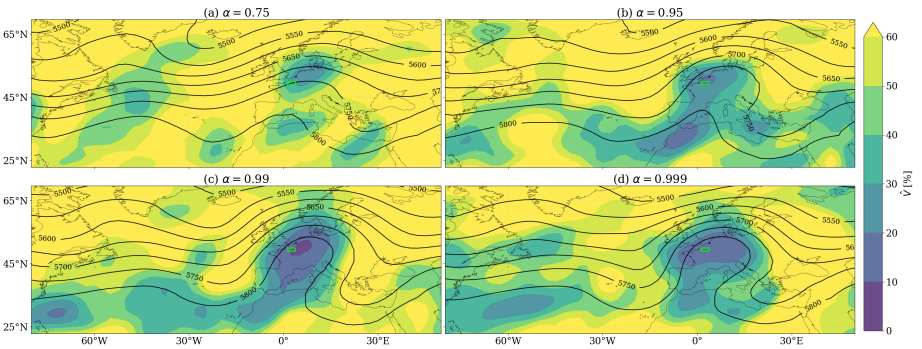

Fig. A5 Composite $\widehat{A}$ (contours) [m] and normalized variance $\widehat{V}$ (colors) of the Z500 field for increasing α-th order of the quantile of the W observable's empirical distribution. The figure is computed for a rolling mean window of $r = 1$ day. The normalized variance is computed after averaging over the rolling window and is expressed in %. The green box displays the location where the observable is computed.

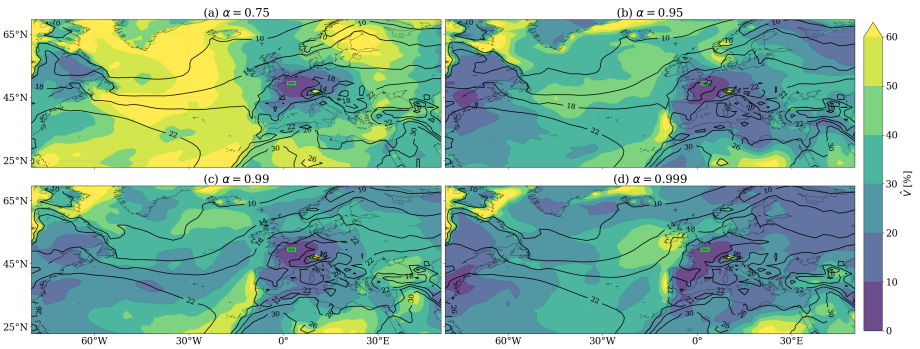

Fig. A6 Composite $\widehat{A}$ (contours) [°C] and normalized variance $\widehat{V}$ (colors) of the T2M field for increasing α-th order of the quantile of the W observable's empirical distribution. The figure is computed for a rolling mean window of $r = 15$ days. The normalized variance is computed after averaging over the rolling window and is expressed in %. The green box displays the location where the observable is computed.

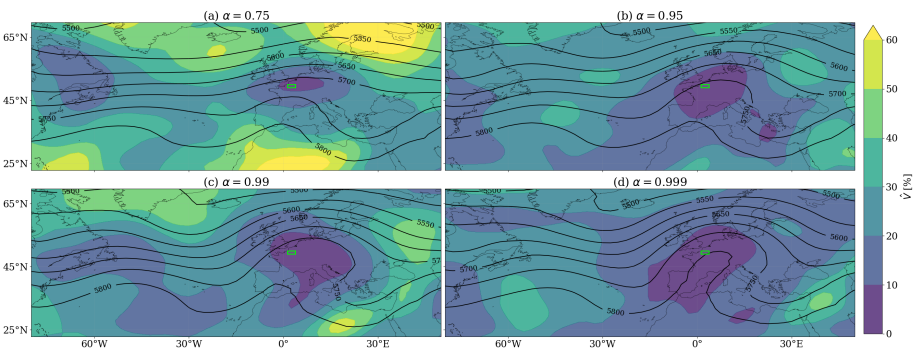

Fig. A7 Composite $\widehat{A}$ (contours) [m] and normalized variance $\widehat{V}$ (colors) of the Z500 field for increasing α-th order of the quantile of the W observable's empirical distribution. The figure is computed for a rolling mean window of $r = 15$ days. The normalized variance is computed after averaging over the rolling window and is expressed in %. The green box displays the location where the observable is computed.

	r=1			r=5			r=15		
Observable	0.95	0.99	0.999	0.95	0.99	0.999	0.95	0.99	0.999
S	22%	30%	22%	22%	16%	24%	24%	22%	16%
W	22%	16%	20%	14%	30%	28%	16%	22%	24%
N	28%	26%	14%	20%	20%	20%	16%	32%	20%
WCE	56%	44%	50%	38%	60%	48%	42%	32%	58%

Table A1 Percentage of cut-off lows in the dynamics leading to extremes.

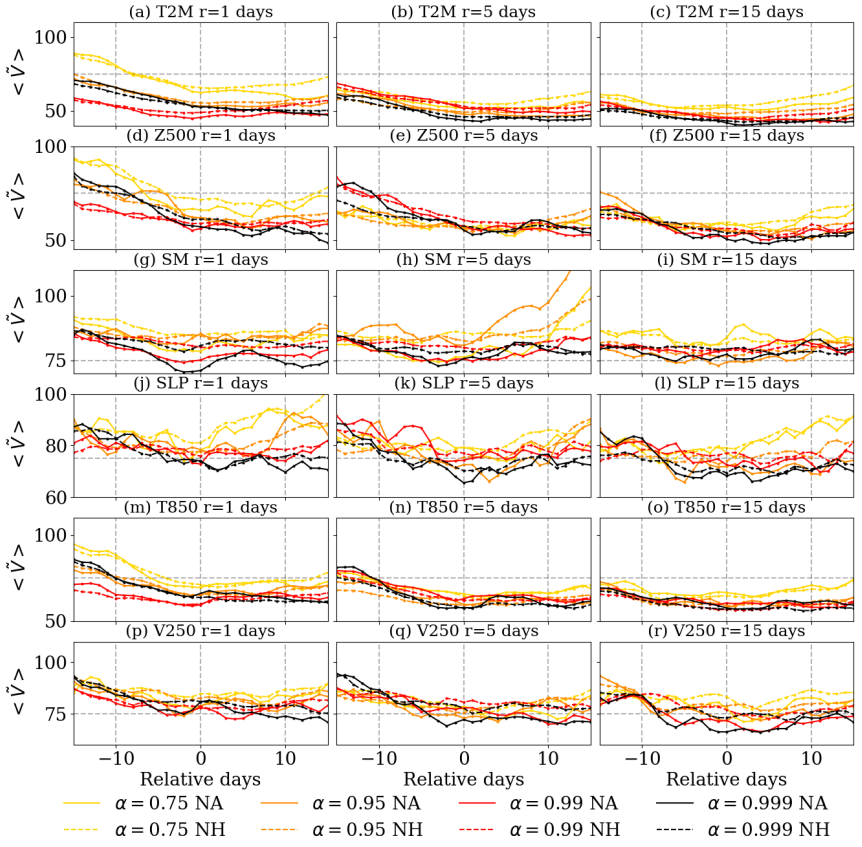

Fig. A8 Evolution of the normalized variance $< \widetilde{V}(t) >$ averaged over the Euro-Atlantic sector (plain lines) and the North-Hemisphere (dashed lines) for the different variables (S observable). The normalized variance is expressed in %. The colors show the α-th order quantile of the observable's empirical distribution. The time is expressed relative to the day when the observable is such that $F_r(x_t) \simeq q_\alpha$. The gray dashed line shows the 75% level.

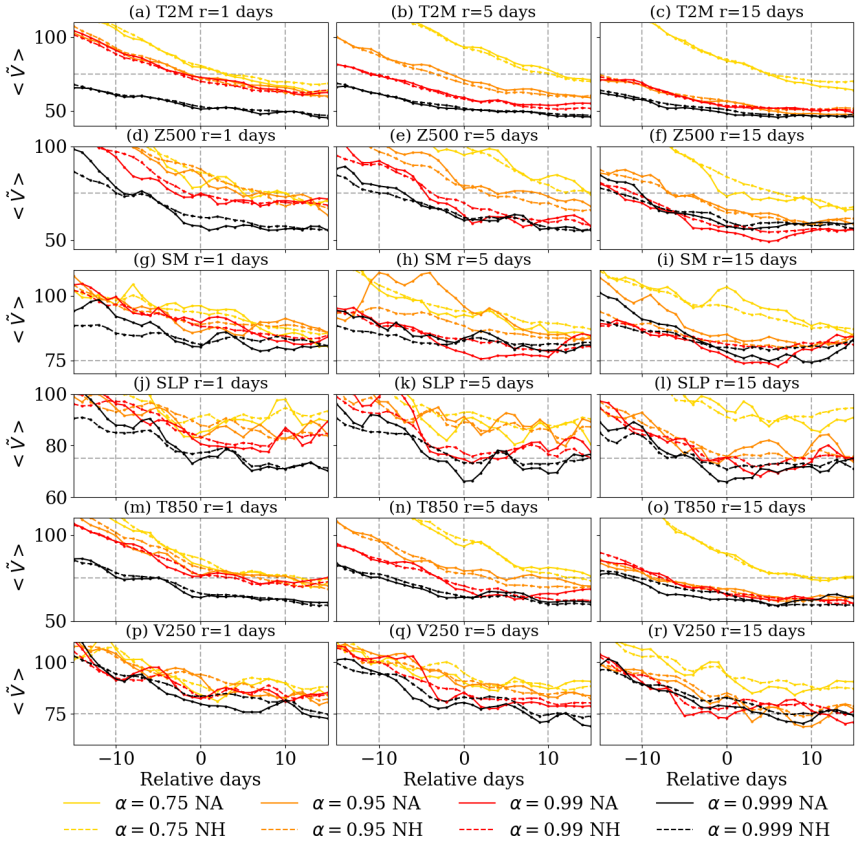

Fig. A9 Evolution of the normalized variance $< \widetilde{V}(t) >$ averaged over the Euro-Atlantic sector (plain lines) and the North-Hemisphere (dashed lines) for the different variables (N observable). The normalized variance is expressed in %. The colors show the α-th order quantile of the observable's empirical distribution. The time is expressed relative to the day when the observable is such that $F_r(x_t) \simeq q_\alpha$. The gray dashed line shows the 75% level.

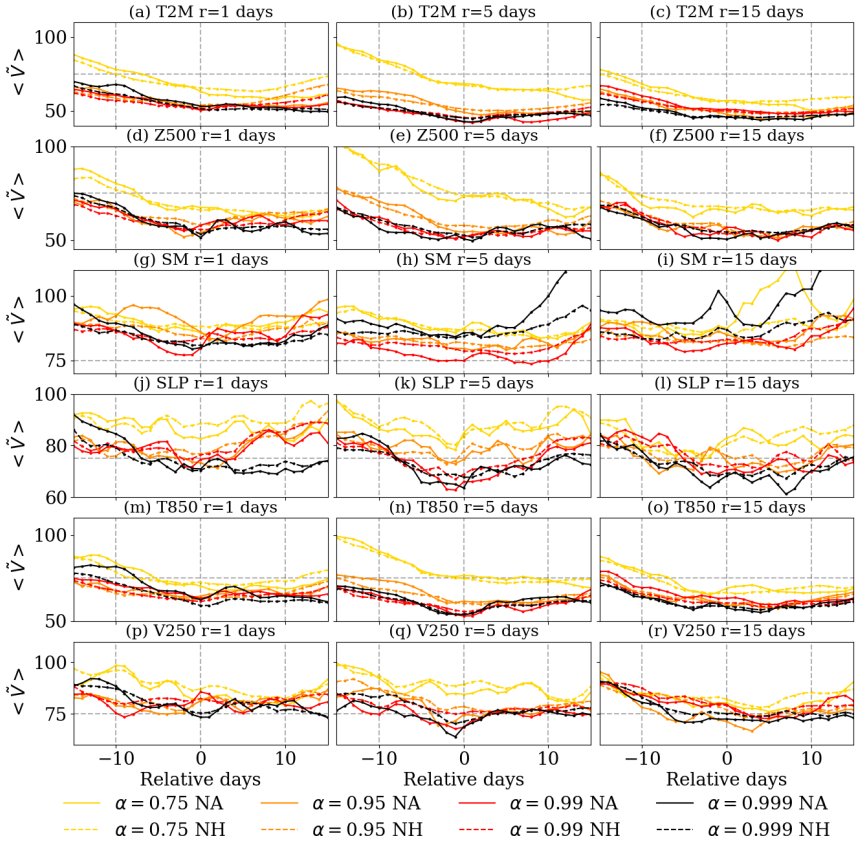

Fig. A10 Evolution of the normalized variance $< \widetilde{V}(t) >$ averaged over the Euro-Atlantic sector (plain lines) and the North-Hemisphere (dashed lines) for the different variables (WCE observable). The normalized variance is expressed in %. The colors show the α-th order quantile of the observable's empirical distribution. The time is expressed relative to the day when the observable is such that $F_r(x_t) \simeq q_\alpha$. The gray dashed line shows the 75% level.

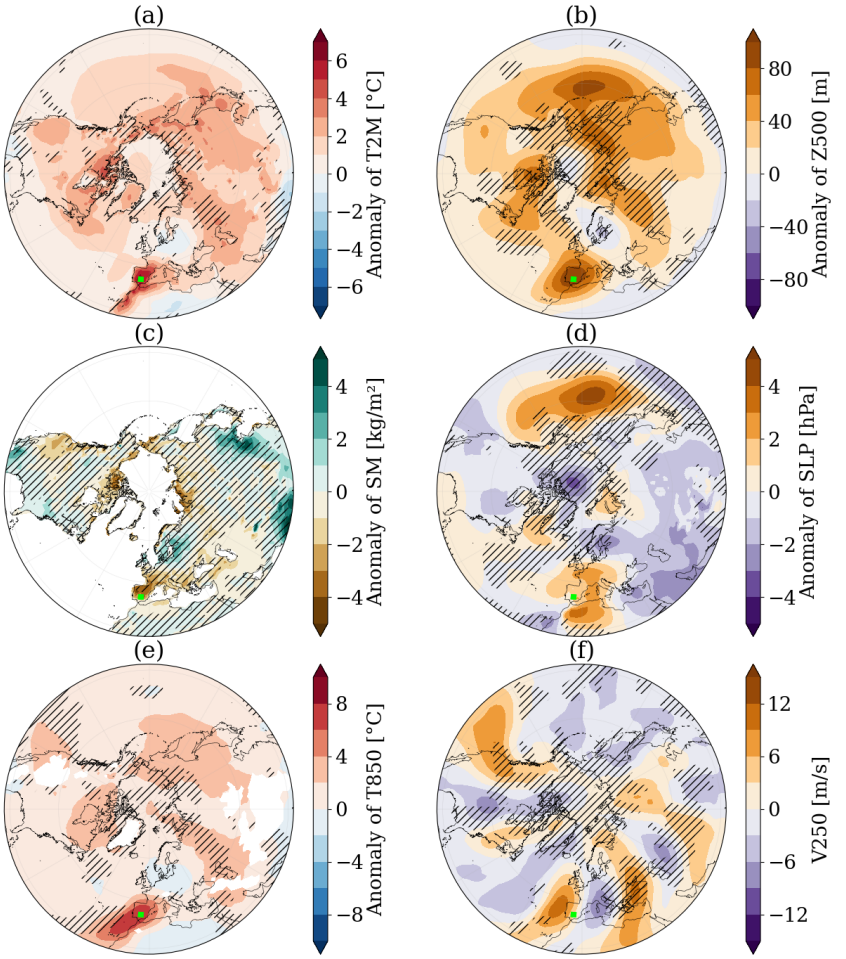

Fig. A11 Composite maps $\widehat{A}$ for T2M, Z500, SM, SLP, T850 and V250 for the quantile of order $\alpha = 0.999$ of the S observable's empirical distribution. The hatch areas correspond to $\widehat{V} > 50\%$. The figure is computed for a rolling mean window of $r = 5$ days. Anomaly of (a) air temperature at 2m (T2M), (b) geopotential height at 500hPa (Z500), (c) soil moisture (SM), (d) sea-level pressure (SLP) and (e) temperature at 850hPa (T850), and (f) meridional wind speed at 250hPa (V250). The anomalies are computed with respect to their average over the summer (JJA). The green box displays the location where the observable is computed.

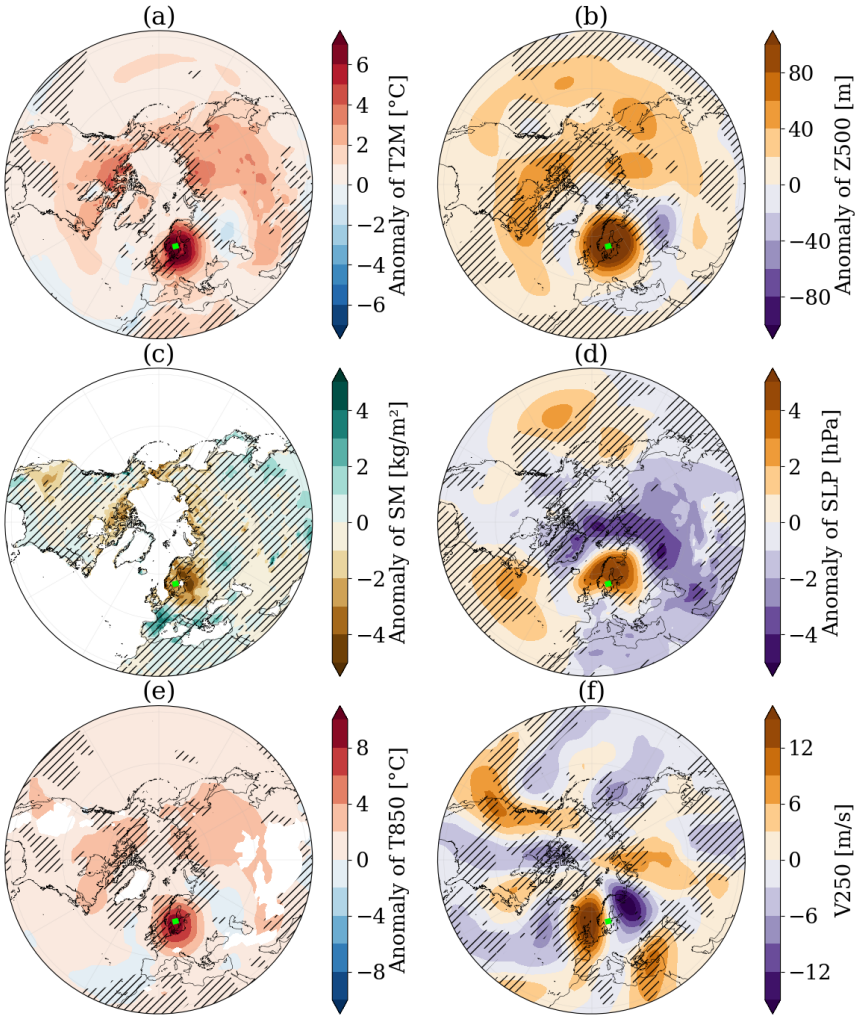

Fig. A12 Composite maps $\widehat{A}$ for T2M, Z500, SM, SLP, T850 and V250 for the quantile of order $\alpha = 0.999$ of the N observable's empirical distribution. The hatch areas correspond to $\widehat{V} > 50\%$. The figure is computed for a rolling mean window of $r = 5$ days. Anomaly of (a) air temperature at 2m (T2M), (b) geopotential height at 500hPa (Z500), (c) soil moisture (SM), (d) sea-level pressure (SLP) and (e) temperature at 850hPa (T850), and (f) meridional wind speed at 250hPa (V250). The anomalies are computed with respect to their average over the summer (JJA). The green box displays the location where the observable is computed.

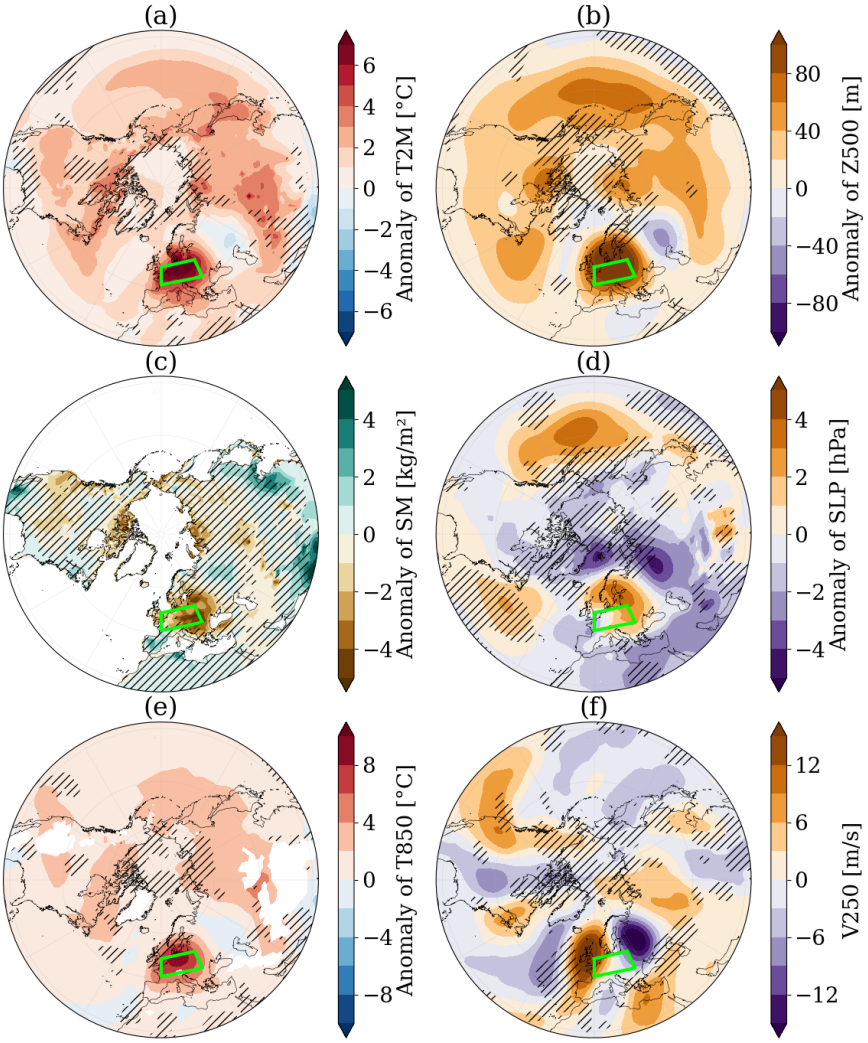

Fig. A13 Composite maps $\widehat{A}$ for T2M, Z500, SM, SLP, T850 and V250 for the quantile of order $\alpha = 0.999$ of the WCE observable's empirical distribution. The hatch areas correspond to $\widehat{V} > 50\%$. The figure is computed for a rolling mean window of $r = 5$ days. Anomaly of (a) air temperature at 2m (T2M), (b) geopotential height at 500hPa (Z500), (c) soil moisture (SM), (d) sea-level pressure (SLP) and (e) temperature at 850hPa (T850), and (f) meridional wind speed at 250hPa (V250). The anomalies are computed with respect to their average over the summer (JJA). The green box displays the location where the observable is computed.

References

Bartusek S, Kornhuber K, Ting M (2022) 2021 north american heatwave amplified by climate change-driven nonlinear interactions. Nature Climate Change pp 1–8

Beguería S, Vicente-Serrano SM, Reig F, et al (2014) Standardized precipitation evapotranspiration index (spei) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. International journal of climatology 34(10):3001–3023

Black E, Sutton R (2007) The influence of oceanic conditions on the hot european summer of 2003. Climate dynamics 28(1):53–66

Boucher O, Servonnat J, Albright AL, et al (2020) Presentation and evaluation of the ipsl-cm6a-lr climate model. Journal of Advances in Modeling Earth Systems 12(7):e2019MS002,010

Butchart N, Haines K, Marshall J (1989) A theoretical and diagnostic study of solitary waves and atmospheric blocking. Journal of Atmospheric Sciences 46(13):2063–2078

Cattiaux J, Ribes A (2018) Defining single extreme weather events in a climate perspective. Bulletin of the American Meteorological Society 99(8):1557–1568

Cheng L, AghaKouchak A, Gilleland E, et al (2014) Non-stationary extreme value analysis in a changing climate. Climatic change 127:353–369

Chetrite R, Touchette H (2015) Nonequilibrium markov processes conditioned on large deviations. In: Annales Henri Poincaré, Springer, pp 2005–2057

Coles S, Bawa J, Trenner L, et al (2001) An introduction to statistical modeling of extreme values, vol 208. Springer

Coumou D, Petoukhov V, Rahmstorf S, et al (2014) Quasi-resonant circulation regimes and hemispheric synchronization of extreme weather in boreal summer. Proceedings of the National Academy of Sciences 111(34):12,331–12,336

Dematteis G, Grafke T, Onorato M, et al (2019a) Experimental evidence of hydrodynamic instantons: the universal route to rogue waves. Physical Review X 9(4):041,057

Dematteis G, Grafke T, Vanden-Eijnden E (2019b) Extreme event quantification in dynamical systems with random components. SIAM/ASA Journal on Uncertainty Quantification 7(3):1029–1059

Di Capua G, Sparrow S, Kornhuber K, et al (2021) Drivers behind the summer 2010 wave train leading to russian heatwave and pakistan flooding. npj Climate and Atmospheric Science 4(1):1–14

Dirmeyer PA, Balsamo G, Blyth EM, et al (2021) Land-atmosphere interactions exacerbated the drought and heatwave over northern europe during summer 2018. AGU Advances 2(2):e2020AV000,283

Dole R, Hoerling M, Perlwitz J, et al (2011) Was there a basis for anticipating the 2010 russian heat wave? Geophysical Research Letters 38(6)

Domeisen DI, Eltahir EA, Fischer EM, et al (2022a) Prediction and projection of heatwaves. Nature Reviews Earth & Environment pp 1–15

Domeisen DI, White CJ, Afargan-Gerstman H, et al (2022b) Advances in the subseasonal prediction of extreme events: Relevant case studies across the globe. Bulletin of the American Meteorological Society

Duchez A, Frajka-Williams E, Josey SA, et al (2016) Drivers of exceptionally cold north atlantic ocean temperatures and their link to the 2015 european heat wave. Environmental Research Letters 11(7):074,004

Eyring V, Bony S, Meehl GA, et al (2016) Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. Geoscientific Model Development 9(5):1937–1958

Faranda D, Pascale S, Bulut B (2022) Persistent anticyclonic conditions and climate change exacerbated the exceptional 2022 european-mediterranean drought

Fragkoulidis G, Wirth V, Bossmann P, et al (2018) Linking northern hemisphere temperature extremes to rossby wave packets. Quarterly Journal of the Royal Meteorological Society 144(711):553–566

Freidlin MI, Wentzell AD (1987) Random perturbations. In: Random perturbations of dynamical systems. Springer, p 15–43

Galfi VM, Lucarini V (2021) Fingerprinting heatwaves and cold spells and assessing their response to climate change using large deviation theory. Physical Review Letters 127(5):058,701

Galfi VM, Lucarini V, Ragone F, et al (2021) Applications of large deviation theory in geophysical fluid dynamics and climate science. La Rivista del Nuovo Cimento 44(6):291–363

García-Herrera R, Díaz J, Trigo RM, et al (2010) A review of the european summer heat wave of 2003. Critical Reviews in Environmental Science and

Technology 40(4):267–306

Ghil M, Yiou P, Hallegatte S, et al (2011) Extreme events: dynamics, statistics and prediction. Nonlinear Processes in Geophysics 18(3):295–350

Gibson PB, Perkins-Kirkpatrick SE, Alexander LV, et al (2017) Comparing australian heat waves in the cmip5 models through cluster analysis. Journal of Geophysical Research: Atmospheres 122(6):3266–3281

Grafke T, Vanden-Eijnden E (2019) Numerical computation of rare events via large deviation theory. Chaos: An Interdisciplinary Journal of Nonlinear Science 29(6):063,118. https://doi.org/10.1063/1.5084025, URL https://aip.scitation.org/doi/abs/10.1063/1.5084025, publisher: American Institute of Physics

Grafke T, Grauer R, Schäfer T (2013) Instanton filtering for the stochastic burgers equation. Journal of Physics A: Mathematical and Theoretical 46(6):062,002

Hirschi M, Seneviratne SI, Alexandrov V, et al (2011) Observational evidence for soil-moisture impact on hot extremes in southeastern europe. Nature Geoscience 4(1):17–21

Horton RM, Mankin JS, Lesk C, et al (2016) A review of recent advances in research on extreme heat events. Current Climate Change Reports 2(4):242–259

Keellings D, Moradkhani H (2020) Spatiotemporal evolution of heat wave severity and coverage across the united states. Geophysical Research Letters 47(9):e2020GL087,097

Koppe C, Kovats S, Jendritzky G, et al (2004) Heat-waves: risks and responses. EUR/03/5036810, World Health Organization. Regional Office for Europe

Kornhuber K, Tamarin-Brodsky T (2021) Future changes in northern hemisphere summer weather persistence linked to projected arctic warming. Geophysical Research Letters 48(4):e2020GL091,603

Kornhuber K, Petoukhov V, Karoly D, et al (2017) Summertime planetary wave resonance in the northern and southern hemispheres. Journal of Climate 30(16):6133–6150

Kornhuber K, Coumou D, Vogel E, et al (2020) Amplified rossby waves enhance risk of concurrent heatwaves in major breadbasket regions. Nature Climate Change 10(1):48–53

Lin H, Mo R, Vitart F (2022) The 2021 western north american heatwave and its subseasonal predictions. Geophysical Research Letters 49(6):e2021GL097,036

Lucarini V, Galfi VM, Messori G, et al (2022) Typicality of the 2021 western north america summer heatwave. arXiv preprint arXiv:220606197

Mann ME, Rahmstorf S, Kornhuber K, et al (2017) Influence of anthropogenic climate change on planetary wave resonance and extreme weather events. Scientific reports 7(1):1–12

Mann ME, Rahmstorf S, Kornhuber K, et al (2018) Projected changes in persistent extreme summer weather events: The role of quasi-resonant amplification. Science advances 4(10):eaat3272

Martija-Díez M, Rodríguez-Fonseca B, López-Parages J (2021) Enso influence on western european summer and fall temperatures. Journal of Climate 34(19):8013–8031

McKinnon KA, Rhines A, Tingley M, et al (2016) Long-lead predictions of eastern united states hot days from pacific sea surface temperatures. Nature Geoscience 9(5):389–394

Miralles DG, Van Den Berg M, Teuling A, et al (2012) Soil moisture-temperature coupling: A multiscale observational analysis. Geophysical Research Letters 39(21)

Miralles DG, Teuling AJ, Van Heerwaarden CC, et al (2014) Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. Nature geoscience 7(5):345–349

Mo R, Lin H, Vitart F (2022) An anomalous warm-season trans-pacific atmospheric river linked to the 2021 western north america heatwave. Communications Earth & Environment 3(1):1–12

Moloney NR, Faranda D, Sato Y (2019) An overview of the extremal index. Chaos: An Interdisciplinary Journal of Nonlinear Science 29(2):022,101

Mondal S, Mishra AK (2021) Complex networks reveal heatwave patterns and propagations over the usa. Geophysical Research Letters 48(2):e2020GL090,411

Muñoz C, Schultz D, Vaughan G (2020) A midlatitude climatology and interannual variability of 200-and 500-hpa cut-off lows. Journal of Climate 33(6):2201–2222

Neal E, Huang CS, Nakamura N (2022) The 2021 pacific northwest heat wave and associated blocking: Meteorology and the role of an upstream cyclone as a diabatic source of wave activity. Geophysical Research Letters 49(8):e2021GL097,699

Otto FE, Massey N, van Oldenborgh GJ, et al (2012) Reconciling two approaches to attribution of the 2010 russian heat wave. Geophysical Research Letters 39(4)

Overland JE (2021) Causes of the record-breaking pacific northwest heatwave, late june 2021. Atmosphere 12(11):1434

Perkins SE (2015) A review on the scientific understanding of heatwaves—their measurement, driving mechanisms, and changes at the global scale. Atmospheric Research 164:242–267

Petoukhov V, Rahmstorf S, Petri S, et al (2013) Quasiresonant amplification of planetary waves and recent northern hemisphere weather extremes. Proceedings of the National Academy of Sciences 110(14):5336–5341

Petoukhov V, Petri S, Rahmstorf S, et al (2016) Role of quasiresonant planetary wave dynamics in recent boreal spring-to-autumn extreme events. Proceedings of the National Academy of Sciences 113(25):6862–6867

Philip SY, Kew SF, van Oldenborgh GJ, et al (2021) Rapid attribution analysis of the extraordinary heatwave on the pacific coast of the us and canada june 2021. Earth System Dynamics Discussions pp 1–34

Pörtner HO, Roberts DC, Adams H, et al (2022) Climate change 2022: Impacts, adaptation and vulnerability. IPCC Sixth Assessment Report

Qian Y, Hsu PC, Yuan J, et al (2022) Effects of subseasonal variation in the east asian monsoon system on the summertime heat wave in western north america in 2021. Geophysical Research Letters 49(8):e2021GL097,659

Rasmijn L, Van der Schrier G, Bintanja R, et al (2018) Future equivalent of 2010 russian heatwave intensified by weakening soil moisture constraints. Nature Climate Change 8(5):381–385

Rousi E, Kornhuber K, Beobide-Arsuaga G, et al (2022) Accelerated western european heatwave trends linked to more-persistent double jets over eurasia. Nature communications 13(1):1–11

Screen JA, Simmonds I (2014) Amplified mid-latitude planetary waves favour particular regional weather extremes. Nature Climate Change 4(8):704–709

Seneviratne SI, Zhang X, Adnan M, et al (2021) 11 Chapter 11: Weather and climate extreme events in a changing climate

Sousa PM, Trigo RM, Barriopedro D, et al (2018) European temperature responses to blocking and ridge regional patterns. Climate Dynamics 50(1):457–477

Stefanon M, D'Andrea F, Drobinski P (2012) Heatwave classification over europe and the mediterranean region. Environmental Research Letters 7(1):014,023

Tawn JA (1990) Modelling multivariate extreme value distributions. Biometrika 77(2):245–253

Thompson V, Kennedy-Asser AT, Vosper E, et al (2022) The 2021 western north america heat wave among the most extreme events ever recorded globally. Science advances 8(18):eabm6860

Touchette H (2009) The large deviation approach to statistical mechanics. Physics Reports 478(1-3):1–69

Trenberth KE, Fasullo JT (2012) Climate extremes and climate change: The russian heat wave and other climate extremes of 2010. Journal of Geophysical Research: Atmospheres 117(D17)

Vitart F, Robertson AW (2018) The sub-seasonal to seasonal prediction project (s2s) and the prediction of extreme events. npj Climate and Atmospheric Science 1(1):1–7

Vitart F, Cunningham C, DeFlorio M, et al (2019) Sub-seasonal to seasonal prediction of weather extremes. In: Sub-seasonal to seasonal prediction. Elsevier, p 365–386

Wang P, Tang J, Wang S, et al (2018) Regional heatwaves in china: a cluster analysis. Climate dynamics 50(5):1901–1917

Yaghmaei N (2020) Human Cost of Disasters: An Overview of the Last 20 Years, 2000-2019. UN Office for Disaster Risk Reduction

Zhang F, Biederman JA, Dannenberg MP, et al (2021) Five decades of observed daily precipitation reveal longer and more variable drought events across much of the western united states. Geophysical Research Letters 48(7):e2020GL092,293

Zhang W, Hari V, S-Y Wang S, et al (2022) Fewer troughs, not more ridges, have led to a drying trend in the western united states. Geophysical Research Letters 49(1):e2021GL097,089

Zschenderlein P, Fink AH, Pfahl S, et al (2019) Processes determining heat waves across different european climates. Quarterly Journal of the Royal Meteorological Society 145(724):2973–2989

Zuo J, Pullen S, Palmer J, et al (2015) Impacts of heat waves and corresponding measures: a review. Journal of Cleaner Production 92:1–12