University of Reading

Typicality of the 2021 Western North America summer heatwave

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Lucarini, V. ORCID: https://orcid.org/0000-0001-9392-1471, Galfi, V. M., Riboldi, J. and Messori, G. (2023) Typicality of the 2021 Western North America summer heatwave. Environmental Research Letters, 18 (1). 015004. ISSN 1748-9326 doi: https://doi.org/10.1088/1748-9326/acab77 Available at https://centaur.reading.ac.uk/109374/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1088/1748-9326/acab77

Publisher: Institute of Physics

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Typicality of the 2021 Western North America summer heatwave

To cite this article: Valerio Lucarini *et al* 2023 *Environ. Res. Lett.* **18** 015004

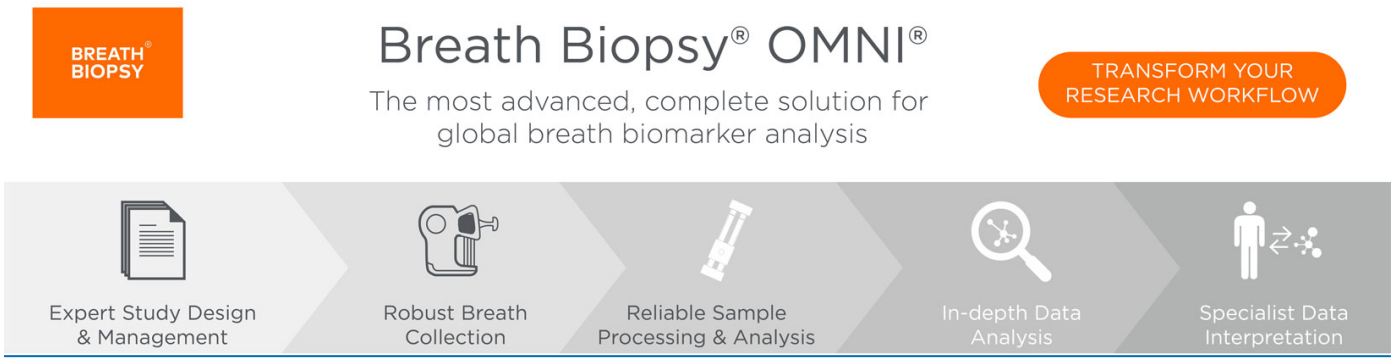
View the article online for updates and enhancements.

You may also like

- Rapidly expanding lake heatwaves under climate change
R Iestyn Woolway, Eric J Anderson and Clément Albergel

- Enhanced nighttime heatwaves over African urban clusters
Eghosa Igun, Xiyan Xu, Zitong Shi et al.

- Interactive influence of ENSO and IOD on contiguous heatwaves in Australia
P Jyoteeshkumar Reddy, Sarah E Perkins-Kirkpatrick and Jason J Sharples



This content was downloaded from IP address 134.225.110.27 on 15/02/2023 at 14:44

ENVIRONMENTAL RESEARCH
LETTERS

CrossMark

OPEN ACCESS

RECEIVED
13 June 2022

REVISED
15 September 2022

ACCEPTED FOR PUBLICATION
14 December 2022

PUBLISHED
3 January 2023

Original Content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

PAPER

Typicality of the 2021 Western North America summer heatwave

Valerio Lucarini[1,*], **Vera Melinda Galfi**[2], **Jacopo Riboldi**[2] and **Gabriele Messori**[3,4]

[1] Department of Mathematics and Statistics and Centre for the Mathematics of Planet Earth, University of Reading, Reading, United Kingdom
[2] Department of Earth Sciences, Uppsala University, Uppsala, Sweden
[3] Department of Earth Sciences and Centre of Natural Hazards and Disaster Science (CNDS), Uppsala University, Uppsala, Sweden
[4] Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden
[*] Author to whom any correspondence should be addressed.

E-mail: v.lucarini@reading.ac.uk

Keywords: heatwave, climate variability, climate change, extreme events, large deviation theory, North America

Abstract

Elucidating the statistical properties of extreme meteo-climatic events and capturing the physical processes responsible for their occurrence are key steps for improving our understanding of climate variability and climate change and for better evaluating the associated hazards. It has recently become apparent that large deviation theory (LDT) is very useful for investigating persistent extreme events, and specifically, for flexibly estimating long return periods and for introducing a notion of dynamical typicality. Using a methodological framework based on LDT and taking advantage of long simulations by a state-of-the-art Earth system model, we investigate the 2021 Western North America summer heatwave. Indeed, our analysis shows that the 2021 event can be seen as an unlikely but possible manifestation of climate variability, whilst its probability of occurrence is greatly amplified by the ongoing climate change. We also clarify the properties of spatial coherence of the 2021 heatwave and elucidate the role played by the Rocky Mountains in modulating hot, dry, and persistent extreme events in the Western Pacific region of North America.

1. Introduction

Investigating the statistical properties of extreme meteo-climatic events and the physical mechanisms behind their occurrence is an area of ever-growing interest in climate science [1]. Additionally, their study allows to access some key aspects of the climate system, since there is in general a fundamental connection between extreme events and the dynamics of the system generating them [2–6]. Long-lasting meteo-climatic extremes are hazards of particular relevance because anomalous and persistent conditions are conducive to extremely damaging impacts on ecosystems and society [1, 7–10]. Moreover, atmospheric phenomena with a long characteristic time scale are usually also characterized by a large spatial extent [11, 12], potentially leading to systemic risk [13]. In general, the occurrence of extremes of surprisingly large magnitude might be indicative of an approaching critical transition [14, 15].

1.1. The 2021 Western North America heatwave

Between late June and mid-July 2021 a large latitudinal band of the western sector of North America faced a very intense heatwave, in the context of the ongoing megadrought in the region [16, 17]. Several new temperature records were established: 49.6 °C were measured in Lytton, which is the new absolute temperature record for 2-m temperature (T2M) for Canada. It is estimated that the heatwave has led to over 1000 directly attributable deaths and damages worth several billion USD, associated with wildfires, crop loss, infrastructural damages, and flooding due to rapid snow melt [18–20]. The record-breaking temperatures were the result of preconditioning combined with an anomalous atmospheric circulation. Prior to the heatwave, the southern portion of the region was anomalously dry [18, 21] and it is known that long-term drought increases the probability of occurrence of heatwaves, as a result of the drying of the soil [22–24], possibly leading to the onset of a cycle of heat, drought, and fire [25, 26]. This was compounded by the presence of an intense atmospheric Ω-block, like the one associated with the 2010 Russian heatwave [27–29]. Atmospheric blocks [30] are often the culprits for individual heatwaves as a result of possibly concurring processes, such as anomalies in clear-sky incoming radiation, warm air advection,

© 2023 The Author(s). Published by IOP Publishing Ltd

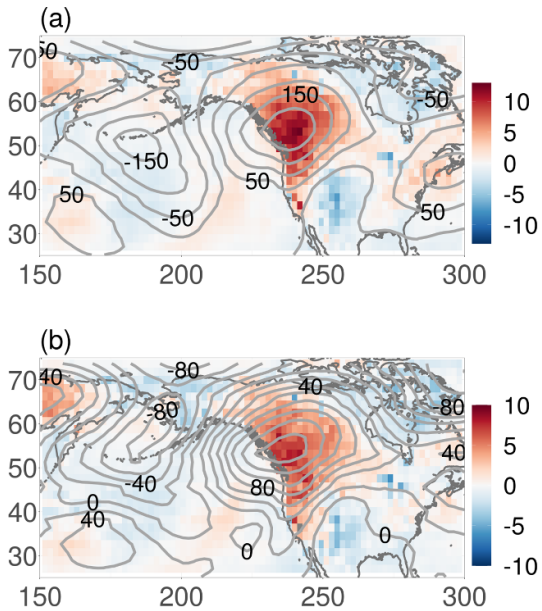


Figure 1. ERA5 anomalies for (a) 24th June–30th June 2021 and (b) 24th June–8th July 2021 average of T2M (colours, in (K) and Z500 (contour lines, in (m) with respect to the 1991–2020 climatology for the corresponding period.

and subsidence [31, 32]. The Western North America heatwave (WNAHW) blocking developed on 25th June in association with a large scale pattern over the mid-latitudes dominated by low zonal wavenumbers [19, 33]. The ridge reached an impressive maximum 500 hPa geopotential height (Z500) of almost 6000 m over western North America. Figure 1 shows the ERA5 [34] T2M and Z500 anomaly fields averaged over periods of 7 and 15 days encompassing the heatwave. The agreement between the large scale patterns in figures 1(a) and (b) clearly shows the persistent and spatially coherent nature of the event. Predicting the onset of blocking is extremely challenging for numerical models [35–37] and, indeed, very few if any of the weather forecasting systems were able to estimate accurately the intensity of the 2021 WNAHW at lead times of a few weeks [20].

Recently, a team from the World Weather Attribution initiative has concluded, following the framework presented in [38] and using a multi-model and multi-method attribution analysis applied to daily maximum temperatures, that it is possible to attribute almost non-ambiguously the 2021 event to climate change, because its likelihood in present climate conditions (factual world) is estimated to be orders of magnitude higher than in past climate conditions (counterfactual world) [18].

1.2. This study

In [6, 29] it was shown that large deviation theory (LDT) [39, 40] allows for constructing a climatology of heatwaves and cold spells over land. This approach is useful because one can go beyond purely descriptive statistics. Indeed, the return times of sufficiently persistent and intense temperature anomalies can be predicted by constructing the so-called rate functions for each given spatial location. For temperature timeseries, such rate functions are to a good approximation quadratic. One can thus construct them, in close agreement with the central limit theorem, by using relatively simple statistics like the variability and integrated autocorrelation time of the daily temperature records.

It has recently become apparent that LDT can also be used to introduce a—somewhat counter-intuitive—notion of typicality of large and persistent temperature fluctuations associated with extreme heatwaves [29, 41, 42] or cold spells [6, 29]. The intense and persistent temperature fluctuations observed in a specific location are associated with a very unusual, rarely seen large-scale atmospheric configuration. Yet, if we subsample our dataset conditional on the occurrence of comparable temperature fluctuations in the same location, such large scale patterns are similar to each other, and are, in the sense discussed below, *typical*.

The main goal of this study is to investigate whether the 2021 WNAHW is dynamically typical according to the LDT viewpoint presented in [6, 29], i.e. whether the corresponding large-scale circulation pattern can be considered to belong to the subset of typical circulation patterns mentioned above. We will also try to better understand the geographical features of the 2021 WNAHW by elucidating its spatial coherence. This study also presents a more detailed and critical account of the methodology presented in [29], which supports its robustness.

Note that, through the use of an Earth system model (ESM) run in preindustrial, steady state configuration as in [29], we will study the event as a fluctuation with respect to a baseline climatology. Hence, our approach is different from that of the attribution studies mentioned before, where what matters is the absolute temperature reading, rather than the anomalies. We will estimate the impact of shifting the climatology in order to roughly account for the impact of historical climate change on the probability of occurrence of events comparable to the 2021 WNAHW. While this is obviously a serious simplification, because one neglects all the possible impacts of shifting climate conditions on climate variability (e.g. changes in the dynamics of the atmosphere, water cycle, soil properties [32]), shifting the climatic mean has been shown to be a good first order approximation of the impact of climate change on the statistics of heatwaves [19].

The 2021 WNAHW featured a particularly acute phase between 27th June and 29th June. Hence, some studies have focused on the statistical investigation

of daily temperature records [18, 19] and taken advantage of the statistical framework provided by extreme value theory [2, 43]. Since persistence plays a major role in determining the impact of heatwaves [9], we prefer to look into cumulative temperature anomalies.

A succinct account of the relevant LDT background that supports our analysis is presented in the next section. This is then followed by the discussion of the main findings of this investigation. A section presenting the concluding remarks and the perspectives for future investigation concludes this contribution. This letter is accompanied by supplementary material (SM) [44] that complements the core findings described here.

2. The mathematical background

We follow below the viewpoint developed for fluid flows in [45, 46], which is based upon the theoretical framework presented in [47]. We consider a continuous-time chaotic dynamical system defined by an ordinary differential equation of the form

$$\dot{x} = G(x) \tag{1}$$

with $x \in \mathbb{R}^N$ with initial condition $x(0) = X$ belonging to the attractor Ω of the system. The solution at time $t \geqslant 0$ is given by $x(t, X) = S^t X$ where S^t is the evolution operator up to time t. We assume that the system possesses a unique physical invariant measure ρ supported on Ω and that our initial condition $X \in \Omega$, i.e. all transients have died out and we are in steady-state conditions. We now introduce a target function of the form $F(\tau, X) = \int_0^\tau \mathrm{d}t f(x(t, X))$ where f is a smooth function of its argument. Such a target function can be tailored to capture the persistent extreme of interest. We define as $\mathcal{P}(F(\tau, X) > z)$ the probability that $F(\tau, X)$ has value larger than z. The working hypothesis is to be able to write $\mathcal{P}(F(\tau, X) > z)$ in the limit of large z as a large deviation law of the form

$$\mathcal{P}(F(\tau, X) > z) \asymp \exp(-\min_{X \in A_{F,\tau,z}} I(X)) \tag{2}$$

where $A_{F,\tau,z}$ is the set of points of Ω such that if $y \in A_{F,\tau,z}$, then $F(\tau, y) > z$ and $\asymp$ indicates that the ratio of the logarithms of the two sides tends to 1. Finally, $I(X)$ is the so-called rate function. As a result, one obtains that

$$\mathcal{P}(F(\tau, X) > z) \asymp \exp(-I(X^*(z))),$$

where

$$X^*(z) = \arg\min_{X \in A_{F,T,z}} (I(X)).$$

This implies that the most likely way the event $\mathcal{P}(F(\tau, X) > z)$ occurs is for $X = X^*(z)$. Namely, as we consider larger and larger thresholds z, the probability measure concentrates more and more around a specific initial condition $X^*(z)$. Let's provide an interpretation of this result following [47]. The initial conditions $X = X^*(z)$ around which the measure concentrates define the forward trajectory $x(t, X^*(z)) = S^t X^*(z)$. It is clear that such initial condition and the ensuing forward trajectory are extremely atypical exactly because there is an association with the desired rare event defined in the limit of large z. Conversely, if one studies the statistics of the system conditional on the occurrence of the rare event (i.e. restricted to the set $A_{F,\tau,z}$), the trajectory departing from $X^*(z)$ becomes typical. Indeed, LDT captures here the least unlikely of all the unlikely ways the rare event of interest can occur [48].

In the previous studies of the 2010 Russian Heatwave, 2010 Mongolian Dzud (an extreme cold spell), and 2019 North American cold spell [6, 29] the model-simulated events selected according to the procedure above, where z was a measure of the observed extreme event at a given location in its core region

(a) resembled each other: the large scale spatial patterns of T2M and Z500, while being abnormal in reference to the climatology, were highly correlated with the mean field computed by averaging over the various events;
(b) resembled the actual historical event: the observed T2M and Z500 fields could be interpreted as being drawn from the same distribution including the various model-generated surrogate events.

The properties 1. and 2. above indicate the usefulness of the mathematical framework given in the previous section: if we consider events selected according to a sufficiently stringent constraint, they can be seen as fluctuations around the optimal $x(t, X^*(z)) = S^t X^*(z)$ case, which can be estimated as the average over all the selected events generated by the model [45, 46].

3. Data

We use here, on top of the T2M and Z500 ERA5 [34] fields for 24th June–8th July 2021, the data for the same variables gathered from a 1000 year simulation performed by the MPI-ESM-LR version 1.2 ESM [49] in standard pre-industrial conditions in terms of atmospheric composition and land-use dataset (*piControl* run). This ESM is among the best performing models participating to the 6th phase of the Climate Model Intercomparison Project (CMIP6) [50], just as its previous version was among the best performing CMIP5 models [51].

After removing the seasonal cycle by subtracting the climatology of each day of the year, we select for each year an extended summer lasting 160 d and beginning on 5th May, as in [29], in order to remove

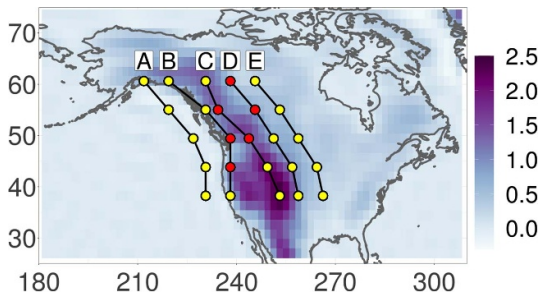


Figure 2. The slanted grid shows the location of the 25 grid points considered in this study. They are arranged in an orography-following fashion and for latitudes ranging from 38 °*N* to 61 °*N*. The red dots indicate the core of the 2021 WNAHW as discussed in the text. The color bar indicates the MPI-ESM-LR model orography in units of Km.

winter heatwaves from the statistics. We consider 25 grid points located in the Western part of North America; see figure 2. These grid points are arranged in an irregular lattice: the North/South direction is distorted in such a way that it approximately follows the orography. The grid points are defined by their latitude (38 ° N–61 ° N in steps of ≈6 °) and by the letters A, B, C, D, and E (from west to east). A refers to points located over the ocean (except in case of latitude 61 ° N), not far from the coast; B to points located near the Pacific coast; C and D to points at the top and at the eastern flank of the Rocky mountains; and E to points located in the continental region to the East of the Rocky mountains.

Following the mathematical framework described above, the first step is to use the following definition for the target function: $F_k(\tau,X) = \int_0^\tau \mathrm{d}t(T_k(t,X) - \bar{T}_k^\tau)$, where k is the index referring to the grid point of interest, $T_k(t,X)$ is the T2M at time t given the initial condition X, and $\bar{T}_k^\tau$ is the long-term climatology of T_k. During the 2021 event, for both cases of $\tau = 15\ d$ and of $\tau = 7\ d$, the target function is largest at the following grid points (indicated in red in figure 2): 44 *B* (near Portland, USA), 49 *B* (near Vancouver, Canada), 49 *C*, 55 *D* (Calgary, Canada, is approximately midway between these two latter locations), 55 *C*, and 61 *D* (near Forth Smith, Canada). See tables 1 and 2.

4. Results

In what follows we show that there is a clear signature of the 2021 WNAHW in the natural variability of the MPI-ESM-LR. We will further test whether there is evidence of a strong signature of the Rocky Mountains on the 2021 WNAHW. Indeed, orography has a key role in catalyzing and determining the location of large scale and persistent weather patterns associated with the low-frequency variability of the atmosphere [52–59] and in determining the properties of heatwaves [31, 60].

The 1000 year MPI-ESM-LR model run features by and large events comparable to the 2021 WNAHW in terms of observed anomalies. The return times for values of the target function at the locations indicated above range from multidecadal to multicentennial. In the case $\tau = 15\ d$, they reach a value of 200 *y* at grid point 49 *C* and at 91 *y* at grid point 55 *D*. In the case $\tau = 7\ d$, one gets multicentennial values at grid points 49 *B* and 49 *C*. Instead, the model cannot find an event of intensity equal or larger than the 2021 WNAHW for grid point 55 *D*. Comparing tables 1 and 2 one finds confirmation of the presence of a very intense phase within the heatwave, as already mentioned in the introduction. For the other grid points shown in figure 2 and not included in tables 1 and 2, instead, the return times of the 2021 value of $F_k(\tau,X)$ estimated through the MPI-ESM-LR model dataset are at most of the order of few years. Hence, they are relatively (or very) common within, e.g. a 30 year time span commonly used to benchmark climatology. We conclude that the red grid points in figure 2 define the (vast) core region of the 2021 event, which results to be larger than the area investigated in [18, 19].

4.1. Robustness of the methodology

We first test whether the mathematical framework presented above provides us with a robust and practically usable methodology. We proceed as follows. As shown in tables 1 and 2, the 50 year return level estimated by the model is by and large a good reference value for the observed anomalies for all of the grid points of the core region of the 2021 WNAHW. The results are weakly dependent on the specific choice of the return time, as long as it is sufficiently long; see discussion below. We then select, for each of the 25 grid points shown in figure 2, as threshold z_k for $F_k(\tau,X)$ the 50 year return level. In such a way, we have a homogeneous statistics throughout the domain of interest for extreme heatwaves, because we identify $J = 20$ events for each of the 25 grid points of interest. Each event corresponds to one initial condition X_k^j compatible with $F_k(\tau,X) > z_k$. In other terms, our procedure aims at defining candidates for weather analogues [61–63] for the extreme event of interest. We present below results for $\tau = 15\ d$. All the main conclusions we draw are valid also for $\tau = 7\ d$, whose corresponding figures are shown in the SM; see further discussion therein.

We first test whether by considering for a given k such a high threshold z_k for the function $F_k(\tau,X)$ we end up identifying sufficiently similar weather patterns, which can be seen as fluctuations around $X^*(z_k)$. This amounts to being able to introduce a notion of typicality for such extreme events that leads to identifying good weather analogues.

The cyan boxplots in figure 3 show the properties of the spatial correlation between the τ averages of T2M and Z500 fields of the $J = 20$ events selected for

Table 1. Surface temperature anomalies ($\Delta T2M$, in °C) averaged over $\tau = 15\,d$ for the 2021 WNAHW at selected grid points (see figure 2) in its core region (third column). For each grid point, we have $\Delta T2M = F_k(\tau, X)/\tau$. The corresponding return times (RTs, in y) estimated with the MPI-ESM model (fourth column) as well as the RTs for the climatology shifted by $\delta = 1.2$ °C (last column) are reported. The model return levels for 50 year return times (RL_{50}, in °C) are reported in the second column.

Grid point	RL_{50}	$\Delta T2M$	$RT(\Delta T2M)$	$RT(\Delta T2M - \delta)$
44B	6.3	5.6	19	6
49B	6.0	5.4	22	5
49C	5.6	6.2	200	22
55C	7.3	6.7	29	8
55D	5.2	5.5	91	17
61D	5.7	5.3	30	6

Table 2. Same as table 1, but for $\tau = 7\,d$.

Grid point	RL_{50}	$\Delta T2M$	$RT(\Delta T2M)$	$RT(\Delta T2M - \delta)$
44B	8.8	8.2	29	8
49B	8.4	9.7	333	53
49C	7.6	8.9	500	62
55C	10	9	21	9
55D	7.2	9.1	>1000	1000
61D	7.9	7.5	36	9

each grid point k and their respective mean. Since we are now testing the robustness of the methodology, we perform this analysis for all the grid points shown in figure 2.

The spatial correlation is computed over the region [20° N, 75° N]×[150° E, 300° E], which includes the whole North America, approximately half of the North Pacific ocean, and a quarter of the North Atlantic ocean. The results shown below are very weakly sensitive to adjustments in the boundaries of such a (vast) region. We clearly see that in most cases the mean of spatial correlation is high, with a moderate standard deviation, which indicates high coherence between the events. In general, the coherence is higher as we focus on higher latitudes, thus reducing the impact of tropical dynamics, and as we move away from the ocean towards the Rockies, in agreement with what is mentioned above regarding the interaction between orography and low-frequency (and large scale) mid-latitude atmospheric variability.

Stronger support to these statements can be found by looking at the Taylor [64] diagrams shown in the SM as figures S3–S4 (S5–S6) for the T2M (Z500) field, 15 d and 7 d averages, respectively. Taylor diagrams allow for a more complete comparison between fields by including, on top of their spatial correlations, also the centered root-mean-square (RMS) difference and the amplitude of their variations (represented by their standard deviations). For most of the 25 grid points of interest, the 20 50-y return level events define a cluster—see below—characterized by high correlation and low centered RMS difference with respect to the mean of such events.

We remark that for the grid points over the ocean (where the temporal correlation of T2M has a slower decay) and in case of the southernmost latitude (where there is strong interaction between tropical and extratropical dynamics), the spread is substantially larger, indicating that it is more difficult to obtain a well-defined typicality [29].

Clearly, because of averaging, the mean tends to have lower standard deviation than the individual events. As we consider longer and longer return times—see figures S7–S10 in the SM—the clusters become better defined as the individual events become more and more similar to their averages. Indeed, in each Taylor diagram the center of the cluster moves in the direction of the reference point in the x-axis. The presence of a single well-defined statistical cluster for the T2M and Z500 fields (separately and taken together) is confirmed in all cases for fields constructed for grid points over land and for return times longer than 20 years through the gap statistics [65] for Gaussian mixture models [66], which has been evaluated using the MATLAB function `evalclusters` [67].

We conclude that the methodology followed here allows us to distill the fundamental properties of heatwaves throughout the domain of interest. The test above is not complete because, for each k, one should look at the agreement between the evolution up to time τ of the various $j = 1, \ldots J$ initial conditions X_k^j. Instead, here we are looking at the agreement between the time-averaged fields $\bar{X}_{k,\tau}^j = 1/\tau \int_0^\tau \mathrm{d}t S^t X_k^j$, and we estimate $\bar{X}^*(z_k)_\tau = 1/J \sum_{l=1}^J \bar{X}_{k\tau}^j$. Nonetheless, the results we obtain are extremely encouraging.

4.2. Fingerprinting the 2021 WNAHW

The next step is to discover whether there is an agreement between the heatwaves constructed as above and the 2021 WNAHW. Indeed, if we repeat the

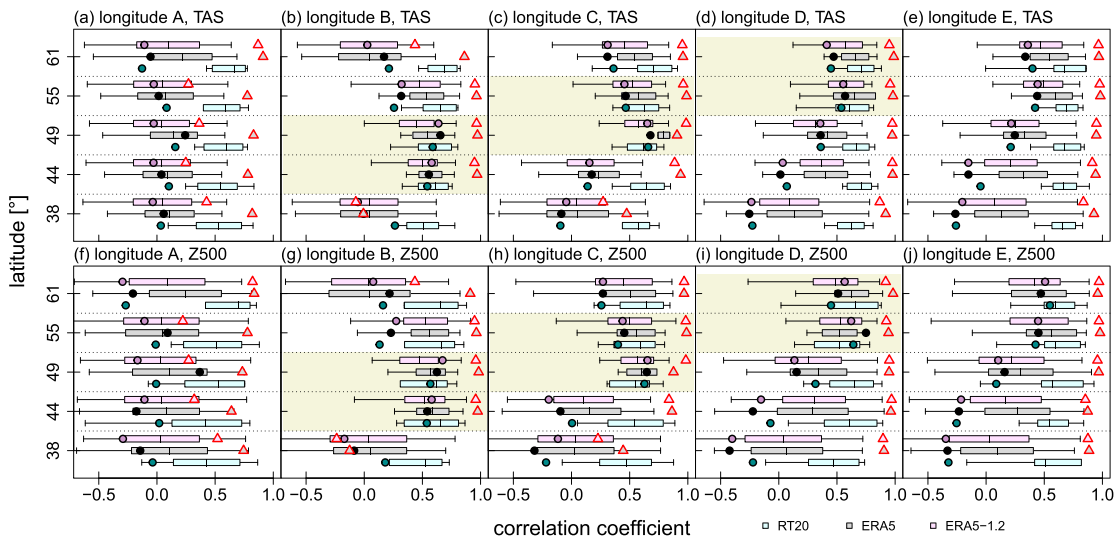


Figure 3. Spatial correlation coefficients between the model-simulated events and their mean for the T2M (a)–(e) and Z500 (f)–(j) fields. The coefficients are obtained for events selected by defining a high threshold in the cumulated temperature anomaly at the corresponding grid point over $\tau = 15\,d$; see indication of the latitude on the side and of the longitude on top. The shading indicates the core region of the 2021 WNAHW, see tables 1 and 2. In each subpanel, the boxes refer to the $\pm 1\sigma$ interval, and the whiskers indicate the full range of results. The dots indicate the spatial correlation coefficient between the mean of the model runs and the corresponding field from ERA5. Additionally, (a) Cyan refers to 50-y return time events, (b) Gray refers to events with return level corresponding to the one of the 2021 WNAHW, and (c) Magenta refers to events where the return level of 2021 WNAHW event is reduced by $\delta = 1.2\,^{\circ}$C. The triangles indicate the spatial coefficient between the mean of the events (a) and the mean of the events (b) and (c), respectively. See the Taylor diagrams in figures S3–S14 in the SM for more complete information on the agreement between the considered patterns.

analysis by using the actual return levels of the 2021 WNAHW—see gray boxplots in figure 3—we obtain for the grid points in the core region results that are in very good agreement with what found using the 50-y return levels. This confirms that for each k the estimate of $X^*(z_k)$ depends weakly on z_k as long as we are choosing a sufficiently large value for z_k.

The full dots in the gray boxplots in figure 3 show, for each k, the value of the spatial correlation coefficient between the estimate of $\bar{X}^*(z_k)_\tau$ and the actual τ-averaged ERA5 fields depicted in figure 1. The spatial correlation coefficient for both the T2M and Z500 fields is highest when considering the events constructed for the grid points defined in tables 1 and 2. For five out of these six grid points, additionally, we have that the value of such spatial correlation coefficients is within the range of values obtained considering the J fields $\bar{X}^j_{k,\tau}$ and $\bar{X}^*(z_k)_\tau$, and in most cases within the one standard deviation interval. Hence, the ERA5 anomaly field can be seen as belonging to the same statistical ensemble as the anomalies produced by the model, given the stringent criteria set for suitably choosing the corresponding events. The only exception comes from the grid point 49 C, basically because we have only few model occurrences of heatwaves of larger intensity than the 2021 WNAHW. Such events closely resemble each other, so that the range of values of correlations with their average is very small. Nonetheless, the spatial correlation between the estimated $\bar{X}^*(z_k)_\tau$ and the actual τ-averaged ERA5 is high. The Taylor diagrams shown in figure S11 (T2M) and figure S13 (Z500) of the SM further support these conclusions. Similar results are obtained when considering $\tau = 7\,d$ (see figures S12 and S14 of the SM), except for the fact that the ensemble is too limited to draw conclusions for the grid points 49 B and 49 C, and no ensemble members are available for the grid point 55 D, possibly as a result of the model's inability to reproduce specific local processes.

The similarity of the six panels of figure 4 among themselves, with figure 1(b), and with the corresponding panels of figure S16 of the SM (which are constructed using the actual return levels of the 2021 WNAHW) provides a visual confirmation of what mentioned above. In meteorological terms, the possibly most striking common features between the model data and the actual 2021 WNAHW are (a) the presence of a wavetrain across the Pacific ocean [20], which closely resembles the anomaly pattern deemed responsible for drought over California [68]; and (b) the powerful ridge in the mid-latitudes extending from the Pacific over the Rockies, which reminds of the pattern that which is associated with the occurrence of intense drought in the Western US [17, 68–70].

We are now able to strengthen the statement made before regarding the fact that the grid points included in tables 1 and 2 correspond to the core of the 2021 WNAHW, which was based purely on the intensity of the local T2M anomalies. For these six grid points we have both high internal coherence

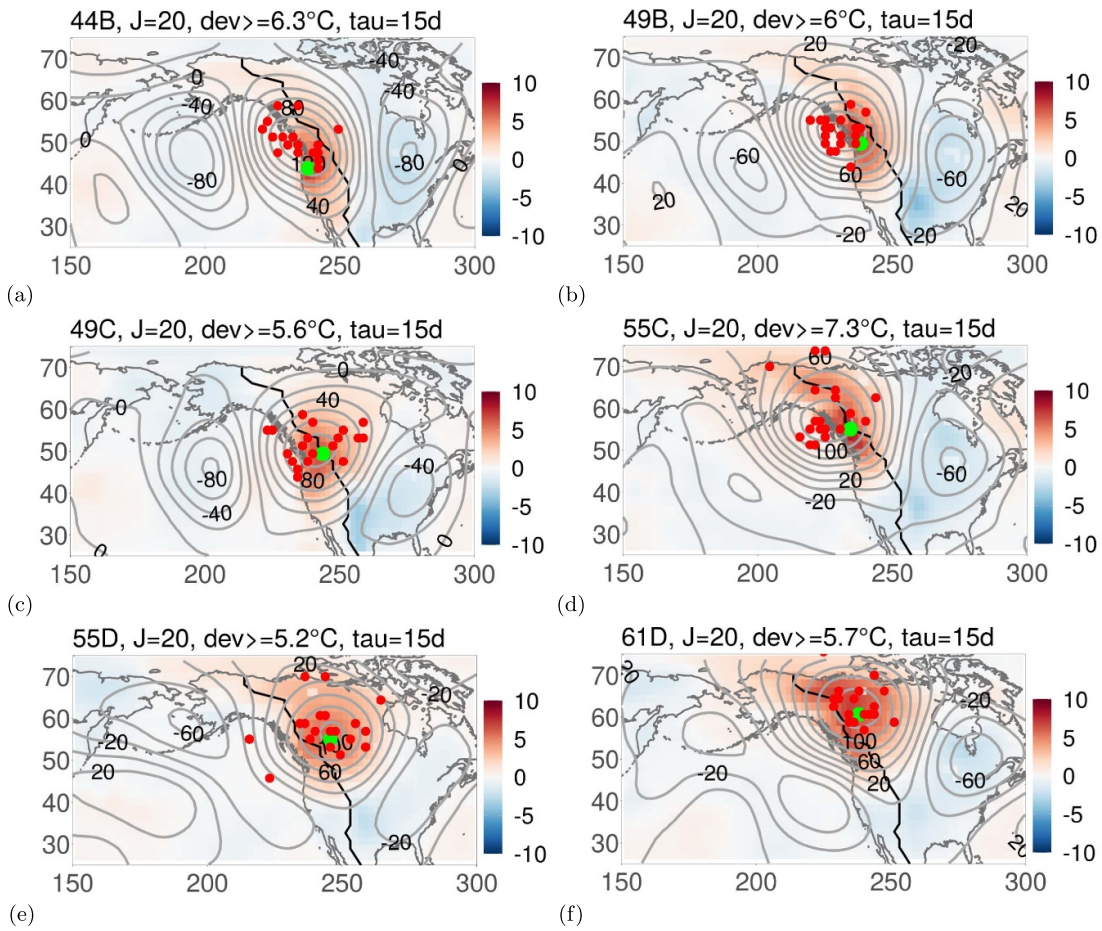


Figure 4. T2M and Z500 anomaly fields (averaged over $\tau = 15$ (d) computed as means over the $J = 20$ most extreme model-simulated events chosen according to a high T2M anomaly thresholds at grid points 44 *B* (a), 49 *B* (b), 49 *C* (c), 55 *C* (d), 55 *D* (e) and 61 *D* (f). The red dots indicate the center of the high pressure for each of the *J* events. All the patterns look remarkably similar, and similar to the 2021 WNAHW, see figure 1.

among the model-generated events and high statistical compatibility with the large-scale fields of the 2021 WNAHW. Geographically, these grid points correspond to a region stretching approximately in the SSW-NNE direction and, hence, shifting from the coastal area to the eastern side of the Rockies as one progresses northwards. These grid points define a coherent region associated with a very special yet very robust feature of the natural variability of the climate, which most recently manifested itself as the 2021 WNAHW. As can be seen by comparing the corresponding panels of figures S2 and S16, for each of these grid points the average pattern of the model-generated events is virtually unchanged if we consider either (a) the 20 events leading to the largest local persistent temperature anomaly (also portrayed in figure 4) or (b) events leading to local persistent temperature anomalies equal or larger than what observed during the 2021 WNAHW. Such an agreement can also be checked by noting that for such grid points the triangles next to the grey boxes in figure 3 are located very close to one.

Let's look at the grid points outside the 2021 WNAHW core region. The gray boxplots, constructed according to the 2021 WNAHW return levels, cover a larger range of values and are substantially shifted towards lower values compared to the corresponding cyan ones. This means that if we select soft extremes (or events that are not extreme at all), the notion of typicality discussed above is lost, because we are outside the regime where LDT applies. The Taylor diagrams indicate a much reduced agreement between the average of the model generated T2M and Z500 fields and the corresponding ERA5 fields as compared to the case of the core region.

Specifically, from figure 3 and figures S3–S6 of the SM, one can see that a good agreement still exists while considering averages constructed for grid points to the NNE of the core region, such as 55 *E* and 61 *E*. Instead, the correlation dramatically decreases when considering grid points geographical close to the core region, such as 44 *C* and 55 *B*, but off-axis with respect to the SSW-NNE direction. Hence, at least in the MPI-ESM-LR model world, heatwaves centered in, e.g. grid points 44 *C* and 55 *B* are due to dynamical processes that are fundamentally different from those responsible for the 2021 WNAHW.

5. Discussion and conclusions

The 2021 WNAHW shocked experts and general audiences around the world and had devastating impacts on a large swath of land in North America. We have investigated this event following the strategy delineated in [6, 29] and based upon the theoretical framework defined in [46, 47]. While the 2021 WNAHW is very rare and unusual according to the overall statistics of weather variability, it becomes *typical* when considering the statistics conditional on the validity of a specific constraint, namely the occurrence of a large and persistent summer temperature anomaly at one specific location, defined by a grid point of our model. Typicality results, against intuition, exactly from *the extreme intensity* of the event. Indeed, we find a good agreement, as shown by figure 3 and by the Taylor diagrams included in the SM, between the large scale meteorological anomaly patterns of the 2021 WNAHW and those of the events selected from the 1000 year-long MPI-ESM-LR preindustrial run according to the procedure above when considering any grid point located in a specific region of Western North America. This region has an extension of 2000 Km in the SSW-NNE direction and defines the core of the 2021 event.

This indicates that the circulation pattern causing heatwaves over this region in the model resembles the one that caused the 2021 WNAHW, at least in the spatial distribution of mid-tropospheric and surface anomalies. This also indicates that, at least in the model climate, there is, roughly speaking, mainly one way to get heatwaves of such large magnitude in that region.

The obtained large scale patterns are in good correspondence with those deemed responsible for the persistent drought in the western US. This supports the existence of a cascading process acting across different spatial and temporal scales that is increasing dramatically climate risk in the western sector of the North American continent, with conditions being and becoming particularly exacerbated in California [71–73].

Modulating the imposed threshold on the local T2M anomaly amounts to taking into account, in first approximation, changing background climatic conditions. Shifting the mean has been recently shown to represent a good first approximation for evaluating the statistics of hot temperature extremes in a changing climate [19, 74]. This can be explained—in the case of persistent extremes—by the fact that the rate function for summer (as opposed to winter) temperature near the region of interest seems to be approximately insensitive to the levels of CO_2 [29]. We find—see tables 1 and 2—that in the core of the 2021 WNAHW the return times of events comparable to the 2021 event decrease substantially (by a factor of about 3 to 10) if one budgets in an offset ($\delta = 1.2\,^{\circ}\mathrm{C}$) that approximately represents the historical shift of temperatures due to the realized climate change between pre-industrial conditions and today. This indicates that the likelihood of events comparable to the 2021 WNAHW is greatly enhanced by climate change: a very or extremely unusual event becomes comparably likelier. Interestingly, in the core region the typicality of the extreme events is maintained even if we relax a bit our procedure: the large scale meteorological patterns associated with events selected according to the less stringent (because of the offset) constraint imposed in the core region are in good agreement with those obtained using the more stringent criterion; see the magenta and gray boxplots in figure 3.

The mathematical setting behind this paper indicates why dynamical similarity exists between the various extreme events, despite their different intensity, as long as they are all chosen according to the same sufficiently stringent constraint. Our analysis is limited in scope by the fact that we look at the data produced by only one state-of-the-art ESM - yet a model with rather good overall skill as compared to the others included in the CMIP6 exercise [50], and at least some of our conclusions could be affected by its biases and deficiencies. It would be important to test whether similar results can be found in other CMIP6 models, also taking into account that ESMs tend to globally underestimate duration, intensity, and frequency of heatwaves, yet having a comparatively better performance in the West North America region [75]. Interestingly, the CMIP6 bias is instead on the warm side for daily temperature maxima [76], which further reinforces the need to deal carefully with the property of persistence when looking at hot events. The methodology adopted here could be replicated for studying the statistics and dynamics of several historical extreme heatwaves, taking advantage of the dataset compiled by [19].

In order to better characterize the features of the event studied here—with the goal, e.g. of better understanding the separate role of dynamical and thermodynamical contributions, preconditioning factors, and the role of teleconnections—one would need larger statistical sets than those considered here, because we construct our ensembles by imposing rather stringent conditions on the fields. A possible strategy is to take advantage of the CMIP6 dataset [50], and use all available ensemble members and various climate scenarios in order to address the problem of the impact of climate change on the statistics and dynamics of heatwaves. A theoretically more challenging yet more promising strategy is, instead, to exploit rare events algorithms to disproportionally populate the exotic configurations of interest and learn a great deal more about them [41, 42]. Nonetheless, both of these directions are beyond the scopes of this investigation as they require extensive data analysis and modeling work.

Our approach can also be seen as leading to the definition of weather analogues for extreme events. In [77] it was shown, by investigating the recurrence of weather analogues, that in the last decades there has been a positive trend in the frequency of occurrence of atmospheric circulation patterns over the North Atlantic driving summertime dryness and heatwaves and a negative trend in those leading to wet, cooler summer conditions across Europe. We envision applying the methodology proposed in [77] to the West North America region to complement the analysis presented here in the direction of understanding the dynamical properties responsible for the exacerbating dry conditions in the region.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

The authors thank D Faranda, T Grafke, F Ragone, J Wouters for many exchanges on extreme events, L Recchia for her comments on a preliminary version of the manuscript, and two anonymous reviewers for constructive criticism. VL acknowledges the support by the Horizon 2020 project TiPES (Grant No. 820970), by the Marie Curie ITN CriticalEarth (Grant No. 956170) and by the EPSRC project EP/T018178/1. G M and J R acknowledge the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Project CENÆ: compound Climate Extremes in North America and Europe: from dynamics to predictability, Grant No. 948309). V M G acknowledges the support of the Air, Water and Landscape Science Programme at the Department of Earth Sciences of Uppsala University.

ORCID iDs

Valerio Lucarini https://orcid.org/0000-0001-9392-1471
Vera Melinda Galfi https://orcid.org/0000-0002-8965-0880
Jacopo Riboldi https://orcid.org/0000-0001-5166-9618
Gabriele Messori https://orcid.org/0000-0002-2032-5211

References

[1] Field C B *et al* (IPCC) (eds) 2012 *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge Univerity Press)
[2] Lucarini V, Faranda D, de Freitas A, de Freitas J, Holland M, Kuna T, Nicol M, Todd M and Vaienti S 2016 *Extremes and Recurrence in Dynamical Systems* (*Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts*) (Hoboken: Wiley)
[3] Gálfi V M, Bódai T and Lucarini V 2017 *Complexity* **2017** 5340858
[4] Faranda D, Messori G and Yiou P 2017 *Sci. Rep.* **7** 41278
[5] Hochman A, Alpert P, Harpaz T, Saaroni H and Messori G 2019 *Sci. Adv.* **5** eaau0936
[6] Gálfi V M, Lucarini V, Ragone F and Wouters J 2021 *Riv. Nuovo Cimento* **44** 291–363
[7] Easterling D R, Meehl G A, Parmesan C, Changnon S A, Karl T R and Mearns L O 2000 *Science* **289** 2068–74
[8] Koppe C, Sari Kovats R, Menne B and Jendritzky G (World Health Organization. Regional Office for Europe and London School of Hygiene and Tropical Medicine and European Commission. Energy, Environment and Sustainable Development and Deutscher Wetterdienst) 2004 *Heat-Waves: Risks and Responses* (Copenhagen : WHO Regional Office for Europe)
[9] Poumadére M, Mays C, Le Mer S and Blong R 2005 *Risk Anal.* **25** 1483–94
[10] Wang S S, Zhao L, Yoon J H, Klotzbach P and Gillies R R 2018 *Environ. Res. Lett.* **13** 054014
[11] Ghil M and Robertson A W 2002 *Proc. Natl Acad. Sci.* **99** 2493–500
[12] Ghil M and Lucarini V 2020 *Rev. Mod. Phys.* **92** 035002
[13] Kornhuber K, Coumou D, Vogel E, Lesk C, Donges J F, Lehmann J and Horton R M 2020 *Nat. Clim. Change* **10** 48–53
[14] Faranda D, Lucarini V, Manneville P and Wouters J 2014 *Chaos Solitons Fractals* **64** 26–35
[15] Franzke C L E, Ciullo A, Gilmore E A, Matias D M, Nagabhatla N, Orlov A, Paterson S K, Scheffran J and Sillmann J 2022 *Environ. Res. Lett.* **17** 015004
[16] Zhang F, Biederman J A, Dannenberg M P, Yan D, Reed S C and Smith W K 2021 *Geophys. Res. Lett.* **48** e2020GL092293
[17] Zhang W, Hari V, Wang S S-Y, LaPlante M D, Garfin G, Affram G and Kumar R 2022 *Geophys. Res. Lett.* **49** e2021GL097089
[18] Philip S Y *et al* 2022 *Earth Syst. Dyn.* **13** 1689–713
[19] Thompson V, Kennedy-Asser A T, Vosper E, Lo Y T E, Huntingford C, Andrews O, Collins M, Hegerl G C and Mitchell D 2022 *Sci. Adv.* **8** eabm6860
[20] Lin H, Mo R and Vitart F 2022 *Geophys. Res. Lett.* **49** e2021GL097036
[21] Osman M, Zaitchik B F and Winstead N S 2022 *Geophys. Res. Lett.* **49** e2022GL099265
[22] Zampieri M, D'Andrea F, Vautard R, Ciais P, de Noblet-Ducoudre N and Yiou P 2009 *J. Clim.* **22** 4747–58
[23] Miralles D G, Teuling A J, van Heerwaarden C C and Vilà-Guerau de Arellano J 2014 *Nat. Geosci.* **7** 345–9
[24] Schumacher D L, Keune J, van Heerwaarden C C, Vilà-Guerau de Arellano J, Teuling A J and Miralles D G 2019 *Nat. Geosci.* **12** 712–7
[25] Sutanto S J, Vitolo C, Di Napoli C, D'Andrea M and Van Lanen H A 2020 *Environ. Int.* **134** 105276
[26] Libonati R *et al* 2022 *Environ. Res. Lett.* **17** 015005
[27] Dole R, Hoerling M, Perlwitz J, Eischeid J, Pegion P, Zhang T, Quan X W, Xu T and Murray D 2011 *Geophys. Res. Lett.* **38** L06702
[28] Lau W K M and Kim K M 2012 *J. Hydrometeorol.* **13** 392–403
[29] Galfi V M and Lucarini V 2021 *Phys. Rev. Lett.* **127** 058701
[30] Woollings T, Barriopedro D, Methven J, Son S W, Martius O, Harvey B, Sillmann J, Lupo A R and Seneviratne S 2018 *Curr. Clim. Change Rep.* **4** 287–300
[31] Jiménez-Esteve B and Domeisen D I 2022 *Q. J. R. Meteorol. Soc.* **148** 2344–67
[32] Chan P W, Catto J L and Collins M 2022 *npj Clim. Atmos. Sci.* **5** 68
[33] Bartusek S, Kornhuber K and Ting M 2022 *Nat. Clim. Change* **12** 1143–50

[34] Hersbach H *et al* 2020 *Q. J. R. Meteorol. Soc.* **146** 1999–2049
[35] Pelly J L and Hoskins B J 2003 *J. Atmos. Sci.* **60** 743–55
[36] Matsueda M and Palmer T 2018 *Q. J. R. Meteorol. Soc.* **144** 1012–27
[37] Lucarini V and Gritsun A 2020 *Clim. Dyn.* **54** 575–98
[38] Hannart A, Pearl J, Otto F E L, Naveau P and Ghil M 2016 *Bull. Am. Meteorol. Soc.* **97** 99–110
[39] Varadhan S 1984 *Large Deviations and Applications* (Philadelphia: SIAM)
[40] Touchette H 2009 *Phys. Rep.* **478** 1–69
[41] Ragone F, Wouters J and Bouchet F 2017 *Proc. Natl Acad. Sci.* **115** 24–29
[42] Ragone F and Bouchet F 2021 *Geophys. Res. Lett.* **48** e2020GL091197
[43] Coles S 2001 *An Introduction to Statistical Modeling of Extreme Values* (New York: Springer)
[44] Lucarini V, Galfi V M, Messori G and Riboldi J 2022 (https://doi.org/10.6084/m9.figshare.21119965.v2)
[45] Dematteis G, Grafke T, Onorato M and Vanden-Eijnden E 2019 *Phys. Rev.* X **9** 041057
[46] Dematteis G, Grafke T and Vanden-Eijnden E 2019 *SIAM/ASA J. Uncertain. Quantification* **7** 1029–59
[47] Chetrite R and Touchette H 2015 *Ann. Henri Poincaré* **16** 2005–57
[48] den Hollander F 2000 *Large Deviations* (Providence, RI: American Mathematical Society)
[49] Mauritsen T *et al* 2019 *J. Adv. Model. Earth Syst.* **11** 998–1038
[50] Eyring V, Bony S, Meehl G A, Senior C A, Stevens B, Stouffer R J and Taylor K E 2016 *Geosci. Model Dev.* **9** 10539–83
[51] Bock L, Lauer A, Schlund M, Barreiro M, Bellouin N, Jones C, Meehl G A, Predoi V, Roberts M J and Eyring V 2020 *J. Geophys. Res.* **125** e2019JD032321
[52] Charney J G and DeVore J G 1979 *J. Atmos. Sci.* **36** 1205–16
[53] Charney J G and Straus D M 1980 *J. Atmos. Sci.* **37** 1157–76
[54] Benzi R, Malguzzi P, Speranza A and Sutera A 1986 *Q. J. R. Meteorol. Soc.* **112** 661–74
[55] Mullen S L 1989 *J. Clim.* **2** 1554–60
[56] Malguzzi P, Speranza A, Sutera A and Caballero R 1997 *J. Atmos. Sci.* **54** 2441–51
[57] Schubert S and Lucarini V 2016 *Q. J. R. Meteorol. Soc.* **142** 2143–58
[58] Ruti P M, Lucarini V, Dell'Aquila A, Calmanti S and Speranza A 2006 *Geophys. Res. Lett.* **33** L06814
[59] Narinesingh V, Booth J F, Clark S K and Ming Y 2020 *Weather Clim. Dyn.* **1** 293–311
[60] Jiménez-Esteve B, Kornhuber K and Domeisen D I V 2022 *Geophys. Res. Lett.* **49** e2021GL096337
[61] Lorenz E N 1969 *J. Atmos. Sci.* **26** 636–46
[62] Zorita E and von Storch H 1999 *J. Clim.* **12** 2474–89
[63] Vautard R and Yiou P 2009 *Geophys. Res. Lett.* **36** L22702
[64] Taylor K E 2001 *J. Geophys. Res.* **106** 7183–92
[65] Tibshirani R, Walther G and Hastie T 2001 *J. R. Stat. Soc.* B **63** 411–23
[66] McLachlan G J and Peel D 2000 *Finite Mixture Models* (New York: Wiley)
[67] MATLAB 2020 *Version R 2020a* (Natick, MA: The MathWorks Inc.)
[68] Teng H and Branstator G 2017 *J. Clim.* **30** 1477–92
[69] Seager R, Hoerling M, Schubert S, Wang H, Lyon B, Kumar A, Nakamura J and Henderson N 2015 *J. Clim.* **28** 6997–7024
[70] Swain D L 2015 *Geophys. Res. Lett.* **42** 9999–10003
[71] AghaKouchak A, Cheng L, Mazdiyasni O and Farahmand A 2014 *Geophys. Res. Lett.* **41** 8847–52
[72] Diffenbaugh N S, Swain D L and Touma D 2015 *Proc. Natl Acad. Sci.* **112** 3931–6
[73] Mann M E and Gleick P H 2015 *Proc. Natl Acad. Sci.* **112** 3858–9
[74] Gessner C, Fischer E M, Beyerle U and Knutti R 2021 *J. Clim.* **34** 6619–34
[75] Hirsch A L, Ridder N N, Perkins-Kirkpatrick S E and Ukkola A 2021 *Geophys. Res. Lett.* **48** e2021GL095161
[76] Di Luca A, Pitman A J and de Elía R 2020 *Geophys. Res. Lett.* **47** e2020GL088031
[77] Faranda D, Messori G, Jézéquel A, Vrac M and Yiou P 2021 Atmospheric circulation compounds anthropogenic warming and its impacts in Europe *Preprint* (available at: https://hal.archives-ouvertes.fr/hal-0345653)