Statistical and dynamical aspects of extremely hot summers in Western Europe sampled with a rare events algorithm

Robin Noyelle, Arnaud Caubel, Yann Meurdesoif, Pascal Yiou, Davide Faranda

Statistical and dynamical aspects of extremely hot summers in Western Europe sampled with a rare events algorithm

Robin Noyelle,[a,b] Arnaud Caubel,[a] Yann Meurdesoif,[a] Pascal Yiou,[a] and Davide Faranda[a,c,d]

[a] *Laboratoire des Sciences du Climat et de l'Environnement, UMR 8212 CEA-CNRS-UVSQ, IPSL & Université Paris Saclay, Gif-sur-Yvette, 91191,France*

[b] *Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland*

[c] *London Mathematical Laboratory, 8 Margravine Gardens London, W6 8RH, London, United Kingdom*

[d] *Laboratoire de Météorologie Dynamique/IPSL, École Normale Supérieure, PSL Research University, Sorbonne Université, École Polytechnique, IP Paris, CNRS, 75005, Paris, France*

Corresponding author: Robin Noyelle, robin.noyelle@lsce.ipsl.fr

ABSTRACT: The study of the statistical and dynamical characteristics of extreme and very extreme events in the climate system is impaired by a strong under-sampling issue. Because extreme events are rare, answering questions about the physical mechanisms from which they arise usually depends on the investigation of just a few cases, either in observations or in models. In this paper we use a rare events algorithm to massively increase the number of extremely hot, dry and anticyclonic summers in Western Europe simulated in the state-of-the-art IPSL-CM6A-LR climate model under pre-industrial anthropogenic forcings. This allows us to reach precise climatological results on the dynamics leading to centennial hot summers. We show that these summers are characterized by both more and longer heatwaves than usual, and are associated to a combination of a local non-barotropic positive geopotential height anomaly and non-local eastward-shifted soil moisture anomalies. The atmospheric dynamics of these summers is mainly driven by a recurrence of short Rossby wave packets of synoptic scale rather than circumglobal wave patterns.

SIGNIFICANCE STATEMENT: Studying very extreme meteorological and climatological events is difficult because these events are rare and therefore seldom observed in both observations and simulations. Here we employ a so-called rare events algorithm to simulate a large number of extremely hot, dry and anticyclonic summers in Western Europe with a climate model. We show that the algorithm is able to simulate efficiently extremely rare events, with return times ranging from one hundred to one million years. This allows us to describe precisely the physical mechanisms leading to these very extreme summers.

1. Introduction

Because of their tremendous impacts on societies and ecosystems, extreme climatological and meteorological events have been increasingly studied by the scientific community (Seneviratne et al. 2021). Besides, understanding properties of the tail of the distribution of climatological variables rather than only the mean, is a legitimate climate science question. Heat events, and especially mid-latitude heatwaves, have been the focus of scientific attention (Perkins 2015; Horton et al. 2016; Domeisen et al. 2023; Barriopedro et al. 2023), in particular since the record breaking 2003 summer in Western Europe (Stott et al. 2004). This attention came as a result of these events being the ones for which there is the strongest consensus on their observed and future increase in both frequency and intensity in response to global warming (Perkins-Kirkpatrick and Lewis 2020; Seneviratne et al. 2021; Barriopedro et al. 2023; Domeisen et al. 2023; Van Loon and Thompson 2023; Huntingford et al. 2024).

A mid-latitude heatwave is primarily an atmospheric-driven event (Horowitz et al. 2022): it occurs because of the presence of a long-lasting quasi-stationary and self-sustaining anticyclonic structure disrupting the climatological westerlies (Hoskins and Woollings 2015; Pfleiderer and Coumou 2018; Kautz et al. 2022). These anticyclones are usually embedded in a slow-moving or quasi-stationary larger scale structure, often an atmospheric block (Xoplaki et al. 2003; Meehl and Tebaldi 2004; Stefanon et al. 2012; Pfahl and Wernli 2012; Schaller et al. 2018; Castañeda and Wang 2024) or a stationary subtropical ridge (Marshall et al. 2014; Sousa et al. 2018; Jiménez-Esteve et al. 2022). Close to the surface, a heat low can be present, creating local weak cyclonic circulation (Della-Marta et al. 2007; Fischer et al. 2007). There is still limited understanding on the factors determining the onset and maintenance of these atmospheric systems in general

and in conjunction with large anomalies of near-surface temperature in particular (Barriopedro et al. 2023). These factors are numerous and may vary regionally, but they usually involve the interaction of atmospheric Rossby waves (Rossby 1939) of different scales. Although the role of amplified Rossby waves is well recognized for the emergence of persistent anticyclonic anomalies, two different mechanisms have been proposed for the formation of blocking anticyclones leading to heatwaves (Horton et al. 2016): amplified quasi-stationary circumglobal Rossby waves (CGW) and high-amplitude transient non-circumglobal Rossby wave packets (RWP).

The hypothesis of amplified CGW by waveguides for explaining heatwaves was first proposed by Petoukhov et al. (2013) and the associated physical mechanism is named quasi-resonant amplification (QRA). The authors suggested that a common mechanism for the generation of persistent planetary-scale high amplitude patterns of the atmospheric circulation with high zonal wave numbers results from the trapping and amplification within mid-latitude waveguides of free synoptic waves. This would create highly persistent configurations in summertime, reducing summer variability and favoring extremes teleconnections. This mechanism may be particularly relevant for concurrent heatwaves around the hemisphere (Coumou et al. 2014; Screen and Simmonds 2014; Petoukhov et al. 2016; Mann et al. 2017; Kornhuber et al. 2019; He et al. 2023). Although the link between preceding patterns of anomalous atmospheric planetary waves and surface heatwaves has been shown by several studies (Teng et al. 2013), the QRA mechanism has been criticized. In general, summer waveguides tend to be weaker than winter waveguides and not circumglobal (Teng and Branstator 2019). The contributions of CGW to temperature extremes is not higher than that of non-circumglobal patterns (Fragkoulidis et al. 2018; Röthlisberger et al. 2016). Although there is a statistically significant link between jet waviness on a hemispheric scale and monthly temperature anomalies (Screen and Simmonds 2014), and on synoptic time scales periods of frequently occurring temperature extremes have been linked to low storm track activity (Lehmann and Coumou 2015; Coumou et al. 2015; Pfleiderer et al. 2019), the link between jet waviness and the occurrence of weather extremes is stronger for regional scale than hemispheric jet waviness (Röthlisberger et al. 2016). For example, Branstator and Teng (2017) analyzed correlation maps of the sub-seasonal variability of 200-hPa meridional wind to quantify the waveguidability and found that waviness in the upper troposphere is typically not stretched out circumglobally but organized in eastward propagating patches of limited spatial extent (Fragkoulidis et al. 2018). Moreover, the

very definition of a waveguide may be problematic and Wirth and Polster (2021) have suggested that jets may arise as the consequence rather than the cause of waves with large amplitudes.

Another explanation for large and persistent anticyclonic anomalies is the recurrence of transient, fast-moving and amplified RWP that can arise from a variety of sources (Zschenderlein et al. 2018; Wirth et al. 2018; Röthlisberger et al. 2019). Recurrent RWPs arising from multiple transient synoptic scale wave packets re-amplifying in the same geographical region may result into a persistent anomalies (Jiménez-Esteve et al. 2022; Tuel and Martius 2024). As such recurrent RWP can trigger persistent surface weather anomalies over multiple synoptic wavelengths, while blocking is more local (one synoptic wavelength). Waves amplifying in the same phase upstream of the block could continuously reinforce the block and lead to recurrent RWP conditions upstream. Diabatic processes and latent heat release associated with a series of transient synoptic cyclones upwind of the block can also contribute to the onset and maintenance of a stationary block and therefore a heatwave at the surface (Pfahl et al. 2015; Steinfeld and Pfahl 2019; Zschenderlein et al. 2020; Schumacher et al. 2022a; Kautz et al. 2022; Neal et al. 2022; Baier et al. 2023).

In addition to specific atmospheric mechanisms, the crucial role of land surface moisture during extreme heat events has been increasingly recognized in the recent years (Miralles et al. 2019). Whenever soil moisture limits latent heat fluxes, more energy is available for sensible heating, inducing an increase of near-surface air temperatures, especially in so-called transitional regime where soil moisture - temperature interactions increases summer temperature variability (Alexander 2011; Seneviratne et al. 2010; Mueller and Seneviratne 2012; Perkins 2015; Gevaert et al. 2018). The interaction of favorable atmospheric conditions with low soil moisture essentially leads to the self-amplification of heatwave events, with a positive feedback loop between high temperatures and low soil moisture via a high evaporative demand in the boundary layer (Fischer et al. 2007; Teuling et al. 2013; Vogel et al. 2017, 2018; Miralles et al. 2019). Studies have also shown that the influence of soil moisture on heatwaves can be non-local through heat advections from dry neighboring regions (Miralles et al. 2014; Schumacher et al. 2019, 2022b), especially dry conditions in the south of Europe spreading in the north of Europe through atmospheric transport of anomalously warm and dry air (Vautard et al. 2007; Zampieri et al. 2009; Quesada et al. 2012). There are nonetheless still substantial discrepancies between models and observations on the partitioning of surface energy into latent and sensible heat fluxes (Hirschi et al. 2011; Dirmeyer

et al. 2018) which constitutes a large source of uncertainties in current and future climate model projections (Vogel et al. 2018; Luo et al. 2022; Al-Yaari et al. 2023).

Even though the understanding of heatwaves physical mechanisms has improved, the mechanisms by which the most extreme events, either at short or long time scales, are reached are still unclear and especially in what measure they differ from those to reach moderate events (Barriopedro et al. 2023; Domeisen et al. 2023). The study of extreme events in general and heatwaves in particular is indeed impaired by a simple but difficult problem: extreme events are rare and therefore under-sampled both in observations and in model simulations. For example, by definition a 100-year run of a climate model will produce — on average — only one heatwave event with a return time of 100 years. As a consequence, a large part of what is known about the dynamics of heatwaves is based on the analysis of case studies of events that have occurred in the last 20 years in the mid-latitudes (e.g. Russo et al. (2015); Hauser et al. (2016); Hoy et al. (2017); Sánchez-Benítez et al. (2018); Wehrli et al. (2019, 2020); Sánchez-Benítez et al. (2022); Tuel et al. (2022); Serrano-Notivoli et al. (2023); Tripathy and Mishra (2023); Rousi et al. (2023)), especially very intense events like 2003 in Western Europe or 2010 in Western Russia. To obtain climatological properties of very extreme events and answer physical questions about their mechanisms, one needs to rely on long and costly runs in order to obtain large sample sizes.

Recently, methods known as *rare events algorithms* have been applied in climate science to simulate more extremes than the models would spontaneously do (Wouters and Bouchet 2016; Plotkin et al. 2019; Webber et al. 2019; Yiou and Jézéquel 2020; Gessner et al. 2021; Finkel and O'Gorman 2024). Most of these methods rely on a form of *importance sampling*, which in this context consists in favoring certain types of trajectories of climate models that lead to extreme events of interest. The seminal paper of Ragone et al. (2018) employed the Giardina-Kurchan-Tailleur-Lecomte (GKTL) (Giardina' et al. 2006; Giardina et al. 2011) algorithm to study extremes of summer temperatures with the intermediate complexity climate model PlaSim (Fraedrich et al. 2005), showing a large efficiency gain in the simulation of extremely hot summers. This algorithm is designed to sample large deviations of a time-averaged functions of interest, hence it simulates extreme seasons. The algorithm was subsequently successfully applied to a state-of-the-art climate model to sample heat extremes (Ragone and Bouchet 2021), and also in an intermediate complexity

model to sample seasonal precipitation extremes (Wouters et al. 2023) and extreme negative summer pan-Arctic sea ice anomalies (Sauer et al. 2024).

In this paper, we address the question of the physical mechanisms to reach extremely rare hot summers in Western Europe in a pre-industrial climate. We apply the GKTL rare events algorithms to the IPSL-CM6A-LR climate model (Boucher et al. 2020) in its atmosphere - land surfaces configuration to sample a large number of centennial hot summers in Western Europe. We investigate the physical mechanisms to reach such low probability events, by favoring very hot, very anticyclonic and very dry conditions over the summer.

This paper is organized as follows. In section 2 we first present the simulations made with the rare events algorithm. Section 3 details the results obtained. The section begins by a statistical description of the extreme summers, then investigate the mean spatial structures of the centennial events sampled by the algorithm and finally details their atmospheric dynamics. Finally, section 4 summarizes and discusses the results obtained.

2. Methods

a. Simulations with a rare events algorithm

In this paper we use the so-called Giardina-Kurchan-Tailleur-Lecomte (GKTL) rare events algorithm to simulate extreme summers in Europe with a climate model (Giardina et al. 2011). As first formulated by Giardina' et al. (2006), the algorithm aims to compute large deviation functions (Touchette 2009a) and is thus adapted to simulate long-lasting events. The name 'GKTL' was employed by Ragone et al. (2018) who were the first to implement it in a climate context and we follow this denomination. The idea of the algorithm is to run an ensemble simulation of a climate model with N members. Contrary to a regular ensemble, members of the rare events algorithm ensemble are not independent from one another and interact at regular resampling times. Between two resampling times t_i and t_{i+1}, we run the N ensemble members $(X_n)_{1\leq n\leq N}$ in parallel. The integration is stopped at t_{i+1} and we compute a score function $\theta(X_n(t)) \in \mathbb{R}$, that allows to assign a weight $w_n(i)$ to each member n:

$$w_n(i) = \frac{e^{k\int_{t_i}^{t_{i+1}} \theta(X_n(t))dt}}{\frac{1}{N}\sum_{i=1}^{N} e^{k\int_{t_i}^{t_{i+1}} \theta(X_n(t))dt}} \tag{1}$$

where k is a control parameter that determines the strength of the selection, and the integrals are computed as empirical sums. If k is positive (negative), the weight is high for high (low) values of the time-averaged score function θ. Then each member is either killed or cloned depending on the value of its weight: the strongest the weight — and therefore the integral of the score function θ — the highest the number of descendants of each member. If the weight of one member is too small, i.e. if it performs poorly according to the score function θ, there is a chance that the member is killed and therefore that it disappears from the ensemble. The number of descendants for each member is stochastic, but on average it is proportional to its weight (Ragone et al. 2018). After several resampling times, the distribution of $\int \theta(X(t))dt$ in the ensemble is biased towards the right ($k > 0$) or left ($k < 0$) tail of the climatological distribution. The higher the absolute value of k, the more members with an extreme value of the time-averaged score function will be favored and therefore the farther they will go towards the tail of the distribution. A schematic of the algorithm is presented in Figure 1.

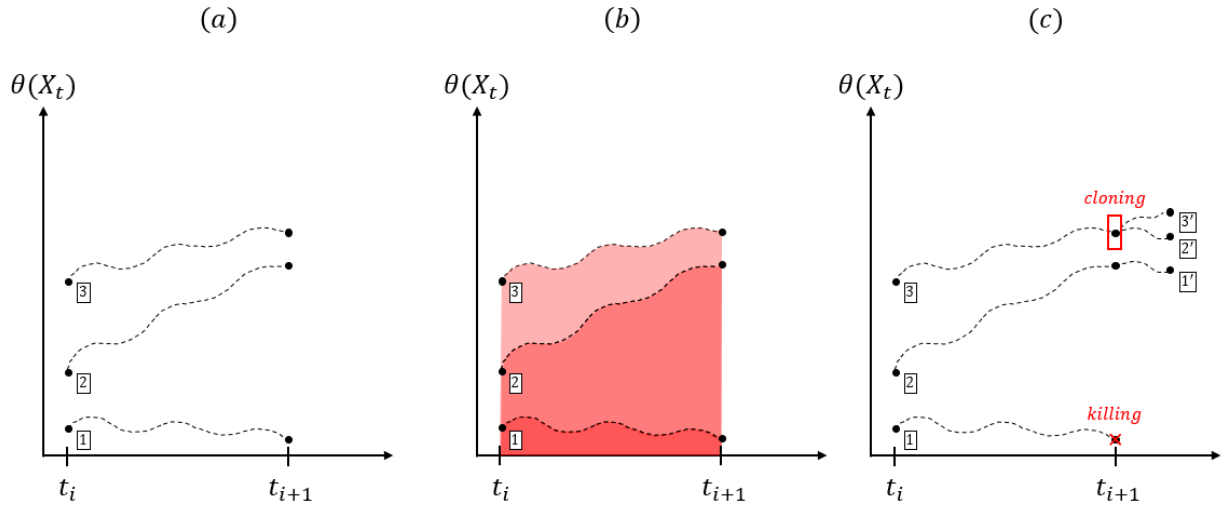


FIG. 1. Schematic representation of the Giardina-Kurchan-Tailleur-Lecomte rare events algorithm. (a) Simulation step where the different members evolve from time t_i to time t_{i+1}. (b) Reweighting step. The red areas under the different curves represent the arguments of the exponential weights updated at this step. (c) Killing and cloning step when the members with the lowest weights are killed (member 1 here) and members with the highest weights are cloned (member 3 here).

Here we employ the same resampling procedure as in Ragone et al. (2018) at each time t_{i+1}. It is important to note that the total number of members N stays constant during the full process, therefore at each resampling time the number of killed members is equal to the number of cloned members. In the simulations presented here the resampling is done every 5 days and a random perturbation on the potential temperature field at all levels is added to let the cloned members diverge from their parent, as the climate model is fully deterministic. The amplitude of the perturbation is random and is at maximum of the order of 0.01% of the initial value. The typical precision obtained with climate models is lower than this value (1-0.1%), therefore in the following we consider that all obtained simulations are physically consistent, i.e. that they could have been reached by the model and are not an artifact of the random perturbations added by the rare events algorithm.

The choice of the score function is crucial because it determines what kind of extremes are selected. By construction the GKTL algorithm does not favor members that will experience an extreme event, but selects retrospectively members which have experienced an extreme event in the last resampling period. Nonetheless, by the choice of the score function and the physics of the system, favoring members which have experienced an extreme may actually favor the appearance of other similar extremes in the future. This is especially the case for heatwaves and the associated feedback loop with soil moisture (Seneviratne et al. 2010). In the simulations presented here we choose four different score functions:

1. 2-m air temperature (T2M) at grid point 1 (49.5°N, 2.5°E) with control parameter k = 0.04/°C/day,
2. geopotential height at 500hPa (Z500) at grid point 1 with control parameter $k = 0.0015$/m/day,
3. upper-level soil moisture at grid point 1 (SM1) with control parameter $k = -0.05$/kg.m^2/day,
4. upper-level soil moisture at grid point 2 (SM2) (49.5°N, 7.5°E) with control parameter k = -0.05/kg.m^2/day.

The rationale for using the first score function is straightforward: 2-m air temperature is the variable classically used to define heatwaves. We therefore aim to select longer and more intense heat events with the algorithm. The rationale for using the second and third score functions is to sample respectively atmospheric and surface conditions which are typically associated with

intense heat events: by maximizing the geopotential height at 500hPa (Z500), we try to sample more anticyclonic structure and by minimizing ($k < 0$) upper-level soil moisture we pre-condition dry soils which would favor more intense sensible heat fluxes and therefore heatwaves. The last score function is not, contrary to the others, computed at the same grid point but at a slightly longitudinally shifted ones which corresponds to the grid point with the strongest correlation with hot summers at the (49.5°N, 2.5°E) grid point as investigated in a previous work on a pre-industrial run of this model (Noyelle et al. 2024). The values of k were chosen according to the formula of Ragone and Bouchet (2020), with the objective to sample events between 2 and 3 standard deviations away from the climatological mean.

In the following, we refer to these simulations as the 'biased simulations', in the sense that they favor a kind of extremes. We call them according to the variable they optimize, respectively biased T2M, Z500, SM1 and SM2. For each score function, we run $N_s = 9$ independent simulations using $N = 100$ members each, for a total of 900 simulated summers for each score function. The simulations are independent in the sense that there is no inter-simulation interaction between the members. Because at each resampling time the number of descendants of each member is stochastic and the added perturbations are also stochastic, each simulation is different. There is a trade-off between the number of members per simulation and the number of simulations one can run (for a fixed computational cost). The number of members cannot be too low, or else the algorithm would have a high variance, but if it is too high then one needs to reduce the number of simulations made, which may also increase the variance. We chose the 9 simulations $\times 100$ members per simulation trade-off by trial and error. We additionally run a control ensemble of 900 independent members with the same starting conditions and a random perturbation for differentiation.

Contrary to previous applications (Ragone et al. 2018; Ragone and Bouchet 2021) we choose grid point score functions rather than extended spatial averages. By doing so we seek to impose a minimum amount of constraint on the full system and recover the spontaneous temporal and spatial scales of heat events in the grid point studied. When using extended spatial averages, one bears the risk of mixing up the dynamics of heat events occurring on one side or on another side of the spatial domain. As a result, composite maps could combine several dynamics that do not occur at the same time for individual events.

Even though the GKTL algorithm favors a certain kind of extremes, it allows to compute the unbiased — i.e. climatological — expectations for these events. For any observable ψ (i.e. a smooth enough function), the climatological expectation of ψ using the rare events algorithm ensemble is (Ragone et al. 2018):

$$\mathbb{E}_\mu[\psi] = \frac{1}{N}\sum_{i=1}^{N} w_n^{-1}\psi(X_n(t)) \tag{2}$$

where $\mathbb{E}_\mu$ is the expectation with respect to the climatological probability distribution μ, $w_n = \Pi_i w_n(i)$ and $X_n(t)$ is the state vector of member n at time t. It is important to note that even if the weights w_n are computed according to the score function θ, the formula is valid for the expectation of *any* other observable ψ.

In particular, if one wants to know the climatological properties of events for which the score function is above a certain level a, one computes:

$$\mathbb{E}_\mu[\psi \mid \theta(X(t)) \geq a] = \frac{\mathbb{E}_\mu[\psi \times \mathbb{1}(\theta(X(t)) \geq a)]}{\mathbb{E}_\mu[\mathbb{1}(\theta(X(t)) \geq a)]}. \tag{3}$$

In plain words, one computes the observable only for members that have reached the level a and divide by the probability to reach the level a ($\mathbb{E}_\mu[\mathbb{1}(\theta(X(t)) \geq a)] = \mathbb{P}_\mu[\theta(X(t)) \geq a]$). These expectations are unprecisely estimated in a regular ensemble simulation as long as a corresponds to a high quantile of the distribution of θ because of an under-sampling of the very extreme events. With the rare events algorithm, there are more members that reach high quantiles, but they come with a certain weight. Therefore the expectations in this equation are computed using Eq. (2), i.e. with the weights w_n. Note that when N_s rare events algorithm simulations are run, each simulation gives an estimation for $\mathbb{E}_\mu[\psi \mid \theta(X(t)) \geq a]$ using Eq. (2). They are then simply averaged to give a final estimation of $\mathbb{E}_\mu[\psi \mid \theta(X(t)) \geq a]$.

It is not clear currently in the literature how to obtain statistical significance estimations with the outputs of the rare events algorithm simulations — especially how to combine the results from the different simulations and take into account the fact that some members are very correlated. In the following we therefore mainly show results normalized by removing the average and dividing by the standard deviation obtained on the control simulation, and we take a high anomaly threshold to interpret the results (> 0.5 std in absolute value). For comparison, if one considers summer

averages, with 900 members the significance at 5% with a Student t-test is obtained as long as the normalized anomaly is greater in absolute value than 0.07 std.

b. Configuration of the model

We use the IPSL-CM6A-LR model (Boucher et al. 2020) in its atmosphere - land surfaces configuration (LMDZOR) under pre-industrial forcing conditions. The model has a horizontal resolution of 2.5° in longitude and 1.27° in latitude and 79 vertical layers. We start by selecting a year in a 2000-y control run of the fully coupled model (i.e. in particular with the ocean). This year is selected randomly and does not present any particular feature. A 4-months spin-up simulation (January to April) of the uncoupled model with oceanic and ice conditions of the initial coupled simulation is then run to let the atmosphere adapt to the uncoupling with the oceans and to provide initial conditions for the rare events simulation in summer. Random perturbations are added at the beginning of May to create a 100-member starting ensemble. The perturbations are added in May to avoid having to simulate the 100 members from January to April while still letting enough time for the different members to separate until the beginning of the rare events simulation in June. All the simulations begin on the 1st of June from those same 100 members and end on the 28th of August (90 days). To create the ensemble, a similar perturbation is added to all potential temperature levels as explained previously. Except specified explicitly, the variables used in the following are daily means.

c. Atmospheric dynamics metrics

We seek to characterize the atmospheric dynamics mechanisms leading to extremely hot summers simulated by the model. To do so we employ metrics used in the meteorological literature to diagnose the transfer of energy from eddies to the mean flow (**E**-vector), Rossby waves dynamics via a space-time spectral analysis and an amplitude-phase decomposition, and finally an algorithm to detect cut-off lows presence. We refer the reader to appendix A for the detail of the computation of these metrics with the outputs of our simulations. To diagnose wave-like behaviors we follow the common practice in the literature (e.g. Jiménez-Esteve et al. (2022)) which considers high-level tropospheric winds between 300 and 200hPa. Here we employ the 200hPa level as it is a standard output of the model.

3. Results

a. Grid point statistics

This section presents statistics for the grid points where the score functions are optimized. Figure 2 shows the empirical probability density functions (PDFs) of summer averaged 2-m air temperature (Fig. 2a), geopotential height at 500hPa (Fig. 2b), upper-level soil moisture (Fig. 2c) at grid point 1 and upper-level soil moisture at grid point 2 (Fig. 2d). When compared to the control PDFs, the PDFs of the score functions that are optimized by the rare events algorithm — i.e. for example T2M for the biased T2M simulation — demonstrate the efficient sampling of the tail of the summer averages distribution. For all score functions the rare events algorithm sample extreme summers that are unprecedented in the 900-members control simulation. The choice of score functions closely related to the dynamics of heatwaves results in the sampling of extreme summers also for observables for which the algorithm does not optimize — i.e. for example Z500 for the biased T2M simulation. Figure 2b shows that the far tail of the extreme Z500 distribution is better sampled when maximizing temperature rather than directly Z500, at least in grid point 1. For the interpretation of Figure 2c and d, one should note that there is a hard coded limit of 8.14 kg/m^2 in the model for the upper-level soil moisture for the grid points considered. Therefore, summer averages close to this value correspond to extremely dry summers, i.e., almost the driest that are possible in the model. Finally, Figure 2a and b show that minimizing soil moisture at grid point 2, i.e. remotely with respect to where the T2M and Z500 variables are maximized, is more efficient for sampling more intense summer averaged T2M and Z500 than minimizing soil moisture at grid point 1. We come back below to this shift between soil moisture and temperature anomalies for extreme summers.

This question of the substructure (Röthlisberger et al. 2020) of extremely hot summers is explored in the simulations by counting the number and mean length of heatwaves during the summer. A heatwave is defined classically as a continuous period of time of at least 3 days for which the daily mean temperature is above the 90% daily climatological quantile of the temperature distribution (Perkins and Alexander 2013). The latter is computed using the control simulation. All such events are computed at grid point 1. Figure 3a shows that for biased T2M simulations the mean number of heatwaves in the summer goes from around 1.3 in the control to around 4.8 in the biased

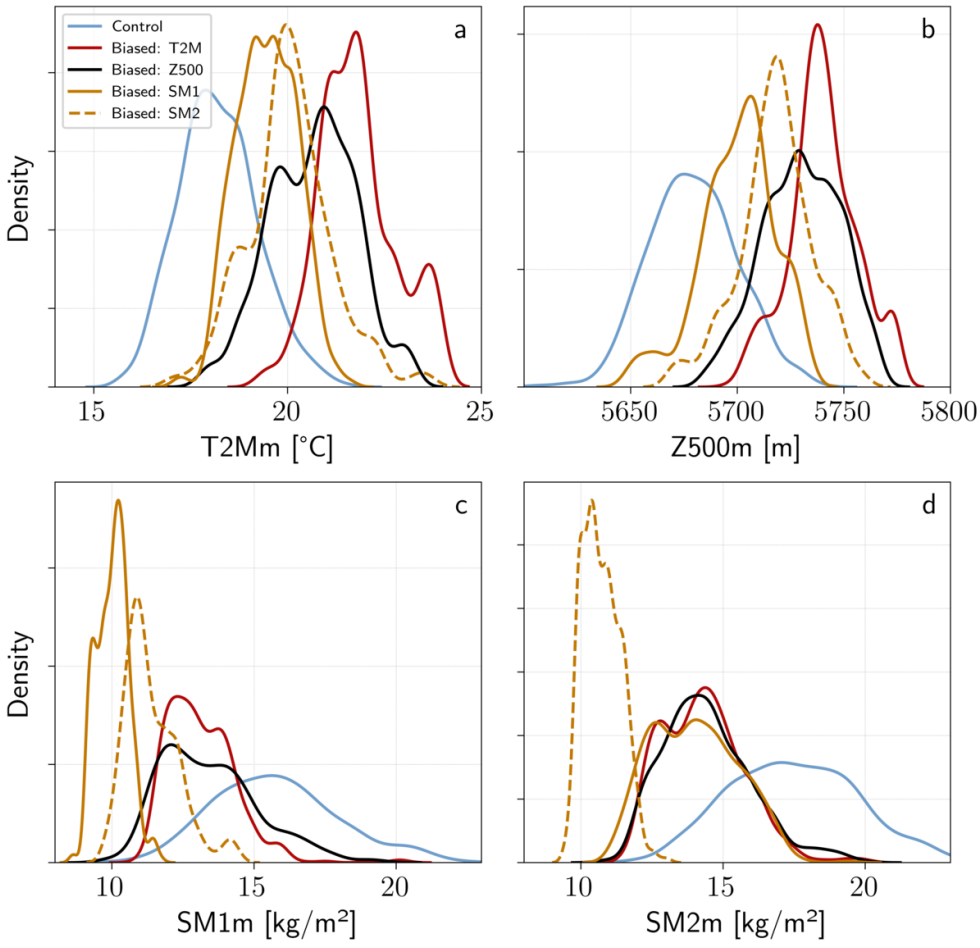


FIG. 2. Empirical probability density functions (PDFs) of summer averaged grid point observables. PDFs of the summer averaged (a) 2-m air temperature, (b) geopotential height at 500hPa, (c) upper-level soil moisture at grid point 1 and (d) upper-level soil moisture at grid point 2. The PDFs are obtained by a kernel smoothing of the empirical histograms.

simulation, while their mean duration goes from 4.6 to 7.3 days. The algorithm therefore selects both more heatwave events and longer heatwave events. When maximizing the geopotential height (Figure 3b) the mean duration increases to 5.7 days while the number of events increases to less than 3.7, illustrating the key role of high geopotential anomalies to reach high temperatures. When minimizing the local soil moisture (Figure 3c) on the other hand, the mean duration of heatwave events does not change much with respect to the control (4.8 vs 4.6), but the number of events increases to around 2.3 per summer. When minimizing soil moisture at grid point 2 however,

there is an increase in both the duration (6.2 days) and the number of heatwaves (3.0 per summer). Neither the biased Z500 nor the SM1 and SM2 biased simulations are enough to reach metrics as high as the ones of the biased T2M simulation. This results is strongly in favor of the need for combined non-local dry soils and high local geopotential height to obtain long and intense heatwaves leading to a hot summer.

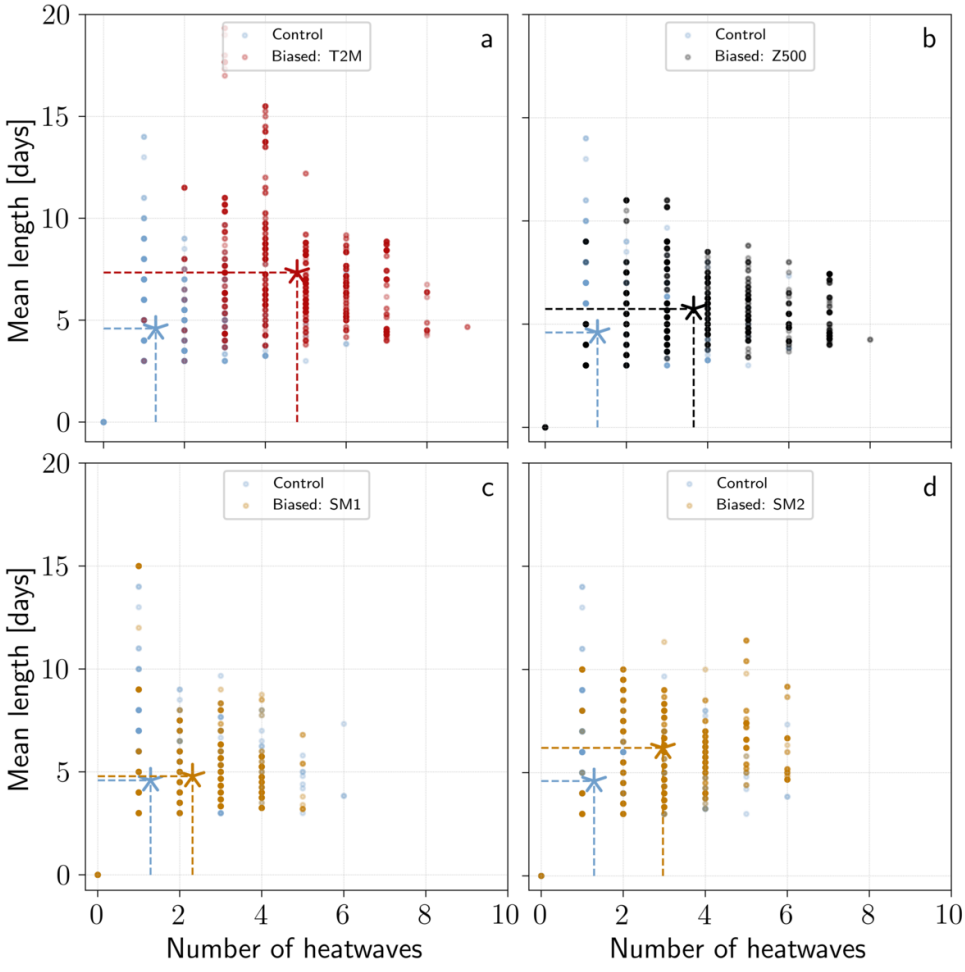


FIG. 3. Number and mean length of heatwaves per summer. All heatwaves are computed at grid point 1. For all plots a heatwave is a continuous period of time of at least 3 days when daily mean temperature is above the 90% climatological quantile. The stars show the ensemble mean.

As explained in section 2, the algorithm provides weights to compute climatological averages of observables of interest. In particular, this allows to recover the probabilities of the simulated extreme summers. Figure 4 shows the probabilities for the simulations run here, each using their

own score function as the observable of interest (i.e. using the observable on which the rare events algorithm is expected to be the most effective). All simulations sample very extreme summers, with probabilities ranging from 10^{-2} to 10^{-6}, i.e. return times between one hundred and one million years. This demonstrates a large efficiency gain compared to the control simulation which, with 900 members, can only sample precisely summers with a ≈ 100 years return time. The variations from one simulation to another of the estimated probability are nonetheless important: around one order of magnitude for example for the biased T2M simulation. Although the algorithm is asymptotically unbiased (Giardina' et al. 2006), one could wonder whether the low probabilities sampled here are really correct. The only way to prove that these probabilities are correct would be to run a much larger number of control members, until even very low probability events happen — but this entails a corresponding high computational cost. Although this is not a formal proof, one can see in Figure 4 that the two results seem at least coherent on the region where they overlap.

Summer averages (i.e. 90 days averages) are considered here, therefore by the central limit theorem, one may expect that the associated distribution is Gaussian. However, the central limit theorem is valid only in the vicinity of the average (Touchette 2009b) and the tails of the distribution — i.e. the regions sampled here — must be described by large deviations theory (Gálfi et al. 2021). To illustrate the deviations from Gaussianity, the mean and standard deviation of a Gaussian distribution are fitted on the control run members and it is shown in dashed line in Figure 4 the associated extrapolation of the probability. Although this extrapolation is correct up to probabilities as low as 10^{-1}, it diverges for lower probabilities. Figure 4a for example shows that an extreme summer with average temperature around 23°C is 10 times more likely than predicted by the Gaussian approximation. The Gaussian approximation is nevertheless on the range of uncertainty of the algorithm when running $N_s = 9$ simulations. For the other simulations (Fig. 4b, c and d), this is the contrary: extreme summers are less likely than predicted by the Gaussian. This is especially clear for soil moisture, which is, contrary to the Gaussian distribution, bounded downwards in the model as explained above. The uncertainty ranges for these simulations also make clear that the tail of the summer average quantities are not Gaussian.

To explore the link between the summer averaged quantities in the different simulations, a scatter plot of their cross distributions is shown in Figure 5. Figure 5b and d show the strong link between high summer-averaged geopotential heights and both high surface temperature and low

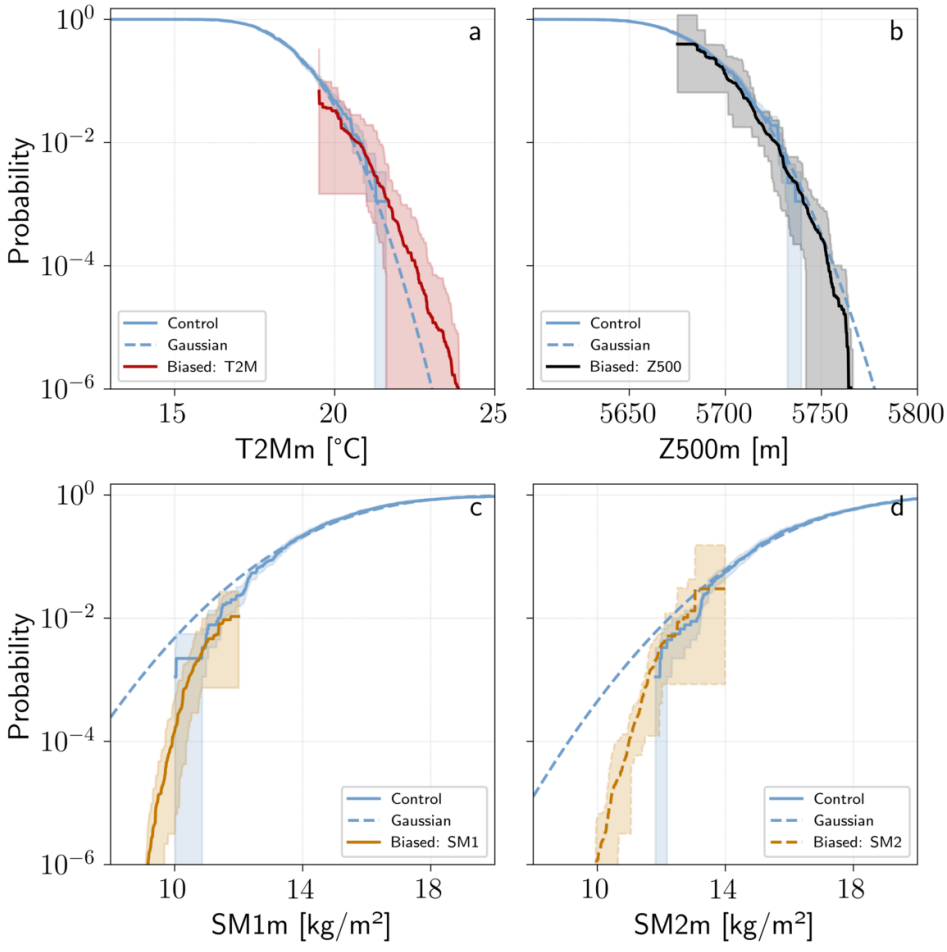


FIG. 4. Probabilities of an extreme summer according to the naive estimator (control simulation) and the rare event estimator. (a) Probability of the summer averaged 2-m surface temperature and (b) probability of the summer averaged geopotential height at 500hPa to be above a certain threshold at grid point 1. (c) Probability of the summer averaged upper-level soil moisture at grid point 1 and (d) probability of the summer averaged upper-level soil moisture at grid point 2 to be below a certain threshold. The blue dashed lines represent the corresponding probabilities for a Gaussian distribution fitted on the control simulation by the method of moments. The shadings for the control simulation show the 5–95% quantiles of the estimated probability obtained using bootstrap on the 900 members. For the biased simulations the shadings show the minimum and maximum of the estimated probability over the $N_s = 9$ simulations.

soil moisture (although the correlation is not as strong in the second case). This link is conserved in the biased simulations, which, when maximizing for either temperature or geopotential height, sample the tail of the joint PDFs of these observables. As illustrated by Figure 5a and c, the

link between soil moisture and temperature is however not as obvious. Although in the control simulation there is a small correlation between those two quantities, the correlation is lost in the biased simulations. Very hot summers sampled by the rare events algorithm are dryer than the average (see also Figure 2c and d), but there is a wide range of different possible temperatures for the same level of dryness. A similar phenomenon is observed for very dry summers. Contrary to geopotential height and surface temperature in Figure 5b, the algorithm does not sample the combined tail of the PDFs (except for a few members in the biased SM2 simulation): there are no both very hot and very dry summers. This suggests that the well-known correlation between these two variables may not be present in the far tail of their joint distribution. One hypothesis to explain such a surprising result could be that when one member reaches very high temperatures at grid point 1, this tends to trigger convection and therefore rain (Zhang and Boos 2023), which is not favored when one wants to minimize soil moisture. Thus, dry summers may have to be hot but not too hot.

b. Summer composites spatial structures

In this section we investigate the summer composites spatial structures associated to the extreme summers sampled by the algorithm. We consider here centennial-like events, i.e. extreme summers that have a probability below 10^{-2} to occur. Contrary to the preceding section, in this section we therefore show climatologically relevant averages, i.e. conditional on reaching a threshold corresponding to centennial events (see methods).

The normalized anomalies of upper-level meridional wind (200hPa) and 2-m air temperature are shown in Figure 6. The normalization is computed by removing the climatological mean and dividing by the climatological standard deviation estimated on the control simulation at each grid point. Figure 6a shows the results for the biased T2M simulation. In an approximately 1000km-diameter circle centered around grid point 1, the summer averaged temperature is 3 standard deviations above the climatology. However, this strong anomaly is concentrated close to grid point 1: although there are corresponding cold and warm anomalies downstream, those are much less intense than in Western Europe. The upper-level wind pattern is also concentrated close to Western Europe, with a large and very anomalous meander around grid point 1. The meander is tilted eastward over Eastern Europe, suggesting a recurrence of wave-breaking phenomena. Similarly,

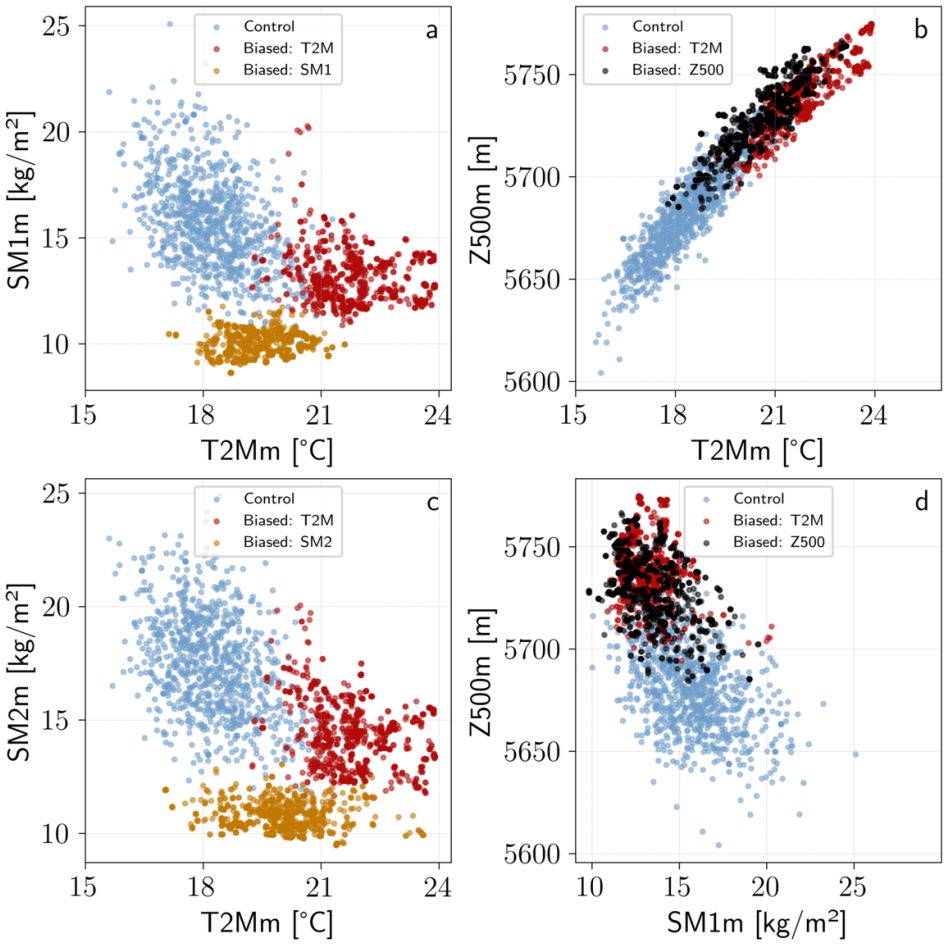


FIG. 5. Scatter plot of summer averaged grid point observables. (a) Summer averaged 2-m air temperature vs. upper-level soil moisture at grid point 1. (b) Summer averaged 2-m air temperature vs. geopotential height at 500hPa at grid point 1. (c) Summer averaged 2-m air temperature vs. upper-level soil moisture at grid point 2. (d) Summer averaged upper-level soil moisture vs. geopotential height at 500hPa at grid point 1.

the positive anomaly above the East Mediterranean region is the result of a northward shift of the subtropical jet (not shown). The anomalies upstream are small and there are no hemispheric pattern. The size of most anomalies is of 1 to 2 synoptic wavelengths. These observations extend to the other simulations, with the exception that the anomalies of both V200 and T2M are smaller when minimizing soil moisture either at grid point 1 or 2. Finally, there is a longitudinal shift of anomalies between the four simulations. For example, contrary to Figure 6a, in Figure 6c only the western facade of France and the Iberian peninsula present positive anomalies of T2M. There is

a corresponding shift in anomalies of V200. It should be also mentioned that there are almost no significant anomalies in the tropics and not at all in the southern hemisphere for the fields displayed (not shown), which suggests a climatological atmospheric dynamics leading to extreme summers mostly confined to the mid-latitudes.

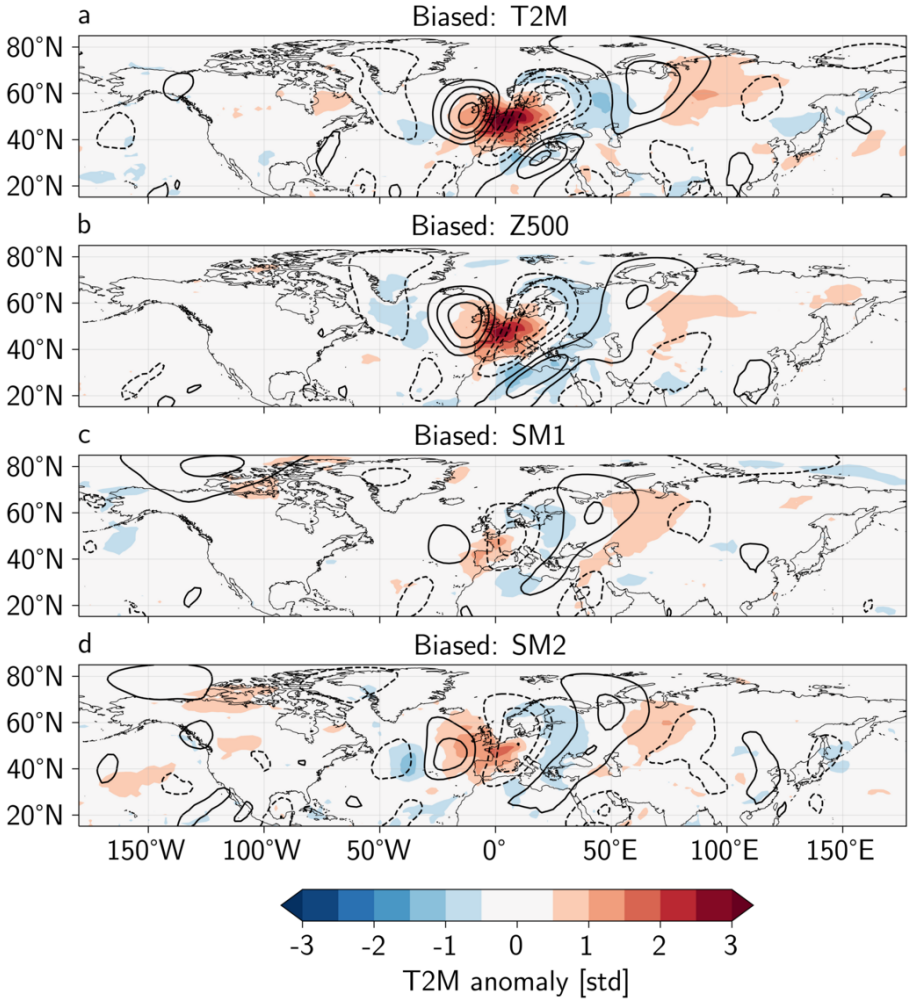


FIG. 6. Summer averaged normalized anomalies of 2-m air temperature (colors) and meridional wind at 200hPa (contours) for centennial events. For both fields, normalized anomalies are computed by removing at each grid point the mean and dividing by the standard deviation computed on the control simulation. The contours are drawn every 0.5 standard deviation starting at +/- 0.5. Plain (dashed) lines represent positive (negative) values of the standard deviation.

Figure 7 shows the normalized anomalies of geopotential height at 500hPa and upper-level soil moisture. All simulations have a large anticyclonic anomaly above Western Europe, but the center of the anomalies is slightly shifted from one simulation to another: for Figure 7b, corresponding to the simulation maximizing Z500, as expected the anomaly is maximum above grid point 1, whereas it is shifted to the east for Figure 7a, to the west for Figure 7d and to the south-west for Figure 7c. Again, the Z500 anomalies are maximum above Western Europe and do not extend around the entire Northern Hemisphere. It should be also noted for all regions, the small negative Z500 anomaly south-east of the anticyclonic region, which corresponds to the isolation of low pressure systems by anticyclonic wave breaking above Eastern Europe. The strongest soil moisture anomalies in Figure 7 are also restricted to the European region and present an eastward shift with respect to the region of maximum anticyclonic anomaly. For Figure 7a in particular, the anomaly of soil moisture at grid point 1 — i.e. where the temperature is maximized — is not as strong as the soil moisture anomalies encompassing Germany and Poland regions.

One hypothesis to explain this shifted pattern is related to the advection of dry and hot air at the core of the heatwave regions. Previous studies (Pfahl et al. 2015; Steinfeld and Pfahl 2019; Zschenderlein et al. 2020) computing Lagrangian backward trajectories in reanalysis and climate model outputs have shown that most of the air parcels during heatwave events in Western Europe come from either the heatwave region itself or its immediate east. This would plead in favor of the shifted low soil moisture pattern to be a precursor of intense heat events at grid point 1 in so far as low soil moisture promote high sensible heat fluxes and diabatic warming. It is however difficult to know whether this pattern is a cause or a result of heat events in our simulations.

An idea of the origin of air parcel in a Eulerian framework can be given by analyzing the vertical structure of the anomalies of geopotential height and sea-level pressure, even though we did not run a Lagrangian particle tracing analysis which would be necessary to validate this result. Figure B1 in appendix shows the anomalies of sea-level pressure (SLP), geopotential height at 850hPa (Z850) and geopotential height at 200hPa (Z200). All panels illustrate the westward tilted vertical structure of the summer anticyclone, with for example for Figure B1a the maxima of SLP anomalies situated above Poland while the maxima of Z200 situated above Southern UK. This vertical structure suggests, at grid point 1, advection from the east on the lower layers of the atmosphere, while

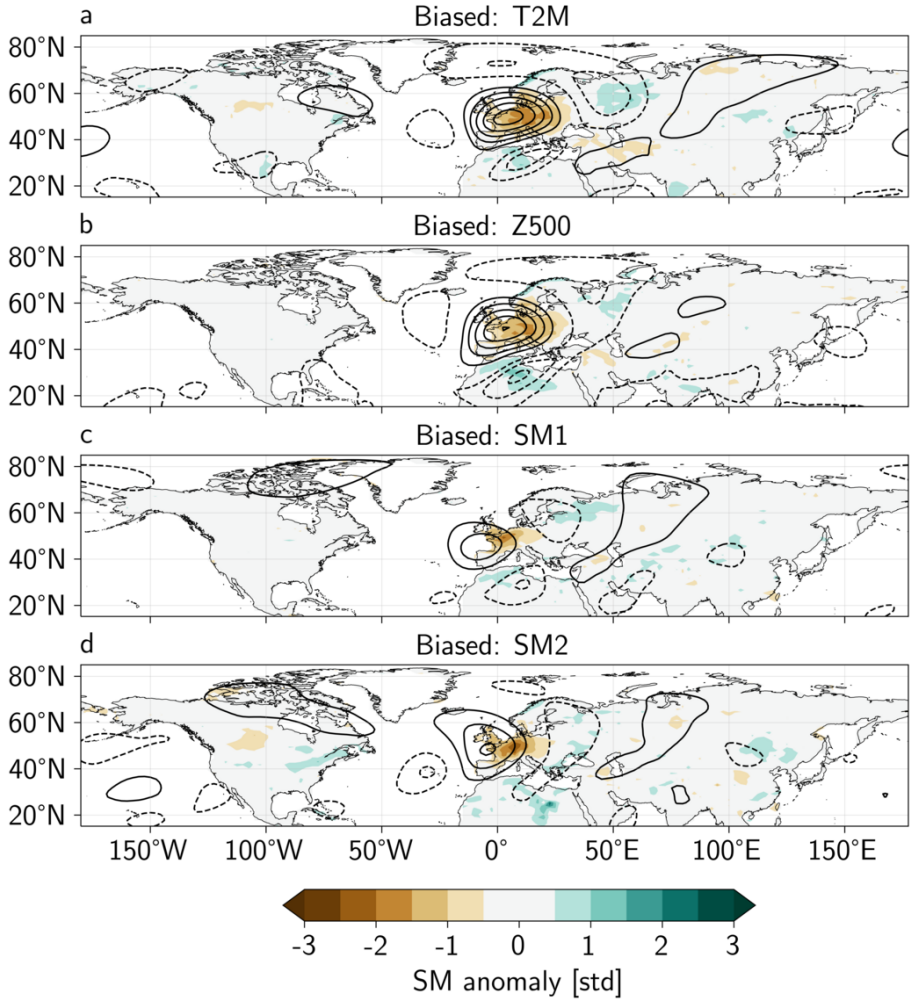


FIG. 7. Summer averaged normalized anomalies of upper-level soil moisture (colors) and geopotential height at 500hPa (contours) for centennial events. For both fields, normalized anomalies are computed by removing at each grid point the mean and dividing by the standard deviation computed on the control simulation. The contours are drawn every 0.5 standard deviation starting at +/- 0.5. Plain (dashed) lines represent positive (negative) values of the standard deviation.

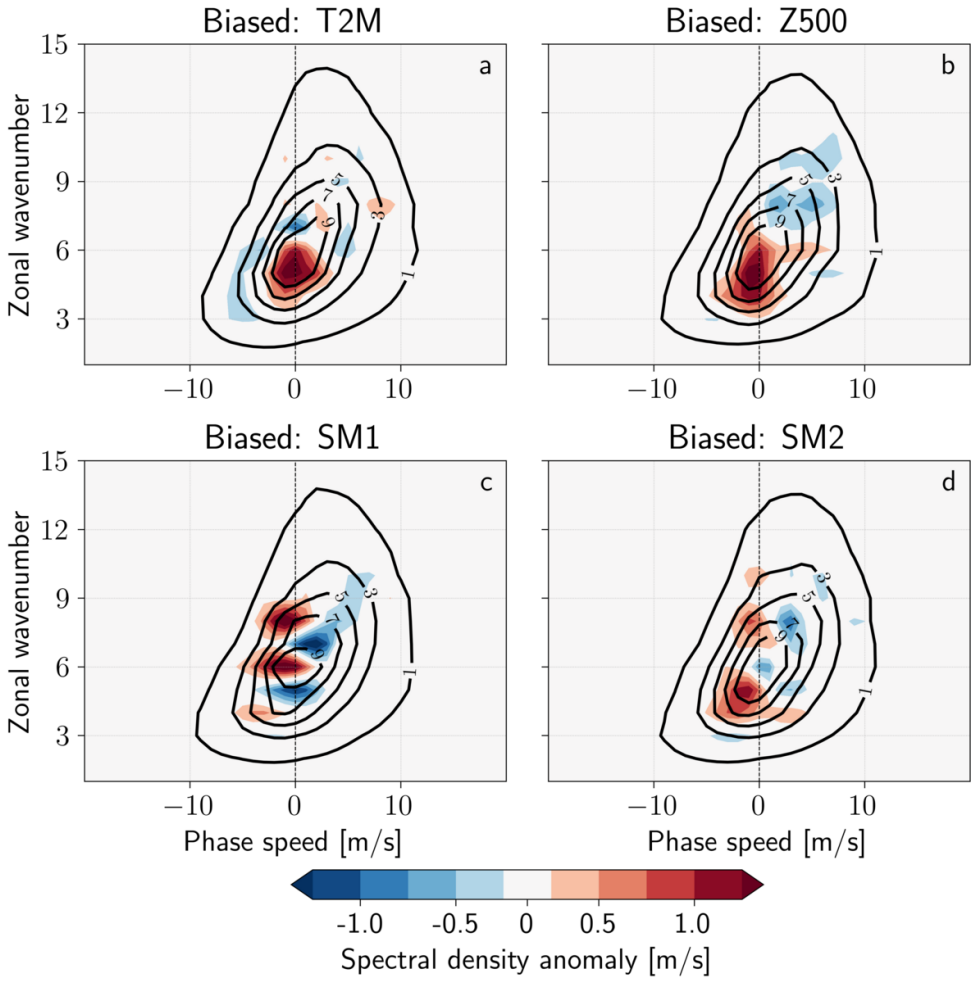
from the south-west on the upper-levels. This reveals a non-barotropic vertical structure of the anticyclone leading to extreme temperature events.

c. Atmospheric dynamics

In this section we investigate the atmospheric dynamics associated with the extreme summers simulated. Figure 8 shows the wavenumber-phase speed spectra at 200hPa obtained for centennial events and the difference with the climatological spectrum. The spectra are different from one biased simulation to another. The biased T2M and Z500 simulations are similar in their increase in stationary and/or westward moving waves with low zonal wavenumbers (4 to 6), highlighting the persistence of anticyclonic blocking patterns. On the other hand, the biased Z500 simulation is associated with a decrease of fast-moving high zonal wavenumbers waves which is not the case for the biased T2M simulation. The results for soil moisture minimizing simulations is more surprising: for both of them there is an increase in stationary/westward-moving waves with high zonal wavenumbers (8 for both, 6 for SM1 and 5 for SM2) and a decrease of eastward moving waves with similar zonal wave numbers. The differences with the climatological spectrum are nevertheless small (around 10%) and the raw spectra are similar from one simulation to another.

We investigate the properties of slow-moving quasi-stationary waves with an amplitude-phase histogram. The results are presented in Figure 9 for meridional wind speed anomalies at 200hPa for centennial events. Simulations with an optimization of the score function do not differ strongly in the quantity of energy distributed in the different zonal wave-numbers: they have approximately the same aggregated spectrum for each of them (not shown). However, they differ strongly in the phase of the different wave numbers. Figure 9b2 for example shows a 2-3 times increase in the frequency of waves with phase around $\pi/2$ compared to the control for wave number 4. There is a similar behavior for the other simulations but the preferred phase changes from a zonal wave-number to another. For example, for $k_\phi = 3-5$ biased simulations have a different preferred phase (which is coherent with the longitudinal shift that was described in Figure 6) but for $k = 6$ they all have the same preferred range of phases. There is no clear shift in neither the amplitude nor the phase of the stationary waves in the biased simulations, except maybe for wavenumbers $k_\phi = 4$ and $k_\phi = 5$ for the biased T2M simulation (Fig. B2b2 and b3).

Whether the excitation of certain modes and their phase-locking behavior is a result of an underlying physical amplification mechanism of the system (Petoukhov et al. 2013) or the simple consequence of the Fourier decomposition of anomalies leading to the patterns of Figure 6 is not clear though. In particular, with this procedure it is assumed that the stationary hemispheric Fourier



488 FIG. 8. Wavenumber-phase speed spectra of meridional wind anomalies at 200hPa for centennial events. Raw
489 spectrum in m/s (contours) and difference with the climatological spectrum (colors).

507 modes exist and are physically relevant. However, when the original field is reconstructed with for
508 example modes $k_\phi = 3$ to $k_\phi = 10$ (not shown), it is found that they — as expected — correspond
509 to a localized anomaly above Western Europe (i.e. the pattern of Figure 6). In other words, even
510 though it is mathematically correct to decompose the field on hemispheric scale patterns, most of
511 the dynamics is local and confined to the North Atlantic in practice.

515 Figure 10 indeed shows the composite Hovmöller plot of anomalies of meridional wind speed
516 at 200hPa averaged between 35°N and 65°N and normalized anomalies of 2-m air temperature
517 averaged between 45°N and 55°N for the heatwave events that occur in each simulations. In all
518 simulations, the atmospheric dynamics is similar: there is a Rossby wave packet of 2-synoptic

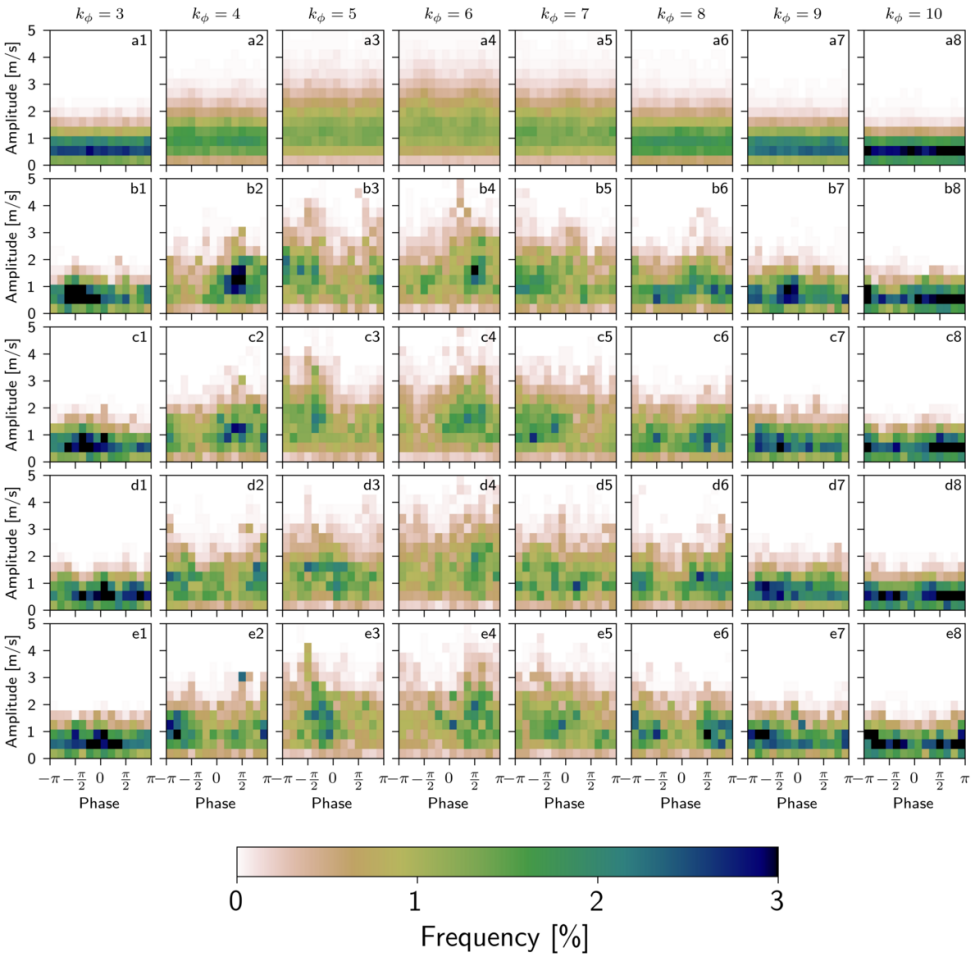
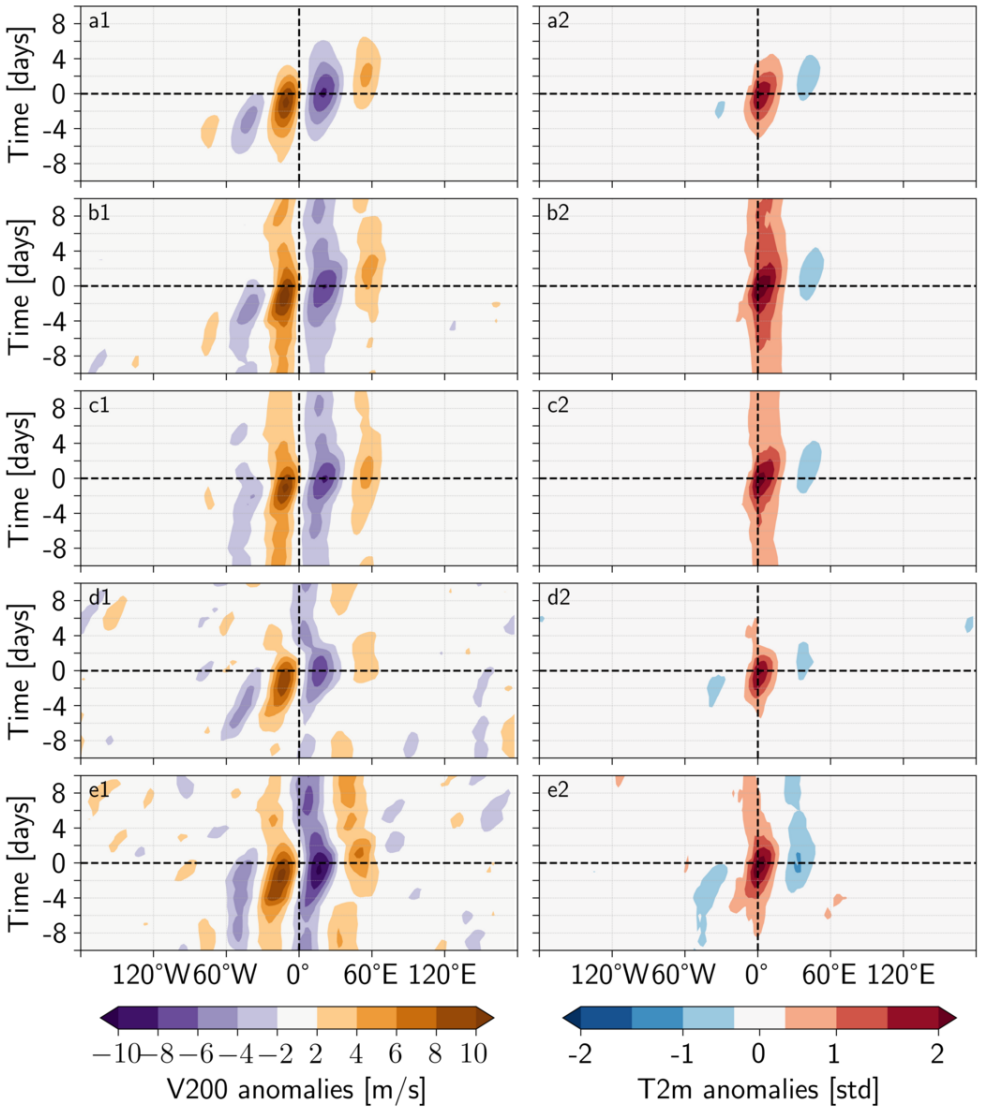


FIG. 9. Amplitude-phase histograms of meridional wind speed anomalies at 200hPa for centennial events. Wave numbers $k_\phi = 3$ to $k_\phi = 10$ in columns and biased simulations in rows: (a) control simulation, (b) biased T2M simulation, (c) biased Z500 simulation, (d) biased SM1 simulation and (e) biased biased SM2 simulation.

wavelength propagating to the east and amplifying locally in Western Europe. Because there are many heatwave events in biased simulations, it is even possible to see the recurrence of new Rossby wave packets at 6-8 days after the heatwave (Fig. 10bce1). The temperature anomalies are restricted to the region of amplification of the Rossby wave packet, i.e. where the score functions of the rare events algorithm are computed. For the biased T2M simulation (Fig. 10b2), the temperature anomalies are larger during the heatwave and extend both before and after, which may be linked to a near zero phase speed of the Rossby wave packet. A similar mechanism, although less intense, may be happening for both the biased Z500 and SM2 simulations (Fig. 10ce2).



527 FIG. 10. Composite Hovmöller plot of the atmospheric and surface dynamics during heatwave events. First
528 column: meridional wind speed at 200hPa anomalies averaged between 35°N and 65°N. Second column: 2-m air
529 surface temperature normalized anomalies averaged between 45°N and 55°N. The time is expressed relative to
530 the hottest day of each heatwave event. Rows: (a) control simulation, (b) biased T2M simulation, (c) biased Z500
531 simulation, (d) biased SM1 simulation and (e) biased biased SM2 simulation. Number of events: (a) $n = 1141$,
532 (b) $n = 4307$, (c) $n = 3290$, (d) $n = 2080$ and (e) $n = 2670$.

533 When looking at daily maps of some of the heatwave events in the simulations, we noticed a
534 frequent occurrence of cut-off lows close to the Iberian peninsula. We investigate further the role
535 of these by computing their frequency and frequency anomaly over the Euro-Atlantic sector for the

biased simulations (Figure 11). All biased simulations have anomalously high frequency of cut-off lows over the Atlantic. Biasing for 2-m air temperature at grid point 1 lead for example to a 2-3 times increase in the frequency of cut-off lows west of the Iberian peninsula. For Z500 and SM1 biased simulations, the frequency anomaly is not as high and is concentrated on the center of the North Atlantic. The biased SM2 simulation has the strongest anomaly, with a 4-5 times increase of cut-off lows frequency with respect to the climatology around the 30°W-40°N grid point.

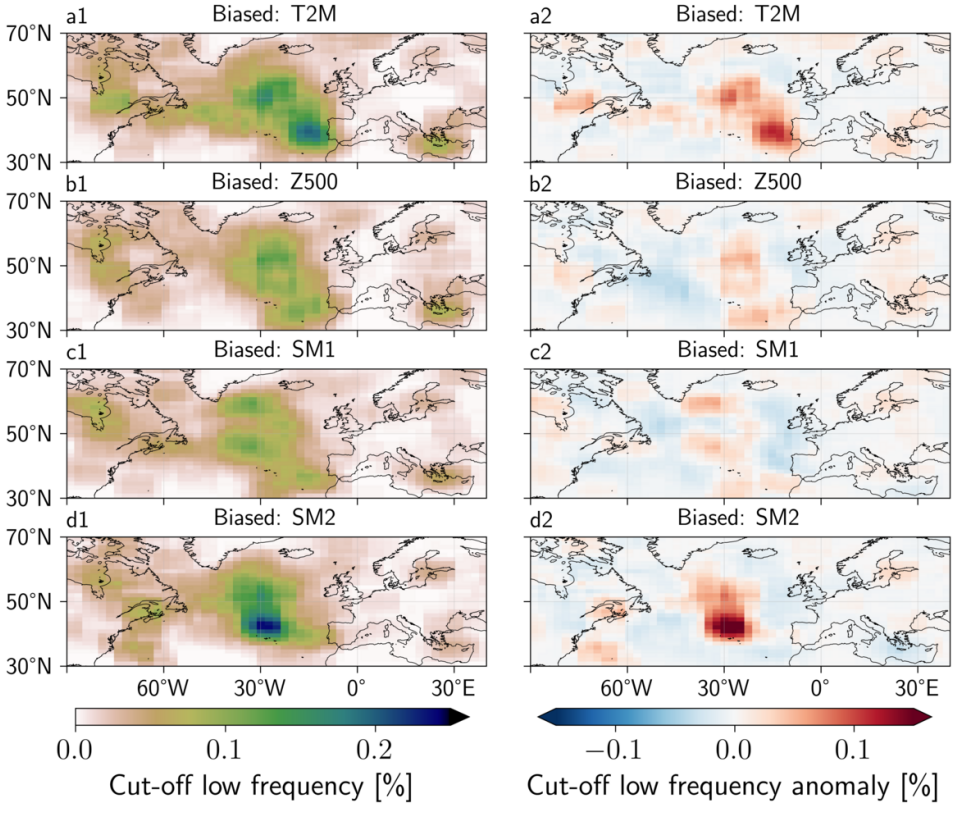


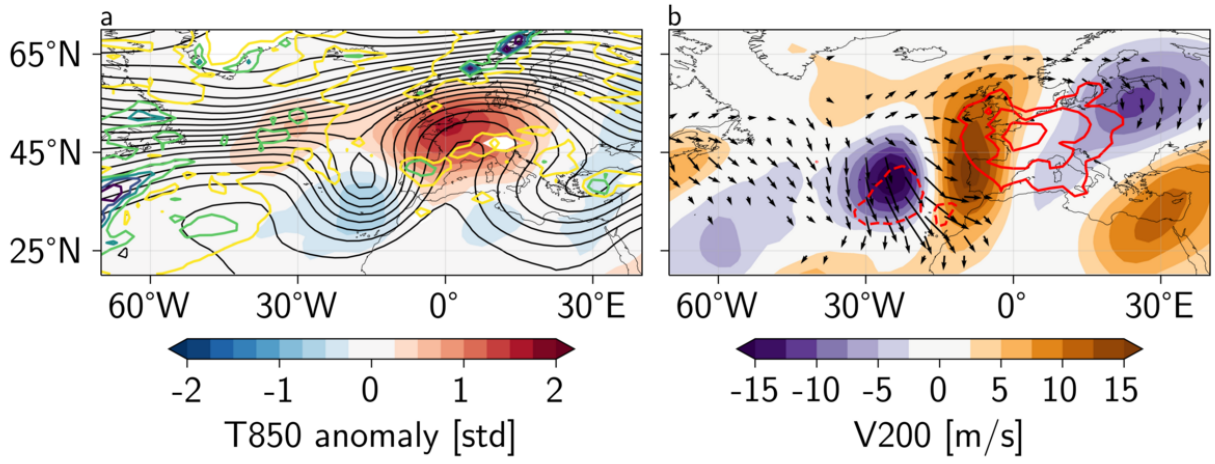
FIG. 11. 500hPa cut-off lows frequency and frequency anomaly. The anomaly is computed with respect to the control simulation.

The composite atmospheric situation associated to the presence of a cut-off low west of the Iberian peninsula is shown in Figure 12 for the biased T2M simulation only. Figure 12a shows the well isolated minimum of Z500, which is not embedded in the jet, the latter being actually mostly zonal and situated much more to the north (50 to 60°N). The meridional wind speed in Figure 12b makes clear that the cut-off low is not embedded in a Rossby wave train (at least at 200hPa) and therefore justifies the use of the term 'vortex' to designate this structure. Although

the **E**-vector composite suggests an anticyclonic wave breaking origin for the cut-off low, it is actually very difficult to validate this hypothesis in so far as, once isolated, cut-off lows behave erratically and because their size is of the same order of magnitude as the variance in their position, they tend to vanish on composite maps. The dynamical consequence of the presence of such a cut-off low is likely to be the advection of negative PV anomalies and hot air to its north-east flank (i.e. above Western Europe) at mid-troposphere favoring the appearance and maintenance of a blocked anticyclone which can subsequently break above Eastern Europe (as suggested also by the southward pointing **E**-vector in the Baltic region).

Above Western Europe, a 1.5 standard deviation localized maximum of air temperature at 850hPa is present, associated with a similar anomaly of T2M. The anomalous temperatures above the boundary layer are the result of south-west advection by the cut-off low, but whether the anomaly is mainly due to advective, adiabatic or diabatic (especially latent heat release by precipitation) mechanisms is not clear. The three mechanisms probably play a role, but a Lagrangian analysis is needed to quantify their respective importance. Although there is a 4mm precipitation contour east of the cut-off low, this value is quite small and probably not enough to explain such large anomalies above Western Europe. Another explanation may be the lifting of the hot boundary layer created over Spain and Morocco above Western Europe. If one plots the same analysis on the control simulation (Figure B3 in supplementary materials), the temperature anomalies are lower, with no strong anomaly over Western Europe. The atmospheric dynamics in particular is quite different, with weaker meridional winds associated to the cut-off and no anticyclonic wave breaking above Eastern Europe. This suggests that the cut-off low-heatwave association is not so straightforward and may occur only when some other conditions — especially with regards to the synoptic dynamics or soil moisture — are present.

We have shown that the presence of an Iberian cut-off low can be linked dynamically to the occurrence of abnormally hot conditions in Western Europe, at least in the biased T2M simulation. On the other hand, Figure B4 shows whether the presence of a heatwave in grid point 1 was preceded by the presence of an Iberian cut-off low. The occurrence of heatwaves are centered by considering the time when the maximum temperature during the heatwave is reached at $t = 0$. For each event and each time t, one counts whether there has been a cut-off west of the Iberian peninsula in the $t + 10$ days before (hence a strictly increasing cumulative frequency). Figure B4a shows that



573 FIG. 12. Atmospheric dynamics associated to cut-off lows west of the Iberian peninsula. The composites are
574 made for all events (n = 1887) in the biased T2M simulation for which a cut-off low is detected in the 33.5°N-
575 42.5°N - 22.5°W-10°W area. (a) Geopotential height at 500hPa (black contours), 3-day average precipitations
576 (colored contours) and 3-day average air temperature at 850hPa anomaly (colors). The geopotential height
577 contours are drawn every 20 meters starting at 5200m. The precipitations contours are drawn every 2mm starting
578 at 2mm. (b) 3-day average 2-m air temperature anomaly (contours), **E**-vector (arrows) and meridional wind
579 speed at 200hPa (colors). The **E**-vector is drawn for norms above 10 m^2/s^2. The anomalies are computed with
580 respect to the daily ensemble mean and standard deviation of the control simulation. For panel (b), the anomalies
581 are drawn every 0.5 standard deviation, starting at +/- 0.5 std. Plain (dashed) lines represent positive (negative)
582 values of the standard deviation.

590 for the control simulation, around 20% of heatwaves have had a cut-off low west of the Iberian
591 peninsula in the 10 days before — which is a proportion similar to the one found in our previous
592 work (Noyelle et al. 2024). For the biased Z500 and SM2 simulations the proportion is similar.
593 However, it reaches 25% for biased T2M and goes as low as 13% for SM1. This suggests different
594 dynamics for the heatwaves in those two cases, probably more driven by advective dynamics and
595 adiabatic warming in the first case and local diabatic warming in the second case.

596 4. Conclusions

597 We have shown that the use of a rare events algorithm optimizing for different score functions
598 allows to vastly increase the amount of extreme summers simulated using the state-of-the-art IPSL

climate model. For a similar computational cost as a control simulation, it is possible to sample events that are 10^2-10^4 times less likely than the ones sampled in the original control simulation. This allows to reach precise climatological results on the dynamics leading to extreme summers. As such the rare events algorithm is clearly an improvement compared to the brute force sampling by a long simulation.

We have shown here that extremely hot summers in the IPSL climate model under pre-industrial forcings are the result of a succession of more persistent heatwaves than in the control simulation. These heatwaves arise as synoptic rather than planetary scale atmospheric anomalies. Centennial-like events correspond to a large positive temperature anomaly that is centered in Western Europe but that does not extend at the planetary scale. Vertically, the anticyclonic structure above the heatwave region is not barotropic: dry soils and high sea-level pressures are situated to the east of the maximum 2-m air anomaly while the 200hPa anticyclone is situated to its west. 20-25% of heatwave events are associated to an Iberian cut-off low, reinforcing the anticyclone in mid-troposphere. Statistically we have additionally shown that extremely hot summers are around 10 times more likely than predicted by a Gaussian approximation, although this estimation comes with a one order of magnitude uncertainty.

Our results on this pre-industrial model simulation strongly favor the hypothesis of localized and recurrent dynamics via Rossby wave packets to explain the appearance of hot and very hot summers in Western Europe (Röthlisberger et al. 2019). The natural question one may ask is how much the physical mechanisms described here would be a faithful representation of the actual climate system if such rare and intense events were to happen. The IPSL-CM6A-LR model is biased with respect to the climatology of the real world (Boucher et al. 2020) — although the biases are small in Europe — and more generally climate models have difficulties to represent faithfully how intense extreme heat events can get (Vautard et al. 2023; Patterson 2023). Moreover, we do not simulate the oceans whereas one should expect a feedback of the oceanic circulation on the atmosphere considering the intensity of the anomalies simulated. It is also difficult to validate our results with observational data in so far as the undersampling issue for very rare and intense events also applies to them: there are not enough cases to compare to. Our results may also be strongly model-dependent and it seems necessary to apply the same methodology at the same place for different models — which could also constitute a test bed for comparing models on their capacity to sample the physical

mechanisms leading to very extreme events. What makes the results obtained here not completely worthless is that we mainly sample many moderately intense heatwaves and the dynamics sampled by the model for these events is coherent with the results found in the literature on heatwaves (Barriopedro et al. 2023; Domeisen et al. 2023). Consequently, it is probably reasonable to assume that the dynamics sampled here could happen in the climate system.

Acknowledgments. RN would like to thank Francesco Ragone and Freddy Bouchet for discussions which led to this work. This paper received support from the grant ANR-20-CE01-0008-01 (SAMPRACE), and from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101003469 (XAIDA), from the European Union's Horizon 2020 Marie Sklodowska-Curie grant agreement No. 956396 (EDIPI). This project was provided with computing HPC and storage resources by GENCI at TGCC thanks to the grant 2023-A0150106877 on the supercomputer Joliot Curie's ROME partition.

Data availability statement. The LMDZOR model is freely distributed at the following link [to be made available after revision]. The model configuration used for this study is available at the following link [to be made available after revision]. The scripts used to run the GKTL rare events algorithm can be found here [to be made available after revision].

APPENDIX A

Computation details

A1. Atmospheric dynamics diagnostics

In this appendix we detail how the atmospheric diagnostic metrics are computed.

(i) **E***-vector.* The propagation and tilt of transient eddies can be diagnosed using the so-called **E**-vector (Hoskins et al. 1983; Trenberth 1986; Schemm et al. 2018):

$$\mathbf{E} = (E_x, E_y) = \left(\frac{1}{2}\overline{v^{*2} - u^{*2}}, -\overline{u^* v^*} \right) \tag{A1}$$

where u^* and v^* are transient wind components and the overbar denotes a temporal average. The horizontal component E_x is proportional to the group speed of transient eddies and therefore diagnose the direction of propagation of wave energy. The meridional component E_y shows the tilt of transient eddies: equatorward (resp. poleward)-pointing **E** indicate anticyclonically (resp. cyclonically) tilted eddies, i.e. also anticyclonic (resp. cyclonic) wave breaking. Up to a good approximation, the divergence of **E** corresponds to the transfer of zonal momentum from the eddies towards the mean flow, and the reverse for convergence. Here the transients u^* and v^* are defined as anomalies with respect to a 7-day rolling average of the wind fields at 200hPa. It is common to use 6-hourly data and consider the daily average to compute the overbar (Schemm et al. 2018; Riboldi et al. 2022). Here we only have access to daily data therefore the overbar is approximated by computing a 3-day rolling average of $v^{*2} - u^{*2}$ and $u^* v^*$.

(ii) Space-time spectral analysis. The characteristics of Rossby waves are diagnosed by performing a longitude-time Fourier decomposition of the meridional wind anomaly field at 200hPa (Randel and Held 1991; Riboldi et al. 2022). The anomalies are computed with respect to the daily ensemble average of the control simulation. With the temporal and spatial resolution of the model, it is possible to resolve harmonics of minimal zonal wavelength of 5° and minimal frequencies of 2 days. This somewhat limits the precision of the analysis proposed here but is still sufficient to resolve the largest Rossby waves of synoptic scales (> 1000km), which are the most relevant for the dynamics of heatwaves. For the sake of representation, in the following the results are interpolated below these two limits.

At each latitude ϕ, the meridional wind anomaly field is decomposed as a linear superposition of monochromatic zonally propagating waves. The Fourier coefficients $\hat{V}(k_\phi, \omega)$ with zonal wavenumber $k_\phi \geq 0$ and angular frequency ω are given by:

$$\hat{V}(k_\phi, \omega) = \frac{\sqrt{2\Delta t}}{N_\lambda \sqrt{N_t}} \sum_{n_t=1}^{N_t} \sum_{n_\lambda=1}^{N_\lambda} V(n_t, n_\lambda) e^{-i(\omega n_t \Delta t + 2\pi k_\phi n_\lambda / N_\lambda)} \tag{A2}$$

where $V(n_t, n_\lambda)$ is the meridional wind anomaly at longitude n_λ and time n_t, $N_\lambda = 144$ is the number of longitudes, $N_t = 90$ is the number of days in the summer and $\Delta t = 1d$ is the temporal resolution of the simulation. The periodogram $P(k_\phi, \omega)$ is obtained by computing the square of the modulus of the Fourier coefficients $\hat{V}(k_\phi, \omega)$ and applying a smoothing Gaussian kernel in the ω dimension. Finally, the power spectral density $\rho(k_\phi, c_p)$ in the wavenumber k_ϕ-phase speed c_p space is recovered as:

$$\rho(k_\phi, c_p) = P(k_\phi, \omega) \frac{k_\phi}{2\pi R_E \cos\phi} \tag{A3}$$

where R_E is the radius of the Earth and the results are interpolated from $c_p = -30m/s$ to $c_p = 30m/s$ in steps of 1 m/s.

This procedure is applied for all latitudes between 35° and 65°N and average the results to obtain a space-time spectrum for the summer. The ensemble average is then computed simply as the average over ensemble members of this procedure (except for the biased simulations for which to each member is applied the weight corresponding to centennial events as explained above).

Additionally, at all latitudes the enveloppe of Rossby wave packets is calculated as the modulus of the complex-valued Hilbert transform of meridional wind anomalies at 200hPa (Zimin et al. 2003).

(iii) Amplitude-phase decomposition. To identify the phase-shift of quasi-stationary Rossby waves, a similar procedure as Jiménez-Esteve et al. (2022) is employed on the meridional wind at 200hPa. The area-weighted latitudinal mean between 35° and 65°N is first computed and a 7-day running mean is applied to filter out transient eddies. Contrary to Jiménez-Esteve et al. (2022), we present results with and without removing the climatological mean of the area-weighted latitudinal mean, computed here using the control simulation ensemble average. Even though anomalies have no preferred phase in the control simulation (by definition of anomalies), this may not be the case

in the biased simulations. We intend to use this analysis to diagnose the shift in the amplitude and/or phase of quasi-stationary Rossby waves between the control and biased simulations.

The temporally and spatially averaged field $\tilde{V}_{200}$ obtained is then decomposed into its Fourier components:

$$\tilde{V}_{200}(\lambda, t) = \sum_{k_\phi=0}^{+\infty} A(t)\cos(k_\phi \lambda + \Phi(t)) \tag{A4}$$

with $A(t)$ the amplitude and $\Phi(t)$ the phase. In the following the 2D amplitude A – phase Φ histograms is computed to diagnose the shift occurring for the extreme summers sampled by the rare events algorithm.

(iv) Cut-off low frequency. Several algorithms exist to detect the presence of so-called cut-off lows — i.e. isolated minima of potential vorticity (PV) or geopotential height in the mid- to high-level troposphere — especially based on PV anomalies (Wernli and Sprenger 2007; Favre et al. 2012; Pinheiro et al. 2017). Here the presence of 500hPa cut-off lows is diagnosed using an adaptation of the detection algorithm proposed by Muñoz et al. (2020). The algorithm does not rely on PV, and is better adapted to mid-level systems which are physically more relevant in the context of surface heatwaves. A cut-off low at 500hPa is considered to be present at a particular grid point if all of the following criteria are fulfilled:

1. Local geopotential height minimum: the 500hPa geopotential height of the grid point is at least 10m lower than the geopotential height in at least six of the eight surrounding grid points,

2. Isolation from the main westerly wind: there is an easterly flow in at least one of the four grid points located poleward of the grid point,

3. Cold core and thickness ridge eastward of the low: the difference between the geopotential height thickness between 500 and 850hPa at the grid point and at its immediate eastward neighbor is positive,

4. Frontal zone on the eastern flanck of the low: the thermal front parameter (TFP) on the immediate eastward neighbor of the grid point is negative. The TFP is computed as the change of the temperature gradient in the direction of the temperature gradient (at 500hPa):

$$\mathrm{TFP} = -\nabla|\nabla T| \cdot \left(\frac{\nabla T}{|\nabla T|} \right). \tag{A5}$$

With the outputs of the model, the algorithm tends to detect the presence of cut-off lows at high latitudes, in particular northward of the eddy-driven jet. Such systems may indeed fill all the preceding criteria but they do not really qualify as an isolated minimum of geopotential height in the middle-high troposphere. Another criteria could have been added to impose that the grid points of cut-offs are below the jet position to correct for this behavior. However, as I am more interested in the dynamics of cut-off lows situated well below the jet, this error does not impact strongly the analysis.

APPENDIX B

Additional figures

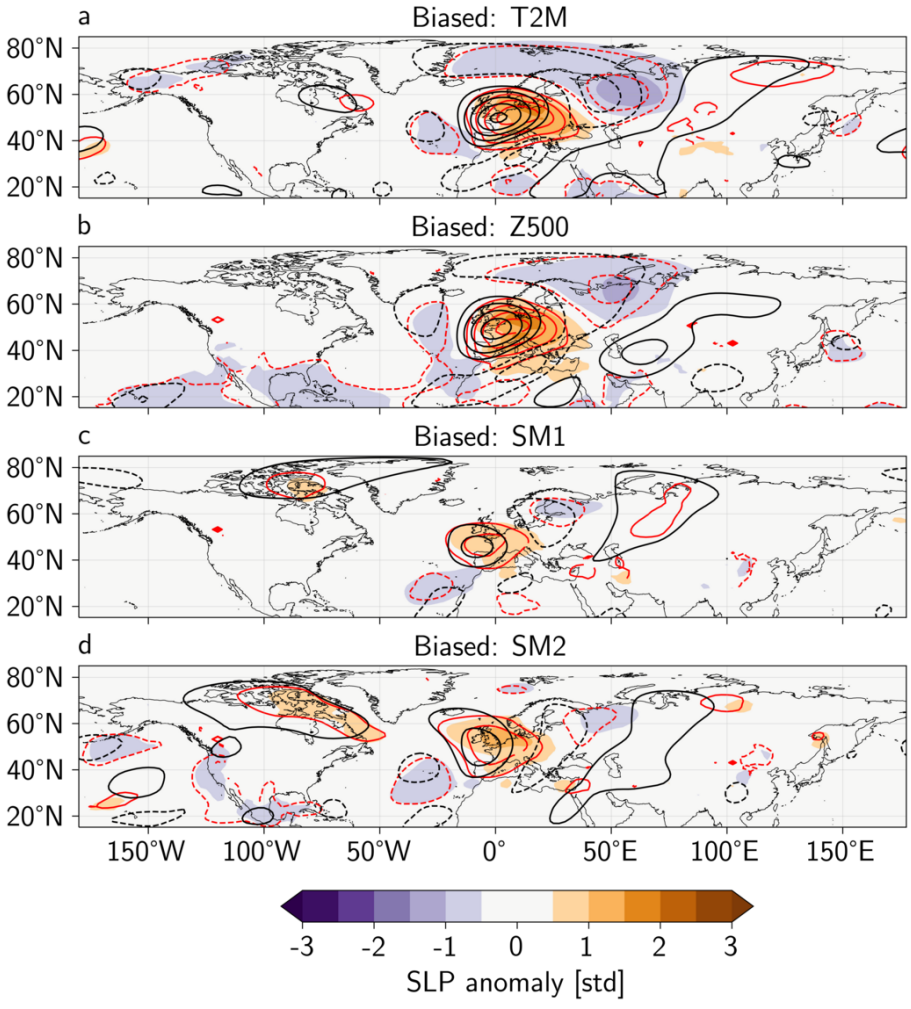


FIG. B1. Summer averaged normalized anomalies of sea-level pressure (colors), geopotential height at 850hPa (red contours) and at 200hPa (black contours) for centennial events. For all fields, normalized anomalies are computed by removing at each grid point the mean and dividing by the standard deviation computed on the control simulation. The contours are drawn every 0.5 standard deviation starting at +/- 0.5. Plain (dashed) lines represent positive (negative) values of the standard deviation.

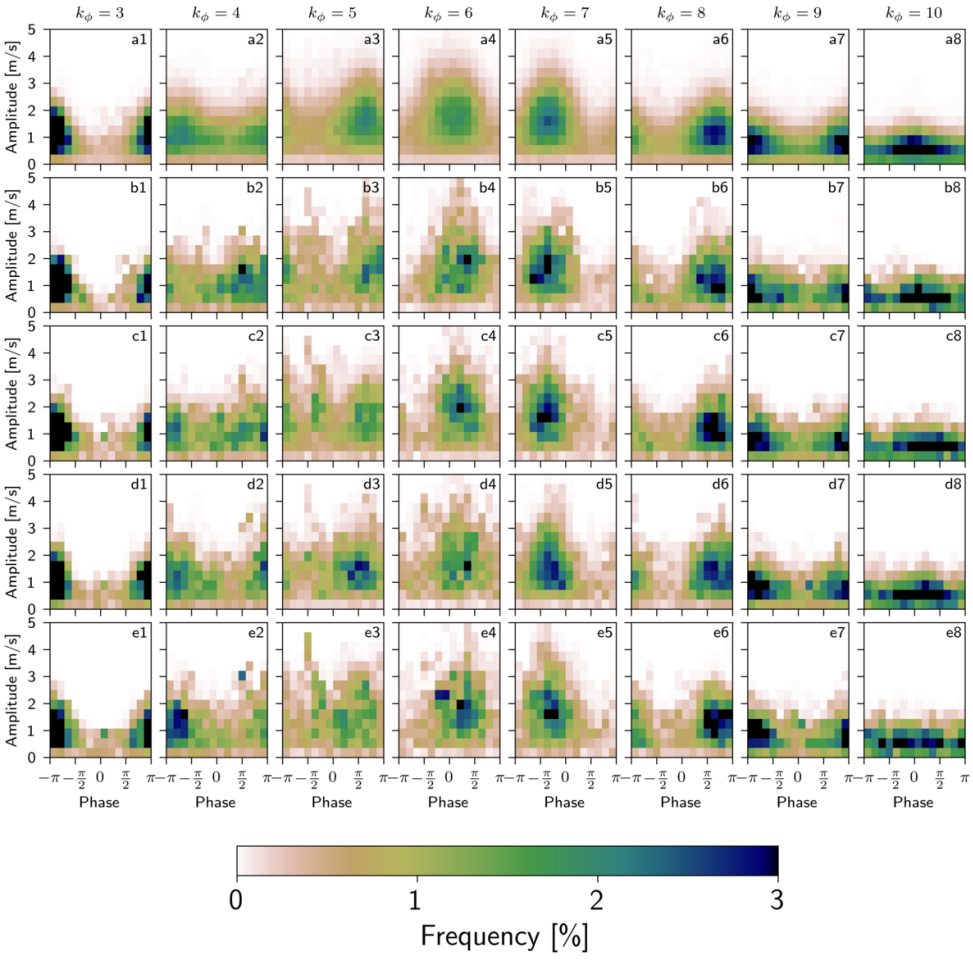


FIG. B2. Amplitude-phase histograms of meridional wind speed at 200hPa for centennial events. Wave numbers $k_\phi = 3$ to $k_\phi = 10$ in columns and biased simulations in rows: (a) control simulation, (b) biased T2M simulation, (c) biased Z500 simulation, (d) biased SM1 simulation and (e) biased SM2 simulation.

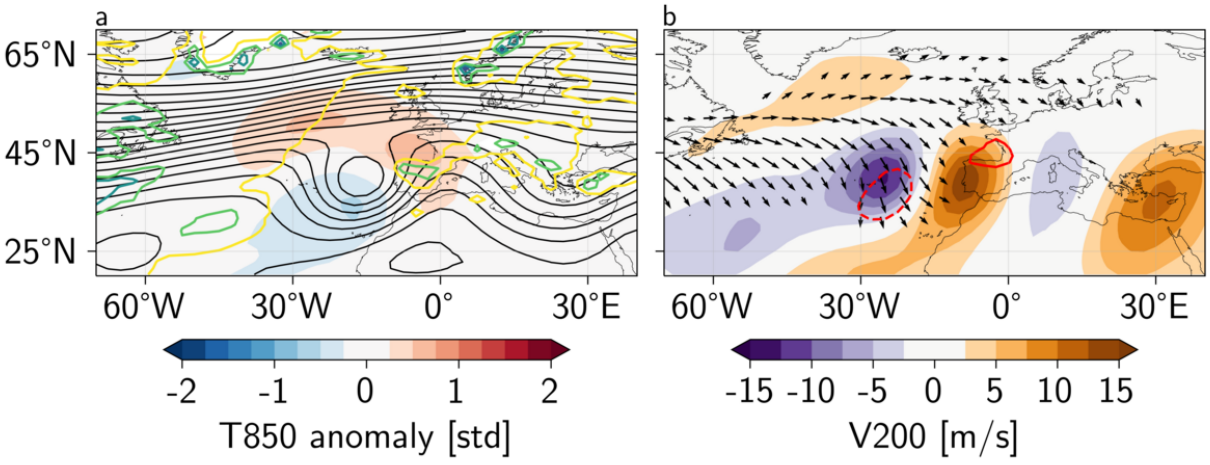


FIG. B3. Atmospheric dynamics associated to cut-off lows west of the Iberian peninsula. The composites are made for all events (n = 1124) in the control simulation for which a cut-off low is detected in the 33.5°N-42.5°N - 22.5°W-10°W area. (a) Geopotential height at 500hPa (black contours), 3-day average precipitations (colored contours) and 3-day average air temperature at 850hPa anomaly (colors). The geopotential height contours are drawn every 20 meters starting at 5200m. The precipitations contours are drawn every 2mm starting at 2mm. (b) 3-day average 2-m air temperature anomaly (contours), **E**-vector (arrows) and meridional wind speed at 200hPa (colors). The **E**-vector is drawn for norms above 10 m^2/s^2. The anomalies are computed with respect to the daily ensemble mean and standard deviation of the control simulation. For panel (b), the anomalies are drawn every 0.5 standard deviation, starting at +/- 0.5 std. Plain (dashed) lines represent positive (negative) values of the standard deviation.

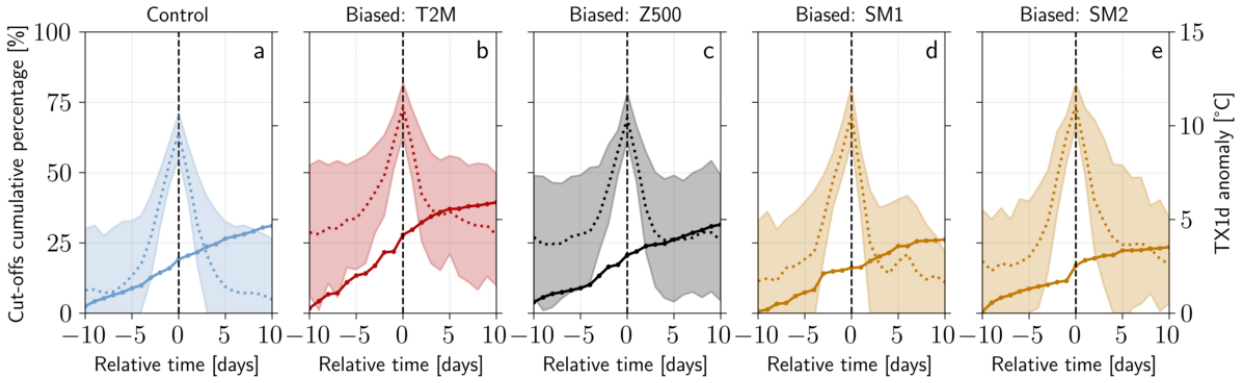


FIG. B4. Occurrence of an Iberian cut-off low when a heatwave is happening at grid point 1 at $t = 0$. The cumulative percentage of cut-offs at time t gives the percentage of heatwave events which have had at least one Iberian cut-off in the last $t + 10$ days. The shadings for the anomaly of maximum daily temperature TX1d shows the 25 and 75% quantiles. Number of events: (a) $n = 1141$, (b) $n = 4307$, (c) $n = 3290$, (d) $n = 2080$ and (e) $n = 2670$.

References

Al-Yaari, A., Y. Zhao, F. Cheruy, and W. Thiery, 2023: Heatwave characteristics in the recent climate and at different global warming levels: A multimodel analysis at the global scale. *Earth's Future*, **11 (9)**, e2022EF003 301.

Alexander, L., 2011: Extreme heat rooted in dry soils. *Nature Geoscience*, **4 (1)**, 12–13.

Baier, K., M. Rubel, and A. Stohl, 2023: The 3-week-long transport history and deep tropical origin of the 2021 extreme heat wave in the pacific northwest. *Geophysical Research Letters*, **50 (24)**, e2023GL105 865.

Barriopedro, D., R. García-Herrera, C. Ordóñez, D. Miralles, and S. Salcedo-Sanz, 2023: Heat waves: Physical understanding and scientific challenges. *Reviews of Geophysics*, e2022RG000780.

Boucher, O., and Coauthors, 2020: Presentation and evaluation of the ipsl-cm6a-lr climate model. *Journal of Advances in Modeling Earth Systems*, **12 (7)**, e2019MS002 010.

Branstator, G., and H. Teng, 2017: Tropospheric waveguide teleconnections and their seasonality. *Journal of the Atmospheric Sciences*, **74 (5)**, 1513–1532.

Castañeda, V., and L. Wang, 2024: The role of climatological state in supporting us heat waves through rossby waves packets. *Journal of Geophysical Research: Atmospheres*, **129 (4)**, e2023JD039 212.

Coumou, D., J. Lehmann, and J. Beckmann, 2015: The weakening summer circulation in the northern hemisphere mid-latitudes. *Science*, **348 (6232)**, 324–327.

Coumou, D., V. Petoukhov, S. Rahmstorf, S. Petri, and H. J. Schellnhuber, 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.

Della-Marta, P. M., J. Luterbacher, H. von Weissenfluh, E. Xoplaki, M. Brunet, and H. Wanner, 2007: Summer heat waves over western europe 1880–2003, their relationship to large-scale forcings and predictability. *Climate Dynamics*, **29**, 251–275.

779 Dirmeyer, P. A., and Coauthors, 2018: Verification of land–atmosphere coupling in forecast models,
780 reanalyses, and land surface models using flux site observations. *Journal of Hydrometeorology*,
781 **19 (2)**, 375–392.

782 Domeisen, D. I., and Coauthors, 2023: Prediction and projection of heatwaves. *Nature Reviews*
783 *Earth & Environment*, **4 (1)**, 36–50.

784 Favre, A., B. Hewitson, M. Tadross, C. Lennard, and R. Cerezo-Mota, 2012: Relationships between
785 cut-off lows and the semiannual and southern oscillations. *Climate dynamics*, **38**, 1473–1487.

786 Finkel, J., and P. A. O'Gorman, 2024: Bringing statistics to storylines: rare event sampling
787 for sudden, transient extreme events. *Journal of Advances in Modeling Earth Systems*, **16 (6)**,
788 e2024MS004 264.

789 Fischer, E. M., S. I. Seneviratne, P. L. Vidale, D. Lüthi, and C. Schär, 2007: Soil moisture–
790 atmosphere interactions during the 2003 european summer heat wave. *Journal of Climate*,
791 **20 (20)**, 5081–5099.

792 Fraedrich, K. F., H. Jansen, E. Kirk, U. Luksch, and F. Lunkeit, 2005: The planet simulator:
793 Towards a user friendly model. *Meteorologische Zeitschrift*, **14 (3)**, 299–304.

794 Fragkoulidis, G., V. Wirth, P. Bossmann, and A. Fink, 2018: Linking northern hemisphere
795 temperature extremes to rossby wave packets. *Quarterly Journal of the Royal Meteorological*
796 *Society*, **144 (711)**, 553–566.

797 Gessner, C., E. M. Fischer, U. Beyerle, and R. Knutti, 2021: Very rare heat extremes: quantifying
798 and understanding using ensemble reinitialization. *Journal of Climate*, **34 (16)**, 6619–6634.

799 Gevaert, A., D. G. Miralles, R. A. de Jeu, J. Schellekens, and A. J. Dolman, 2018: Soil moisture-
800 temperature coupling in a set of land surface models. *Journal of Geophysical Research: Atmo-*
801 *spheres*, **123 (3)**, 1481–1498.

802 Giardina, C., J. Kurchan, V. Lecomte, and J. Tailleur, 2011: Simulating Rare Events in
803 Dynamical Processes. *Journal of Statistical Physics*, **145 (4)**, 787–811, https://doi.org/
804 10.1007/s10955-011-0350-4.

Giardina', C., J. Kurchan, and L. Peliti, 2006: Direct evaluation of large-deviation functions. *arXiv:cond-mat/0511248*, https://doi.org/10.1103/PhysRevLett.96.120603.

Gálfi, V. M., V. Lucarini, F. Ragone, and J. Wouters, 2021: Applications of large deviation theory in geophysical fluid dynamics and climate science. *La Rivista del Nuovo Cimento*, **44 (6)**, 291–363, https://doi.org/10.1007/s40766-021-00020-z.

Hauser, M., R. Orth, and S. I. Seneviratne, 2016: Role of soil moisture versus recent climate change for the 2010 heat wave in western russia. *Geophysical Research Letters*, **43 (6)**, 2819–2826.

He, Y., X. Zhu, Z. Sheng, and M. He, 2023: Resonant waves play an important role in the increasing heat waves in northern hemisphere mid-latitudes under global warming. *Geophysical Research Letters*, **50 (14)**, e2023GL104 839.

Hirschi, M., and Coauthors, 2011: Observational evidence for soil-moisture impact on hot extremes in southeastern europe. *Nature Geoscience*, **4 (1)**, 17–21.

Horowitz, R. L., K. A. McKinnon, and I. R. Simpson, 2022: Circulation and soil moisture contributions to heatwaves in the united states. *Journal of Climate*, **35 (24)**, 8031–8048.

Horton, R. M., J. S. Mankin, C. Lesk, E. Coffel, and C. Raymond, 2016: A review of recent advances in research on extreme heat events. *Current Climate Change Reports*, **2 (4)**, 242–259.

Hoskins, B., and T. Woollings, 2015: Persistent extratropical regimes and climate extremes. *Current Climate Change Reports*, **1**, 115–124.

Hoskins, B. J., I. N. James, and G. H. White, 1983: The shape, propagation and mean-flow interaction of large-scale weather systems. *Journal of Atmospheric Sciences*, **40 (7)**, 1595–1612.

Hoy, A., S. Haensel, P. Skalak, Z. Ustrnul, and O. Bochníček, 2017: The extreme european summer of 2015 in a long-term perspective. *International Journal of Climatology*, **37 (2)**, 943–962.

Huntingford, C., P. M. Cox, P. D. Ritchie, J. J. Clarke, I. M. Parry, and M. S. Williamson, 2024: Acceleration of daily land temperature extremes and correlations with surface energy fluxes. *npj Climate and Atmospheric Science*, **7 (1)**, 84.

Jiménez-Esteve, B., K. Kornhuber, and D. Domeisen, 2022: Heat extremes driven by amplification of phase-locked circumglobal waves forced by topography in an idealized atmospheric model. *Geophysical Research Letters*, **49 (21)**, e2021GL096 337.

Kautz, L.-A., O. Martius, S. Pfahl, J. G. Pinto, A. M. Ramos, P. M. Sousa, and T. Woollings, 2022: Atmospheric blocking and weather extremes over the euro-atlantic sector–a review. *Weather and climate dynamics*, **3 (1)**, 305–336.

Kornhuber, K., S. Osprey, D. Coumou, S. Petri, V. Petoukhov, S. Rahmstorf, and L. Gray, 2019: Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern. *Environmental Research Letters*, **14 (5)**, 054 002.

Lehmann, J., and D. Coumou, 2015: The influence of mid-latitude storm tracks on hot, cold, dry and wet extremes. *Scientific reports*, **5 (1)**, 17 491.

Luo, F., and Coauthors, 2022: Summertime rossby waves in climate models: substantial biases in surface imprint associated with small biases in upper-level circulation. *Weather and Climate Dynamics*, **3 (3)**, 905–935.

Mann, M. E., S. Rahmstorf, K. Kornhuber, B. A. Steinman, S. K. Miller, and D. Coumou, 2017: Influence of anthropogenic climate change on planetary wave resonance and extreme weather events. *Scientific reports*, **7 (1)**, 1–12.

Marshall, A., D. Hudson, M. Wheeler, O. Alves, H. Hendon, M. Pook, and J. Risbey, 2014: Intra-seasonal drivers of extreme heat over australia in observations and poama-2. *Climate dynamics*, **43**, 1915–1937.

Meehl, G. A., and C. Tebaldi, 2004: More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*, **305 (5686)**, 994–997.

Miralles, D. G., P. Gentine, S. I. Seneviratne, and A. J. Teuling, 2019: Land–atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges. *Annals of the New York Academy of Sciences*, **1436 (1)**, 19–35.

Miralles, D. G., A. J. Teuling, C. C. Van Heerwaarden, and J. Vilà-Guerau de Arellano, 2014: Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature geoscience*, **7 (5)**, 345–349.

Mueller, B., and S. I. Seneviratne, 2012: Hot days induced by precipitation deficits at the global scale. *Proceedings of the national academy of sciences*, **109 (31)**, 12 398–12 403.

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

Neal, E., C. S. Huang, and N. Nakamura, 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.

Noyelle, R., P. Yiou, and D. Faranda, 2024: Investigating the typicality of the dynamics leading to extreme temperatures in the ipsl-cm6a-lr model. *Climate Dynamics*, **62 (2)**, 1329–1357.

Patterson, M., 2023: North-west europe hottest days are warming twice as fast as mean summer days. *Geophysical Research Letters*, **50 (10)**, e2023GL102 757.

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

Perkins, S. E., and L. V. Alexander, 2013: On the measurement of heat waves. *Journal of climate*, **26 (13)**, 4500–4517.

Perkins-Kirkpatrick, S., and S. Lewis, 2020: Increasing trends in regional heatwaves. *Nature communications*, **11 (1)**, 3357.

Petoukhov, V., S. Petri, S. Rahmstorf, D. Coumou, K. Kornhuber, and H. J. Schellnhuber, 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.

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

Pfahl, S., C. Schwierz, M. Croci-Maspoli, C. M. Grams, and H. Wernli, 2015: Importance of latent heat release in ascending air streams for atmospheric blocking. *Nature Geoscience*, **8 (8)**, 610–614.

Pfahl, S., and H. Wernli, 2012: Quantifying the relevance of atmospheric blocking for co-located temperature extremes in the northern hemisphere on (sub-) daily time scales. *Geophysical Research Letters*, **39 (12)**.

Pfleiderer, P., and D. Coumou, 2018: Quantification of temperature persistence over the northern hemisphere land-area. *Climate Dynamics*, **51 (1)**, 627–637.

Pfleiderer, P., C.-F. Schleussner, K. Kornhuber, and D. Coumou, 2019: Summer weather becomes more persistent in a 2 °C world. *Nature Climate Change*, **9 (9)**, 666–671, https://doi.org/10.1038/s41558-019-0555-0.

Pinheiro, H. R., K. I. Hodges, M. A. Gan, and N. J. Ferreira, 2017: A new perspective of the climatological features of upper-level cut-off lows in the southern hemisphere. *Climate Dynamics*, **48**, 541–559.

Plotkin, D. A., R. J. Webber, M. E. O'Neill, J. Weare, and D. S. Abbot, 2019: Maximizing Simulated Tropical Cyclone Intensity With Action Minimization. *Journal of Advances in Modeling Earth Systems*, **11 (4)**, 863–891, https://doi.org/10.1029/2018MS001419.

Quesada, B., R. Vautard, P. Yiou, M. Hirschi, and S. I. Seneviratne, 2012: Asymmetric european summer heat predictability from wet and dry southern winters and springs. *Nature Climate Change*, **2 (10)**, 736–741.

Ragone, F., and F. Bouchet, 2020: Computation of Extreme Values of Time Averaged Observables in Climate Models with Large Deviation Techniques. *Journal of Statistical Physics*, **179 (5)**, 1637–1665, https://doi.org/10.1007/s10955-019-02429-7.

Ragone, F., and F. Bouchet, 2021: Rare Event Algorithm Study of Extreme Warm Summers and Heatwaves Over Europe. *Geophysical Research Letters*, **48 (12)**, e2020GL091 197, https://doi.org/10.1029/2020GL091197.

Ragone, F., J. Wouters, and F. Bouchet, 2018: Computation of extreme heat waves in climate models using a large deviation algorithm. *Proceedings of the National Academy of Sciences*, **115 (1)**, 24–29, https://doi.org/10.1073/pnas.1712645115.

Randel, W., and I. Held, 1991: Phase speed spectra of transient eddy fluxes and critical layer absorption. *Journal of the atmospheric sciences*, **48 (5)**, 688–697.

Riboldi, J., E. Rousi, F. d'Andrea, G. Rivière, and F. Lott, 2022: Circumglobal rossby wave patterns during boreal winter highlighted by space–time spectral analysis. *Weather and Climate Dynamics*, **3 (2)**, 449–469.

Rossby, C.-G., 1939: Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *J. mar. Res.*, **2**, 38–55.

Röthlisberger, M., L. Frossard, L. F. Bosart, D. Keyser, and O. Martius, 2019: Recurrent synoptic-scale rossby wave patterns and their effect on the persistence of cold and hot spells. *Journal of Climate*, **32 (11)**, 3207–3226.

Röthlisberger, M., S. Pfahl, and O. Martius, 2016: Regional-scale jet waviness modulates the occurrence of midlatitude weather extremes. *Geophysical Research Letters*, **43 (20)**, 10–989.

Röthlisberger, M., M. Sprenger, E. Flaounas, U. Beyerle, and H. Wernli, 2020: The substructure of extremely hot summers in the northern hemisphere. *Weather and Climate Dynamics*, **1 (1)**, 45–62.

Rousi, E., and Coauthors, 2023: The extremely hot and dry 2018 summer in central and northern europe from a multi-faceted weather and climate perspective. *Natural Hazards and Earth System Sciences*, **23 (5)**, 1699–1718.

Russo, S., J. Sillmann, and E. M. Fischer, 2015: Top ten european heatwaves since 1950 and their occurrence in the coming decades. *Environmental Research Letters*, **10 (12)**, 124 003.

Sánchez-Benítez, A., R. García-Herrera, D. Barriopedro, P. M. Sousa, and R. M. Trigo, 2018: June 2017: the earliest european summer mega-heatwave of reanalysis period. *Geophysical Research Letters*, **45 (4)**, 1955–1962.

Sánchez-Benítez, A., H. Goessling, F. Pithan, T. Semmler, and T. Jung, 2022: The july 2019 european heat wave in a warmer climate: storyline scenarios with a coupled model using spectral nudging. *Journal of Climate*, **35 (8)**, 2373–2390.

Sauer, J., J. Demaeyer, G. Zappa, F. Massonnet, and F. Ragone, 2024: Extremes of summer arctic sea ice reduction investigated with a rare event algorithm. *Climate Dynamics*, 1–19.

Schaller, N., J. Sillmann, J. Anstey, E. M. Fischer, C. M. Grams, and S. Russo, 2018: Influence of blocking on northern european and western russian heatwaves in large climate model ensembles. *Environmental Research Letters*, **13 (5)**, 054 015.

Schemm, S., G. Rivière, L. M. Ciasto, and C. Li, 2018: Extratropical cyclogenesis changes in connection with tropospheric enso teleconnections to the north atlantic: Role of stationary and transient waves. *Journal of the Atmospheric Sciences*, **75 (11)**, 3943–3964.

Schumacher, D., M. Hauser, and S. I. Seneviratne, 2022a: Drivers and mechanisms of the 2021 pacific northwest heatwave. *Earth's Future*, **10 (12)**, e2022EF002 967.

Schumacher, D. L., J. Keune, P. Dirmeyer, and D. G. Miralles, 2022b: Drought self-propagation in drylands due to land–atmosphere feedbacks. *Nature geoscience*, **15 (4)**, 262–268.

Schumacher, D. L., J. Keune, C. C. Van Heerwaarden, J. Vilà-Guerau de Arellano, A. J. Teuling, and D. G. Miralles, 2019: Amplification of mega-heatwaves through heat torrents fuelled by upwind drought. *Nature Geoscience*, **12 (9)**, 712–717.

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

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, **99 (3-4)**, 125–161.

Seneviratne, S. I., and Coauthors, 2021: Weather and climate extreme events in a changing climate (chapter 11).

Serrano-Notivoli, R., E. Tejedor, P. Sarricolea, O. Meseguer-Ruiz, M. de Luis, M. Á. Saz, L. A. Longares, and J. Olcina, 2023: Unprecedented warmth: A look at spain's exceptional summer of 2022. *Atmospheric Research*, 106931.

Sousa, P. M., R. M. Trigo, D. Barriopedro, P. M. Soares, and J. A. Santos, 2018: European temperature responses to blocking and ridge regional patterns. *Climate Dynamics*, **50 (1)**, 457–477.

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

Steinfeld, D., and S. Pfahl, 2019: The role of latent heating in atmospheric blocking dynamics: a global climatology. *Climate Dynamics*, **53 (9)**, 6159–6180.

Stott, P. A., D. A. Stone, and M. R. Allen, 2004: Human contribution to the european heatwave of 2003. *Nature*, **432 (7017)**, 610–614.

Teng, H., and G. Branstator, 2019: Amplification of waveguide teleconnections in the boreal summer. *Current Climate Change Reports*, **5**, 421–432.

Teng, H., G. Branstator, H. Wang, G. A. Meehl, and W. M. Washington, 2013: Probability of us heat waves affected by a subseasonal planetary wave pattern. *Nature Geoscience*, **6 (12)**, 1056–1061.

Teuling, A. J., and Coauthors, 2013: Evapotranspiration amplifies european summer drought. *Geophysical Research Letters*, **40 (10)**, 2071–2075.

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

Touchette, H., 2009b: The large deviation approach to statistical mechanics. *Physics Reports*, **478 (1)**, 1–69, https://doi.org/10.1016/j.physrep.2009.05.002.

Trenberth, K. E., 1986: An assessment of the impact of transient eddies on the zonal flow during a blocking episode using localized eliassen-palm flux diagnostics. *Journal of the Atmospheric Sciences*, **43 (19)**, 2070–2087.

Tripathy, K. P., and A. K. Mishra, 2023: How unusual is the 2022 european compound drought and heatwave event? *Geophysical Research Letters*, **50 (15)**, e2023GL105 453.

Tuel, A., and O. Martius, 2024: Persistent warm and cold spells in the northern hemisphere extratropics: regionalisation, synoptic-scale dynamics and temperature budget. *Weather and Climate Dynamics*, **5 (1)**, 263–292.

Tuel, A., D. Steinfeld, S. M. Ali, M. Sprenger, and O. Martius, 2022: Large-scale drivers of persistent extreme weather during early summer 2021 in europe. *Geophysical Research Letters*, **49 (18)**, e2022GL099 624.

Van Loon, S., and D. W. Thompson, 2023: Comparing local versus hemispheric perspectives of extreme heat events. *Geophysical Research Letters*, **50 (24)**, e2023GL105 246.

Vautard, R., and Coauthors, 2007: Summertime european heat and drought waves induced by wintertime mediterranean rainfall deficit. *Geophysical Research Letters*, **34 (7)**.

Vautard, R., and Coauthors, 2023: Heat extremes in western europe increasing faster than simulated due to atmospheric circulation trends. *Nature Communications*, **14 (1)**, 6803.

Vogel, M. M., R. Orth, F. Cheruy, S. Hagemann, R. Lorenz, B. J. van den Hurk, and S. I. Seneviratne, 2017: Regional amplification of projected changes in extreme temperatures strongly controlled by soil moisture-temperature feedbacks. *Geophysical Research Letters*, **44 (3)**, 1511–1519.

Vogel, M. M., J. Zscheischler, and S. I. Seneviratne, 2018: Varying soil moisture–atmosphere feedbacks explain divergent temperature extremes and precipitation projections in central europe. *Earth System Dynamics*, **9 (3)**, 1107–1125.

Webber, R. J., D. A. Plotkin, M. E. O'Neill, D. S. Abbot, and J. Weare, 2019: Practical rare event sampling for extreme mesoscale weather. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **29 (5)**, 053 109, https://doi.org/10.1063/1.5081461.

Wehrli, K., B. P. Guillod, M. Hauser, M. Leclair, and S. I. Seneviratne, 2019: Identifying key driving processes of major recent heat waves. *Journal of Geophysical Research: Atmospheres*, **124 (22)**, 11 746–11 765.

Wehrli, K., M. Hauser, and S. I. Seneviratne, 2020: Storylines of the 2018 northern hemisphere heatwave at pre-industrial and higher global warming levels. *Earth System Dynamics*, **11 (4)**, 855–873.

Wernli, H., and M. Sprenger, 2007: Identification and era-15 climatology of potential vorticity streamers and cutoffs near the extratropical tropopause. *Journal of the atmospheric sciences*, **64 (5)**, 1569–1586.

Wirth, V., and C. Polster, 2021: The problem of diagnosing jet waveguidability in the presence of large-amplitude eddies. *Journal of the Atmospheric Sciences*, **78 (10)**, 3137–3151.

Wirth, V., M. Riemer, E. K. Chang, and O. Martius, 2018: Rossby wave packets on the midlatitude waveguide—a review. *Monthly Weather Review*, **146 (7)**, 1965–2001.

Wouters, J., and F. Bouchet, 2016: Rare event computation in deterministic chaotic systems using genealogical particle analysis. *Journal of Physics A: Mathematical and Theoretical*, **49 (37)**, 374 002.

Wouters, J., R. K. Schiemann, and L. C. Shaffrey, 2023: Rare event simulation of extreme european winter rainfall in an intermediate complexity climate model. *Journal of Advances in Modeling Earth Systems*, **15 (4)**, e2022MS003 537.

Xoplaki, E., J. F. González-Rouco, J. Luterbacher, and H. Wanner, 2003: Mediterranean summer air temperature variability and its connection to the large-scale atmospheric circulation and ssts. *Climate dynamics*, **20**, 723–739.

Yiou, P., and A. Jézéquel, 2020: Simulation of extreme heat waves with empirical importance sampling. *Geoscientific Model Development*, **13 (2)**, 763–781, https://doi.org/10.5194/gmd-13-763-2020.

Zampieri, M., F. D'andrea, R. Vautard, P. Ciais, N. de Noblet-Ducoudré, and P. Yiou, 2009: Hot european summers and the role of soil moisture in the propagation of mediterranean drought. *Journal of Climate*, **22 (18)**, 4747–4758.

Zhang, Y., and W. R. Boos, 2023: An upper bound for extreme temperatures over midlatitude land. *Proceedings of the National Academy of Sciences*, **120 (12)**, e2215278 120.

Zimin, A. V., I. Szunyogh, D. Patil, B. R. Hunt, and E. Ott, 2003: Extracting envelopes of rossby wave packets. *Monthly weather review*, **131 (5)**, 1011–1017.

Zschenderlein, P., G. Fragkoulidis, A. H. Fink, and V. Wirth, 2018: Large-scale rossby wave and synoptic-scale dynamic analyses of the unusually late 2016 heatwave over europe. *Weather*, **73 (9)**, 275–283.

1043 Zschenderlein, P., S. Pfahl, H. Wernli, and A. H. Fink, 2020: A lagrangian analysis of upper-
1044 tropospheric anticyclones associated with heat waves in europe. *Weather and Climate Dynamics*,
1045 **1 (1)**, 191–206.