

# Rare event algorithm study of extreme warm summers and heatwaves over Europe

Francesco Ragone<sup>1</sup> and Freddy Bouchet<sup>2</sup>

<sup>1</sup>Royal Meteorological Institute of Belgium and Université Catholique de Louvain

<sup>2</sup>ENS de Lyon and CNRS

November 24, 2022

## Abstract

The analysis of extremes in climate models is hindered by the lack of statistics due to the computational costs required to run simulations long enough to sample rare events. We demonstrate how rare event algorithms can improve the statistics of extreme events in state-of-the-art climate models. We study extreme warm summers and heatwaves over France and Scandinavia with CESM1.2.2 in present-day climate. The algorithm concentrates the simulations on events of importance, and shifts the probability distributions of regional temperatures such that warm summers become common. We estimate return times of extremes orders of magnitude larger than what feasible with direct sampling, and we compute statistically significant composite maps of dynamical quantities conditional on the occurrence of the extremes. We show that extreme warm summers are associated to wavenumber 3 hemispheric teleconnection patterns, and that the most extreme summers are related to the succession of rare subseasonal heatwaves.

# Rare event algorithm study of extreme warm summers and heatwaves over Europe

F. Ragone<sup>1,2,3</sup> and F. Bouchet<sup>1</sup>

<sup>1</sup>Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, Lyon, France

<sup>2</sup>Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium

<sup>3</sup>Royal Meteorological Institute of Belgium, Brussels, Belgium

## Key Points:

- The rare event algorithm increases by several orders of magnitude the number of warm summers and heatwaves sampled by the model.
- Warm summers over either France or Scandinavia are linked to wavenumber 3 hemispheric teleconnection patterns.
- Warm summers in Scandinavia show bimodality due to different distribution of subsequent subseasonal heatwaves.

---

Corresponding author: Francesco Ragone, [francesco.ragone@uclouvain.be](mailto:francesco.ragone@uclouvain.be)

## Abstract

The analysis of extremes in climate models is hindered by the lack of statistics due to the computational costs required to run simulations long enough to sample rare events. We demonstrate how rare event algorithms can improve the statistics of extreme events in state-of-the-art climate models. We study extreme warm summers and heatwaves over France and Scandinavia with CESM1.2.2 in present-day climate. The algorithm concentrates the simulations on events of importance, and shifts the probability distributions of regional temperatures such that warm summers become common. We estimate return times of extremes orders of magnitude larger than what feasible with direct sampling, and we compute statistically significant composite maps of dynamical quantities conditional on the occurrence of the extremes. We show that extreme warm summers are associated to wavenumber 3 hemispheric teleconnection patterns, and that the most extreme summers are related to the succession of rare subseasonal heatwaves.

## Plain Language Summary

The impact of extreme climatic events is often dominated by the rarest events. These events have return times (a measure of how often they occur on average) of hundreds of years or more, but they could, and do, happen anytime. These events are poorly understood because of lack of statistics. Climate models are computationally expensive, and can not be run for long enough to study events with return times longer than a few decades. We use a new computational technique that allows simulations to focus only on trajectories leading to extreme heatwaves over a target region, optimizing the use of computational resources. We thus gather robust statistics for seasonal heatwaves with return times of hundreds or thousands of years, and observe even rarer events impossible to observe otherwise. We find that extreme warm summer over France or Scandinavia are synchronised with extreme warm summers in specific regions of Asia and North-America by a teleconnection pattern extending over the entire Northern hemisphere. We also find suggestions that extreme warm summers over Scandinavia may occur in two distinct ways. This new method can be used to better study the impact of global warming on the risk of catastrophic events, and to improve their predictability.

## 1 Introduction

Recent decades have seen a number of exceptionally warm summers and record breaking heatwaves at the Northern hemisphere midlatitudes (Luterbacher et al., 2004; García-Herrera et al., 2010; Barriopedro et al., 2011; Otto et al., 2012; Agha Kouchak, 2012; IPCC, 2012; Russo et al., 2015; Coumou & Rahmstorf, 2012). The Intergovernmental Panel on Climate Change (IPCC) has concluded that hot days and heavy precipitation events have become more frequent since 1950 (IPCC, 2013). The precise extent to which heat extremes will become more common in the future and under the target scenarios of +1.5 °C and +2 °C global surface temperature with respect to preindustrial levels is an active field of research (Mueller et al., 2016; Dosio et al., 2018; Suarez-Gutierrez et al., 2018).

A phenomenological theory exists on the causes of heatwaves (Schubert et al., 2014; Perkins, 2015; Horton et al., 2016), including large scale atmospheric circulation patterns, (Della-Marta et al., 2007; Cassou et al., 2005; Jezequel et al., 2018), lack of precipitation and soil moisture (Vautard et al., 2007; Zampieri et al., 2014; D’Andrea et al., 2016), and sea surface temperature anomalies (Della-Marta et al., 2007; Cassou et al., 2005), with different levels of importance in different regions (Stéfanon et al., 2012). Long lasting heatwaves and warm summers are among the most relevant events in terms of impacts (Camargo & Seth, 2016; Trenberth, 2012). Although they are known to be related to persistent weather regimes (Lau & Kim, 2012; Teng et al., 2013; Hoskins & Woollings, 2015; Petoukhov et al., 2016; Boers et al., 2019; Kornhuber et al., 2019), their dynamics is not fully understood. Some studies show stronger increase with global warming of the frequency

of persistent temperature extremes (Pfleiderer et al., 2019). However, the response of the global circulation to climate change is extremely complex and not well understood (Hoskins & Woollings, 2015), in particular for the type of rare dynamics that lead to extreme events.

The impact of climate extremes is often dominated by the rarest events. For instance, the death toll of the Western European heatwave in 2003, about 70.000 casualties, exceeded the sum of the fatalities due to all other heatwaves during the last three decades (Barriopedro et al., 2011). Moreover, extreme events with very long return times (also called return periods) always occur. For instance, an extreme event with a return time of 1000 years, has a chance 1/1000 to occur next year. Among the huge number of possible extreme events with a return time of 1000 years, a few of them will certainly occur somewhere sometimes next year. It is thus crucial to study the most extreme and rare events. However, the study of such events faces strong scientific limitations: we can not rely on historical data for events with return times of 100 years or more, because often no similar events have ever been observed, and it is extremely difficult to study these events using climate models because of the required computational costs.

In this paper we address the problem of computational cost limitations. In order to face this problem, we use a rare event algorithm that concentrates ensemble simulations on trajectories of importance for the extreme events, optimizing the use of the computational resources. Rare event algorithms have recently been applied to turbulence problems (Grafke et al., 2015; Laurie & Bouchet, 2015; Ebener et al., 2019; Bouchet et al., 2019; Lestang et al., 2020) and climate applications (Ragone et al., 2018; Webber et al., 2019; Ragone & Bouchet, 2020; Plotkin et al., 2019). Based on the phenomenology of the dynamics for each family of extreme events, an appropriated type of rare event algorithm should be carefully chosen. Here we use a genealogical algorithm (Ragone et al., 2018; Ragone & Bouchet, 2020) adapted from (Del Moral et al., 2005; Giardina et al., 2011), that is efficient to study long lasting events. This rare event algorithm has been used to study heatwaves with an intermediate complexity model (Ragone et al., 2018) in perpetual summer conditions. Here we use it to sample rare heatwaves in CESM 1.2.2 (Hurrell et al., 2013), a fully realistic climate model, in presence of daily and seasonal cycles.

The goal of this paper is to demonstrate the applicability of rare event algorithms to state-of-the-art climate models, and to highlight properties of extreme long lasting heatwaves that can not be assessed with traditional sampling strategies. We study extreme warm summers and heatwaves over France and Scandinavia in CESM1.2.2, in present-day climate. We compute return times orders of magnitude larger than what feasible with direct sampling, and statistically significant composite maps of dynamical quantities. Our results show that extreme warm summers are associated to wavenumber 3 teleconnection patterns in the Northern hemisphere, and suggest that the most extreme summers are related to the succession of multiple rare subseasonal heatwaves.

## 2 Data and methods

**Model.** All simulations are performed with the Community Earth System Model (CESM) version 1.2.2 (Hurrell et al., 2013). We use an atmosphere and land only setup, whose active components are the Community Atmospheric Model version 4 (CAM4) and the Community Land Model version 2 (CLM2). The model is run at statistically stationary state, with sea surface temperature (SST), sea ice cover, and the concentration of atmospheric CO<sub>2</sub> and other greenhouse gases fixed at values representative of present day climate (year 2000). Contrary to (Ragone et al., 2018), the model features daily and seasonal cycle, and reproduces a fully realistic climate. See the SI for more details. We study the statistics of a 1000 years long control run, and the statistics of several sets of simulations using the rare event algorithm.

**Heatwaves and physical observable definition.** We consider the surface temperature  $T_s(\vec{r}, t)$ , where  $\vec{r}$  is the space variable and  $t$  is time, and define the mean surface temperature  $\mathbb{E}(T_s)(\vec{r}, t)$ , which varies in time (because of the seasonal cycle) and in space. In practice,  $\mathbb{E}(T_s)$  will be approximated by the climatological average of  $T_s$  computed from a 1000 years long control run. We study the statistics of the spatially and temporally averaged surface temperature anomaly defined by

$$a(t_I) = \frac{1}{T} \int_{t_I}^{t_I+T} A(t) dt \quad \text{where} \quad A(t) = \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} (T_s - \mathbb{E}(T_s))(\vec{r}, t) d\vec{r}, \quad (1)$$

where  $A(t)$  is the instantaneous spatial average,  $\mathcal{D}$  is a spatial domain area,  $t_I$  is the starting date of the time average,  $T$  is the averaging time, for instance several days up to one season. Heatwaves will be defined as extreme values of the observable  $a(t_I)$ , as heatwave indices developed for dynamical studies often use anomalies rather than absolute values (Perkins, 2015). We are specifically interested in long events, for instance warm summers with  $T=3$  months. For summer anomalies, we consider  $a_{JJA}$  defined by (1) with  $t_I$  being June 1st and  $t_I + T$  August 29. In the following, we consider  $\mathcal{D}$  as the area over either France or Scandinavia (see the SI for the definition of the areas).

**Rare event algorithm experiments and importance sampling.** Rare event algorithms allows a numerical model to produce very efficiently rare trajectories of importance for the study of extremes. We simulate an ensemble of  $N$  model trajectories, and at constant intervals of a resampling time  $\tau$  we perform trajectory selection, pruning and cloning trajectories in order to favour those leading to the extreme events. See the SI for a more detailed description and (Ragone et al., 2018). The algorithm also allows to compute probabilities for the obtained trajectories. Let  $\vec{X}(t)$  be the vector of values of all the model variables at time  $t$  (which includes but is not limited to the temperature  $T_s$  defined above). We consider the trajectory  $\{\vec{X}(t)\}$ , with all values of  $X$  along all the interval  $t_a \leq t \leq t_a + T_a$ , where  $t_a$  is the starting date and  $T_a$  the trajectory duration in the algorithm. We denote  $\mathbb{P}_k(\{\vec{X}(t)\})$  the probability distribution function of the trajectories obtained in the algorithm. Then the rare event algorithm produce trajectories distributed according to

$$\mathbb{P}_k(\{\vec{X}(t)\}) = \frac{e^{k \int_{t_a}^{t_a+T_a} A(u) du}}{Z} \mathbb{P}_0(\{\vec{X}(t)\}), \quad (2)$$

where  $\mathbb{P}_0(\{\vec{X}(t)\})$  is the probability distribution function of the trajectories in the model climate,  $Z$  is a normalization term computed by the algorithm,  $A$  is the selection function (with the dependance on the trajectory implicit for simplicity,  $A(t) = A(\{\vec{X}(t)\})$ ), and  $k$  controls the strength of the algorithm selection. Equation 2 is called an importance ratio formula. Using equation 2 we can compute the actual probabilities of the rare trajectories obtained in the simulations with the algorithm. See the SI and (Ragone et al., 2018) for more details.

We use as selection function  $A$  the same function used to define a heatwave in formula (1). We see from equation (2) that if the parameter  $k$  is positive, then trajectories with large values of the time average of the surface temperature anomaly will be much more probable in simulations obtained with the algorithm rather than in simple simulations with the model. The larger  $k$ , the stronger the selection, and the more probable trajectories with extreme values of the time averaged surface temperature. We analyse extremely warm summers over France and Scandinavia. For each case, we perform  $K = 10$  ensemble simulations with the algorithm, each with  $N=100$  trajectories, biasing parameter  $k=30$ , with  $T_a=90$  days from June 1st to August 29th, and resampling time  $\tau=5$  days (see the SI). A 1000 years long control run is used to provide initial conditions for the experiments with the algorithm, and as a benchmark for the statistics. The computational cost of the experiments with the algorithm is equivalent to simulating 1000 summers in the control run, but they allow to gather a much richer statistics for the extreme events of interest.

### 3 Results

#### 3.1 Importance sampling of extreme warm summers

The main goal of the algorithm is to perform importance sampling for the distribution of the summer temperature anomalies over the target region. Figures 1a and 1b show the probability distribution functions of the summer temperature anomalies  $a_{JJA}$  for the control run and the rare event algorithm. The control distribution shows a similar variance of about 1 °K for both France and Scandinavia. The algorithm is very effective in performing importance sampling and populating the upper tails of the distributions. The typical value of the seasonal anomaly  $a_{JJA}$  in the rare event algorithm experiments is around 4 °K for both cases, which are values never observed in the control run. Thanks to the algorithm most of the computational power is indeed used to simulate extremely warm summers, rather than trajectories belonging to the bulk of the distribution. In the case of Scandinavia the algorithm statistics in figure 1b shows a bimodality that we will discuss in section 3.3.

In figures 1c,d we compare return times of  $a_{JJA}$  in the control run and in the experiments with the rare event algorithm. A description of the computation of return times both for direct sampling and importance sampling using formula (2) is presented in (Lestang et al., 2018). The black curves are obtained from the control run using 1000 years of data. In order to estimate uncertainty ranges, we compute also an estimate dividing the control run in  $K=10$  samples of 100 years each, computing 10 estimates of the return times curve. We then take their average (blue curves), and compute the empirical standard deviation (blue shaded areas). The red line and shaded areas are obtained in the same way, but using the estimates from  $K = 10$  independent experiments with the algorithm.

With the rare event algorithm we reach return times up to  $10^5$ - $10^6$  years with uncertainty ranges comparable with the ones with the control run for return times of order  $10^2$  years. Given the large value of  $k$  we chose, we have a small range of overlap for estimates of the return times from the control run and the rare event algorithm. When they do overlap, the values are consistent with each other within the uncertainty ranges. After about  $10^6$ - $10^7$  years the return time curves reach a plateau. Such plateaux are due to undersampling, as discussed in the SI and in (Lestang et al., 2018). A full demonstration of the reliability of the results obtained with the rare event algorithm in similar simulations can be found in (Ragone et al., 2018; Ragone & Bouchet, 2020).

#### 3.2 Teleconnection patterns for extreme warm summers

We study the dynamical properties of extremely warm summers with return time larger than 100 years (called 100-year warm summers or seasonal heatwaves from now on). We compute composite maps of anomalies of the local JJA surface temperature and 500 hPa geopotential height conditional on the occurrence of 100-year warm summers, for the control run and the rare event algorithm experiments. One of the key advantage of the rare event algorithm is that it gives much better results than the control run for composite statistics for large return times. Moreover the rare event algorithm gives access to composite statistics for return times larger than 1000 years, which is impossible with the control run.

Figure 2a shows composite statistics of surface temperature and 500 hPa geopotential height of 100-year warm summers over France in the control run. The map hints at the presence of a wavenumber 3 teleconnection pattern for both observables. However, how much of these patterns, which were obtained by averaging over only 10 warm summers, is statistically significant? Figure 2b shows the  $t$ -value map for the geopotential height (see the SI for details). Only the geopotential height anomalies over Europe pass a statistical significance test with  $|t| > 2$ , which means then we can not really assess the reality of the teleconnection pattern using a 1000 years control run: we just do not have enough data. Figures 2c,d show the same composite maps and statistical significance analysis for 100-year warm summers over France, but computed using the rare event algorithm. The algorithm

results are globally significant, except over the northern part of the Pacific area. Transition areas between cyclonic and anticyclonic anomalies have a  $|t|$  value smaller than 2, because they have a low value of the conditional average. They are however rather narrow, so that their location can be still considered rather precise. The rare event algorithm thus allows to properly assess the existence of teleconnections of warm summers in Europe, North America and Central Asia.

The algorithm also gives a much better estimate of the amplitude of the anomalies, since it gives access to large amplitude events unavailable in the control run, because they are too rare. In the same way, the rare event algorithm gives also access to composite statistics that totally unavailable with the control run. For instance figure 3a shows composite statistics of surface temperature and 500 hPa geopotential height for 1000-year warm summers over Scandinavia, which also show a teleconnection pattern with a different spatial pattern but similar broad stroke features.

Based on the rare event algorithm results, we obtain a better overview of the characteristics of the synoptic dynamics occurring during extremely warm summers in the considered regions. A warming pattern centered over the target region is present, encompassing a larger area on a spatial scale of a few thousands km, which coincide with central-Western Europe for France and Northern-Eastern Europe for Scandinavia. Persistent anticyclonic synoptic scale structures centered over the target area are associated to this warming pattern. Locally these structures are consistent with the observed synoptic conditions for the occurrence of European heatwaves. In particular, for the North Atlantic and European area, the local patterns are very close to the observed patterns for the Western-European and the Scandinavian heatwave clusters obtained from reanalysis data in (Stéfanon et al., 2012).

The application of the rare event algorithm allows to assess in a statistically robust way that these local dynamics are part of hemispheric structures of approximately wavenumber 3, which induce a temperature teleconnection pattern, with persistent regional temperature anomalies of alternating signs around the hemisphere. The patterns for the two regions are in broad strokes similar, although they differ in the exact location of positive and negative anomalies. Warm summers over France occur systematically with warm summers over Siberia and North-East America. The corresponding tripolar structure of anticyclonic anomalies is accompanied by a localized low over Central-Asia and a general lower pressure over the Arctic, with minimum over Greenland. The structure related to Scandinavian warm summers is similar on the North-Atlantic sector, but over Asia it is quite different, with a negative temperature anomaly (and related low pressure) extending from Southern Europe to the whole central Asia, with the positive anomaly constrained over the far East. The circulation over the Arctic is also different, with a strong low over the Pole.

### 3.3 Bimodality of warm Scandinavian summers and subseasonal heatwaves

Figure 1b a bimodality of the Scandinavia summer temperatures in the rare event algorithm data. This suggests that two types of distinct dynamical events might lead to Scandinavian extreme summers. We note that, because of the nonlinear relation (2), a bimodality in the algorithm distribution does not necessarily imply a bimodality in the tail of the model distribution. Still it is interesting to test the hypothesis of two types of dynamics.

Figure 3a shows composite statistics of surface temperature and 500 hPa geopotential height for 1000-year warm summers over Scandinavia. Figures 3b and 3c show the composite maps from the rare event algorithm computed separately for  $a_{JJA} < 4.2^\circ K$  and  $a_{JJA} > 4.2^\circ K$ , where  $4.2^\circ K$  corresponds to the local minimum between the two peaks of the distribution in red in figure 1b. While the overall structure is the same as the total composite map in figure 3a, the map of the first range (figure 3b) shows a weaker teleconnection pattern, and an anticyclonic anomaly over the North-Atlantic that is not present in the map of the second range (figure 3c). The Scandinavian warm summer dynamics is compatible with a



northward shift of the jet stream along the entire hemisphere. In the case of the first range (figure 3b), this shift is less clear over the North-Atlantic, where a different dynamics seems to be in place. It is however likely that, if two distinct types of dynamics are occurring, a selection based only on the range of seasonal regional temperature fluctuations is not enough to differentiate between different dynamics, and therefore simple composite maps show from that point of view mixed information.

The study of relation between subseasonal fluctuations of surface temperature and extreme warm summers can give an explanation of this bimodality. To visualize this relation, we look at the genealogical structure of the trajectories in an experiment with the rare event algorithm for Scandinavia in figure 4a. The black lines correspond to trajectories belonging to the first range ( $a_{JJA} < 4.2^\circ\text{K}$ ), while the red lines to trajectories belonging to the second range ( $a_{JJA} > 4.2^\circ\text{K}$ ). In the experiment represented in figure 4a, trajectories belonging to each of the two ranges of the bimodal distribution coexist. In figure 4b, we show the same genealogical tree, but indicate with different colors the values of the 5-days temperature anomaly in each segment. These values are computed using equation 1 with  $T=5$  days and  $t_I$  corresponding to the date of each resampling event. In particular red segments indicate a 5-day anomaly above  $4.5^\circ\text{K}$ . The chosen threshold of  $4.5^\circ\text{K}$  is the median of the distribution of the 5-day temperature anomalies rare event algorithm experiments for Scandinavia (shown in the SI). This value corresponds to the 96.6th percentile of the distribution of the 5-day temperature anomalies in the control run. We can thus consider 5-days periods with temperature anomaly larger than this threshold as heatwave periods, and a succession of 2 or more consecutive heatwave periods as a subseasonal heatwave.

With this definitions, during warmer summers belonging to the second range (red in figure 4a), we see a first subseasonal heatwave that lasts about 20 days in late June-early July, and then a second subseasonal heatwave that last again about 15 days in August. Less extreme summers (black in figure 4a) have instead only one subseasonal heatwave in June-July. The bimodality of the algorithm statistics therefore corresponds to two qualitatively different types of warm summer, with either one or two subseasonal heatwaves. We note that the 2003 warm summer in France was characterized by two separate subseasonal heatwaves in June and August (García-Herrera et al., 2010).

## 4 Conclusions

We have shown how simulations with the algorithm produce hundreds of times more extremes than a control run for the same computational cost, and allow us to estimate return times of extreme events orders of magnitude larger. This allows to compute precisely statistically significant composite maps of dynamical quantities for warm summers with extremely strong seasonal surface temperature anomalies. In this way we are able to identify rigorously the occurrence of persistent teleconnection patterns of wavenumber 3 during very warm summers. These patterns and the related dynamics can not be properly studied with direct sampling, as the corresponding local anomalies of temperature and geopotential height are statistically significant only over the target region. The improved statistics given by the rare event algorithm instead allows to obtain statistically significant maps and a very clear signal also away from the target region.

Teleconnection patterns similar to what we have obtained here were observed in the first application of the rare event algorithm to a climate model (Ragone et al., 2018). In that case the model was an intermediate complexity model, run in perpetual summer setup, and the target area was the whole Europe. In this case we use a much more realistic model with seasonal cycle, calibrated to reproduce present day climate, and we target areas consistent with observed heatwave clusters (Stéfanon et al., 2012) and recent cases of very intense heatwaves or warm summer, like the heatwave over France of 2003 and the warm summer over Northern Europe of 2018. It is striking that we obtain similar qualitative results. This supports the idea that these teleconnections at seasonal scale and the corresponding

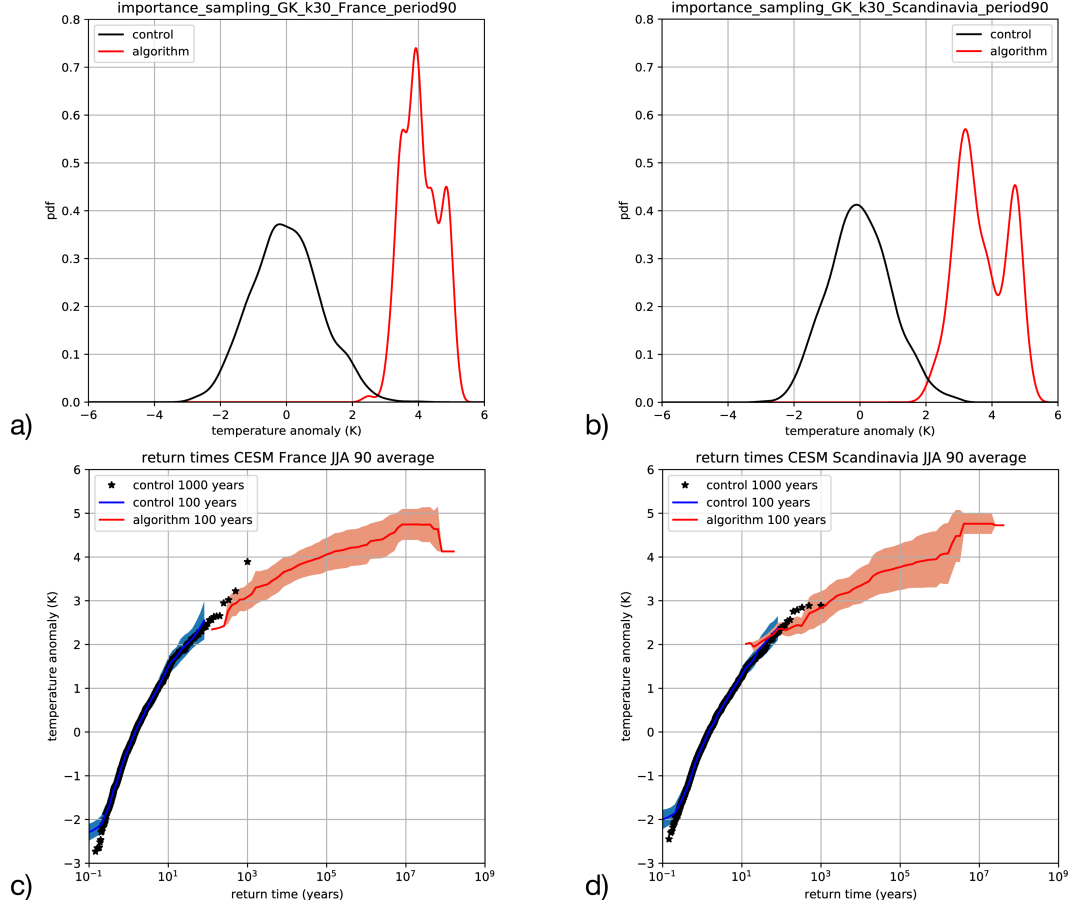


wavenumber 3 dynamics are robust features, mainly dynamical and little affected by the details of the physical parameterizations.

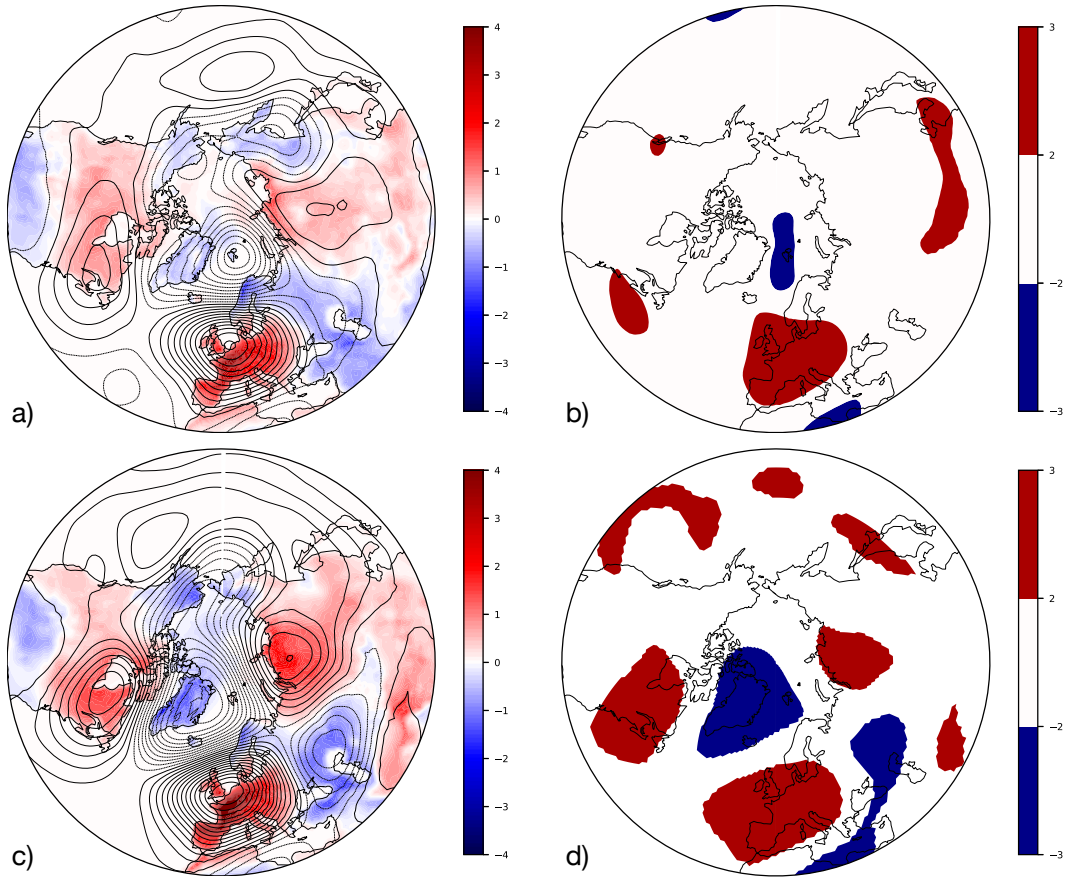
Comparison with observations is not straightforward, due to the lack of data and the fact that most studies have focused on events at subseasonal scale. Several authors have recently highlighted the role of Rossby waves in determining teleconnection of extreme events (Schubert et al., 2011; Lau & Kim, 2012; Petoukhov et al., 2013, 2016), with particular emphasis on strong heatwaves, e.g. cases over the central USA (Teng et al., 2013), Alberta (Petoukhov et al., 2018) and Western Europe (Kornhuber et al., 2019). These studies typically find patterns with wavenumber 5 to 7 and analyse subseasonal temperature fluctuations. Several different detection methods are used in the literature to identify the dynamics leading to heatwaves, including projections of heatwave states on typical variability patterns obtained with cluster analysis (Cassou et al., 2005) or empirical orthogonal functions (Teng et al., 2013), spectral analysis (Petoukhov et al., 2018; Kornhuber et al., 2019) and/or indicators based on resonance models (Petoukhov et al., 2013, 2016). Merging these approaches with rare event simulations should lead to very promising future studies, aimed at investigating the relation between the seasonal scale structures we find here and the possible role of higher wavenumber stationary Rossby waves proposed in the literature, as well as the possible multimodality in the way of occurrence of some extreme events.

### Acknowledgments

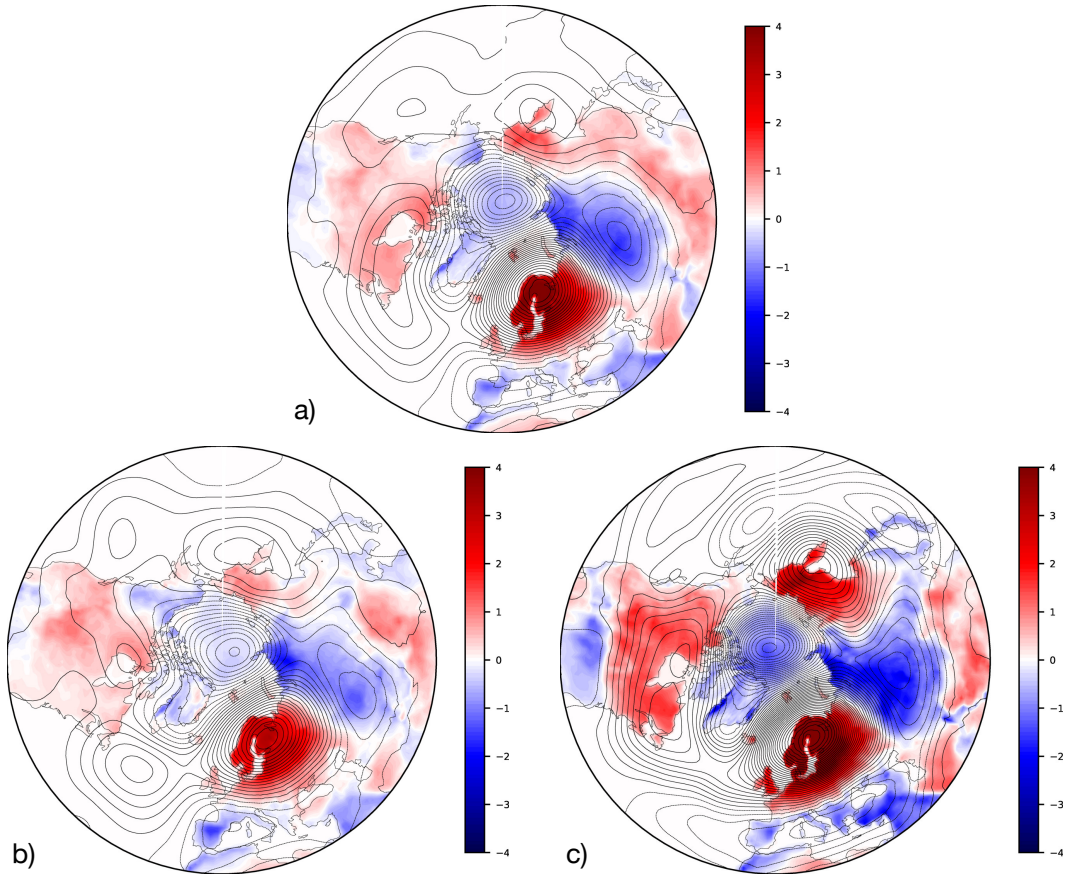
This work was granted access to the HPC resources of CINES under the allocation 2019-A0070110575 made by GENCI. The computation of this work were partially performed on the PSMN platform of ENS de Lyon. This work has received funding through the ACADEMICS grant of the IDEXLYON, project of the Université de Lyon, PIA operated by ANR-16-IDEX-0005. The research leading to these results has received funding from the European Research Council under the European Union’s seventh Framework Programme (FP7/2007-2013 Grant Agreement No. 616811). The data related to the figures presented in this paper have been uploaded with the submission for review purposes. We are discussing with CINES the long term storage and open access of the full dataset of our simulations according to the Enabling FAIR data Project guidelines.



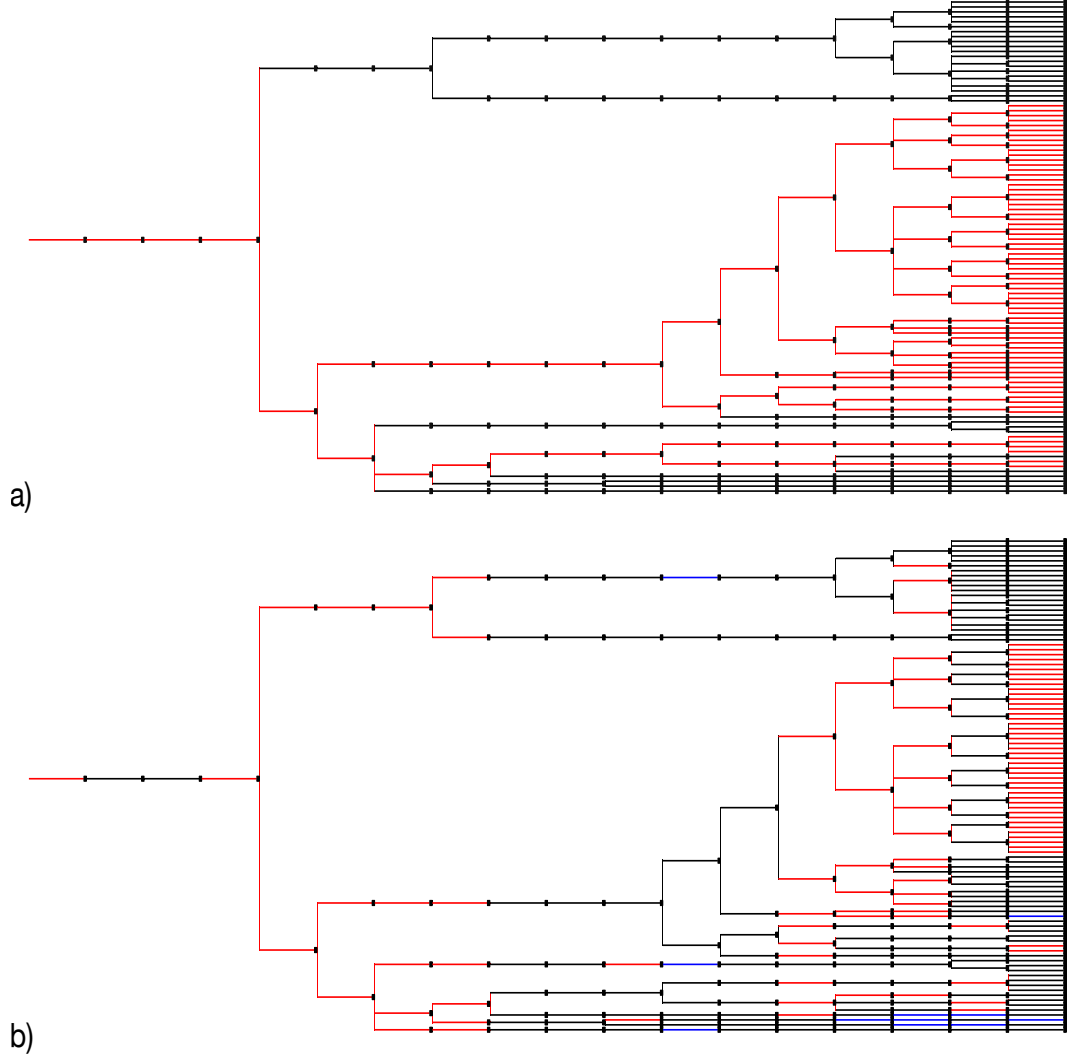
**Figure 1.** Distribution of seasonal JJA temperatures anomalies averaged over France (a) and Scandinavia (b) from the control run (black) and the rare event algorithm (red). Return times for the seasonal temperature anomalies for France (c) and Scandinavia (d), from the control run (black) and the rare event algorithm (red). The shaded areas, light blue for the control run and light red for the algorithm, correspond to 1 standard deviation of the sample used to compute the estimate (see SI).



**Figure 2.** Northern hemisphere composite maps (conditional statistics) for the JJA anomalies of the surface temperature (colors) and 500 hPa geopotential height (contours) for 100-year warm summers over France, in the control run (a) and in the rare event algorithm statistics (c). Panels b) and d) show for the control run and the rare event algorithm respectively the corresponding maps of the  $t$  values of the 500 hPa geopotential height anomalies (see the SI).



**Figure 3.** Composite maps for the anomalies of the surface temperature (colors) and 500 hPa geopotential height (contours) for warm summers over Scandinavia with return times larger than 1000 years (a). Panel b) and c) show respectively composite maps for the same variables with the condition of JJA temperature anomaly smaller or higher than 4.2°K.



**Figure 4.** Panel a) shows the genealogical tree of a rare event algorithm experiment, with  $N = 100$  ensemble members, for selection of Scandinavia seasonal heatwaves. Each broken line, composed by one trunk and its branches up to the last leaf, represents a trajectory. The horizontal axis represents time in blocks of  $\tau = 5$  days. The genealogical trees have been created with the help of the ETE Python toolkit (Huerta-Cepas et al., 2016). Black lines represent trajectories with  $a_{JJA} < 4.2^\circ\text{K}$ , red lines trajectories with  $a_{JJA} > 4.2^\circ\text{K}$ . Panel b): the same, but with different colors for each 5 day period: blue corresponds to negative 5-day time averaged anomalies, black between 0 and  $4.5^\circ\text{K}$ , and red above  $4.5^\circ\text{K}$ .

## References

- Agha Kouchak, A. (2012). *Extremes in a changing climate detection, analysis and uncertainty*. Dordrecht; New York: Springer.
- Barriopedro, D., Fischer, E., Luterbacher, J., Trigo, R., & Garcia-Herrera, R. (2011). Redrawing the temperature record map of Europe. *Science*, *332*, 220–224.
- Boers, N., Goswami, B., & Rheinwalt, A. e. a. (2019). Complex networks reveal global pattern of extreme-rainfall teleconnections. *Nature*, *566*, 373–377. doi: 10.1038/s41586-018-0872-x
- Bouchet, F., Rolland, J., & Simonnet, E. (2019). Rare event algorithm links transitions in turbulent flows with activated nucleations. *Physical Review Letters*, *122*(7), 074502.
- Camargo, S. J., & Seth, A. (2016). Hottest summers the new normal. *Environ. Res. Lett.*, *11*, 081001.
- Cassou, C., Terray, L., & Phillips, A. S. (2005). Tropical atlantic influence on European heat waves. *Journal of Climate*, *18*(15), 2805–2811. doi: 10.1175/JCLI3506.1
- Coumou, D., & Rahmstorf, S. (2012). A decade of weather extremes. *Nature Climate Change*, *2*(7), 491–496. doi: 10.1038/nclimate1452
- D’Andrea, F., Drobinski, P., & Stefanon, M. (2016). European heat waves: the effect of soil moisture, vegetation, and land use. In *Dynamics and predictability of large-scale, high-impact weather and climate events* (p. 185–197). Cambridge University Press.
- Della-Marta, P. M., Luterbacher, J., von Weissenfluh, H., Xoplaki, E., Brunet, M., & Wanner, H. (2007). Summer heat waves over western Europe 1880–2003, their relationship to large-scale forcings and predictability. *Climate Dynamics*, *29*(2–3), 251–275. doi: 10.1007/s00382-007-0233-1
- Del Moral, P., Garnier, J., et al. (2005). Genealogical particle analysis of rare events. *The Annals of Applied Probability*, *15*(4), 2496–2534.
- Dosio, A., Mentaschi, L., Fischer, E. M., & Wyser, K. (2018). Extreme heat waves under 1.5 c and 2 c global warming. *Environ. Res. Lett.*, *13*, 054006. doi: 10.1088/1748-9326/aab827
- Ebener, L., Margazoglou, G., Friedrich, J., Biferale, L., & Grauer, R. (2019). Instanton base importance sampling for rare events in stochastic pdes. *Chaos: An Interdisciplinary Journal of Nonlinear Science*.
- García-Herrera, R., Díaz, J., Trigo, R. M., Luterbacher, J., & Fischer, E. M. (2010). A Review of the European Summer Heat Wave of 2003. *Critical Reviews in Environmental Science and Technology*, *40*(4), 267–306. doi: 10.1080/10643380802238137
- Giardina, C., Kurchan, J., Lecomte, V., & Tailleur, J. (2011). Simulating Rare Events in Dynamical Processes. *Journal of Statistical Physics*, *145*(4), 787–811. doi: 10.1007/s10955-011-0350-4
- Grafke, T., Grauer, R., & Schäfer, T. (2015). The instanton method and its numerical implementation in fluid mechanics. *J.Phys.A:Math.Theor.*, *48*, 333001.
- Horton, R., Mankin, J., & Lesk, C. e. a. (2016). A review of recent advances in research on extreme heat events. *Curr Clim Change Rep*, *2*, 242–259. doi: 10.1007/s40641-016-0042-x
- Hoskins, B., & Woollings, T. (2015). Persistent extratropical regimes and climate extremes. *Current Climate Change Reports*, *1*(3), 115–124. doi: 10.1007/s40641-015-0020-8
- Huerta-Cepas, J., Serra, F., & Bork, P. (2016). Ete 3: Reconstruction, analysis, and visualization of phylogenomic data. *Molecular Biology and Evolution*, *33*(6), 1635–1638. doi: 10.1093/molbev/msw046
- Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., ... Marshall, S. (2013). The community earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, *94*(9), 1339–1360. doi: 10.1175/BAMS-D-12-00121.1
- IPCC. (2012). *Managing the risks of extreme events and disasters to advance climate change adaption: special report of the Intergovernmental Panel on Climate Change*. New York, NY: Cambridge University Press.



- IPCC. (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.
- Jezequel, A., Yiou, P., & Radanovics, S. (2018). Role of circulation in european heatwaves using flow analogues. *Clim Dyn*, 50, 1145–1159. doi: 10.1007/s00382-017-3667-0
- Kornhuber, K., Osprey, S., Coumou, D., Petri, S., Petoukhov, V., Rahmstorf, S., & Gray, L. (2019). Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern. *Environmental Research Letters*, 14(5), 054002. doi: 10.1088/1748-9326/ab13bf
- Lau, W., & Kim, K.-M. (2012, 02). The 2010 pakistan flood and russian heat wave: Teleconnection of hydrometeorological extremes. *Journal of Hydrometeorology*, 13, 392-403. doi: 10.1175/JHM-D-11-016.1
- Laurie, J., & Bouchet, F. (2015). Computation of rare transitions in the barotropic quasi-geostrophic equations. *New J. of Phys.*, 17, 015009.
- Lestang, T., Bouchet, F., & L  v  que, E. (2020). Numerical study of extreme mechanical force exerted by a turbulent flow on a bluff body by direct and rare-event sampling techniques. *Journal of Fluid Mechanics*, 895.
- Lestang, T., Ragone, F., Br  hier, C.-E., Herbert, C., & Bouchet, F. (2018). Computing return times or return periods with rare event algorithms. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(4), 043213. doi: 10.1088/1742-5468/aab856
- Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., & Wanner, H. (2004). European seasonal and annual temperature variability, trends, and extremes since 1500. *Science*, 303(5663), 1499-1503. doi: 10.1126/science.1093877
- Mueller, B., Zhang, X., & Zwiers, F. W. (2016). Historically hottest summers projected to be the norm for more than half of the world’s population within 20 years. *Environ. Res. Lett.*, 11, 044011.
- Otto, F., Massey, N., Van Oldenborgh, G. J., Jones, R., & Allen, M. (2012, 02). Reconciling two approaches to attribution of the 2010 russian heat wave. *Geophysical Research Letters*, 39, 4702-. doi: 10.1029/2011GL050422
- 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. doi: 10.1016/j.atmosres.2015.05.014
- Petoukhov, V., Petri, S., Kornhuber, K., Thonicke, K., Coumou, D., & Schellnhuber, H. (2018). Alberta wildfire 2016: Apt contribution from anomalous planetary wave dynamics. *Scientific Reports*, 8. doi: 10.1038/s41598-018-30812-z
- Petoukhov, V., Petri, S., Rahmstorf, S., Coumou, D., Kornhuber, K., & Schellnhuber, H. J. (2016). Role of quasi-resonant planetary wave dynamics in recent boreal spring-to-autumn extreme events. *Proc Natl Acad Sci USA*, 113(25), 6862–6867.
- Petoukhov, V., Rahmstorf, S., Petri, S., & Schellnhuber, H. J. (2013). Quasiresonant amplification of planetary waves and recent northern hemisphere weather extremes. *Proc Natl Acad Sci USA*, 110(4), 5336–5341.
- Pfleiderer, P., Schleussner, C., & Kornhuber, K. e. a. (2019). Summer weather becomes more persistent in a 2 c world. *Nat. Clim. Chang.*, 9, 666-671. doi: 10.1038/s41558-019-0555-0
- Plotkin, D. A., Webber, R. J., O’Neill, M. E., Weare, J., & Abbot, D. S. (2019). Maximizing simulated tropical cyclone intensity with action minimization. *Journal of Advances in Modeling Earth Systems*, 11(4), 863–891.
- Ragone, F., & Bouchet, F. (2020). Computation of extreme values of time averaged observables in climate models with large deviation techniques. *J Stat Phys*, 179, 1637–1665. doi: 10.1007/s10955-019-02429-7
- Ragone, F., Wouters, J., & Bouchet, F. (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. doi: 10.1073/pnas.1712645115
- Russo, S., Sillmann, J., & Fischer, E. M. (2015). Top ten european heatwaves since 1950



- and their occurrence in the coming decades. *Environmental Research Letters*, 10(12), 124003. doi: 10.1088/1748-9326/10/12/124003
- Schubert, S., Wang, H., Koster, R., Suarez, M., & Groisman, P. (2014). Northern Eurasian Heat Waves and Droughts. *Journal of Climate*, 27(9), 3169-3207. doi: 10.1175/JCLI-D-13-00360.1
- Schubert, S., Wang, H., & Suarez, M. (2011). Warm season subseasonal variability and climate extremes in the northern hemisphere: the role of stationary rossby waves. *J Clim*, 24(8), 4773-4792.
- Stéfanon, M., D’Andrea, F., & Drobinski, P. (2012). Heatwave classification over europe and the mediterranean region. *Environmental Research Letters*, 7, 014023. doi: 10.1088/1748-9326/7/1/014023
- Suarez-Gutierrez, L., Li, C., Müller, W. A., & Marotzke, J. (2018). Internal variability in european summer temperatures at 1.5 c and 2 c of global warming. *Environmental Research Letters*, 13(6), 064026. doi: 10.1088/1748-9326/aaba58
- Teng, H., Branstator, G., Wang, H., Meehl, G., & Washington, W. (2013, 10). Probability of us heat waves affected by a subseasonal planetary wave pattern. *Nature Geoscience*, 6, 1056-1061. doi: 10.1038/ngeo1988
- Trenberth, K. (2012). Framing the way to relate climate extremes to climate change. *Clim. Change*, 115, 283-290.
- Vautard, R., Yiou, P., D’Andrea, F., de Noblet, N., Viovy, N., Cassou, C., ... Fan, Y. (2007). Summertime european heat and drought waves induced by wintertime mediterranean rainfall deficit. *Geophysical Research Letters*, 34(7). doi: 10.1029/2006GL028001
- Webber, R. J., Plotkin, D. A., O’Neill, M. E., Abbot, D. S., & Weare, J. (2019). Practical rare event sampling for extreme mesoscale weather. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(5), 053109. doi: 10.1063/1.5081461
- Zampieri, M., D’Andrea, F., Vautard, R., Ciais, P., de Noblet-Ducoudré, N., & Yiou, P. (2014). Hot European summers and the role of soil moisture in the propagation of mediterranean drought. *J. Climate*, 22, 4747-4758. doi: 10.1175/2009JCLI2568.1

# Supporting Information for ”Rare event algorithm study of extreme warm summers and heatwaves over Europe”

Francesco Ragone<sup>1</sup> and Freddy Bouchet<sup>1</sup>

<sup>1</sup>Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, Lyon, France

**Introduction** In these Supporting Informations we present a more detailed description of the theory and the implementation of the algorithm, a more detailed description of the climate model we use and of the experiments we have performed, and some comments on data analysis. We also show additional figures for the definition of the areas we apply the rare event algorithm to, and for the importance sampling for subseasonal fluctuations.

---

Corresponding author: Francesco Ragone, Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, Lyon, France (francesco.ragone@uclouvain.be)

**Description of the algorithm** Rare event algorithms have been devised long ago (Kahn & Harris, 1951; Del Moral, 2004) in order to lower the numerical cost of rare events. In a dynamical context, they have been first applied for complex and bio-molecules, see for instance (Metzner et al., 2009; Hartmann et al., 2013). More recently, in order to progressively go towards genuine geophysical applications, they have been applied to Lorenz models (Wouters & Bouchet, 2016), partial differential equations (Rolland et al., 2016), turbulence problems (Grafke et al., 2015; Laurie & Bouchet, 2015; Ebener et al., 2019; Bouchet et al., 2019; Lestang et al., 2018, 2020), geophysical fluid dynamics (Bouchet et al., 2019), and climate applications (Ragone et al., 2018; Webber et al., 2019; Ragone & Bouchet, 2020; Plotkin et al., 2019). Some approaches through minimum action methods, related to large deviation theory, are reviewed in (Grafke & Vanden-Eijnden, 2019).

Based on the phenomenology of the dynamics for each family of extreme events, an appropriated type of rare event algorithm should be carefully chosen. In this paper, in order to study long-lasting events, we use an algorithm of the family of genealogical algorithms, very close in spirit to the initial one (Kahn & Harris, 1951), and the one described by (Del Moral et al., 2005) and mathematically comprehensively studied in (Del Moral, 2004). The idea is to perform an ensemble simulation, stop regularly every resampling time, and based on a score function select some trajectories to clone them and discard some others. The ensemble statistics is then tilted towards the events of interest.

Choosing well the score function in order to define the selection criteria is critical. For this algorithm, we use a score function which is the time integral of an observable, as for instance in (Giardina et al., 2011). In (Giardina et al., 2011), the algorithm was specifically designed for studying long time large deviations. This is suited for long-lasting events. In this letter, we use a setup similar to the one in (Giardina et al., 2011), with the difference that our simulations, like in (Del Moral et al., 2005) are transient ones (they depend from

the initial state and extend over finite time) rather than statistically stationary. We call this algorithm the Del-Moral Garnier, or Giardina-Kurchan-Tailleur-Lecomte algorithm, or simply the rare event algorithm.

We have first proposed this rare event algorithm to study heatwaves and generally long-lasting extreme events in (Ragone et al., 2018). Briefly, the algorithm consists in running an ensemble simulation of  $N$  ensemble members or trajectories with a climate model starting from different initial conditions. Let us indicate with  $\vec{X}_n(t)$  the state vector of trajectory  $n$  at time  $t$ , and with  $A(t) = A(\{\vec{X}(t)\})$  an observable of interest function of the state of the system, of whose time average we want to study the extremes (e.g. the surface temperature over a region). At constant intervals of a fixed resampling time  $\tau$ , each trajectory is assigned a weight, which determines if that trajectory is killed or if it continues its evolution, possibly spawning copies of itself. The weight  $w_i^n$  of trajectory  $n$  at time  $t_i = i\tau$  is computed as

$$w_i^n = \frac{e^{k \int_{t_{i-1}}^{t_i} A(\{\vec{X}_n(t)\}) dt}}{\frac{1}{N} \sum_{n=1}^N e^{k \int_{t_{i-1}}^{t_i} A(\{\vec{X}_n(t)\}) dt}}, \quad (1)$$

where  $k$  is a parameter that control how stringent is the selection. At times  $t_i$  the simulation is stopped, the weights are computed, and each trajectory generates a number of copies of itself proportional to the weight assigned to it. A small random perturbation is added to each copied trajectory, so that they evolve differently from the original trajectory and from each other. Trajectories with weights smaller than 1 are killed, and do not continue their evolution. Killed trajectories are substituted in the ensemble by the copies of the surviving trajectories. In practice this resampling is done in such a way that the the number of trajectories in the ensemble always remains constant and equal to  $N$ . See (Ragone et al., 2018; Ragone & Bouchet, 2020) for details about the technical implementation.

The weights are defined in such a way that trajectories featuring large values of the time average of the observable are the ones that will populate more densely the ensemble. After a total running time  $T_a$ , the probability  $\mathbb{P}_k(\{\vec{X}(t)\})$  of observing a trajectory in the ensemble generated by the algorithm is related to the probability  $\mathbb{P}_0(\{\vec{X}(t)\})$  of observing the same trajectory in a normal ensemble simulation with no resampling as

$$\mathbb{P}_k(\{\vec{X}(t)\}) = \frac{e^{k \int_{t_a}^{t_a+T_a} A(u) du}}{Z} \mathbb{P}_0(\{\vec{X}(t)\}), \quad (2)$$

where  $Z$  is a normalization term, and the equation is valid in the limit of large  $N$  with relative errors on the computation of expectation values of the order of  $1/\sqrt{N}$  (Del Moral, 2004). In ensemble simulations performed with the rare event algorithm, trajectories featuring a large value of the time average of the control observable over the entire simulation period are thus much more likely to be observed than in a normal simulation, and viceversa. The value of  $k$  determines how strong is the tilt of the probability distribution. This is called *importance sampling*. More details on the theory behind the method and the reconstruction of the effective ensemble of trajectories, as well as on the choice of  $N$ ,  $k$  and  $\tau$ , can be found in (Ragone et al., 2018; Ragone & Bouchet, 2020).

**Description of the climate model and of the control run** The simulations are performed with the Community Earth System Model (CESM) version 1.2.2 (Hurrell et al., 2013). We use an atmosphere and land only setup, whose active components are the Community Atmospheric Model version 4 (CAM4) and the Community Land Model version 2 (CLM2). CAM4 is a widely used model of the Earth’s atmosphere. The dynamical core is a finite volume discretisation of the Navier-Stokes equations for a planetary fluid envelope in hydrostatic approximation. The physics suite consists of parameterisations of moist precipitation processes, clouds and radiation, turbulent mixing, and surface fluxes. CLM2 is a model of the Earth’s land surface. It represents several aspects of the land surface, including surface heterogeneity, and consists of components related to land biogeophysics, hydrologic cycle, biogeochemistry, human dimensions, and ecosystem dynamics. The horizontal resolution is 0.9 and 1.25 degrees in latitude and longitude respectively, and we use 26 vertical layers in hybrid pressure coordinates. The output is sampled every 3 hours. The model is run at statistically stationary state, with sea surface temperatures (SST), sea ice cover, and the concentration of atmospheric CO<sub>2</sub> and other greenhouse gases fixed at values representative of present day climate (year 2000). Contrary to our previous proof of concept (Ragone et al., 2018), the model features daily and seasonal cycles. We have performed a control run of 1000 years, as a set of  $K=10$  independent 100 years long simulations. The control run is used as a benchmark for the statistics, in order to evaluate the performances of the algorithm.

**Description of the experiments with the rare event algorithm** We perform two sets of simulations, setting  $D$  as the area over France and Scandinavia. For both cases we perform  $K=10$  ensemble simulations, each with  $N=100$  trajectories, biasing parameter set to  $k=30$ , and running for  $T_a=90$  days from June 1st to August 29th. The resampling time  $\tau$  in general should be of the order of magnitude of the decorrelation time of the dynamics (see the discussion in (Ragone et al., 2018; Ragone & Bouchet, 2020)). In this case the dominant dynamics relevant for heatwaves acts at synoptic timescale, thus we set  $\tau=5$  days. In order to allow copies of the same trajectory to evolve differently, we add a small perturbation to each trajectory after the resampling. The perturbation is added multiplying at each vertical level the amplitudes of the spherical harmonics of order  $(m, n)$  of the potential temperature field by  $\gamma = 1 + \epsilon r_{m,n}$ , where  $\epsilon = 10^{-4}$  and the  $r_{m,n}$  are uniform random numbers extracted between -1 and 1. The initial conditions for the 1000 trajectories are taken from the 1st of June of the 1000 years of the control run. For a choice of  $D$  each ensemble starts therefore from a set of 100 different initial conditions.



**Computation of error bars** In order to compute error bars for a quantity  $f$ , except otherwise stated, we perform a set of  $K$  independent experiments and compute an empirical standard deviation, denoted  $\sigma(f)$ , based on these  $K$  samples. For instance, from the control run, we divide the 1000-year data in  $K = 10$  samples of 100-years. For the rare event algorithm, we have  $K = 10$  independent realizations. We plot composite maps which are the averages of a quantity  $f$ , conditional on a heat wave to occur ( $f$  will be either surface temperature or geopotential height). We estimate the statistical significance of the average of a quantity  $f$  by computing its  $t$  value:  $t = \sqrt{K}\mathbb{E}(f)/\sigma(f)$ , where  $\mathbb{E}(f)$  is the empirical average.  $t$  is a test of the null hypothesis that  $\mathbb{E}(f) \neq 0$ . For instance if  $t > 2$ , our interpretation of the test is that the probability that  $\mathbb{E}(f) \neq 0$  and positive is larger or equal to 97%.

**Plateau on return time plots.** The plateau observed on return time plots for return times larger than  $10^6$  years are due to undersampling by the algorithm. Such a plateau is caused by the increasing multiplicity of trajectories with the same ancestor, as the amplitude  $a$  increases. Indeed, because of the selection procedure involved in the rare event algorithm, a subset of trajectories can share the same ancestor. Henceforth, they are likely to differ only by a small time-interval at the end of their whole duration. As a consequence, this subset of trajectories will contribute the same value to the set of observable maxima from which return times are computed, leading to a plateau. This feature is further discussed in (Lestang et al., 2018).

**Lack of independence.** As said, a consequence of the cloning mechanics is that the trajectories tend to have common ancestors in the first part of the simulation. This is visible in figure 4 of the main text. From a statistical point of view, the efficiency of the algorithm comes from a trade-off between the reduction of the independence among the trajectories and the shift of the distribution. In this example all the trajectories come

from a common ancestor, and differentiation starts after the first 20 days. All the 10 sets of experiments show a similar behaviour, with the number of original ancestors varying between 1 and 3. Since the 10 sets of experiments start from pools of independent initial conditions, this means that even at the beginning of the simulation we always have at least 10-20 independent trajectories when we take composite averages.

## References

- Bouchet, F., Rolland, J., & Simonnet, E. (2019). Rare event algorithm links transitions in turbulent flows with activated nucleations. *Physical Review Letters*, 122(7), 074502.
- Del Moral, P. (2004). *Feynman-Kac Formulae Genealogical and Interacting Particle Systems with Applications*. New York, NY: Springer New York.
- Del Moral, P., Garnier, J., et al. (2005). Genealogical particle analysis of rare events. *The Annals of Applied Probability*, 15(4), 2496–2534.
- Ebener, L., Margazoglou, G., Friedrich, J., Biferale, L., & Grauer, R. (2019). Instanton base importance sampling for rare events in stochastic pdes. *Chaos: An Interdisciplinary Journal of Nonlinear Science*.
- Giardina, C., Kurchan, J., Lecomte, V., & Tailleur, J. (2011). Simulating Rare Events in Dynamical Processes. *Journal of Statistical Physics*, 145(4), 787–811. doi: 10.1007/s10955-011-0350-4
- Grafke, T., Grauer, R., & Schäfer, T. (2015). The instanton method and its numerical implementation in fluid mechanics. *J.Phys.A:Math.Theor.*, 48, 333001.
- Grafke, T., & Vanden-Eijnden, E. (2019). Numerical computation of rare events via large deviation theory. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(6), 063118.
- Hartmann, C., Banisch, R., Sarich, M., Badowski, T., & Schütte, C. (2013, December). Characterization of Rare Events in Molecular Dynamics. *Entropy*, 16(1), 350–376. Retrieved 2015-11-18, from <http://www.mdpi.com/1099-4300/16/1/350> doi: 10.3390/e16010350
- Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., ... Marshall, S. (2013). The community earth system model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, 94(9), 1339–

1360. doi: 10.1175/BAMS-D-12-00121.1

Kahn, H., & Harris, T. (1951). Estimation of particle transmission by random sampling.

*Natl. Bur. Stand., Appl. Math. Ser.*, 12, 27–30.

Laurie, J., & Bouchet, F. (2015). Computation of rare transitions in the barotropic quasi-geostrophic equations. *New J. of Phys.*, 17, 015009.

Lestang, T., Bouchet, F., & L  v  que, E. (2020). Numerical study of extreme mechanical force exerted by a turbulent flow on a bluff body by direct and rare-event sampling techniques. *Journal of Fluid Mechanics*, 895.

Lestang, T., Ragone, F., Br  hier, C.-E., Herbert, C., & Bouchet, F. (2018). Computing return times or return periods with rare event algorithms. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(4), 043213. doi: 10.1088/1742-5468/aab856

Metzner, P., Sch  tte, C., & Vanden-Eijnden, E. (2009). Transition path theory for markov jump processes. *Multiscale Modeling & Simulation*, 7(3), 1192–1219.

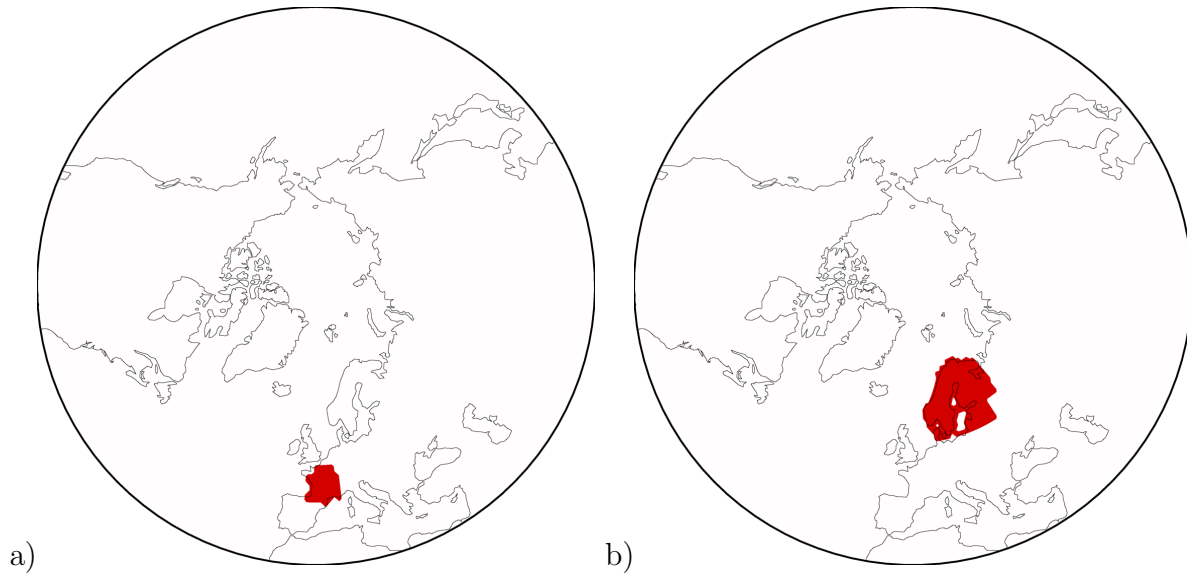
Plotkin, D. A., Webber, R. J., O’Neill, M. E., Weare, J., & Abbot, D. S. (2019). Maximizing simulated tropical cyclone intensity with action minimization. *Journal of Advances in Modeling Earth Systems*, 11(4), 863–891.

Ragone, F., & Bouchet, F. (2020). Computation of extreme values of time averaged observables in climate models with large deviation techniques. *J Stat Phys*, 179, 1637–1665. doi: 10.1007/s10955-019-02429-7

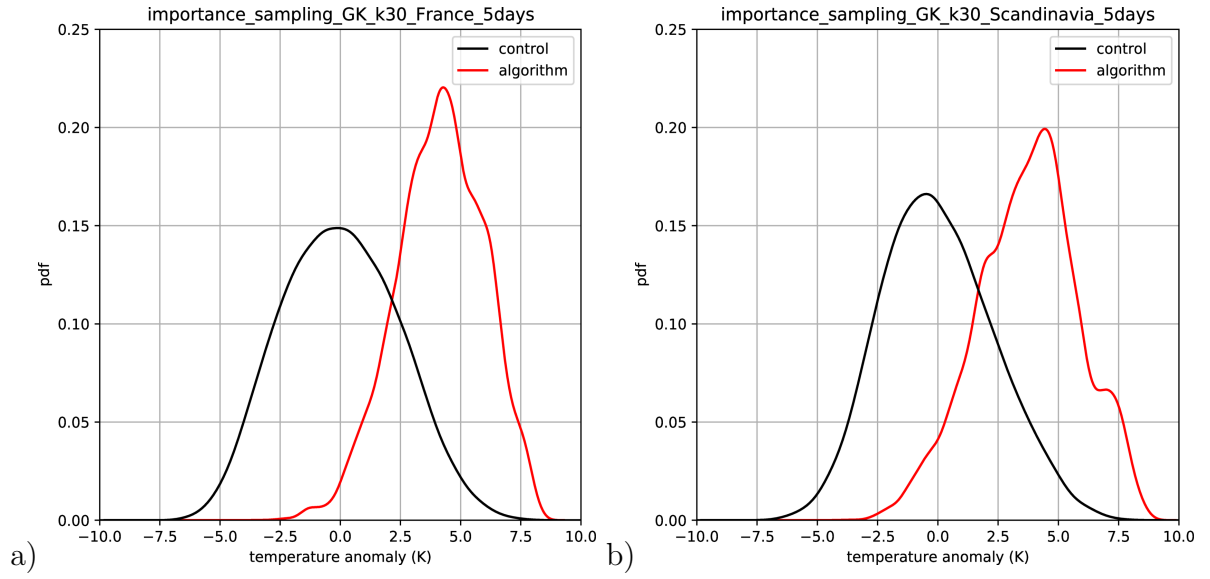
Ragone, F., Wouters, J., & Bouchet, F. (2018). Computation of extreme heat waves in climate models using a large deviation algorithm. *Proceedings of the National Academy of Sciences*, 115(1), 24-29. doi: 10.1073/pnas.1712645115

Rolland, J., Bouchet, F., & Simonnet, E. (2016). Computing transition rates for the 1-D stochastic Ginzburg–Landau–Allen–Cahn equation for finite-amplitude noise with a rare event algorithm . *J. Stat. Phys.*, 162, 277–311.

- Webber, R. J., Plotkin, D. A., O'Neill, M. E., Abbot, D. S., & Weare, J. (2019). Practical rare event sampling for extreme mesoscale weather. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(5), 053109. doi: 10.1063/1.5081461
- Wouters, J., & Bouchet, F. (2016). Rare event computation in deterministic chaotic systems using genealogical particle analysis. *Journal of Physics A: Mathematical and Theoretical*, 49(37), 374002.



**Figure S1.** Domain  $D$  used to define the area integral in the selection function of the algorithm, for France (a) and Scandinavia (b).



**Figure S2.** Distribution of 5 days running mean temperatures anomalies averaged over France (a) and Scandinavia (b) from a control run of 1000 years (black) and a set of 10 experiments with the algorithm with target domain the corresponding region, with  $k=30$ , trajectories  $N=100$ , resampling time 5 days, and running time 90 days starting from June 1st (red).