



HAL
open science

A statistical physics and dynamical systems perspective on geophysical extreme events

Davide Faranda, Gabriele Messori, Tommaso Alberti, M Carmen Alvarez-Castro, Théophile Caby, Leone Cavicchia, Erika Coppola, Reik Donner, Berengere Dubrulle, Vera Melinda Galfi, et al.

► **To cite this version:**

Davide Faranda, Gabriele Messori, Tommaso Alberti, M Carmen Alvarez-Castro, Théophile Caby, et al.. A statistical physics and dynamical systems perspective on geophysical extreme events. 2023. hal-04214322v1

HAL Id: hal-04214322

<https://hal.science/hal-04214322v1>

Preprint submitted on 21 Sep 2023 (v1), last revised 8 May 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A statistical physics and dynamical systems perspective on geophysical extreme events

D. Faranda*

*Laboratoire des Sciences du Climat et de l'Environnement,
UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay & IPSL,
CE Saclay l'Orme des Merisiers, 91191, Gif-sur-Yvette, France
London Mathematical Laboratory, 8 Margravine Gardens, London, W6 8RH, UK and
Laboratoire de Météorologie Dynamique/IPSL, École Normale Supérieure, PSL Research University,
Sorbonne Université, École Polytechnique, IP Paris, CNRS, Paris, France*

G. Messori*

*Department of Earth Sciences and Centre of Natural Hazards and
Disaster Science (CNDS), Uppsala University, Uppsala, Sweden and
Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden.*

T. Alberti

Istituto Nazionale di Geofisica e Vulcanologia, via di Vigna Murata 605, 00143, Rome, Italy

C. Alvarez-Castro

*Universidad Pablo de Olavide, Seville, Spain and
Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici, Bologna, Italy*

T. Caby

*CMUP, Departamento de Matemática, Faculdade de Ciências,
Universidade do Porto, Rua do Campo Alegre s/n, 4169007 Porto, Portugal.*

L. Cavicchia

Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici, Bologna, Italy

E. Coppola

The Abdus Salam International Center for Theoretical Physics, Trieste, Italy

R. Donner

*Department of Water, Environment, Construction and Safety,
Magdeburg-Stendal University of Applied Sciences,
Breitscheidstraße 2, 39114 Magdeburg, Germany and
Potsdam Institute for Climate Impact Research (PIK)—Member of
the Leibniz Association, Telegrafenberg A31, 14473 Potsdam, Germany*

B. Dubrulle

Université Paris-Saclay, CEA, CNRS, SPEC, CEA Saclay 91191 Gif-sur-Yvette CEDEX, France

V. M. Galfi

Department of Earth Sciences, Uppsala University, Uppsala, Sweden

V. Lembo

Istituto di Scienze dell'Atmosfera e del Clima, Consiglio Nazionale delle Ricerche (CNR-ISAC), Bologna, Italy

R. Noyelle and P. Yiou

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

B. Spagnolo and D. Valenti

*Dipartimento di Fisica e Chimica "E. Segrè", Group of Interdisciplinary Theoretical Physics,
Università degli Studi di Palermo, I-90128 Palermo, Italy*

S. Vaienti

Aix Marseille Université, Université de Toulon, CNRS, CPT, 13009 Marseille, France

C. Wormell

Mathematical Sciences Institute, The Australian National University, Canberra ACT 2601, Australia

(Dated: September 21, 2023)

Statistical physics and dynamical systems theory are key tools to study high-impact geophysical events such as temperature extremes, cyclones, thunderstorms, geomagnetic storms and many more. Despite the intrinsic differences between these events, they all originate as temporary deviations from the typical trajectories of a geophysical system, resulting in well-organised, coherent structures at characteristic spatial and temporal scales. While statistical extreme value analysis techniques are capable to provide return times and probabilities of occurrence of certain geophysical events, they are not apt to account for their underlying physics. Their focus is to compute the probability of occurrence of events that are large or small with respect to some specific observable (e.g. temperature, precipitation, solar wind), rather than to relate rare or extreme phenomena to the underlying anomalous geophysical regimes. This paper delineates this knowledge gap, presenting some related challenges and new formalisms which arise in the study of geophysical extreme events and may help better understand them.

I. OPEN CHALLENGES IN THE STUDY OF GEOPHYSICAL EXTREME EVENTS

The aim of this perspective is to bridge statistical physics, statistics, dynamical systems theory and geophysics to provide an overview of current techniques suitable to study high-impact events in the earth system. Examples include temperature extremes, cyclones, thunderstorms or geomagnetic storms, all of which can be interpreted as rare states of the underlying geophysical dynamical systems.

We build upon the evidence that traditional statistical extreme value analysis techniques (see, e.g. [1]) only partly elucidate this class of events. Such techniques successfully define a framework for determining the probability of occurrence of events that are large or small relative to some specific observable (e.g. temperature, precipitation, solar wind). Yet, they are not suitable for describing spatially heterogeneous phenomena or phenomena issuing from largely unprecedented dynamics, such as may be the case for geophysical extreme events, nor to account for their underlying physical drivers. This has motivated the introduction of a number of new mathematical formalisms based on defining extreme events as rare recurrences in the phase space of high-dimensional systems [2–13]. We specifically highlight the work by Lucarini et al. [2], which presents a rigorous framework for the link between extreme value theory and dynamical systems, and [6] which brings together the dynamical systems theory and nonequilibrium statistical physics perspectives. Among the main achievements of the above approaches we note: demonstrating how rare geophysical events correspond to unstable fixed points of the attractor [2, 7, 8], identifying special sets of trajectories corresponding to rare events [3, 5, 12], detecting partial synchronization of system variables [4, 10, 11] and elucidating the spatio-temporal variability of dynamical properties of complex systems [9, 13]. These approaches have recently provided information on a number of geophysical

phenomena, including climate extremes and their atmospheric drivers [14–19], ocean dynamics [20, 21], ecosystems [22], paleoclimate [23, 24], solar wind turbulence and space weather [13, 25], turbulence [26] and earthquake dynamics [27, 28].

Here, we seek to provide a detailed perspective of the above techniques and their links to established concepts from dynamical systems theory, which are applied to a range of data from idealised systems to experimental or observational datasets. Examples include the Lorenz-63 attractor [29], state-of-the-art climate data such as the ERA5 reanalysis [30] or global climate models [31, 32], geomagnetic data [33] and turbulence data [34]. Considering such a wide range of datasets illustrates how approaches developed for low- and medium-dimensional systems can provide insights into high-dimensional chaotic data. The overview also considers the importance of finite size effects (including space and time resolution of the datasets). We thus bridge the gap between theoretical tools often defined with asymptotic limit theorems and complex datasets.

While we do not perform a holistic review of all available mathematical and statistical tools for geophysics and geophysical extremes, we briefly outline here some relevant review studies. For example, Lucarini et al [35] emphasizes the need for a comprehensive understanding of the structural and multiscale properties of climate dynamics. It explores various mathematical and theoretical approaches, including the Nambu formulation of fluid dynamics and statistical mechanics, to construct numerical models and study phenomena like geophysical turbulence and the nonequilibrium nature of the climate system. Another promising theoretical framework for studying geophysical extremes is proposed by [36]. The authors apply Large Deviation Theory, which issues from statistical physics, to problems in geophysical fluid dynamics and climate science. A broader overview of the physics of climate science, including connections to astrophysics is provided by Ghil and Lucarini[6] who emphasize the complexity and non-equilibrium nature of the climate system, highlighting the importance of natural variability, external forcings, and advanced modeling techniques. In the

* These authors have equally contributed to the manuscript

present perspective, we primarily focus on the application of dynamical systems theory and nonequilibrium statistical physics to understand geophysical extremes. As is the nature of a perspective, we highlight selected geophysical extreme events and the associated relevant formalisms.

Our perspective is structured as follows:

- In Section 2 we introduce a general framework for the study of rare events in dynamical systems, the concept of rarity of a recurrence and its link to the asymptotic theory of extreme events. We further discuss the theoretical problems arising in finite datasets and introduce local dynamical systems indicators for rare events.
- In Section 3 we explore how stochastic dynamics and noise-induced phenomena can help explaining geophysical rare events.
- In Section 4 we apply dynamical indicators to understand, model and simulate a variety of geophysical rare events.
- In Section 5 we present opportunities and open challenges for the study of geophysical extreme events. We specifically outline lines of research enabled by recent findings, but that are yet to be explored in the literature.

We conclude with a short summary of the key perspectives and findings outlined in the paper (Sections 6 and 7).

II. HOW TO DEFINE AND TRACK RARE EVENTS IN GEOPHYSICAL DATASETS

Mathematically robust limit theorems allow the characterisation of extreme events in idealised systems. Examples include the use of Extreme Value Theory (EVT), including Generalised Extreme Value (GEV) and Generalised Pareto distributions (GPD), to fit block maxima or threshold exceedances. This assumes an infinite timeseries of i.i.d. random variables. Extremes arising in real-world dynamical systems however, cannot in principle satisfy these requirements. When studying these systems it is thus important to go beyond a pure statistical analysis and also include information on the dynamics. This leads to a distinction between extreme events as defined by classical EVT, and extreme events in a dynamical systems theory context, which can be equated to rare events. In the latter sense, the two terms will hereafter be used interchangeably.

In order to quantify the rarity of an event, a key tool in dynamical systems theory is the Poincaré recurrence theorem. This theorem states that for a conservative dynamical system with a finite phase space, if the system evolves for a sufficiently long time, it will return arbitrarily close to any initial position. In other words, the system will exhibit recurrent behavior and, although rare,

certain states of the system will recur infinitely often. More specifically, let $\mu(X, T)$ be a measure μ -preserving dynamical system, where X is the phase space and T is the time evolution map. The Poincaré recurrence theorem states that for almost every point x in X and any neighborhood U of x , there exists a positive integer n such that $T^n(x)$ (the point x after n iterations of the time evolution map) belongs to U . Although the theorem was formulated for conservative dynamical systems, it was later shown that similar conclusions hold for chaotic forced dissipative systems having a strange attractor (e.g. [37]).

The Poincaré recurrence theorem has applications in various fields, including statistical mechanics and ergodic theory. The time required for a trajectory to return to its initial state, known as the recurrence time, is related to the system's phase space volume and the speed of evolution along the trajectory, and can be extremely long for rare states. By studying the recurrence times of trajectories near these rare points, it is possible to gain insights into the underlying dynamics of the system and to identify the most likely pathways leading to the rare events.

A. Probabilistic theories for extreme events in geophysical time series

The two main probabilistic approaches to analyse geophysical extreme events are EVT and Large Deviation Theory (LDT). Whereas EVT deals with tails of probability distributions, LDT considers, in its basic formulation, probabilities of sample averages. From a mathematical perspective, these theories are formulated for i.i.d. random variables. However, from a practical perspective, the main requirement is that the extreme values or the sample averages are uncorrelated and homogeneous. This is a reasonable assumption if the system is chaotic enough (see Sect. 3.2 in [2]), correlations decay rapidly, and trivial non-stationarities [38, 39], such as seasonal or other regular cycles or anthropogenic climate change signals, are accounted for.

EVT contemplates two main approaches, the *block maxima* and the *peak over threshold* approaches. Both are asymptotic theories of probability distributions of extreme values, with the limit being approached by selecting more and more extreme states. However, the two approaches differ in the way extreme values are selected.

In case of the *block maxima* approach, extreme values are defined as maxima $M_n = \max\{X_1, X_2, \dots, X_n\}$ of a series of i.i.d. random variables X_1, X_2, \dots divided into blocks of equal length n . If certain conditions are fulfilled, the distribution of properly normalised block maxima M_n converges to a so-called GEV distribution for $n \rightarrow \infty$ [40]. In practice, one selects the maxima over fixed time periods, e.g. 1 year, and verifies whether their distribution is properly described by a GEV distribution. The GEV distribution has three parameters: location parameter (κ), scale parameter (σ), and shape parameter

(ξ). The probability density function (PDF) of the GEV

$$f(x; \kappa, \sigma, \xi) = \frac{1}{\sigma} \left\{ 1 + \xi \left(\frac{x - \kappa}{\sigma} \right) \right\}^{-1/\xi - 1} \exp \left\{ - \left[1 + \xi \left(\frac{x - \kappa}{\sigma} \right) \right]^{-1/\xi} \right\}. \quad (1)$$

The sign of the shape parameter determines the type of GEV distribution:

- $\xi > 0$: The distribution has a finite upper endpoint, and is called the Fréchet distribution. It is used to model distributions with a heavy tail.
- $\xi < 0$: The distribution has a finite lower endpoint, and is called the Weibull distribution. It is used to model distributions with a light tail.
- $\xi = 0$: The distribution reduces to the Gumbel distribution, which is used to model the extreme value distribution of a normal random variable.

In case of the *peak over threshold* approach, we consider again i.i.d. random variables X_1, X_2, \dots and select extreme values as values exceeding a certain high threshold u . If u is high enough, the distribution of threshold exceedances $X - u$ follows a GPD [40]. In practice, similarly to the block maxima approach, one verifies that the distribution of exceedances of a high threshold is properly approximated by a GPD. The probability of exceedance of a high threshold u by X_i , i.e. $Y_i = X_i - u$ with $Y_i > 0 \forall i$, is given by a cumulative distribution function that converges to the GPD family:

$$H(y) = 1 - \left(1 + \frac{\xi y}{\sigma} \right)^{\frac{1}{\xi}} \quad \text{for } \xi \neq 0, \quad (2)$$

$$H(y) = 1 - \exp^{-\frac{y}{\sigma}} \quad \text{for } \xi = 0. \quad (3)$$

where $1 + \frac{\xi y}{\sigma} > 0$ for $\xi \neq 0$, $y > 0$ and $\sigma > 0$. In a similar fashion as the GEV distribution, if $\xi = 0$ the tail decays exponentially, if $\xi > 0$ it decays polynomially and if $\xi < 0$ it has an upper bound. The scale parameter describes how wide the distribution is, and is related to the variability of the extremes.

If the GEV or GPD are adequate models of the distribution of extreme values, one can estimate the return periods of very rare and even unobserved extreme events.

LDT deals with the probability of sample averages $A_m = \frac{1}{m} \sum_{i=1}^m X_i$ over blocks of equal length m . Under adequate conditions, the probability of A_m decays exponentially for $m \rightarrow \infty$:

$$\mathbb{P}(A_m = a) \approx e^{-mI(a)}, \quad (4)$$

where $I(a) \geq 0$ is the so-called rate function [41]. If (4) holds and $I(a)$ has a unique global minimum, the probability of averages decays exponentially everywhere with increasing m , except at the mean $\mathbb{E}[A_m] = \bar{a}$ where

distribution is given by:

$I(\bar{a}) = 0$ and $P(A_m = \bar{a}) = 1$. This points to the convergence of sample averages to the real mean as described by the law of large numbers [42]. In case of applications to geophysical time series, one verifies whether (4) holds for long enough averaging periods $m \geq m^*$, i.e. whether $I(a_{m > m^*}) \approx I(a_{m^*})$. If this is the case, it is possible to estimate the probabilities of averages far away from \bar{a} over time periods longer than m^* based on $I(a_{m^*})$. Hence, one obtains the probability of very rare, even unobserved anomalies from the long-term mean in terms of both their average intensity and duration [12, 19].

EVT primarily focuses on the intensity or magnitude of extreme values. The duration of extreme events can be incorporated in form of the mean cluster size, based on the extremal index (see Eq. 7). LDT considers temporally averaged anomalies, and includes the event duration via the averaging time length. Thus, it is especially useful to study persistent extreme events, such as heatwaves or cold spells, and periods during which extreme values in some observable are unusually frequent [12, 19].

The above description of these methods relies on time series of observables and takes a stochastic or probabilistic perspective. A number of geophysical systems, such as the atmosphere or ocean, are (dissipative) chaotic deterministic systems. If the system is chaotic enough (see Sect. 3.2 in [2]), the stochastic perspective is reasonable and the above methods work. In the following sections, we show how by combining these probabilistic methods with dynamical systems theory, new concepts and methods arise. These have the potential to provide a deeper understanding of the structure of chaotic attractors and the dynamics of geophysical systems.

B. An asymptotic framework to study extreme events in dynamical systems

In the context of dynamical systems, an application of EVT informed by the dynamics of the system, allows formalizing two questions related to the recurrence statistics of specific system states. Suppose we have a set U in the phase space, of small measure, which we accordingly qualify as a *rare set* or as the location of an *extreme event*. A first interesting question is: What is the probability that the first visit to U of our physical system is larger than some prescribed time n ? Suppose now that the system entered the set U . A second question is: what is the probability that it resides there k times in a prescribed time interval? These two questions allow quantifying the

rarity of an event and its persistence – both essential characteristics for understanding geophysical extremes.

We next attempt to give mathematically workable answers to the above questions. To do so, we adopt the framework of discrete dynamical systems, which allow for a simpler formalism. We will consider the temporal evolution given by a discrete dynamical system, or map, T . This acts on some compact metric space M , with distance $\text{dist}(\cdot, \cdot)$, and carrying a T -invariant measure μ which will be the underlying (stationary) probability describing the statistical properties of the system. For a dissipative system, this defines the system's attractor. The rare set U will change with n and we will denote it with U_n , $n \geq 1$; moreover $\mu(U_n) \rightarrow 0$ when n goes to infinity. In order to get rigorous limit theorems, the sequence U_n is usually taken as monotonically decreasing and converging to a null set Λ . We now fix a positive number t and we choose the sets U_n in such a way that

$$\mu(U_n) = \frac{t}{n}. \quad (5)$$

This is usually possible whenever μ is not atomic, namely a measure concentrated at a single point, and the sets U_n are sufficiently regular, for instance they are balls shrinking around a point or strips collapsing on a smooth submanifold. We then define, for $x \in M$:

$$\tau_{U_n}(x) := \inf\{k \geq 1, T^k x \in U_n\},$$

which is the first time the initial point x enters the set U_n . To address our first question, we compute:

$$\mu(\tau_{U_n} > n).$$

The use of EVT combined with strong chaotic properties of the dynamical systems (T, μ) allows us to prove the existence of a positive number $\theta \in [0, 1]$ such that

$$\lim_{n \rightarrow \infty} \mu(\tau_{U_n} > n) = e^{-\theta t}. \quad (6)$$

This is Gumbel's law. The number θ is called the *extremal index (EI)*. We refer the reader to [2] and [9] for a detailed discussion of the dynamical systems for which the previous result holds; in a few words: they are hyperbolic dynamical systems with exponential decay of correlations.

The EI can be related to a local persistence indicator [43], suitable to estimate the average cluster size of the trajectories within the neighborhood of the null set Λ . Such an index is in fact less than one when clusters of successive recurrences happen, which is the case, for instance, when the set U_n shrinks around a periodic point or, more generally, around an invariant submanifold. Given a time interval Δ , the distribution of λ_l for a cluster of size l , can be defined as the frequency of visiting l times the set U_n . We give in the Appendix the analytic

expression of λ_l . This holds in the limit of $n \rightarrow \infty$ and $\Delta \rightarrow \infty$. It can then be shown that [10]:

$$\theta^{-1} = \sum_l l \lambda_l. \quad (7)$$

In other words, the EI can be interpreted as the inverse of the expected cluster size of recurrences about U_n . For the strongly chaotic systems quoted above, the EI can be computed by estimating suitable return time functions. We in fact have the formula:

$$\theta = 1 - \sum_{k=0}^{\infty} q_k, \quad (8)$$

with $q_k = \lim_{n \rightarrow \infty} q_{k,n}$, provided the limit exist, and where the analytic expression of the $q_{k,n}$ is given in the Appendix. They have however a clear interpretation, being the probability to start in U_n and then return to it after exactly $k + 1$ steps. This result is based on a perturbative technique introduced by Keller and Liverani [44, 45]. It has been used in the context of EVT for both deterministic and random systems, providing very efficient explicit formulae for the computation of the extremal index [9, 46, 47].

Let us now suppose that the set U_n is a ball $B(z, e^{-u_n})$ of center $z \in M$ and radius e^{-u_n} with $\mu(B(z, e^{-u_n})) = \frac{t}{n}$ and $u_n \rightarrow \infty, n \rightarrow \infty$. If we define the process $X_i(x) := -\log \text{dist}(T^i x, z)$, we can then elucidate the connection between hitting times and EVT as:

$$\lim_{n \rightarrow \infty} \mu(M_n \leq u_n) = \lim_{n \rightarrow \infty} \mu(\tau_{B(z, e^{-u_n})} > n) = e^{-\theta t}, \quad (9)$$

where $M_n = \max(X_0, X_1, \dots, X_{n-1})$. The condition $\mu(B(z, e^{-u_n})) = \frac{t}{n}$ now becomes, for large n ,

$$e^{-u_n d_\mu(z)} \sim \frac{t}{n},$$

where $d_\mu(z)$ is an estimate of the local dimension at the point z . Leveraging GEV theory [48], it is possible to fit the boundary level u_n , and therefore the dimension $d_\mu(z)$, provided that the convergence in (9) holds. The EI θ computed with formula (8) will also be a function of the point z , and we therefore write $\theta(z)$. The couple $(d_\mu(z), \theta(z))$ allows us to characterise the local structure of attractors. In [49] it was shown that an alternative approach for the selection of the maxima is to use the peak over threshold approach, allowing to estimate the couple $(d_\mu(z), \theta(z))$ by leveraging the GPD. This enables computing the couple for a larger class of dynamical systems than those satisfying the conditions for the block maxima approach [2]. In particular, the GPD approach is suitable for periodic and quasi-periodic systems, with the exclusion of periodic points, and non-axiom A chaotic dynamical systems with multifractal properties. The above include a large number of geophysical systems (see the examples provided in the Introduction).

We now come to our second question. The orbit of our dynamical system visits infinitely often the set U_n . We could therefore expect that the exponential law $e^{-\theta t}$ given by the extreme value distribution describes the time between successive events in a Poisson process. To formalize this, we introduce the random variable

$$\mathcal{N}^{(n)}(t) := \sum_{j=0}^{\frac{t}{\mu(U_n)}} \mathbf{1}_{U_n} \circ T^j, \quad (10)$$

and we consider the following distribution:

$$\mu(\mathcal{N}^{(n)}(t) = k). \quad (11)$$

For a large class of dynamical systems, it has been proven in [50] and also in [10] that

$$\mu(\mathcal{N}^{(n)}(t) = k) \rightarrow \nu(\{k\}),$$

as $n \rightarrow \infty$, where ν is the compound Poisson distribution for the parameters $s\lambda_l$, where $s = \theta t$. The generating function of a compound Poisson distribution ν with parameters $s\lambda_k$ reads:

$$\varphi_\nu(z) = e^{\sum_{k \geq 1} s\lambda_k (z^k - 1)}.$$

When $\lambda_1 = \lambda$ and $\lambda_k = 0$, $k \geq 2$, we obtain the Poisson distribution with parameter $t\lambda$. An integer valued random variable W is compound Poisson distributed if there are i.i.d. integer valued random variables $\{X_i\}_{i \geq 1}$, and an independent Poisson distributed random variable N so that $W = \sum_{j=1}^N X_j$. The Poisson distribution of N describes the distribution of clusters whose sizes are given by the random variables X_j whose probability densities are the values $\lambda_l = \mathbb{P}(X_j) = l$, $l = 1, 2, \dots$. Therefore, for the class of chaotic systems considered here, we get the distribution of clusters which occur on a large timescale, and the number of returns in each cluster is given by the random variables X_j . These returns are on a fixed timescale and nearly independent of the size of the return set as its measure is shrunk to zero.

For instance, when $\lambda_\ell = (1-p)p^{\ell-1}$ we have that $\theta = 1-p$ and $N \sim \text{Poisson}(t(1-p))$. This specific case is termed *Pólya-Aeppli* distribution, with parameter $t(1-p)$. The mass distribution $\nu(\{k\})$ of such a distribution is recalled in the Appendix. It is interesting to note that we have two families of indices to compute respectively the EI and the compound Poisson, namely the q_k and the λ_l . The two are related since $\theta^{-1} = (1 - \sum_{k=0}^{\infty} q_k)^{-1} = \sum_l l\lambda_l$, which shows, as we already pointed out, that the EI is the inverse of the expected cluster size. This leads to a whole spectrum of λ_l , or alternatively of the associated compound Poisson statistics, providing insights into the distribution of the clusters even in systems with multiple timescales [9, 10, 51]. For example, [9] studied the statistics of the number of visits of a given observable in the neighborhood of a particular value, for a vector containing the

values of sea-level pressure on a grid of ≈ 100 locations over the North Atlantic. The resulting distribution was very close to Pólya-Aeppli (see Fig. 11 in [9]), with an excellent agreement with the observational data.

C. Finite time approaches to study rare events in dynamical systems

The results of EVT are stated with respect to the physical invariant measure μ supported on the system's attractor M , and are asymptotic results. Yet when it comes to the study of geophysical systems, the measuring process often allows to collect only a limited amount of data. These limitations can be due to short time series, yielding pre-asymptotic estimates, or to a measurement process that is partial, i.e. that does not account for all the variables of the system. In this section, we discuss how such observational data can still be relevant for the study of high-dimensional complex systems.

The term *penultimate attractor* was coined to describe the union of the finite-size trajectories along which data are collected. The characteristics of the actual attractor of the system can be deduced from those of its penultimate attractor, using the framework of large deviations. We will here state results concerning the local dimensions as a paradigmatic example, but this framework can be applied to other dynamical quantities of interest, such as recurrence times, which are closely related objects [46]. The local dimension at a point x in phase space is defined as

$$d_\mu(x) = \lim_{r \rightarrow 0} \frac{\log \mu(B(x, r))}{\log r}, \quad (12)$$

where $B(x, r)$ as before denotes a ball centered at x of radius r . Its pre-asymptotic version

$$d_{\mu,r}(x) = \frac{\log \mu(B(x, r))}{\log r}, \quad (13)$$

converges for almost all states $x \in M$ to the so-called information dimension D_1 of the system as $r \rightarrow 0$. Depending on the length of the available time series, one can fix a small radius $r > 0$ and evaluate $d_{\mu,r}$ at this finite resolution, for example by fitting the empirical distribution of the maximum associated with a suitable observable. When the radius r is small, but not zero, deviations from the typical value D_1 can be observed, and are modulated by a rate function [52]:

$$Q(s) = \sup_{q \in \mathbb{R}} \{-qs + qD_{q+1}\}, \quad (14)$$

which is convex and vanishes at D_1 . Q depends only on the generalized dimensions of the system, defined for $q \neq 1$ by:

$$D_q = \lim_{r \rightarrow 0} \frac{\log \int_M \mu(B(x, r))^{q-1} d\mu(x)}{(q-1) \log r}.$$

This spectrum of dimensions is non trivial for systems exhibiting a wide variety of scaling behaviors, as expected for geophysical systems such as the climate system. A large deviation relation is explicit in the following [46]:

$$\mu(\{z \in M \text{ s.t. } d_\mu(z, r) \in I\}) \underset{r \rightarrow 0}{\sim} r^{\inf_{s \in I} Q(s)}. \quad (15)$$

Here I denotes an open interval. In the sense of the above formula, the wide distribution of (finite-resolution) local dimensions over the phase space originates from the multi-fractal structure of the attractor, which exhibits a wide variety of scaling behaviors. It is easy to see from (13) that for a given radius r , the value of $d_{\mu, r}(x)$ is determined by the measure of a ball centered at x . Smaller values of $d_{\mu, r}$ correspond to regions of the attractor that are more visited by the dynamics, whereas higher values correspond to less dense regions of the attractor.

Partial observation of the system can be modeled by an observable $f : M \rightarrow \mathbb{R}^k$ that acts as a projection from the phase space of the system to the observational space \mathbb{R}^k . Examples of such f include gridded observables, delay-coordinate observables used in embedding techniques, or a concatenation of various scalar observables of interest. The measuring process consists in collecting the values of f along a typical trajectory of the system. The obtained observational data are sampled with respect to the image measure μ_f . This measure is such that for all $A \subset \mathbb{R}^k$,

$$\mu_f(A) = \mu(f^{-1}(A)).$$

The statistical properties of the observable f are governed by the geometric structure of the support of this measure. In particular, the recurrence properties are modulated by the local dimensions, and the synchronization properties (several or all variables of the system attain similar values) are governed by the generalized dimensions of μ_f [9, 53].

It is crucial to understand the relations between the properties of these observations and the ones of the underlying physical system that is being observed. The following result, by [54], provides an insight into the matter. For a generic, smooth observable $f : M \rightarrow \mathbb{R}^k$ and for a generic point $x \in M$, the local dimension of the image measure at $f(x)$ is given by

$$d_{\mu_f}(f(x)) = \min(k, d_\mu(x)). \quad (16)$$

From here, there are two possibilities:

- If $k < d(x)$, then $d_{\mu_f}(f(x)) = k$, and the information on the local structure of underlying attractor is lost.
- If k is large enough, the information dimension of the underlying system is preserved, since for μ -almost all $x \in M$:

$$d_{\mu_f}(f(x)) = d_\mu(x)$$

It is therefore possible to access the geometric structure of the underlying attractor by using observational data. For instance, in the case of a gridded observable one can increase the spatial resolution k of the grid, and estimate for each k the dimensions d_{μ_f} until they stabilize around some values that should correspond to the dimensions of the underlying attractor.

D. Local dynamical indicators for rare events

The estimation of the above-mentioned local dimension d rests on the application of extreme value theory to Poincaré recurrences in dynamical systems. This approach was first introduced in [2] and [14]. For a given point z in phase space (e.g., a given sea-level pressure map over a given geographical domain), one computes the probability that the system returns within a ball of radius e^{-u} centered on the point z . The Freitas *et al.* [55] theorem, modified by Lucarini *et al.* [49], states that the negative logarithmic returns

$$g(x(t)) = -\log(\text{dist}(x(t), z)) \quad (17)$$

yield a probability distribution such that

$$\Pr(z > s(q)) \simeq \exp \left[-\theta(z) \left(\frac{z - \kappa(z)}{\sigma(z)} \right) \right], \quad (18)$$

where $z = g(x(t))$ and u is a high threshold associated to a quantile q of the series $g(x(t))$. Requiring that a point on the orbit falls within a ball of radius e^{-u} around the point z is equivalent to asking that the corresponding value of the series $g(x(t))$ is above the threshold s ; therefore, the ball radius r is simply $e^{-u(q)}$. The resulting distribution is the exponential member of the GPD family. The parameters κ and σ , namely the location and the scale parameter of the distribution, depend on the point z in phase space. $\kappa(z)$ corresponds to the threshold $u(q)$ while the local dimension $d(z)$ can be obtained via the relation $\sigma = 1/d(z)$. This is the local dimension introduced in Sect. II C. A metric of persistence can be obtained as the inverse of the extremal index, dimensionalised by the timestep of the data being used.

The same framework can also be applied in a bivariate context. Given two observables $x(t)$ and $y(t)$, one can define the state $z = \{z_x, z_y\}$. The joint negative logarithmic returns can then be defined as:

$$g(x(t), y(t)) = -\log \left[\text{dist} \left(\frac{x(t)}{\|x\|}, \frac{z_x}{\|z_x\|} \right)^2 + \text{dist} \left(\frac{y(t)}{\|y\|}, \frac{z_y}{\|z_y\|} \right)^2 \right]^{\frac{1}{2}} \quad (19)$$

where $\|\cdot\|$ is the average root mean square norm of the coordinates of a vector. For example, $\|x\| = E_t \left(\left[\sum_i^K x_i(t)^2 \right]^{\frac{1}{2}} \right)$, where K is the number of components of x and E_t is an average over time t . Based on Eq. (19), one can then compute the co-dimension $d_{x,y}$.

A similar approach can be applied to persistence. One can define an adimensional inverse co-persistence $\theta_{x,y}$ based on series of consecutive days on which both observables display recurrences of z_x and z_y . $\theta_{x,y}$ is a weighted average of θ_x and θ_y , where the weights depend on the density of observed neighbours of z .

Finally, in a bivariate setting one can define an additional indicator, which we term co-recurrence ratio, as:

$$\alpha(z) = \frac{\nu[g(x(t)) > s_x(q) \cap g(y(t)) > s_y(q)]}{\nu[g(x(t)) > s_x(q)]} \quad (20)$$

with $0 \leq \alpha \leq 1$, provided that the same number of recurrences is defined for both observables. Here, $\nu[-]$ is the number of events satisfying condition $[-]$, and all other variables are defined as before. Thus, α quantifies how often the two observables have joint recurrences, namely their co-recurrence. If the same number of recurrences is defined for both observables, then by definition α is symmetric with respect to the choice of variable (x or y), since $\nu[g(x(t)) > s_x(q)] \equiv \nu[g(y(t)) > s_y(q)]$.

The local dimension, persistence and co-recurrence ratio can be computed at each point in the phase space of a physical system, including when the phase space is imperfectly sampled. For the local dimension, averaging over the whole attractor gives an estimation of the information dimension of the underlying attractor. In the limit of an infinite number of infinitely long trajectories, the local dimension of all points on the attractor will almost surely tend to the information dimension of the attractor. Hence, all local dimensions would be equal, with the exception of a negligible set (which nevertheless includes objects such as periodic points and chaotic saddles). With a finite size sample, this is not the case and the value of the local dimension computed at each point gives information on the recurrences of these points in the phase space, namely on the geometric characteristics of the trajectory [2].

The above dynamical systems metrics are typically estimated without taking explicitly into account the continuous nature of the trajectories representing the evolution of geophysical systems (see e.g. [14]). We propose here an estimator for the local dimension d targeted to continuous dynamical systems, and investigate its second order properties (i.e. we compute its gradient in the phase space). The new estimator uses the N closest re-

currences of z within a radius e^{-u} , with u given. The local dimension $d_u(z)$ is given by:

$$d_u(z) = \frac{\sum_{n=1}^N v_n^{-1} l_n}{\sum_{n=1}^N v_n^{-1} (l_n - r_n \cos^{-1}(e^u r_n))}, \quad (21)$$

where v_n is the speed in the phase space at the n -th closest recurrence, r_n is the distance between the point z and the n -th closest recurrence and $l_n = \sqrt{e^{-2u} - r_n^2}$.

We emphasise that even assuming infinite data, this is only a consistent estimator as $u \rightarrow 0$ for measures with a continuous density on a submanifold, for example those generated by stochastic processes. For fractal attractors it will converge only in the sense of Césaro averaging, and even in that setting is not known to be consistent. Nevertheless, at finite resolution it provides useful information about the scaling of the system.

In a simple low-dimensional attractor such as for the Lorenz-63 system [29], one can visualize the regions with low and high dimensions (cf. Fig. 1a, b). The two estimates of local dimension based on the GPD and on Eq. 21 are essentially indistinguishable. We further note that for a point situated in the low-dimensional part of one of the wings of the attractor, its local dimension can be increased by moving either towards the exterior of the wing or towards the fixed point in the middle of the wing.

However, in a high-dimensional system, it is far from easy to know which direction in the phase space would increase or decrease the local dimension, i.e. which are the common and rare structures with respect to the sampling provided by the data set. We therefore propose an estimator of the gradient in the phase space of the quantity $d_u(z)$ defined above.

The gradient $\nabla_z d_u(z)$ is a vector in a phase space which has the point z as its origin. It therefore points towards the direction of increasing local dimension relative to the point z . This gradient is computed on the Lorenz-63 model. The directions of the normalized gradients $\frac{\nabla_z d_u(z)}{\|\nabla_z d_u(z)\|}$ in Fig. 1c at each phase space point match the intuition with regards to the distribution of local dimensions in Fig. 1b.

As discussed in Sect. II C, applications of dynamical systems theory to geophysical data must account for short or incomplete sets of data. When computed on a data set with finite size, i.e. undersampling the true attractor of the dynamical system studied, the local dimension provides information on the clustering of nearest neighbors in the phase space. Here we proposed to expand this characterization by computing a second order statistics, namely the gradient in phase space of local dimension. We illustrate how this can inform on high-dimensional geophysical systems in Sect. IV B.

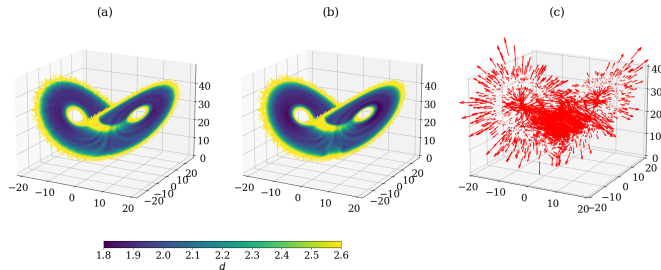


FIG. 1. **Gradient of local dimension for the Lorenz-63 system.** (a) Local dimension of the Lorenz-63 model using the GPD estimator; and (b) the estimates based on Eq. 21. (c) Corresponding gradient of the local dimension.

E. Applying local dynamical indicators to rare events in turbulent flows

Virtually all geophysical flows present an ubiquitous property: they are turbulent. One of the properties of geophysical turbulent flows is their chaotic and unpredictable nature which suggested to explore the connection between turbulence, chaos, and dynamical systems. In Sections II B-II D we have presented local dynamical systems indicators which are suitable for systems where the chaotic components dominate over turbulence. Here, we investigate how to treat systems where chaos and turbulence coexist and shape the system dynamics. Nevertheless, a *universal* underlying turbulent attractor has not been proven to exist for turbulent flows. A promising way to characterized highly turbulent flows in constrained geometries is the use of stochastic low-dimensional dynamical systems [56]. This approach suggests that turbulent flows live on non-hyperbolic strange stochastic attractors [57]. This property, i.e., non-hyperbolicity, is exactly strictly connected with non-homogeneity in terms of instantaneous properties of turbulent fluctuations, closely related with the universality of the statistics of fluctuations with respect to the mechanisms producing turbulence.

Turbulence, which is one of the unsolved problems of classical physics, manifests itself via vortices (or eddies) of different sizes hierarchically organized in a self-similar way [58]. Turbulent flows are usually described in a statistical way derived from symmetry assumptions (homogeneity, stationarity, self-similarity) of the Navier-Stokes equations describing the dynamics of averaged quantities over spatial scales larger than the mean free path length of molecules [59]. We begin here by outlining the statistical theory that has been developed by Kolmogorov [58], based on the *universality* of the energy spectrum when normalized to the fluid viscosity ν and the energy transfer rate ϵ , i.e.,

$$S_2(\ell) \propto \epsilon^{2/3} \ell^{2/3}. \quad (22)$$

By assuming that vortices are hierarchically organized,

i.e., they are derived from a cascade mechanism, the Kolmogorov universality concept should extend to higher moments field fluctuations. This means that if δu_ℓ is a measure of the velocity at scale ℓ , then Eq. (22) can be generalized as

$$S_q(\ell) = \langle \delta u_\ell^q \rangle \propto \epsilon^{q/3} \ell^{q/3}, \quad (23)$$

where $\langle \dots \rangle$ stands for time average. Unfortunately, the global self-similarity assumption was one of the main failures of Kolmogorov theory, as shown by experimental evidence [60, 61]. The breaking of this assumption resides in the *local* nature of the energy transfer rate ϵ , i.e., the phenomenon of intermittency arising from localized bursts of activity, both in time and in space. This was the starting point of the multifractal formalism developed by Parisi and Frisch [60], based on the assumption that the energy transfer rate depends on the scale we are focusing on, i.e.,

$$S_q(\ell) \propto \epsilon_\ell^{q/3} \ell^{q/3}. \quad (24)$$

Under the assumption of *local self-similarity*, i.e., $\epsilon_\ell \sim \ell^{\mu(q/3)}$, Eq. (24) can be written as

$$S_q(\ell) \propto \ell^{\zeta(q)} \quad (25)$$

where the scaling exponent $\zeta(q)$ accounts for all possible (infinite) rescaling symmetries of the Navier-Stokes Equations, i.e., accounts for the existence of singularities in the energy cascade mechanism [61, 62]. This refined theory has successfully described the statistical properties of the velocity field in experiments [63–68]. However, this approach only provides time-averaged statistical information on velocity field fluctuations at different scales, lacking an instantaneous description of the dynamics, the latter being particularly helpful for exploring the local statistics of velocity field fluctuations [26, 61].

Very recently, starting from the non-universality of the statistics of small-scale fluctuations, Alberti *et al.* [69] proposed a time-dependent and scale-dependent framework for retrieving information on the symmetries of a turbulent steady state. The main idea is to overcome the limitation given by previous measures of singularities (as the scaling exponents $\zeta(q)$ or the generalized dimensions D_q [46]), tracing time-averaged features of field fluctuations, by introducing two measures which are time-dependent. In this way, one answers Landau's objection to the Kolmogorov theory on the time and space fluctuations of the energy transfer rate, providing an instantaneous view of scale-dependent fluctuations. Furthermore, this allows us to deal with the problem of intermittency and its interpretation as a time- or space-dependent phenomenon. The newly introduced measures can be derived by searching for extremes of the field fluctuations. We assume, for simplicity, a 1-D velocity field $u(t)$ (nonetheless, the method is suitable for any n -D field) which can be written as:

$$u(t) = \langle u \rangle + \sum_{\ell} \delta u_{\ell}(t) \quad (26)$$

where $\langle u \rangle$ is the mean field value (sometimes, the notation in turbulence is U_0) and $\delta u_\ell(t)$ is the fluctuating component of the velocity field at the scale ℓ [70].

An additional feature associated with non-universality are intermittent phenomena. These are ubiquitous in turbulent flows and play an important role in dissipative energetics [71]. Intermittency can be described using the multifractal formalism [72], which among other things showed that turbulence lives on a strange attractor with sudden, large-amplitude, low-frequency, localized bursts (i.e., rare events), induced by intermittency. An appropriate framework to deal with this behaviour is EVT, where state(time)-dependent metrics can be derived [2]. We specifically seek to determine time-dependent EVT-based metrics for each velocity fluctuating component at each scale ℓ . To this purpose, we use two metrics from EVT: the local dimension ($d(t)$, introduced in Section IID) and the extremal index ($\theta(t)$, introduced in Section IIB). In this way, instead of having a single pair ($d(t), \theta(t)$) of descriptors of the turbulent flow we have a hierarchy of pairs, each associated with the dynamics of the fluctuating field at a scale ℓ , i.e., ($d_\ell(t), \theta_\ell(t)$). Thus, a time- and scale-dependent view of the system is obtained.

A widely studied laboratory setup which displays the joint properties of chaoticity and turbulence is the von Karman swirling flow [73]. For such a flow under high Reynolds number turbulent conditions, the above indicators provide evidence of a scale-dependent underlying attractor whose geometric and topological properties depend on the large-scale forcing and also affect the distribution of the energy across the inertial range of scales [56]. This mirrors the role of singularities that break the global self-similarity and that can be considered the main reason of the failure to find a *universal* attractor for turbulent flows. These results also shed new light on the role of the inertial range of scales, where the turbulent cascade takes place, in distributing the energy injected at larger scales. Indeed, around the injection scales the individual scales are in quasi-equilibrium with low values of the local dimensions $d_\ell(t) \sim 2 - 3$. Conversely at smaller scales than the injection one the mean energy transfer is positive and the out-of-equilibrium energy cascade transfers energy to the small scales where viscous effects become relevant. Thus, while the the statistical equilibrium at large scales is driven by a few degrees of freedom, generating a well defined low-dimensional attractor, the dynamics at scales smaller than the injection scale restore the symmetry broken by the cascade and generate a stochastic attractor. Beyond turbulent flows, this formalism can be applied to any time series and/or dynamical system. As an example, Alberti *et al.* [74] have shown that the formalism can be used to disentangle the role and the nature of noise in dynamical systems, allowing to clearly distinguish between a purely noise-like contribution, being characterized by an ergodic coverage of the available phase-space with dimensions fluctuating around 3 (as expected), and

a more forcing-like contribution for a multiplicative noise, with dimensions larger than 3 which are distributed differently across the attractor.

III. STOCHASTIC DYNAMICS FOR UNDERSTANDING AND SIMULATING GEOPHYSICAL RARE EVENTS

The modeling of chaotic systems is a challenging task because these systems are highly sensitive to initial conditions and external forcings, and small perturbations can lead to significant changes in their behavior. Examples include modulating the fine structure of the attractor [13, 75]; or leading to (abrupt) shifting between different basins of attraction [76, 77], underlying major changes in the dynamics of the system. The latter shifts are often termed critical phenomena [78] or tipping points [79] in the climate and physical sciences, while they are referred to as bifurcations leading to shifts between metastable basins of attraction in a mathematical context. Thus, the presence of noise can change in a fundamental way the physics of a complex system [76]. Noise-induced phenomena also exist in the Earth system, and the possibility of high-impact climate tipping points has garnered increasing attention in recent years [79, 80].

Indeed, complex systems are multistable systems characterized by metastability and noise usually manifests itself in a nonmonotonic dependence of the average lifetime of a metastable state on the noise intensity (or temperature). The interplay between the nonlinearity of complex systems and the environmental stochastic fluctuations can lead to new, counterintuitive dynamical behavior, such as noise-induced phase transitions, emergence of ordered spatial structures [81–86], stochastic resonance [87, 88], stochastic resonant activation [89], noise-induced synchronization [90], noise-induced switching [91], and noise enhanced stability [89, 92]. The latter phenomenon is a consequence of the interplay between the random fluctuations and the nonlinearity of the underlying complex system.

Dynamical systems theory is concerned with the study of systems that evolve deterministically over time, while statistical physics deals with the study of systems that evolve stochastically due to random fluctuations. The stochastic dynamical systems approach combines these two perspectives by modeling the deterministic evolution of a system, yet incorporating stochastic fluctuations due to environmental noise – for example using stochastic differential equations. This overcomes the limitations of traditional modeling techniques, such as differential equations and numerical simulations, that may not capture the full complexity of physical systems. Stochastic dynamics are relevant to many real-world systems across virtually all scales, from electrons or quarks to stars in a galaxy and galaxies in the Universe. Stochastic dy-

ynamical systems theory has consequently emerged as a useful tool for understanding the behavior of complex systems that also exhibit chaotic dynamics. By incorporating randomness and noise, stochastic models can capture the complexity and unpredictability of chaotic systems, including the complex behavior that arises from small perturbations. Dynamical indicators, such as local dimensions, persistence and Lyapunov exponents, can then be used to estimate the rate of divergence of nearby trajectories in stochastic dynamical systems. In this section, we will explore how the application of stochastic dynamical systems theory can serve as a connection between the approaches to extreme events described earlier and the modeling of chaotic systems.

A. Simulating Noise-induced Phenomena in Nonequilibrium Complex Systems

We present three examples of how environmental noise plays a crucial role in understanding the dynamics of non-equilibrium complex systems. We first consider the mean switching time of a Josephson junction (JJ), directly coupled to a dark matter axion field and subject to both a continuous bias current and thermal fluctuations. A JJ is composed of two or more superconductors coupled by a weak link, while the axion is a hypothetical particle that was originally proposed to solve a problem in the theory of strong nuclear interactions, but it has since been suggested as a candidate for dark matter. One way to search for axions is to look for their effects on various physical systems, for example detecting the axion presence through the analysis of the escape times from the initial superconducting state of a JJ (Fig. 2). By studying the switching time statistics we find a resonant activation phenomenon based on the plasma frequency, induced on the JJ by the axion (Fig. 3, [93]). This approach is widely applicable, including to coupled qubit architectures for quantum computing and other technological systems.

Next, we consider a sine-Gordon (SG) breather. The latter is a type of non-linear wave solution that arises in the study of the sine-Gordon equation. These breathers are characterized by their ability to form localized, periodic wave packets that oscillate in time and space. We explore numerically the influence of both dissipative and stochastic perturbing terms on the dynamics of the breather, starting from a stationary breather with a random phase value. We find that a spatially homogeneous, noisy force can make the oscillating state last far beyond its well-known radiative decay lifetime, which is akin to noise-enhanced stability. Considering the degree of universality of the SG model, insights into its fundamental excitations possess an inherent interdisciplinary scope [e.g. 94].

Finally, we analyse the effects of randomly fluctuating solar irradiance on the population dynamics of a marine ecosystem [22]. The solar irradiance data comes from

the Boussole buoy located in the Gulf of Lion, and spans 2004 to 2013 [95]. By exploiting a 0-dimensional stochastic biogeochemical flux model, we find a nonmonotonic behaviour of the variance of the marine populations' biomass with respect to both the intensity and the auto-correlation time of the solar irradiance. This indicates a noise-induced transition of the ecosystem towards an out-of-equilibrium steady state. Moreover, we see evidence of noise-induced effects on the organic carbon cycling processes underlying the food web dynamics. These results clearly show the profound impact the stochastic environmental variables can have on both the populations and the biogeochemistry at the basis of a marine trophic network. Non-trivial, noise-induced dynamics can push the ecosystem away from the deterministic attractor and drive it towards a new non-equilibrium steady state. This and the previous examples highlight the interdisciplinary links between stochastic dynamical systems theory and a variety of complex systems, both geophysical and technological.

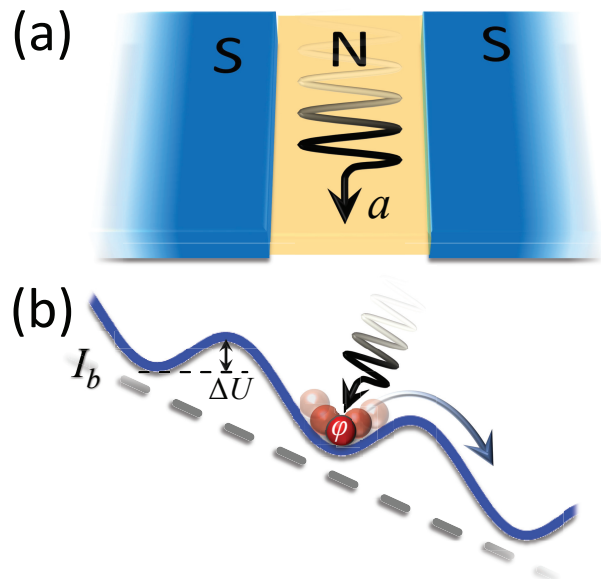


FIG. 2. **Josephson Junction (JJ) device and washboard potential.** (a) Schematic illustration of the JJ device discussed in the text: two superconductors and an axion field through the weak link. (b) Phase particle in a minimum of the washboard potential tilted by a bias current I_b . The phase can overcome the potential barrier ΔU , rolling down along the potential because of the combined effect of thermal noise and axion-JJ coupling.

B. Stochastic approaches to simulate ensembles of unprecedented events

Data limitations (in observations and even numerical model simulations) are a major challenging in modelling geophysical extremes. This makes direct application of

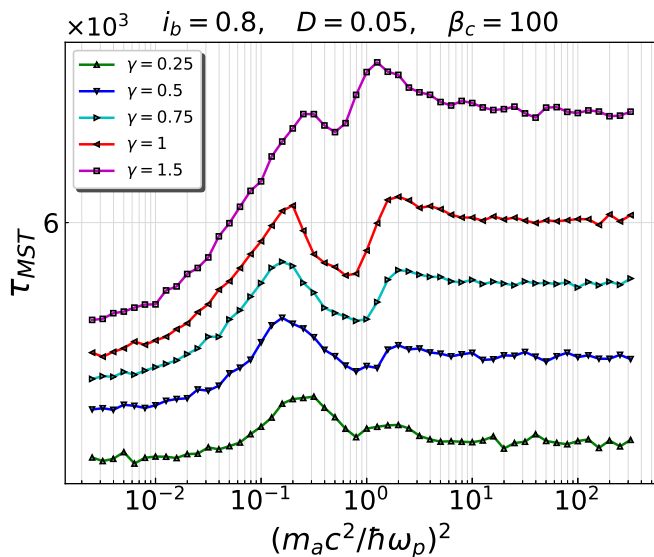


FIG. 3. **Mean switching time in a Josephson Junction (JJ).** Log-log plot of the dependence of the mean switching time τ_{MST} and the ratio $\varepsilon = (m_a c^2 / \hbar \omega_p)^2$ between the axion energy (ω_a is the characteristic frequency) and the Josephson plasma energy in underdamped regime for a JJ subject to thermal noise and coupled with an axion field against γ , the coupling constant between axion and Josephson junction.

many mathematical analysis tools sometimes difficult to implement. Indeed, many methods rely on the assumption of an unlimited amount of data on the system (cf. Sect. II). Observations of geophysical systems instead provide one finite trajectory. Even numerical simulations, such as climate model simulations, provide at most hundreds of $O(10^2)$ year long trajectories [32], which limits the extremeness of the events that can be studied. The problem is made more acute by the presence of low frequency trends (including, but not limited to climate change) leading to time-scale interactions that may affect extreme events [96].

Record breaking events, like the European heatwave of 2003 [97] or the Canadian heatwave in 2021 [98], have opened questions specifically focussed on the non-stationarity of the underlying system. Indeed, one can wonder how likely such extreme events are in the present climate, and how this may change in the future. We consider here two types of extreme events, where time t corresponds to the publication of this study:

- Type 1: Events that occur at time t , but never occurred before t (or during a reference period of observations). Hence there are documented examples and data for type 1 events, albeit small samples.
- Type 2: Events that have *never* occurred before time t . By definition, no observation exists for type 2 events at t . Note that the 2003 European and 2021 Canadian summer heatwaves were type 2

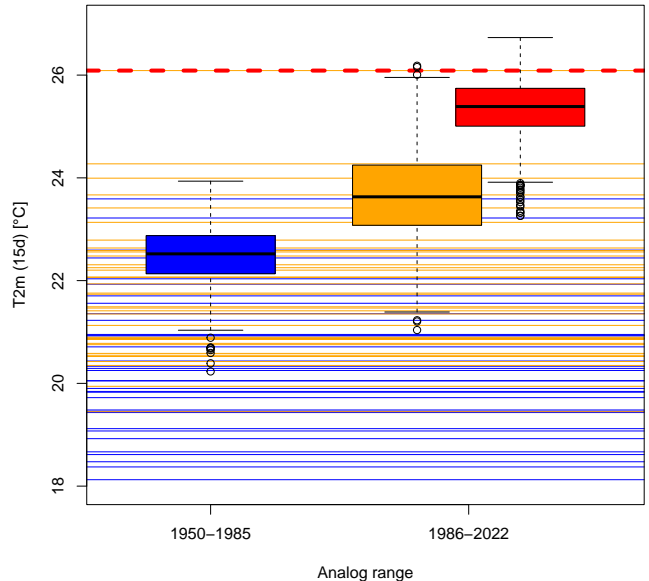


FIG. 4. **Empirical probability distribution of extreme 15 day heatwaves in France.** The horizontal lines represent the maximal 15 day averages of 2m temperature (in $^{\circ}\text{C}$) in the summer, in France, from the ERA5 reanalysis. The blue lines are between 1950 and 1986. The orange lines are between 1986 and 2002. The red dashed line is the record value obtained in 2003. The boxplots represent the empirical probability distributions of 1000 simulations of the SWG with importance sampling, with an initial condition in 2003. The blue boxplot is for analogs in 1950–1985. The orange boxplot is for analogs in 1986–2022, excluding information from 2003. The red boxplot is for analogs in 1986–2022, including information from 2003.

events before 2003 and 2021, respectively. Therefore, type 2 events can become type 1 with time, as observational databases grow.

From the mathematical point of view, there are no fundamental differences between the two types of events, when infinite trajectories are available. The difference stems from the limited amount of “real world” observations or numerical data. We then ask the following question: can type 2 events be deduced from (known) type 1 events, and if so how?

A heuristic and pragmatic approach for simulation of type 2 events is to identify a set of initial conditions leading to intense type 1 events, and run ensembles of a climate model with perturbations around those initial conditions [99]. This so-called “ensemble boosting” approach can provide samples of type 2 events based on existing type 1 events. In this vein, Ragone *et al.* [3] pioneered the application of rare event algorithms to the simulation of extreme European heatwaves in climate models. Such approaches helped exploring type 2 heatwaves, and showed that the statistics of type 2 (heat-

wave) events can be deduced from type 1 events. Such approaches have been applied to "long lasting" events (e.g. 90 days or more), i.e. in order to optimize averages over seasons. Lucarini *et al.* [98], Gálfi *et al.* [100] used the connected principle of large deviation theory to assess the *typicality* of extreme seasonal heatwaves and cold spells.

As climate models yield their own biases and often underestimate climate variability and extremes, observation-based approaches are also needed to solve the type 2 anticipation challenge, given data on type 1 events. For example, [101, 102] developed an observation-based stochastic model called a Stochastic Weather Generator (SWG). This SWG is a Markov chain (e.g. temperature) with hidden states (which can be the large-scale atmospheric circulation) whose transition probabilities drive the system. By modifying transition probabilities, the SWG can emulate through observations the application of rare event algorithms in a climate model [102]. Indeed, the SWG simulates a variable $Y(t)$ from sampling analogs of the hidden states $Z(t)$. The observable to be optimized is $f(Y)$.

Let us assume that the most extreme event (with the variable Y) in the "real world" occurs at time t_x . Then the challenge identified above can be solved in a rather natural way by simulating the observable $f(Y)$ with analogs of the hidden variable Z that include the knowledge of t_x (*cum data*) or *not* (*sine data*): can the record value of $f(Y(t_x))$ be reached or exceeded when we simulate $f(Y)$ with analogs of Z that never consider information at t_x ?

We illustrate this statistical approach on record 15 day heatwaves in France. We consider the daily mean temperature averaged over France as Y , taken from the ERA5 reanalysis [30] between 1950 and 2022. We determine the warmest 15 day spell (horizontal lines in Fig. 4). Therefore, the observable f is a 15 day average of Y . The record value is obtained in 2003. This value exceeds the preceding record by almost 2°C. The hidden variable Z is geopotential height at 500 hPa (Z500) over the North Atlantic region. For each day in 1950–2022, we compute the 20 best analogs of Z500 in 1950–1985 and 1986–2022. We then simulate 1000 trajectories of the SWG with importance sampling, starting in August 2003 (beginning of the hottest 15 day heatwave), using analogs in each period in turn. For the 1986–2022 period, we can exclude analogs in 2003 or include them, which allows determining the weight of this record event in the probability distribution of the simulated extremes, and evaluate to what extent the 2003 event could be anticipated without information on 2003. The result is reported in Figure 4.

We find that if an event is initiated in August 2003, but with meteorological conditions of 1950–1985, the worst heatwaves cannot reach the value of 2003 (although the SWG is close to other recent heatwaves, blue boxplot). If we consider information from 2003, then the 2003 record can easily be exceeded (red boxplot). If information on 2003 (apart from the initial condition) are removed in

the SWG simulations, the 2003 value can be reached, although with low probability (orange boxplot). This means that the record-shattering event of 2003 could be anticipated. As a caveat, we note that in this example we used information from other extreme heatwaves, three of which occurred after 2003.

IV. APPLICATIONS OF DYNAMICAL INDICATORS TO GEOPHYSICAL DATA

In this section, we provide examples of a number of applications of the analysis frameworks described in Sect. II to geophysical rare events. We specifically consider meridional energy transport in the atmosphere, extremes associated with persistent large-scale atmospheric circulation states typically termed "weather regimes", and spatially compounding weather extremes, namely geographically remote weather extremes that co-occur and are physically connected.

A. Statistical properties, temporal and spatial scales of meridional energy transport extremes

The atmosphere transports large amounts of energy meridionally, from the tropical regions to the polar regions. These meridional energy transports are intrinsically intermittent in time and space [103–110]. We hereby review an EVT-based methodology to detect extreme events in meridional energy transports in the mid-latitudes, as first presented in [110].

Meridional energy transports are obtained at every spatial point and pressure level as the scalar product of meridional velocity v and total energy E . The total energy is given by

$$E = H + \frac{1}{2}\mathbf{v}^2 = c_p T + L_v q + g z_g \quad (27)$$

where \mathbf{v} is the horizontal velocity vector, T the tropospheric air temperature, q the specific humidity and z_g the geopotential height, while constants c_p , L_v and g are specific heat at constant pressure, latent heat of vaporization and gravitational acceleration, respectively. The zonal and vertical mean of this quantity is obtained by spatial integration:

$$\oint \int_{p_s}^0 v E \frac{dp}{g} dx \quad (28)$$

where p_s is surface pressure. Meridional energy transports are positive when poleward-directed, negative otherwise. We consider ERA5 Reanalysis data for 1979–2014 period [30], in the Northern Hemisphere. We detrend our time series and remove the seasonal and daily cycles for every latitude, and analyse winter (DJF) and summer (JJA) separately.

In order to account for clusters in time that would possibly bias the estimates of the GPD parameters, we

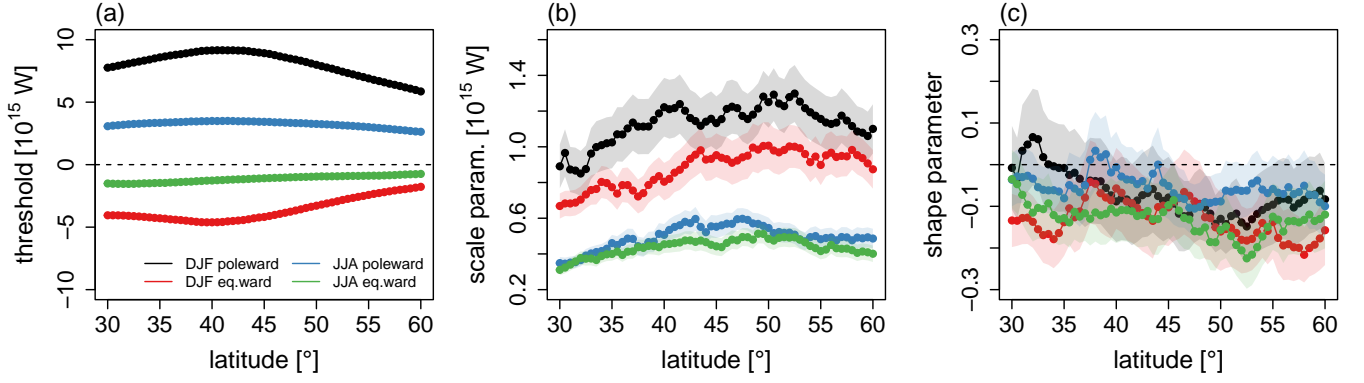


FIG. 5. **GPD parameters for meridional atmospheric energy transport.** Meridional profiles of the GPD (a) threshold, (b) scale parameter and (c) shape parameter in case of DJF poleward (black), DJF equatorward (red), JJA poleward (blue) and JJA equatorward (green) meridional energy transport extremes. The shading represents 95% maximum likelihood confidence intervals of the respective parameters.

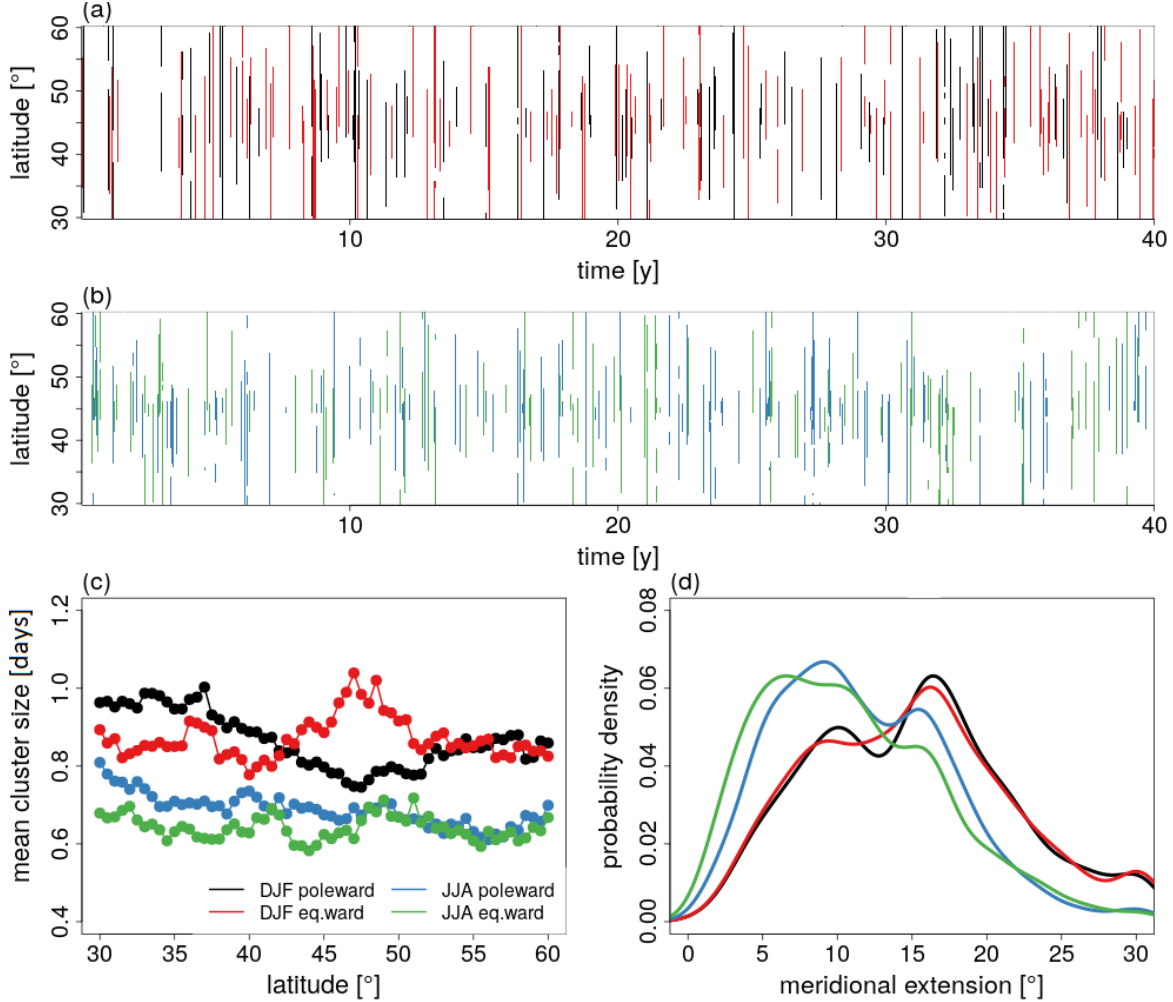


FIG. 6. **Hovmoller plots and scale of meridional atmospheric energy transport extremes.** Hovmoller plots indicating the occurrence of: (a) DJF poleward (black) and equatorward (red) and (b) JJA poleward (blue) and equatorward (green) meridional energy transport extremes. (c) Mean cluster size and (d) probability density of the meridional extension of meridional energy transport extremes.

apply a declustering technique based on the above mentioned extremal index (Eq. 7) [111, 112]. We estimate

the extremal index based on the intervals method [111] and use the obtained information to decluster the time series, thus ensuring that the selected extreme values are independent as required by EVT.

In the following, we apply EVT to both tails of the meridional energy transport distribution, namely poleward and equatorward extremes. Following Lembo *et al.* [110], these are defined as events beyond the 90 % and 10 % percentiles of the meridional energy transport distribution. Based on the peak-over-threshold GPD scale parameters (Eq. 3), we notice that the magnitude and variability of extremes is generally larger in winter than in summer, and it is more intense for poleward extremes than for equatorward extremes (Fig. 5a, b). The largest magnitude and variability occur for DJF poleward extremes. The threshold values for the left tails at each latitude are below zero; these are thus equatorward energy transport extremes and not weak poleward extremes. The shape parameters are predominantly negative, suggesting that meridional energy transport extremes have an upper bound (Fig. 5c). Only in the case of poleward extremes, do we observe shape parameter values close to or above zero for the latitudes south of 37 °N in DJF, while in JJA the largest shape parameters occur between latitudes 37 °N and 47 °N. This suggests that, in these regions and seasons, there is a chance for “surprises”, i.e. exceptionally large poleward energy transport extremes. The shape parameter values, however, have a substantial uncertainty. Thus, we cannot exclude that shape parameters related to other types of extremes or in other regions reach values close or above zero as well, nor that the near-zero or positive values we highlight above are indeed such.

The above-discussed GPD parameters do not provide information on the temporal or spatial scales of the energy transport extremes. To illustrate these scales, we mark the occurrence of extremes in the DJF (JJA) energy transport time series. For illustration purposes, we identify the timesteps of the extreme events using latitude 45 °N, and then verify whether the transport attains extreme values also at other latitudes (Fig. 6a, b). We summarise the information regarding the duration and meridional extension of the extremes in Fig. 6c, d, respectively. The duration is expressed in terms of the mean cluster size, which is computed as the inverse of the extremal index. The meridional extension is given in degrees latitude, and is defined as the number of consecutive latitudes affected by an extreme event at a given time step. In case of multiple extremes occurring at a given time, marked by vertical lines with interruptions, we compute the size of the largest one. DJF extremes often last close to one day, and their most common meridional extension is around 17°. Their distribution is skewed to the right, and we notice that the probability of large events extending over the full latitude band is substantially higher than the probability of small events of a few degrees latitude. JJA extremes are slightly shorter-lived – with mean duration shorter

than 1 day but longer than 0.5 days – and smaller in meridional extent. Their distribution is skewed to the left, and the most common event has a meridional extent of approximately 9 °N in case of poleward extremes and is slightly smaller in case of equatorward extremes. The probability of events extending over the full latitudinal band is much smaller than in winter. However, the distribution has a long right tail, thus the probability of large events is not negligible. The temporal and spatial scales of meridional energy transport extremes discussed here underlines their intermittent nature, and hints at their large-scale coherence and their signature in extratropical atmospheric dynamics.

B. Identifying rare features in the large-scale atmospheric circulation associated with extreme weather events

We consider here the application of the estimator of the local dimensions and their gradients, discussed in Sect. IID, to sea-level pressure (SLP) data over the Euro-Atlantic region. We specifically consider the years 1950–2020 in the NCEP/NCAR reanalysis data [113] and take two extreme events as examples: the 1987 Great Storm [114] and the June 2019 French heatwave [115]. The former was a severe extratropical cyclone that affected primarily France and the United Kingdom in October 1987. The latter was a heatwave which set record high temperatures in multiple European countries, including France. The aim is to obtain information on the dynamical characteristics of the large-scale atmospheric circulation features associated with these extremes, as diagnosed through SLP. Each point z in the dataset we use is defined by a daily SLP map over the chosen geographical domain. The maps for the two extremes being analysed are shown in Fig. 7a, b.

We then compute the normalised gradient in phase space of the local dimension. At a point z , this is given by: $\frac{\nabla_z d_u(z)}{\|\nabla_z d_u(z)\|}$. This is a vector of the same dimension as z , and can therefore be represented on a map. The two gradient vectors for the 1987 Great Storm and the June 2019 French heatwave are shown in Fig. 7c, d. Scalars at each grid point are components of the full vector. Since $\frac{\nabla_z d_u(z)}{\|\nabla_z d_u(z)\|}$ has an origin, which is z itself, care should be taken in the interpretation of the directions in the phase space indicated by this gradient. Indeed, they must be read with respect to the point where the gradient is computed. Negative (blue) gradients mean that if one wants to increase d , then one should decrease the SLP values at that grid point. Positive (red) gradients means that one should increase SLP to increase d . As a consequence, the geographical regions where the absolute values of the normalized gradient are the largest, correspond to directions in the phase space where d would change the most for a given change in SLP. In a finite data set, the geographical regions with the strongest gradients correspond

to phase-space directions in which z has the least number of analogues (or analogues which are further away from z for a given distance metric, if a fixed number of analogues is identified). The geographical regions with very high gradients hence correspond to uncommon features of an SLP map.

Fig. 7b shows that the key structure in the SLP field of the 1987 Great Storm that makes the atmospheric situation uncommon is not the low pressure situated just south of Iceland, but rather the storm itself (which also corresponds to a low pressure) situated south-east of Brittany. Similarly for the June 2019 French heatwave, Fig. 7d shows that the atmospheric feature which determines the rarity of the situation is the low pressure situated above Russia. The low pressure to the east of Ireland and the high pressure in Northern Europe are also uncommon atmospheric features.

For each point in the phase space of some geophysical system, the gradient of the local dimension thus indicates structures which are rare or common with respect to the sampling done by the data set. We foresee that investigating the link between this gradient and the dynamics of the system may be a fruitful future line of research.

C. From transition probabilities to the predictability of midlatitude extreme weather

Atmospheric motions are chaotic, which implies an unavoidable loss of predictability with increasing lead time. However, in chaotic systems it is possible to identify recurrent patterns. These patterns have often been conceptualized as so-called weather regimes, namely clusters of similar atmospheric states [116]. Weather regimes are then recurrent, quasi-stationary, and persistent large-scale atmospheric circulation states [117] usually defined as typical clusters of atmospheric flows that are observed in a specific geographical region. Weather regimes appear as sticky regions of the phase-space where the trajectories slow down, possible due to the vicinity of stationary or quasi-stationary solutions (e.g. [116, 118]).

Weather regimes have been defined over a number of different geographical regions, but their use has been most widespread in the Euro-Atlantic sector. There, the optimal number of weather regimes is typically taken to be four [119], although fewer [120] or more [121] have been proposed. Figure 8 shows the four canonical regimes using sea-level pressure (SLP) over the region $[80^{\circ}\text{W}-50^{\circ}\text{E}, 20^{\circ}-70^{\circ}\text{N}]$. These are, ordered by decreasing frequency of occurrence:

(a) the Atlantic Ridge (AR), with a high pressure anomaly over the center of the North Atlantic; (b) the positive phase of North-Atlantic Oscillation (NAO+) with a dipole of anomalously low pressure in the northern North Atlantic and anomalously high pressure to the south; (c) Scandinavian Blocking (BLO), with an anomalous high pressure center over Scandinavia; (d) the negative phase of North-Atlantic Oscillation (NAO-), showing

a dipole which is roughly inverse to that of the NAO+.

Weather regimes can establish conditions favorable for large-scale extremes ([122–128]), and are used in the context of medium-range and sub-seasonal to seasonal weather forecasts and climate projections ([129–131]). This motivates an interest in studying the large-scale atmospheric evolution during extreme weather events in a weather regimes perspective.

The weather regimes and their transitions have also been connected to the dynamics of the underlying atmospheric system. Faranda *et al.* [14], showed a statistical link between anomalies in d and θ computed on the SLP in the Euro-Atlantic sector and the four canonical weather regimes (see labels in Fig. 9 here). Later, [118] and [132] found that d and θ reflect the regimes' life cycles. Specifically, d and θ decrease when a regime is established, and increase during transitions between different regimes, albeit with some regime dependence.

The two dynamical systems metrics have also been connected to the intrinsic predictability of the atmosphere, making the argument that low-dimension, high persistence atmospheric patterns should afford a higher predictability than high-dimension, low persistence cases [14, 15, 133–135].

We illustrate here how combining an analysis based on d and θ as described in Sect. IID with one of regime transitions can shed light on weather extremes. As example, we pick storm Filomena (Fig. 9): a storm which affected the Iberian Peninsula in January 2021, bringing heavy snowfall to large regions of Spain [136, 137].

The first days of storm Filomena corresponded to an NAO- type regime regime. This transitioned very rapidly to a rare atmospheric configuration, which was very unusual in terms of d and θ values and which did not correspond closely to any of the four canonical weather regimes. The very low persistence (high θ) and high d on these days suggests that the atmosphere was in a low predictability situation. It was during these days that intense snowfall occurred in Spain. Next, the atmosphere transitioned towards AR and from AR to BLO.

The example of Filomena shows that rare and low-predictability weather regimes as diagnosed by d and θ can be associated with extreme events. This was also noted by [138] when investigating snowfall events in the middle-east. This type of analysis in d - θ space can be applied to a range of different extreme event categories, as a complement to a categorical study of weather regimes and their transitions. Indeed, the case of Filomena shows that the circulation associated with extremes does not always fit neatly in one of the canonical regimes. The proposed analysis also informs on the temporal evolution of the large-scale atmospheric predictability during the extreme event from a dynamical systems perspective.

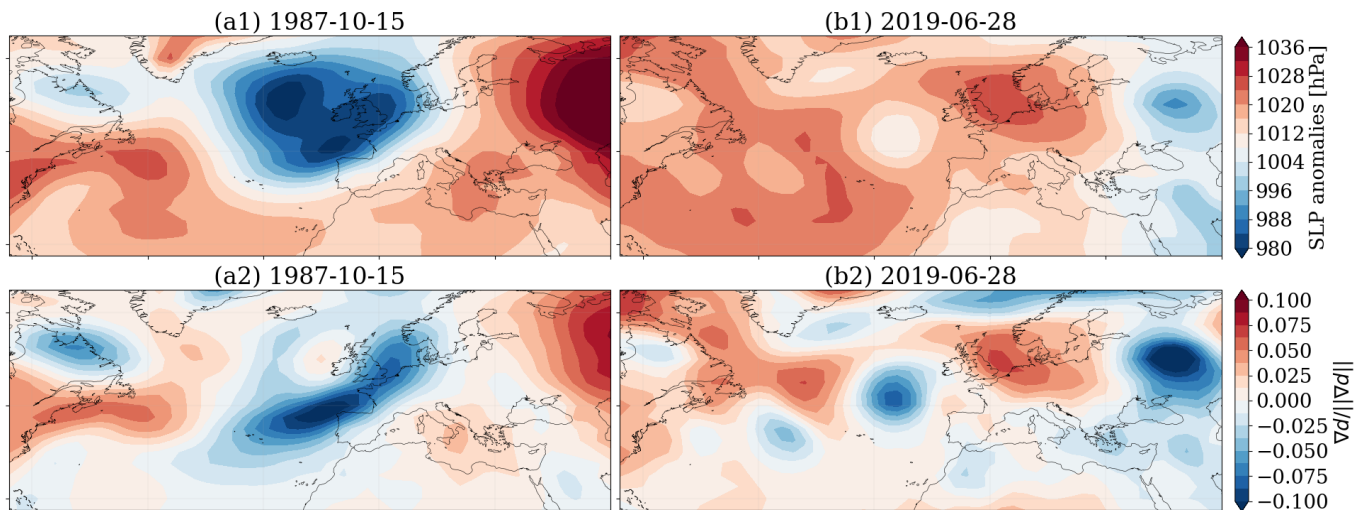


FIG. 7. Sea level pressure and gradient of local dimension for two extreme case studies: the 1987 Great Storm and the June 2019 French Heatwave. SLP (hPa) for (a1) the 1987 Great Storm and (b1) the June 2019 French Heatwave. (a2, b2) Corresponding direction of the normalized gradient of local dimension on the dates indicated in the panel titles. Scalars at each grid point are components of the full vector. The figure uses NCEP/NCAR reanalysis data.

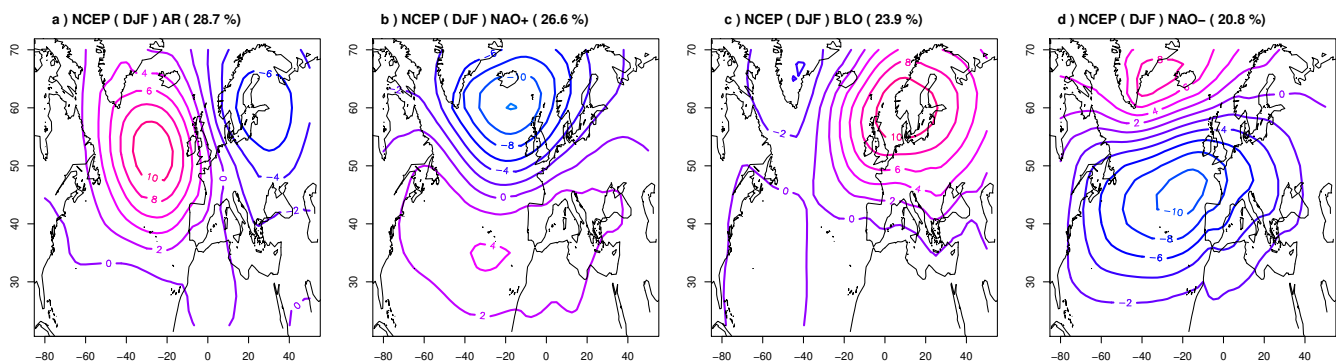


FIG. 8. Euro-Atlantic Weather regimes. Weather regimes in the Euro-Atlantic sector $[80^{\circ}\text{W}-50^{\circ}\text{E}^{\circ}, 20^{\circ}-70^{\circ}\text{N}]$ computed using sea level Pressure from NCEP/NCAR reanalysis [113] over 1948–2021, following the methodology of [122].

D. Co-recurrence statistics for spatially compounding weather extremes

Understanding the drivers and dynamics of weather extremes is facilitated by a multivariate analysis framework. Indeed, different weather extremes are often associated to common physical drivers (for example a cyclone causing both strong winds and heavy rain), and even geographically remote extremes can be physically connected [e.g. 139] – so-called spatially compounding extremes. Even extreme events that relate to a single impact variable benefit from a multivariate analysis. They can be associated with multiple drivers (e.g. a heatwave may be driven by both the atmospheric circulation and soil moisture characteristics [140]) or require investigating the impact variable and a possible driving variable jointly. In all these cases, a multivariate analysis framework is advantageous.

Here, we consider both a monivariate extreme and the case of spatially compounding extremes. Such extremes can pose a greater threat to human societies than single, isolated extremes, as their effects may exacerbate each other and lead to correlated losses [e.g. 141, 142]. We specifically illustrate the application of multivariate dynamical systems indicators to first study drivers of cold spells in North America and next their co-occurrence with cold spells in North America and wet or windy extremes in Europe – sets of extreme weather events which have been highlighted in the literature as being spatially compounding [139, 143–145].

We implement here the previously discussed estimation of d , θ and α in a bivariate context (Sect. IID). We illustrate in Fig. 10b these indicators computed for SLP and 2-meter temperature (t2m) over North America. The atmospheric configurations with a high co-recurrence ratio α between SLP and t2m tend to be those with relatively

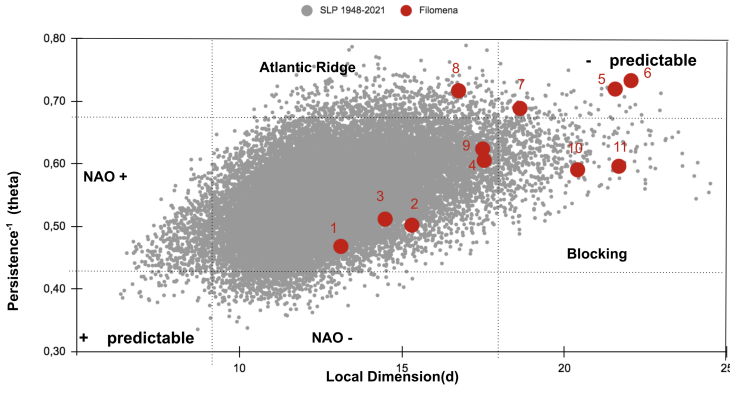


FIG. 9. **A dynamical systems and weather regime analysis of storm Filomena.** The grey points in the scatter-plot show daily d and θ values computed using SLP from the NCEP/NCAR reanalysis over the period 1948–2021, in the region $[80^{\circ}\text{W}-50^{\circ}\text{E}^{\circ}, 20^{\circ}-70^{\circ}\text{N}]$. Red dots show the values during storm Filomena, numbered consecutively starting from the 5th of January 2021. The weather regime labels in the different quadrants of the plot are based on [14].

low d and θ . When a given pair of SLP and t2m patterns recurs, it thus tends to coincide with low-dimensional and persistent configurations.

We next focus specifically on days displaying high α in winter, and look at the corresponding SLP and t2m anomalies (Fig. 11) – the focus of our analysis here. High co-recurrence ratio days display a tripolar SLP anomaly pattern, favouring the meridional advection of cold air and below average temperatures over a large part of the North American continent. Coupled with the information provided by Fig. 10b, this points to cold wintertime spells being favoured by persistent circulation patterns, and to the fact that whenever cold days occur, similar large-scale t2m and SLP patterns are found. This connects back to the concept of large-scale meteorological patterns, whereby specific regional climate extremes are associated with recurrent atmospheric configurations [e.g. 147, 148].

We next consider how the co-recurrence ratio can provide information on spatially compounding climate extremes. We again analyse wintertime cold spells over North America, but now investigate their connection to wet or windy European extremes. We specifically compute the co-recurrence ratio between SLP over North America and Europe. We then select the 50 winter days with the highest co-recurrence ratio and investigate whether they correspond to local climate extremes (defined here as t2m temperature anomalies below the local 5th percentiles and precipitation and 10-metre wind anomalies above the local 95th percentiles). There is a clear signal of heightened frequency of cold extremes over south-eastern North America and wet or windy extremes over western-continental Europe on high-coupling days (Fig. 12). This supports previous literature arguing for

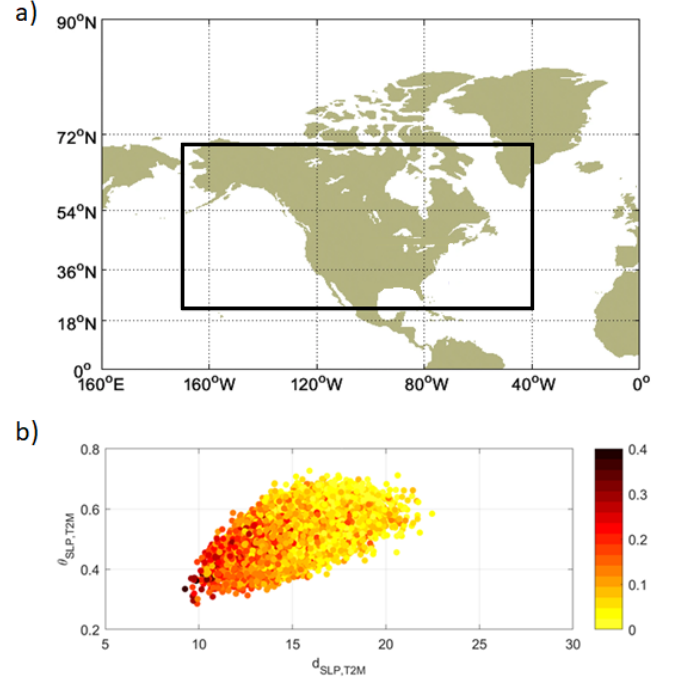


FIG. 10. **Geographical domains and scatter plot for the dynamical systems indicators computed on SLP and t2m.** a) The black rectangle highlights the domain over which the dynamical systems metrics are computed. b) d – θ scatter plot coloured with the values of the co-recurrence ratio α . The figure is adapted from [146] and uses NCEP/NCAR reanalysis data [113] over 1948–2018.

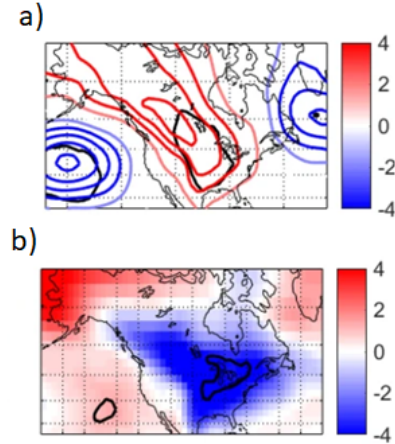


FIG. 11. **Composite SLP and t2m anomalies for winter days with a high co-recurrence ratio.** Composite SLP anomalies in hPa (a) and t2m anomalies in K (b) corresponding to the 10% most positive anomalies of α relative its seasonal cycle, during DJF. The black lines bound regions where at least 2/3 of the composited anomalies have the same sign. The figure is adapted from [146] and uses NCEP/NCAR reanalysis data [113] over 1948–2018.

a systematic physical and statistical connection between

these sets of extremes [17, 139, 143, 145]. We thus conclude that the application of the above dynamical systems metrics in a bivariate context can provide useful information of compound climate extremes. We foresee that a full multivariate implementation could be profitably applied to the study of a range of different geophysical systems.

V. OPPORTUNITIES AND OPEN CHALLENGES FOR THE STUDY OF GEOPHYSICAL RARE EVENTS

While several advances have been made in the understanding, modeling and simulation of geophysical extreme or rare events, several challenges still limit our capabilities of modeling small-scale or sporadic extreme events. Two examples of geophysical extreme events having these characteristics are heavy thunderstorms and tropical cyclones (TCs). Both are highly localized convective events, which are affected by physical processes on a large number of spatial scales. This makes it challenging to monitor and predict their occurrence.

A final challenge in studying all these geophysical extremes is the impact of climate change, which can alter the frequency, duration, and intensity of these events (see also Sect. III B). Predicting the impact of climate change requires the integration of non-stationarity in statistical and dynamical models, to separate the role of forced versus natural variability in the occurrence of these phenomena. In this perspective, a growing attention has been devoted to so-called "attribution" studies, namely determining whether a specific extreme geophysical event may be ascribed to anthropogenically-forced climate change [149].

In this section we discuss in more detail the recent advances and challenges in studying the above-discussed geophysical extremes, with a focus on the opportunities issuing from dynamical systems theory and statistical physics, not least for extreme event attribution. We specifically outline lines of research that have been made possible by recent findings, but are yet to be explored in the literature.

A. Physics and dynamics of convective precipitation phenomena

Convective events are some of the most intense precipitation events on the planet, associated with rising plumes of moist air in response to instability in the atmosphere. Convection triggers updraft and downdraft

and as a result is the main process of vertical exchange of water vapour, heat and chemical species between the lower and upper troposphere.

Convection plays an important role in both the mid-latitudes and the tropics. In the mid-latitudes it can give rise to heavy thunderstorms; in the tropics, it is a key ingredient of TCs. TCs are synoptic-scale ($O(10^3)$ km) vortices fueled by the thermal energy accumulated in the ocean. This is transferred to the atmosphere through evaporation and convective processes, and finally released through precipitation and partly converted into kinetic energy (extreme winds). Specifically, the aggregation of convective towers, or clusters of thunderstorms, can lead to the formation of a TC as the energy released from the towers fuels the development of a low-pressure center, which can then intensify into a cyclonic system.

The study of thunderstorms and TCs is an active area of research, with significant progress having been made in understanding the physical processes that drive these events. Examples include the role of atmospheric instability in the development of thunderstorms [150] and the role of large-scale atmospheric circulation patterns in the formation and intensification of thunderstorms and TCs e.g. [151–153]. A further area of research is the role of kinetic energy, moisture and heat budgets in the development of thunderstorms and TCs. These budgets describe the exchanges of energy and moisture between the atmosphere and the surface, span a continuum of spatial scales and are important for understanding how storms and TCs develop and intensify [26, 154]. We begin here by considering statistical physics approaches to the study of energy transfers in convective systems.

Diagnostics of inertial energy transfers typically rely on filtering approaches that separate resolved fields from subfilter-scale fields [155]. Filtering is broadly used in the study of turbulent geophysical flows, including both atmospheric and oceanographic flows [156–160]. In all these studies, a filter is applied to the Navier-Stokes equation, and this is used to infer the energy transfers from the reference filtered lengthscale ℓ to larger and smaller scales arising due to non-linear interactions [157]. Using this approach, [161] diagnosed organised regions of upscale and downscale inertial energy transfers in the hurricane boundary layer, based on remotely sensed wind observations during Hurricane Rita (2005).

In an alternative approach based on the weak solution formalism, [162] showed that energy transfers in a fluid at an arbitrary scale ℓ satisfy a local energy balance equation:

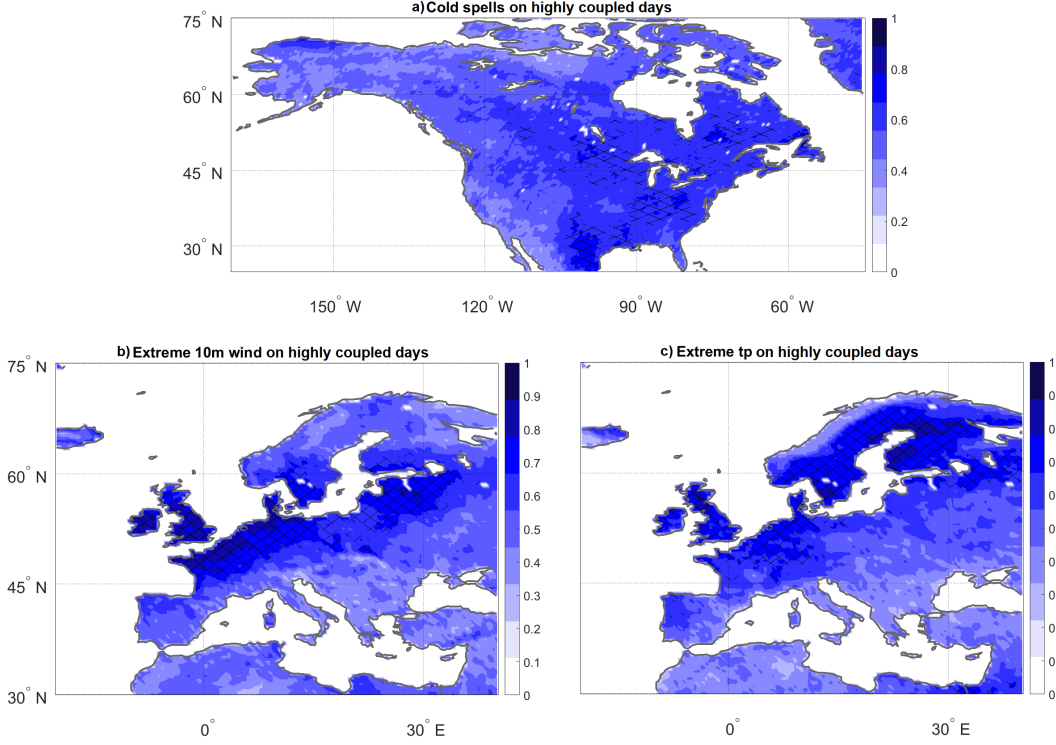


FIG. 12. **Normalised occurrence of extreme climate events for winter days with a high co-recurrence ratio.** Normalised occurrence of (a) cold spells, (b) extreme 10-metre wind and (c) extreme precipitation during the 50 winter days with the highest co-recurrence ratio between SLPs over North America and Europe. Cross-hatching marks regions where the occurrence of extremes during high co-recurrence days is significantly higher than climatology at the 5% one-sided level, determined using 1000 random sampling iterations. The figure is taken from [144] and uses ERA5 reanalysis data [30] over 1979–2020.

$$\partial_t E^\ell + \partial_j \left(u_j E^\ell + \frac{1}{2} (u_j \hat{p} + \hat{u}_j p) + \frac{1}{4} (\widehat{u^2 u_j} - \widehat{u^2} u_j) - \nu \partial_j E^\ell \right) = -\nu \partial_j u_i \partial_j \hat{u}_i - D_\ell, \quad (29)$$

where u_i are the components of the velocity field, p the pressure, \hat{u} and \hat{p} their coarse-grained components at scale ℓ , and $E^\ell = \sum_i \frac{\hat{u}_i \hat{u}_i}{2}$ is the kinetic energy per unit mass at scale ℓ (such that $\lim_{\ell \rightarrow 0} E^\ell = u^2/2$). The term D_ℓ is expressed in terms of velocity increments $\delta \vec{u}(\vec{r}, \vec{x}) \stackrel{def}{=} \vec{u}(\vec{x} + \vec{r}) - \vec{u}(\vec{x}) \equiv \delta \vec{u}(\vec{r})$ (the dependence on ℓ and \vec{x} is kept implicit) as:

$$D_\ell(\vec{u}) = \frac{1}{4\ell} \int_{\mathcal{V}} d\vec{r} (\vec{\nabla} G_\ell)(\vec{r}) \cdot \delta \vec{u}(\vec{r}) |\delta \vec{u}(\vec{r})|^2. \quad (30)$$

In this definition, G is a smooth filtering function, non-negative, spatially localized and such that $\int d\vec{r} G(\vec{r}) = 1$, and $\int d\vec{r} |\vec{r}|^2 G(\vec{r}) \approx 1$. The function G_ℓ is rescaled for a given volume with ℓ as $G_\ell(\vec{r}) = \ell^{-3} G(\vec{r}/\ell)$.

The 2D filter is adapted from [163] as a circular symmetric filtering function of the scalar increment r given

by

$$G(r) = \begin{cases} \frac{1}{N} \exp\left(-\frac{1}{1-(r/2a)^2}\right), & \text{if } r < 2a, \\ 0, & \text{otherwise.} \end{cases} \quad (31)$$

where N is a normalization constant such that $\int d^2r G(r) = 1$

We provide here an example of applying the framework from [162] to the spatial distribution of instantaneous energy transfers in TCs. We specifically consider Typhoon Jolina, which hit Japan in September 2005. We use the NCEP/NCAR reanalysis data [113] and a scale $\ell = 250$ km. Positive D_ℓ indicate transfers from the mesoscale (≈ 250 km) towards the smaller convective scales, while negative values correspond to transfers towards the larger synoptic scales. Fig. 13 suggests that energy transfers are organized around the eye of the cyclone and that intense downscale energy transfers occur in a relatively small region of the planetary boundary layer, where hazards such as heavy rainfall and wind-gusts

are also concentrated. However, both the direct and inverse energy cascades coexist, and transfers towards the synoptic scales are found in cyclone’s outflow in the upper troposphere. Following such energy transfers over time along the trajectory of TCs would offer a way to study, in scale spaces, rapid intensification and rapid weakening phenomena. The above approach could also be used to investigate other convective events such as medicanes, mesocyclones, squall lines or derechoes. However, such systematic investigations are still absent from the literature.

B. Statistics and modelling of convective precipitation phenomena

A complementary approach to studying the energy cascades in individual TCs is to take a bulk view of cyclones as point vortices and investigate the environmental conditions that favour or suppress their growth. In fact, TCs can be viewed as a type of heat engine, where the warm ocean surface provides the fuel for the storm to extract energy and convert it into wind. Despite the advances in understanding the dynamics of individual TCs, key open questions remain on TC bulk statistics. These include:

1. What controls the yearly number of observed TCs (global rate of about 70–90 events per year [164])?
2. What is the response of TC activity to variations in the mean state of climate?

Using a bulk perspective, [165] developed the potential intensity theory, which predicts the maximum possible wind speed that a TC can attain based on the underlying ocean and atmospheric conditions. This theory has helped researchers to explain the occurrence of TCs, and has led to the development of a number of semi-empirical indices known as Genesis Potential Indices (GPIs) [166–171]). However, GPIs are by construction not able to address question 1 above, since they do not take into account the small-scale dynamics leading to the “seeding”. In other words, the process or family of processes responsible for the formation of the precursor disturbances, from which self-organized convection and then the cyclones originate [172]. Indeed, the actual number of TCs is only partially accounted for by the large-scale factors entering the GPI, the remaining part depending on the amount of seeds [172]. The GPI approach also struggles in addressing question 2. Indeed, when using high-resolution climate model simulations to estimate future TC activity, it is found that the trends of GPI are not consistent with those of simulated cyclones, often disagreeing even in the sign of change [173].

These limitations have motivated new lines of research, such as synthetic tracking. This involves using dynamical systems and statistical methods to generate

large numbers of virtual TCs, which are then used to study the bulk statistical properties of the cyclones [e.g. 174, 175]. Specifically, a combination of deterministic and stochastic approaches are used to iteratively test “seeding” cyclone cores at different locations with different environmental conditions and then identifying how different possible tracks affect the development of the cores. Synthetic tracking has proven to be a valuable tool for predicting the paths and impacts of TCs, which can have significant implications for disaster planning and response efforts [e.g. 176, 177]. The very large number of tracks generated by synthetic tracking algorithms has supported a better understanding of the global rate of TC occurrence, for example by informing on the likelihood of TC occurrences in regions that may not have experienced any in the historical period [178, 179]. Similarly, it has aided in exploring the implications of climate change for TC frequency and characteristics [180, 181]. These points are of direct relevance to the two questions listed above.

Like TCs, convective precipitation and thunderstorms have also been investigated leveraging concepts from dynamical systems theory and statistical physics. An example is a framework for predicting the onset and evolution of thunderstorms using large scale environmental conditions [151, 182, 183]. This approach involves using statistical mechanics to derive probabilistic models of the atmospheric processes that govern the behavior of thunderstorms, including the interaction between convection and the large-scale atmospheric circulation. Another approach that has been used to study thunderstorms is the application of Ising models [184]. Ising models are statistical physics models that are typically used to describe the behavior of interacting spins in a lattice. However, they have been adapted to describe the atmosphere as a grid of interacting variables, with each variable representing a particular atmospheric state or parameter. By analyzing the statistical properties of the variables and their interactions, researchers can gain insights into the behavior of thunderstorms, including the onset, growth, and decay of convective activity. While Ising models are a relatively new approach to studying thunderstorms, they have shown promise in capturing some of the complex dynamics of these storms and could potentially lead to more accurate predictions of their behavior in the future [185].

Modelling convection presents several challenges. Perhaps the most pressing is the scale mismatch between the small scales at which convection occurs and the larger scales that can be directly resolved by numerical climate models. To address this challenge, parameterization schemes are used to represent the effects of convection on larger scales. However, these schemes are a major source of uncertainty in climate simulations. Convection also involves interactions between the atmosphere, the surface, and other physical processes such as radiation

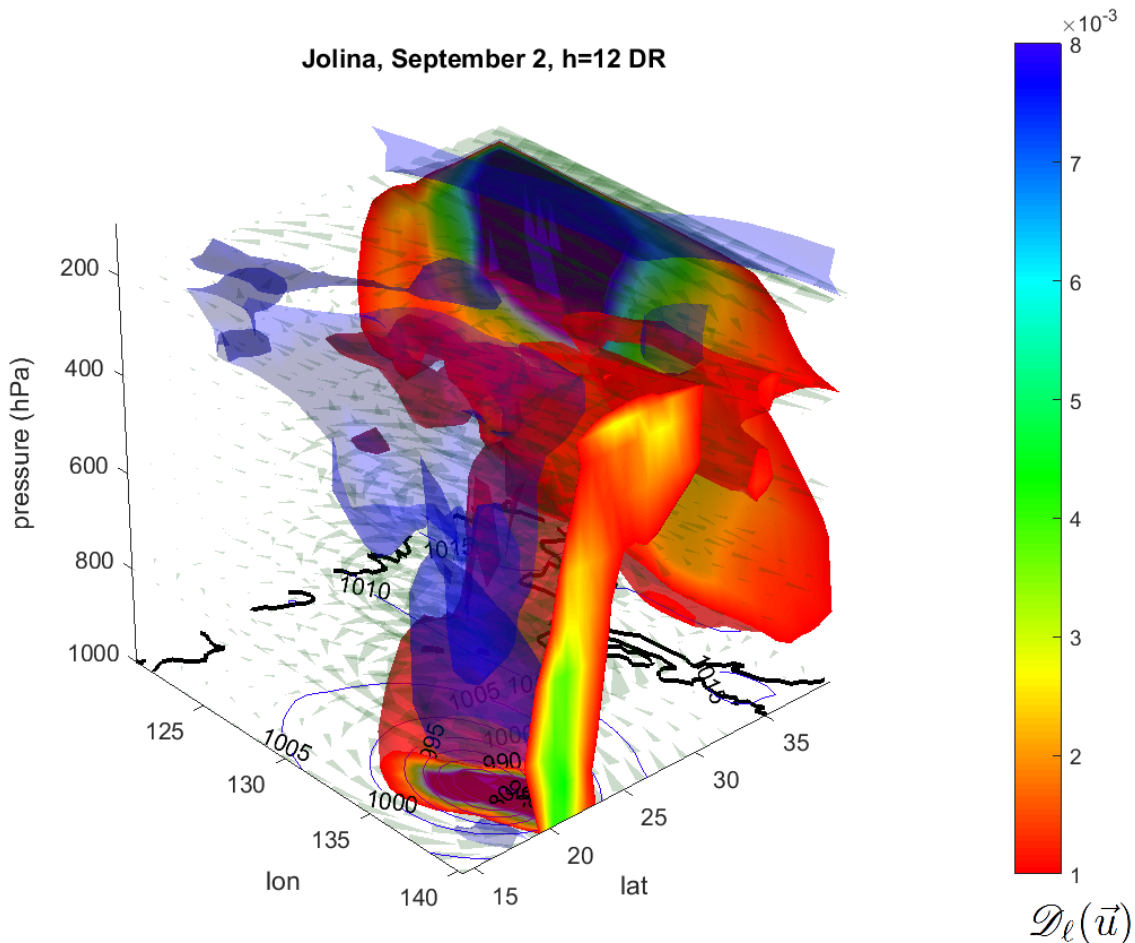


FIG. 13. **Mesoscale to convective scale energy transfer in Typhoon Jolina.** Energy transfer D_ℓ , with $\ell = 250$ km computed using the 3D velocity fields from the NCEP/NCAR reanalysis dataset for the cyclone Jolina (2nd September 2005, 12h UTC, when the cyclone reached peak intensity). The colorscale indicates energy transfers directed towards smaller scales. The dark blue patch the energy transfers towards scales larger than ℓ .

and moisture transport. Additionally, the behaviour of clouds and their evolution is affected by factors such as aerosols and atmospheric dynamics, which further complicates the modelling process. Accurately capturing all of these processes precludes a clear separation of scales as often performed in the study of stochastic dynamical systems. In fact, while ad-hoc studies employing scale decomposition approaches are possible [13, 154, 162], they cannot account for the full range of processes leading to convective precipitation.

A possible fashion to address this challenge, and reduce the uncertainties associated with parametrisation schemes, is to develop numerical climate models with km-scale resolution. These allow to directly resolve the small scales of convection. Both single-model and multi-model ensembles of such models are important for improving robustness of results, rather than relying on

individual model simulations. Single-model ensembles involve running multiple simulations with slightly different initial or boundary conditions, or with slightly different model parameter settings, while multi-model ensembles refer to the use of different models, typically using identical initial or boundary conditions. These approaches allow us to account for the uncertainty inherent in the climate system's chaotic evolution and reduce the uncertainty associated with model structure and parameters. Over the past years, several coordinated efforts have emerged to develop such ensembles (e.g. [186–189]).

C. Dynamical systems attribution of extreme weather events under climate change

Attribution studies seek to answer the question: how has human-induced climate change modulated the likelihood and/or physical characteristics of a given extreme event? A wide range of statistical and modelling tools have been used in attribution studies, including a range of modelling and statistical approaches [149]. Relatively recently, dynamical systems theory has also been used in attribution studies [153, 190]. Since the atmosphere is a complex system whose dynamics is influenced by many factors, including the ocean, land surface, and human activities, dynamical systems theory provides a framework for understanding how changes in one part of the system can affect the behavior of the entire system, including the occurrence of extreme weather events. One key concept in dynamical systems theory is the use of "recurrences" (see Sect. II). This involves comparing the observed weather patterns during an extreme event to those that have occurred in the past, in order to identify whether the event is more or less likely to occur, or has occurred with a different intensity, as a result of human-induced climate change. For example, if a given heatwave and the associated circulation is found to have a high probability of recurrence in a world without human-induced climate change, then it is less likely that the heatwave is solely due to climate change. However, if the probability of recurrence is found to be much higher in a world with human-induced climate change, then it is more likely that climate change played a significant role in the event. Dynamical systems theory therefore helps to condition attribution to a specific circulation. The approach can also support diagnosing the physical processes associated with the attribution. For example, if there is a persistent atmospheric circulation pattern that is associated with extreme weather events, such as a blocking high-pressure system, then researchers can use dynamical systems theory to understand how changes in the climate system can affect the persistence and intensity of the circulation pattern, and how this can modulate the associated extreme weather events. [153, 191, 192]. The above analogue approaches can be complemented by narratives that describe the physical mechanisms that led to an extreme weather event, namely storylines [193–195]. Storylines support a detailed understanding of the factors that contributed to an event, and how they may have been influenced by human-induced climate change. For example, a storyline for a heatwave event might describe how a persistent high-pressure system led to hot and dry conditions, which were then exacerbated by human-induced climate change.

By combining analogue approaches and storylines, researchers can provide a complete picture of the link between climate change and extreme weather events [196]. Analogue approaches can quantify the probability of the event occurring, while storylines contextualise the physical mechanisms behind the event. This can help policy-

makers and the public to better understand the risks of climate change, and to develop strategies for adapting to and mitigating these risks.

Despite the success of dynamical systems attribution-based techniques, the mathematical justification of these approaches is still largely missing. For several physical processes it is unclear whether there exists some invariant or stationary measure which could be used for computing statistics of extreme or rare events. In this context it can be useful to return to discrete dynamical systems, for which recent results pave the way for a rigorous application of EVT-based dynamical systems approaches in a non-stationary setting. We specifically consider the following two non-stationary systems: *sequential* and *non-autonomous*. They are defined by concatenating maps chosen in some set, usually in the close neighborhood of a given map. As a probability measure one typically takes some ambient measure like Lebesgue.

The extreme value theory and the extremal index must then be redefined. Sect. 4.5 of [38] presents an example of a sequential system modeled on maps chosen in the neighborhood \mathcal{U} of a given β -transformation, say T_{β_0} . According to suitable choices of \mathcal{U} it is possible to show either that the EI of the unperturbed map and of the sequential system are the same, or the two differ, in particular when the elements of the concatenation are far enough from T_{β_0} . In this case the EI is simply 1. A second relevant example is in the Appendix of [9]. There, sequentially composed maps of the one-dimensional torus of the type $T_i = 2x + b_i \text{-mod } 1$, are considered. The constants b_i and the Dirac masses p_i associated to each T_i vary randomly respectively in some spaces $B_k, P_k, k = 1, 2, \dots$ which in turn change every 10 temporal steps. The empirical distribution of the number of visits in a ball centered at the origin is in perfect agreement with a Pòlya-Aeppli distribution.

The second set of systems we consider are non-autonomous random (quenched) dynamical systems. They are constructed by taking a deterministic driving map σ on the space E and preserving a probability measure \mathbb{P} which codes a family of transformations T_ω for $\omega \in E$ on the space M via the composition rule $T_\omega^n(x) = T_{\sigma^{n-1}\omega} \circ \dots \circ T_\omega(x)$. One can prove the existence of sample measure μ_ω , verifying, for each measurable set $A \in M : \mu_\omega(T_\omega^{-1}(A)) = \mu_{\sigma\omega}(A)$. These measures describe the statistical properties in M and they do not give rise to stationary processes. The first applications of EVT to these random systems are given in [38, 197]. A recent article [39] further developed a new spectral approach for a quenched extreme value theory that considers random dynamics with general ergodic invertible driving σ , and random observations. It also provided a general formula for the computation of the extremal index which will be also random. To give a flavour of such results, we quote here the following ex-

ample, taken from [39]. Consider the bi-infinite sequence $\omega = \{\dots, \omega_{-1}, \omega_0, \omega_1, \dots\}$, where each ω_l takes values in the finite alphabet $\{1, \dots, d\}$, and move it with the shift σ . Each symbol will be taken with the same probability $1/d$. We then take d maps T_1, \dots, T_d , and random *rare* sets of the form $B_{\omega,n} := B(v(\omega_0), e^{-z_n(\omega)})$. Notice that the centers of these balls, $v(\omega_0)$, are random (ω_0 denotes the 0-th coordinate of ω), and we also allow random radii satisfying the scaling $2e^{-z_n(\omega)} = \frac{t+\xi_{\omega,n}}{n}$, t being a non-random constant and $|\xi_{\omega,n}|$ bounded uniformly in n and ω and going to zero when $n \rightarrow \infty$ and for \mathbb{P} -a.a. ω . We then define the first random hitting time as

$$\tau_{\omega,n}(x) = \inf\{k \geq 1, T_{\omega}^k(x) \in B_{\sigma^k \omega, n}\},$$

which gives the first time the random orbit of x enters an element of the sequence of random balls $B_{\sigma^k \omega, n}$. This covers, for instance, the interesting case of rare sets which are known with a limited precision, or cases when the iterations at each step are affected by some disturbance. For particular choices of the maps T for which all the sample measures coincide and are equal to η , we have

$$\lim_{n \rightarrow \infty} \eta\{\tau_{\omega,n} > n\} = e^{-t \int \theta_{\omega} d\mathbb{P}(\omega)},$$

where θ_{ω} is the random EI, defined by suitable random generalizations of formulae (8) and (33), see [39] for the details. It is interesting to note that the previous example can be worked out in such a way that the expectation of θ_{ω} is strictly less than 1, thus showing that we could have the formation of clusters even in the absence of periodic or invariant structures, which could not persist forever in the presence of noise.

The above idealised examples show that the behaviour of non-stationary systems can be described by EI distributions, which depend on the systems' non-stationarity characteristics and on the (stochastic) system's realisation. This underscores the existence of an extreme value law for such systems, which opens the possibility of computing other dynamical quantities beyond the EI. These notions may in the future be expanded to more complex real-world systems.

VI. PERSPECTIVES

Geophysical extremes, spanning phenomena as diverse as thunderstorms, tropical cyclones, earthquakes, and geomagnetic storms, encompass a broad array of spatial and temporal scales. Thunderstorms exhibit fine-scale microphysical processes occurring at the micrometer level, juxtaposed with the kilometer-scale dynamics of towering storm clouds or the scales of up to thousands of kilometers of tropical cyclones. The pursuit of understanding geophysical extremes thus exemplifies the mathematical intricacies and dynamical complexities inherent to multiscale dynamics. To elucidate the multiscale fabric of these events, mathematical models must

bridge multiple magnitudes in both spatial and temporal scales. We envision developments in mathematical methods such as coupled dynamical systems and network theory as instrumental to capture complex inter-scale and inter-variable dependencies. Multiscale decomposition techniques may also play a key role in further research on the topic. These involve breaking down complex geophysical systems into multiple scales, each with its own set of governing equations and dynamics, supporting an improved understanding of how small-scale processes interact with large-scale phenomena. For example, in studying thunderstorms, multiscale decomposition can reveal how microphysical processes related to nucleation within clouds interact with the broader atmospheric circulation, or how accessory phenomena such as tornadoes and hail form in the convective clouds. This type of knowledge is essential for improving the accuracy of climate projections and weather forecasts, and more broadly for the understanding of the physical mechanisms driving multiple geophysical extremes.

The mathematical modeling of multiscale geophysical dynamics is rooted in partial differential equations (PDEs). Nevertheless, the severe numerical constraints posed by the direct simulations of partial differential equations necessitate multiscale modeling approaches. As we have shown in this perspective, stochastic parameterizations are a useful tool to this effect, and are able to reconcile the interactions between disparate geophysical scales. Stochastic systems further provide a probabilistic framework for modeling the inherent randomness and uncertainty in geophysical extremes, making them valuable for risk assessment and prediction. Incorporating these mathematical and analytical approaches into large deviation theory, multivariate extreme value analysis, analogue-based methods and multiscale decomposition techniques will further enhance our ability to quantify, predict, and manage geophysical extreme events, ultimately improving our resilience to their impacts. We specifically view developments in multivariate extreme value analysis as crucial to capture the joint behavior of multiple geophysical variables. In the context of geophysical extremes, this would advance the study of compound geophysical extreme events, such as the co-occurrence of high wind speeds and heavy precipitation during a (tropical) cyclone, enable devising novel bias correction techniques, or improved downscaling of extreme weather events.

In addition to the mathematical challenges mentioned above, the study of geophysical extremes presents several other dimensions. One such aspect is the need to develop methods that can accurately separate the natural variability of the climate system from the effects of human-induced climate change in engendering extreme geophysical events – namely the so-called *extreme event attribution*. This requires statistical techniques that can distinguish between short-term, internal fluctuations of geophysical systems and long-term, forced trends. The above-mentioned mathematical advances for

non-stationary systems can support this effort.

Another important challenge is the need to account for the spatial and temporal dependence of extreme events. Extreme weather events often occur in clusters or specific spatial patterns, and their occurrence in one location can be associated with the occurrence of extremes in other regions or influenced by remote drivers (see Sect. IV D). Understanding the origin of spatially co-occurring extremes requires the use of advanced spatial and temporal statistical models that can account for complex dependencies within the climate system and other geophysical systems. As we have shown when discussing co-recurrence statistics (Sect. II D, IV D) understanding the dependence structure of extreme events is crucial for assessing compound risks, where multiple extreme events co-occur, potentially leading to cascading impacts. In this sense the analogues-based approach presented in Sect. II D could be combined with multivariate extreme value analyses to assess the likelihood and potential impacts of extreme events. More generally, the work outlined in this perspective provides some initial steps in this direction, laying the bases for a more general mathematical treatment of spatially or temporally clustered geophysical extremes.

VII. CONCLUSIONS

This paper has provided an overview of techniques for studying geophysical extreme events, focusing on the interplay between statistical physics, dynamical systems theory, and geophysics. By identifying the limitations of traditional statistical extreme value analysis techniques to study geophysical rare events, we have motivated the introduction of new mathematical formalisms based on rare recurrences in high-dimensional systems.

The application of these techniques to various geophysical phenomena, such as temperature extremes, cyclones, thunderstorms and geomagnetic storms, has provided valuable insights into the underlying dynamics and physical drivers of these events. Additionally, the examination of diverse data, including climate reanalysis, numerical climate models, geomagnetic data and turbulence measurements has illustrated how approaches developed for low-dimensional systems can be applied to high-dimensional chaotic data. While significant insights have been gained regarding the dynamics of rare events in geophysical systems, further investigation is necessary to fully comprehend the underlying mechanisms and drivers.

Understanding and analyzing high-impact events in the Earth system involves addressing several challenges, spanning both the mathematical and physical sciences. In terms of mathematical challenges, one major concern is the non-stationarity of geophysical systems. The multifarious interactions and feedback loops within the Earth system introduce a further level of complexity. Another critical challenge lies in accounting for finite size

effects, including the spatial and temporal resolution of the datasets used for analysis. Moreover, geophysical systems exhibit high dimensionality, necessitating an examination of how data resolution affects the identification and characterization of rare events. Properly addressing finite size effects is essential to avoid potential biases or artifacts in the analysis [10, 12].

Finally, the limited availability of high-quality observational data poses a significant challenge, for example for the study of tropical cyclones that often occur over remote ocean regions. Insufficient observations make it difficult to validate and improve the accuracy of models, as well as to gain a robust understanding of the physical processes that drive these events.

To overcome these challenges, interdisciplinary research efforts combining mathematics, physics, and geophysics are crucial. Advances in mathematical techniques, continuing to build upon the novel indicators, approaches and modeling frameworks presented in this perspective, will provide a stepping stone for future advances. In parallel with this, a comprehensive understanding of the physics behind rare geophysical events requires a combination of observational data and theoretical insights. In this context, it is crucial to devise stochastic dynamic models, which account for the effects of environmental random fluctuations and can serve as a connection between the extreme events approach and the dynamical systems modeling. We view collaborative endeavors among researchers from various disciplines as instrumental in this respect. We specifically encourage continued collaboration between researchers in statistical physics, statistics, dynamical systems theory and geophysics, whose expertise this perspective has sought to gather. In the longer term, we envision the integration of tailored machine learning algorithms and advanced data assimilation techniques with dynamical systems and statistical physics approaches to further enhance our understanding of geophysical extremes and improve their prediction capabilities. Ultimately, our hope is that these collective efforts of researchers across various disciplines will contribute to building a more resilient and sustainable future.

ACKNOWLEDGEMENTS

This perspective has issued from discussions held during the UNDERPIN symposium held in Erice, Italy, with the support of MITI-CNRS and the Ettore Majorana Foundation and Centre for Scientific Culture. SV thanks the Mathematical Research Institute MATRIX, the Sydney Mathematical Research Institute (SMRI), the University of New South Wales, and the University of Queensland for their support and hospitality. DF, PY, RN and EC received support from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101003469 (XAIDA). DF, PY, RN and GM received further support from the European

Union's Horizon 2020 Marie Skłodowska-Curie grant agreement No. 956396 (EDIPI) and the LEFE-MANU-INSU-CNRS grant "CROIRE". The research of LC was supported by the European Union's Horizon 2020 Marie Skłodowska-Curie grant agreement No. 101065985 (CYCLOPS). We also acknowledge useful discussions with the MedCyclones COST Action (CA19109) community. The research of SV was supported by the project *Dynamics and Information Research Institute* within the agreement between UniCredit Bank and Scuola Normale Superiore di Pisa and by the Laboratoire International Associé LIA LYSM, of the French CNRS and INdAM (Italy). SV was also supported by the project MATH-AmSud TOMCAT 22-Math-10, N. 49958WH, du French CNRS and MEAE. PY and DF received support from the grant ANR-20-CE01-0008-01 (SAMPRACE)

VIII. APPENDIX

We collect here a few formulas and quantities introduced in section II.B. First of all the quantities λ_l given in (7) are formally defined as

$$\lambda_l := \lim_{\Delta \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\mu(x; \sum_{j=0}^{\Delta} \mathbf{1}_{U_n}(T^j(x)) = l)}{\mu(x; \sum_{j=0}^{\Delta} \mathbf{1}_{U_n}(T^j(x)) \geq 1)}, \quad (32)$$

provided the limits exist.

We now quote the analytic expression of the quantities $q_{k,n}$ entering the definition of the extremal index (8):

$$q_{k,n} = \frac{\mu(x \in U_n, T(x) \in U_n^c, \dots, T^k(x) \in U_n^c, X_{k+1} \in U_n)}{\mu(U_n)}. \quad (33)$$

We finally quote the mass distribution of the Pölya-Aepli distribution

$$\tilde{\nu}(\{k\}) = e^{-(1-p)t} \sum_{j=1}^k \binom{k-1}{j-1} \frac{((1-p)^2 t)^j}{j!} p^{k-j}. \quad (34)$$

-
- [1] L. de Haan and A. Ferreira, *Extreme Value Theory: An Introduction (Springer Series in Operations Research and Financial Engineering)*, 1st ed. (Springer, 2010).
- [2] V. Lucarini, D. Faranda, J. M. M. de Freitas, M. Holland, T. Kuna, M. Nicol, M. Todd, S. Vaienti, *et al.*, *Extremes and recurrence in dynamical systems* (John Wiley & Sons, 2016).
- [3] F. Ragone, J. Wouters, and F. Bouchet, Computation of extreme heat waves in climate models using a large deviation algorithm, *Proceedings of the National Academy of Sciences* **115**, 24 (2018).
- [4] D. Faranda, H. Ghoudi, P. Guiraud, and S. Vaienti, Extreme value theory for synchronization of coupled map lattices, *Nonlinearity* **31**, 3326 (2018).
- [5] F. Bouchet, J. Rolland, and J. Wouters, Rare event sampling methods, *Chaos: An interdisciplinary journal of nonlinear science* **29**, 080402 (2019).
- [6] M. Ghil and V. Lucarini, The physics of climate variability and climate change, *Reviews of Modern Physics* **92**, 035002 (2020).
- [7] M. Abadi, A. C. M. Freitas, and J. M. Freitas, Dynamical counterexamples regarding the extremal index and the mean of the limiting cluster size distribution, *Journal of the London Mathematical Society* **102**, 670 (2020).
- [8] A. C. M. Freitas, J. M. Freitas, F. B. Rodrigues, and J. V. Soares, Rare events for cantor target sets, *Communications in Mathematical Physics* **378**, 75 (2020).
- [9] T. Caby, D. Faranda, S. Vaienti, and P. Yiou, Extreme value distributions of observation recurrences, *Nonlinearity* **34**, 118 (2020).
- [10] N. Haydn and S. Vaienti, Limiting entry and return times distribution for arbitrary null sets, *Communications in Mathematical Physics* **378**, 149 (2020).
- [11] F. M. E. Pons, G. Messori, M. C. Alvarez-Castro, and D. Faranda, Sampling hyperspheres via extreme value theory: implications for measuring attractor dimensions, *Journal of statistical physics* **179**, 1698 (2020).
- [12] V. M. Gálfi and V. Lucarini, Fingerprinting heatwaves and cold spells and assessing their response to climate change using large deviation theory, *Physical Review Letters* **127**, 058701 (2021).
- [13] T. Alberti, D. Faranda, R. V. Donner, T. Caby, V. Carbone, G. Consolini, B. Dubrulle, and S. Vaienti, Small-scale induced large-scale transitions in solar wind magnetic field, *The Astrophysical journal letters* **914**, L6 (2021).
- [14] D. Faranda, G. Messori, and P. Yiou, Dynamical proxies of north atlantic predictability and extremes, *Scientific reports* **7**, 1 (2017).
- [15] G. Messori, R. Caballero, and D. Faranda, A dynamical systems approach to studying midlatitude weather extremes, *Geophysical Research Letters* **44**, 3346 (2017).
- [16] D. Faranda, M. C. Alvarez-Castro, G. Messori, D. Rodrigues, and P. Yiou, The hammam effect or how a warm ocean enhances large scale atmospheric predictability, *Nature communications* **10**, 1 (2019).
- [17] P. De Luca, G. Messori, D. Faranda, P. J. Ward, and D. Coumou, Compound warm-dry and cold-wet events over the mediterranean, *Earth System Dynamics* **11**, 793 (2020).
- [18] F. Ragone and F. Bouchet, Rare event algorithm study of extreme warm summers and heatwaves over europe, *Geophysical Research Letters* **48**, e2020GL091197 (2021).
- [19] V. M. Gálfi and G. Messori, Persistent anomalies of

- the north atlantic jet stream and associated surface extremes over europe, *Environmental Research Letters* **18** (2023).
- [20] K. Giamalaki, C. Beaulieu, D. Faranda, S. Henson, S. Josey, and A. Martin, Signatures of the 1976–1977 regime shift in the north pacific revealed by statistical analysis, *Journal of Geophysical Research: Oceans* **123**, 4388 (2018).
- [21] K. Giamalaki, C. Beaulieu, S. Henson, A. Martin, H. Kassem, and D. Faranda, Future intensification of extreme aleutian low events and their climate impacts, *Scientific reports* **11**, 1 (2021).
- [22] R. Grimaudo, P. Lazzari, C. Solidoro, and D. Valenti, Effects of solar irradiance noise on a complex marine trophic web, *Scientific Reports* **12**, 1 (2022).
- [23] G. Messori and D. Faranda, Characterising and comparing different palaeoclimates with dynamical systems theory, *Climate of the Past* **17**, 545 (2021).
- [24] M. Brunetti, J. Kasparian, and C. V erard, Co-existing climate attractors in a coupled aquaplanet, *Climate Dynamics* **53**, 6293 (2019).
- [25] T. Alberti, D. Faranda, G. Consolini, P. De Michelis, R. V. Donner, and V. Carbone, Concurrent effects between geomagnetic storms and magnetospheric substorms, *Universe* **8**, 226 (2022).
- [26] B. Dubrulle, F. Daviaud, D. Faranda, L. Mari e, and B. Saint-Michel, How many modes are needed to predict climate bifurcations? lessons from an experiment, *Nonlinear Processes in Geophysics* **29**, 17 (2022).
- [27] A. Gualandi, J.-P. Avouac, S. Michel, and D. Faranda, The predictable chaos of slow earthquakes, *Science advances* **6**, eaaz5548 (2020).
- [28] A. Gualandi, D. Faranda, C. Marone, M. Cocco, and G. Mengaldo, Stochastic chaos and predictability in laboratory earthquakes, arXiv preprint arXiv:2204.02763 (2022).
- [29] E. N. Lorenz, Deterministic nonperiodic flow, *Journal of atmospheric sciences* **20**, 130 (1963).
- [30] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Hor anyi, J. Mu oz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, *et al.*, The era5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society* **146**, 1999 (2020).
- [31] K. E. Taylor, R. J. Stouffer, and G. A. Meehl, An overview of cmip5 and the experiment design, *Bulletin of the American meteorological Society* **93**, 485 (2012).
- [32] V. Eyring, S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization, *Geoscientific Model Development* **9**, 1937 (2016).
- [33] A. R. Jonkers, A. Jackson, and A. Murray, Four centuries of geomagnetic data from historical records, *Reviews of Geophysics* **41** (2003).
- [34] L. Cappanera, P. Debue, H. Faller, D. Kuzay, E. Saw, C. Nore, J.-L. Guermond, F. Daviaud, C. Wiertel-Gasquet, and B. Dubrulle, Turbulence in realistic geometries with moving boundaries: when simulations meet experiments, *Computers & Fluids* **214**, 104750 (2021).
- [35] V. Lucarini, R. Blender, C. Herbert, F. Ragone, S. Pascale, and J. Wouters, Mathematical and physical ideas for climate science, *Reviews of Geophysics* **52**, 809 (2014).
- [36] V. M. G alfi, V. Lucarini, F. Ragone, and J. Wouters, Applications of large deviation theory in geophysical fluid dynamics and climate science, *La Rivista del Nuovo Cimento* **44**, 291 (2021).
- [37] D. Ruelle, *Turbulence, strange attractors and chaos*, Vol. 16 (World Scientific, 1995).
- [38] A. C. M. Freitas, J. M. Freitas, and S. Vaienti, Extreme value laws for non stationary processes generated by sequential and random dynamical systems, in *Annales de l'Institut Henri Poincar e (B) Probabilit es et Statistiques*, Vol. 53 (2017) pp. 1341–1370.
- [39] J. Atnip, G. Froyland, C. Gonzalez-Tokman, and S. Vaienti, Perturbation formulae for quenched random dynamics with applications to open systems and extreme value theory, arXiv preprint arXiv:2206.02471 (2022).
- [40] S. Coles, J. Bawa, L. Trenner, and P. Dorazio, *An introduction to statistical modeling of extreme values*, Vol. 208 (Springer, 2001).
- [41] H. Touchette, The large deviation approach to statistical mechanics, *Physics Reports* **478**, 1 (2009).
- [42] F. den Hollander, *Large Deviations* (American Mathematical Society, 2000).
- [43] N. R. Moloney, D. Faranda, and Y. Sato, An overview of the extremal index, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **29**, 022101 (2019).
- [44] G. Keller and C. Liverani, Rare events, escape rates and quasistationarity: some exact formulae, *Journal of Statistical Physics* **135**, 519 (2009).
- [45] G. Keller, Rare events, exponential hitting times and extremal indices via spectral perturbation, *Dynamical Systems* **27**, 11 (2012).
- [46] T. Caby, D. Faranda, G. Mantica, S. Vaienti, and P. Yiou, Generalized dimensions, large deviations and the distribution of rare events, *Physica D: Nonlinear Phenomena* **400**, 132143 (2019).
- [47] T. Caby, D. Faranda, S. Vaienti, and P. Yiou, On the computation of the extremal index for time series, *Journal of Statistical Physics* **179**, 1666 (2020).
- [48] D. Faranda, V. Lucarini, G. Turchetti, and S. Vaienti, Numerical convergence of the block-maxima approach to the generalized extreme value distribution, *Journal of statistical physics* **145**, 1156 (2011).
- [49] V. Lucarini, D. Faranda, and J. Wouters, Universal behaviour of extreme value statistics for selected observables of dynamical systems, *Journal of statistical physics* **147**, 63 (2012).
- [50] N. Haydn, M. Nicol, A. T or ok, and S. Vaienti, Almost sure invariance principle for sequential and non-stationary dynamical systems, *Transactions of the American Mathematical Society* **369**, 5293 (2017).
- [51] S. Gallo, N. Haydn, and S. Vaienti, Number of visits in arbitrary sets for ϕ -mixing dynamics, arXiv preprint arXiv:2107.13453 (2021).
- [52] A. Coutinho, J. Rousseau, and B. Saussol, Large deviation for return times, *Nonlinearity* **31**, 5162 (2018).
- [53] T. Caby, Matching of observations of dynamical systems, with applications to sequence matching, *Physica D: Nonlinear Phenomena* **440**, 133456 (2022).
- [54] B. R. Hunt and V. Y. Kaloshin, How projections affect the dimension spectrum of fractal measures, *Nonlinearity* **10**, 1031 (1997).
- [55] A. C. M. Freitas, J. M. Freitas, and M. Todd, Hitting time statistics and extreme value theory, *Probability Theory and Related Fields* **147**, 675 (2010).

- [56] D. Faranda, Y. Sato, B. Saint-Michel, C. Wiertel, V. Padilla, B. Dubrulle, and F. Daviaud, Stochastic chaos in a turbulent swirling flow, *Physical review letters* **119**, 014502 (2017).
- [57] F. Flandoli and B. Schmalfuss, Random attractors for the 3d stochastic navier-stokes equation with multiplicative white noise, *Stochastics: An International Journal of Probability and Stochastic Processes* **59**, 21 (1996).
- [58] A. Kolmogorov, The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers, *Akademiia Nauk SSSR Doklady* **30**, 301 (1941).
- [59] C. Foias, O. Manley, R. M. S. Rosa, and R. Temam, *Navier-Stokes Equations and Turbulence* (Cambridge University Press, Cambridge, 2001).
- [60] G. Parisi and U. Frisch, Turbulence and predictability in geophysical fluid dynamics and climate dynamics (North-Holland Publ. Co., Amsterdam/New York, 1985) p. 449.
- [61] B. Dubrulle, Beyond kolmogorov cascades, *Journal of Fluid Mechanics* **867** (2019).
- [62] R. Benzi, G. Paladin, A. Vulpiani, and G. Parisi, On the multifractal nature of fully developed turbulence and chaotic systems, *Journal of Physics A: Mathematical General* **17**, 3521 (1984).
- [63] J. F. Muzy, E. Bacry, and A. Arneodo, Wavelets and multifractal formalism for singular signals: Application to turbulence data, *Physical Review Letters* **67**, 3515 (1991).
- [64] R. Benzi, L. Biferale, G. Paladin, A. Vulpiani, and M. Vergassola, Multifractality in the statistics of the velocity gradients in turbulence, *Physical Review Letters* **67**, 2299 (1991).
- [65] L. Biferale, G. Boffetta, A. Celani, B. J. Devenish, A. Lanotte, and F. Toschi, Multifractal Statistics of Lagrangian Velocity and Acceleration in Turbulence, *Physical Review Letters* **93**, 064502 (2004), arXiv:nlin/0403020 [nlin.CD].
- [66] G. Boffetta, A. Mazzino, and A. Vulpiani, TOPICAL REVIEW: Twenty-five years of multifractals in fully developed turbulence: a tribute to Giovanni Paladin, *Journal of Physics A: Mathematical General* **41**, 363001 (2008), arXiv:0809.0196 [nlin.CD].
- [67] R. Benzi, L. Biferale, R. T. Fisher, L. P. Kadanoff, D. Q. Lamb, and F. Toschi, Intermittency and Universality in Fully Developed Inviscid and Weakly Compressible Turbulent Flows, *Physical Review Letters* **100**, 234503 (2008), arXiv:0709.3073 [nlin.CD].
- [68] A. Arnéodo, R. Benzi, J. Berg, L. Biferale, E. Bodenschatz, A. Busse, E. Calzavarini, B. Castaing, M. Cencini, L. Chevillard, R. T. Fisher, R. Grauer, H. Homann, D. Lamb, A. S. Lanotte, E. Lévêque, B. Lüthi, J. Mann, N. Mordant, W. C. Müller, S. Ott, N. T. Ouellette, J. F. Pinton, S. B. Pope, S. G. Roux, F. Toschi, H. Xu, and P. K. Yeung, Universal Intermittent Properties of Particle Trajectories in Highly Turbulent Flows, *Physical Review Letters* **100**, 254504 (2008), arXiv:0802.3776 [nlin.CD].
- [69] T. Alberti, F. Daviaud, R. V. Donner, B. Dubrulle, D. Faranda, and V. Lucarini, Chameleon attractors in turbulent flows, *Chaos, Solitons & Fractals* **168**, 113195 (2023).
- [70] U. Frisch, *Turbulence* (1995).
- [71] P. Debue, V. Shukla, D. Kuzzay, D. Faranda, E.-W. Saw, F. Daviaud, and B. Dubrulle, Dissipation, intermittency, and singularities in incompressible turbulent flows, *Physical Review E* **97**, 053101 (2018).
- [72] L. Biferale, G. Boffetta, A. Celani, B. Devenish, A. Lanotte, and F. Toschi, Multifractal statistics of lagrangian velocity and acceleration in turbulence, *Physical review letters* **93**, 064502 (2004).
- [73] P. Zandbergen and D. Dijkstra, Von kármán swirling flows, *Annual review of fluid mechanics* **19**, 465 (1987).
- [74] T. Alberti, D. Faranda, V. Lucarini, R. V. Donner, B. Dubrulle, and F. Daviaud, Scale dependence of fractal dimension in deterministic and stochastic lorenz-63 systems, *Chaos: An Interdisciplinary Journal of Nonlinear Science* **33**, 023144 (2023), <https://doi.org/10.1063/5.0106053>.
- [75] D. Faranda, J. M. Freitas, V. Lucarini, G. Turchetti, and S. Vaienti, Extreme value statistics for dynamical systems with noise, *Nonlinearity* **26**, 2597 (2013).
- [76] D. Hristopulos, B. Spagnolo, and D. Valenti, Open challenges in environmental data analysis and ecological complex systems (a), *Europhysics Letters* **132**, 68001 (2021).
- [77] E. Forgoston and R. O. Moore, A primer on noise-induced transitions in applied dynamical systems, *SIAM Review* **60**, 969 (2018).
- [78] P. C. Hohenberg and B. I. Halperin, Theory of dynamic critical phenomena, *Reviews of Modern Physics* **49**, 435 (1977).
- [79] T. M. Lenton, J. Rockström, O. Gaffney, S. Rahmstorf, K. Richardson, W. Steffen, and H. J. Schellnhuber, Climate tipping points—too risky to bet against, *Nature* **575**, 592 (2019).
- [80] N. Wunderling, J. F. Donges, J. Kurths, and R. Winkelmann, Interacting tipping elements increase risk of climate domino effects under global warming, *Earth System Dynamics* **12**, 601 (2021).
- [81] D. Valenti, G. Fazio, and B. Spagnolo, Stabilizing effect of volatility in financial markets, *Physical Review E* **97**, 062307 (2018).
- [82] C. Guarcello, D. Valenti, A. Augello, and B. Spagnolo, The role of non-gaussian sources in the transient dynamics of long josephson junctions, *Acta Physica Polonica B* **44**, 997 (2013).
- [83] C. Guarcello, G. Filatrella, V. Pierro, B. Spagnolo, and D. Valenti, Voltage drop across josephson junctions for lévy noise detection, *Physical Review Research* **2**, 043332 (2020).
- [84] C. Guarcello and F. Bergeret, Cryogenic memory element based on an anomalous josephson junction, *Physical Review Applied* **13**, 034012 (2020).
- [85] B. Lisowski, D. Valenti, B. Spagnolo, M. Bier, and E. Gudowska-Nowak, Stepping molecular motor amid lévy white noise, *Physical Review E* **91**, 042713 (2015).
- [86] C. Guarcello, D. Valenti, B. Spagnolo, V. Pierro, and G. Filatrella, Josephson-based threshold detector for lévy distributed current fluctuations, *Physical Review Applied* **11**, 044078 (2019).
- [87] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani, Stochastic resonance in climatic change, *Tellus* **34**, 10 (1982).
- [88] A. Mikhaylov, D. Guseinov, A. Belov, D. Korolev, V. Shishmakova, M. Koryazhkina, D. Filatov, O. Gorskov, D. Maldonado, F. Alonso, *et al.*, Stochastic resonance in a metal-oxide memristive device, *Chaos, Solitons & Fractals* **144**, 110723 (2021).

- [89] C. Guarcello, D. Valenti, A. Carollo, and B. Spagnolo, Stabilization effects of dichotomous noise on the lifetime of the superconducting state in a long josephson junction, *Entropy* **17**, 2862 (2015).
- [90] K. Yoshida, K. Sato, and A. Sugamata, Noise-induced synchronization of uncoupled nonlinear systems, *Journal of sound and vibration* **290**, 34 (2006).
- [91] D. Filatov, D. Vrzheschch, O. Tabakov, A. Novikov, A. Belov, I. Antonov, V. Sharkov, M. Koryazhkina, A. Mikhaylov, O. Gorshkov, *et al.*, Noise-induced resistive switching in a memristor based on zro2 (y)/ta2o5 stack, *Journal of Statistical Mechanics: Theory and Experiment* **2019**, 124026 (2019).
- [92] N. Agudov, A. Safonov, A. Krichigin, A. Kharcheva, A. Dubkov, D. Valenti, D. Guseinov, A. Belov, A. Mikhaylov, A. Carollo, *et al.*, Nonstationary distributions and relaxation times in a stochastic model of memristor, *Journal of Statistical Mechanics: Theory and Experiment* **2020**, 024003 (2020).
- [93] R. Grimaudo, D. Valenti, B. Spagnolo, G. Filatrella, and C. Guarcello, Josephson-junction-based axion detection through resonant activation, *Physical Review D* **105**, 033007 (2022).
- [94] D. De Santis, C. Guarcello, B. Spagnolo, A. Carollo, and D. Valenti, Generation of travelling sine-gordon breathers in noisy long josephson junctions, *Chaos, Solitons & Fractals* **158**, 112039 (2022).
- [95] P. Lazzari, S. Salon, E. Terzić, W. W. Gregg, F. d’Ortenzio, V. Vellucci, E. Organelli, and D. Antoine, Assessment of the spectral downward irradiance at the surface of the mediterranean sea using the radiative ocean-atmosphere spectral irradiance model (oasim), *Ocean Science* **17**, 675 (2021).
- [96] S. Seneviratne, X. Zhang, M. Adnan, W. Badi, C. Dereczynski, A. Di Luca, S. Ghosh, I. Iskandar, J. Kossin, S. Lewis, F. Otto, I. Pinto, M. Satoh, S. Vicente-Serrano, M. Wehner, and B. Zhou, Weather and Climate Extreme Events in a Changing Climate, in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2021) pp. 1513–1766, type: Book Section.
- [97] R. García-Herrera, J. Díaz, R. M. Trigo, J. Luterbacher, and E. M. Fischer, A review of the european summer heat wave of 2003, *Critical Reviews in Environmental Science and Technology* **40**, 267 (2010).
- [98] V. Lucarini, V. Melinda Galfi, J. Riboldi, and G. Messori, Typicality of the 2021 western north america summer heatwave, *Environmental Research Letters* **18**, 015004 (2023).
- [99] C. Gessner, E. M. Fischer, U. Beyerle, and R. Knutti, Very rare heat extremes: quantifying and understanding using ensemble reinitialization, *Journal of Climate* **34**, 6619 (2021), iISBN: 0894-8755.
- [100] V. M. Gálfí, V. Lucarini, and J. Wouters, A large deviation theory-based analysis of heat waves and cold spells in a simplified model of the general circulation of the atmosphere, *Journal of Statistical Mechanics: Theory and Experiment* **2019**, 033404 (2019), iISBN: 1742-5468 Publisher: IOP Publishing.
- [101] P. Yiou, AnaWEGE: a weather generator based on analogues of atmospheric circulation, *Geoscientific Model Development* **7**, 531 (2014).
- [102] P. Yiou and A. Jézéquel, Simulation of extreme heat waves with empirical importance sampling, *Geoscientific Model Development* **13**, 763 (2020).
- [103] K. L. Swanson and R. T. Pierrehumbert, Lower-tropospheric heat transport in the pacific storm track, *Journal of the atmospheric sciences* **54**, 1533 (1997).
- [104] G. Messori and A. Czaja, On the sporadic nature of meridional heat transport by transient eddies, *Quarterly Journal of the Royal Meteorological Society* **139**, 999 (2013).
- [105] G. Messori and A. Czaja, Some considerations on the spectral features of meridional heat transport by transient eddies, *Quarterly Journal of the Royal Meteorological Society* **140**, 1377 (2014).
- [106] G. Messori and A. Czaja, On local and zonal pulses of atmospheric heat transport in reanalysis data, *Quarterly Journal of the Royal Meteorological Society* **141**, 2376 (2015).
- [107] G. Messori, R. Geen, and A. Czaja, On the spatial and temporal variability of atmospheric heat transport in a hierarchy of models, *Journal of the Atmospheric Sciences* **74**, 2163 (2017).
- [108] V. Lembo, G. Messori, R. Graversen, and V. Lucarini, Spectral decomposition and extremes of atmospheric meridional energy transport in the northern hemisphere midlatitudes, *Geophysical Research Letters* **46**, 7602 (2019).
- [109] J. H. Rydsaa, R. Graversen, T. I. H. Heiskanen, and P. Stoll, Changes in atmospheric latent energy transport into the arctic: Planetary versus synoptic scales, *Quarterly Journal of the Royal Meteorological Society* **147**, 2281 (2021).
- [110] V. Lembo, F. Fabiano, V. M. Galfi, R. G. Graversen, G. Messori, *et al.*, Meridional-energy-transport extremes and the general circulation of northern hemisphere mid-latitudes: dominant weather regimes and preferred zonal wavenumbers, *Weather and Climate Dynamics* **3**, 1037 (2022).
- [111] C. A. Ferro and J. Segers, Inference for clusters of extreme values, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **65**, 545 (2003).
- [112] V. M. Gálfí, T. Bódai, and V. Lucarini, Convergence of extreme value statistics in a two-layer quasi-geostrophic atmospheric model, *Complexity* **2017** (2017).
- [113] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, *et al.*, The ncep/ncar 40-year reanalysis project, *Bulletin of the American meteorological Society* **77**, 437 (1996).
- [114] S. D. Burt and D. Mansfield, The great storm of 15–16 october 1987, *Weather* **43**, 90 (1988).
- [115] D. Mitchell, K. Kornhuber, C. Huntingford, and P. Uhe, The day the 2003 european heatwave record was broken, *The Lancet Planetary Health* **3**, e290 (2019).
- [116] P.-A. Michelangeli, R. Vautard, and B. Legras, Weather regimes: Recurrence and quasi stationarity, *Journal of the atmospheric sciences* **52**, 1237 (1995).
- [117] A. Hannachi, D. M. Straus, C. L. Franzke, S. Corti, and T. Woollings, Low-frequency nonlinearity and regime

- behavior in the northern hemisphere extratropical atmosphere, *Reviews of Geophysics* **55**, 199 (2017).
- [118] A. Hochman, G. Messori, J. F. Quinting, J. G. Pinto, and C. M. Grams, Do atlantic-european weather regimes physically exist?, *Geophysical Research Letters* **48**, e2021GL095574 (2021).
- [119] R. Vautard, Multiple weather regimes over the north atlantic: Analysis of precursors and successors, *Monthly weather review* **118**, 2056 (1990).
- [120] J. Dorrington, K. Strommen, and F. Fabiano, Quantifying climate model representation of the wintertime euro-atlantic circulation using geopotential-jet regimes, *Weather and Climate Dynamics* **3**, 505 (2022).
- [121] C. M. Grams, R. Beerli, S. Pfenniger, I. Staffell, and H. Wernli, Balancing europe’s wind-power output through spatial deployment informed by weather regimes, *Nature climate change* **7**, 557 (2017).
- [122] P. Yiou and M. Nogaj, Extreme climatic events and weather regimes over the north atlantic: when and where?, *Geophysical Research Letters* **31** (2004).
- [123] C. Lavaysse, J. Vogt, A. Toreti, M. L. Carrera, and F. Pappenberger, On the use of weather regimes to forecast meteorological drought over europe, *Natural Hazards and Earth System Sciences* **18**, 3297 (2018).
- [124] R. Beerli and C. M. Grams, Stratospheric modulation of the large-scale circulation in the atlantic–european region and its implications for surface weather events, *Quarterly Journal of the Royal Meteorological Society* **145**, 3732 (2019).
- [125] J. Pasquier, S. Pfahl, and C. M. Grams, Modulation of atmospheric river occurrence and associated precipitation extremes in the north atlantic region by european weather regimes, *Geophysical Research Letters* **46**, 1014 (2019).
- [126] P. Ruggieri, M. C. Alvarez-Castro, P. Athanasiadis, A. Bellucci, S. Materia, and S. Gualdi, North atlantic circulation regimes and heat transport by synoptic eddies, *Journal of Climate* **33**, 4769 (2020).
- [127] S. Hauser, S. Mueller, X. Chen, T.-C. Chen, J. G. Pinto, and C. M. Grams, The linkage of serial cyclone clustering in western europe and weather regimes in the north atlantic-european region in boreal winter, *Geophysical Research Letters* **50**, e2022GL101900 (2023).
- [128] C. Alvarez-Castro, D. Faranda, P. Yiou, *et al.*, Atmospheric dynamics leading to west european summer hot temperatures since 1851., in *Complex.*, Vol. 2018 (2018) pp. 2494509–1.
- [129] N. Cortesi, V. Torralba, N. González-Reviriego, A. Soret, and F. J. Doblas-Reyes, Characterization of european wind speed variability using weather regimes, *Climate Dynamics* **53**, 4961 (2019).
- [130] N. Cortesi, V. Torralba, L. Lledó, A. Manrique-Suñén, N. Gonzalez-Reviriego, A. Soret, and F. J. Doblas-Reyes, Yearly evolution of euro-atlantic weather regimes and of their sub-seasonal predictability, *Climate Dynamics* **56**, 3933 (2021).
- [131] W. J. Merryfield, J. Baehr, L. Batté, E. J. Becker, A. H. Butler, C. A. Coelho, G. Danabasoglu, P. A. Dirmeyer, F. J. Doblas-Reyes, D. I. Domeisen, *et al.*, Current and emerging developments in subseasonal to decadal prediction, *Bulletin of the American Meteorological Society* **101**, E869 (2020).
- [132] P. Platzer, B. Chapron, and P. Tandeo, Dynamical properties of weather regime transitions, *Stochastic Transport in Upper Ocean Dynamics*, 223 (2023).
- [133] S. Scher and G. Messori, Predicting weather forecast uncertainty with machine learning, *Quarterly Journal of the Royal Meteorological Society* **144**, 2830 (2018).
- [134] A. Hochman, S. Scher, J. Quinting, J. G. Pinto, and G. Messori, A new view of heat wave dynamics and predictability over the eastern mediterranean, *Earth System Dynamics* **12**, 133 (2021).
- [135] A. Hochman, S. Scher, J. Quinting, J. G. Pinto, and G. Messori, Dynamics and predictability of cold spells over the eastern mediterranean, *Climate Dynamics* **58**, 2047 (2022).
- [136] F. J. Tapiador, A. Villalba-Pradas, A. Navarro, R. Martín, A. Merino, E. García-Ortega, J. L. Sánchez, K. Kim, and G. Lee, A satellite view of an intense snowfall in madrid (spain): The storm ‘filomena’ in january 2021, *Remote Sensing* **13**, 2702 (2021).
- [137] P. Zschenderlein and H. Wernli, The unusually long cold spell and the snowstorm filomena in spain in january 2021, *Natural Hazards and Earth System Sciences Discussions*, 1 (2022).
- [138] A. Hochman, P. Alpert, T. Harpaz, H. Saaroni, and G. Messori, A new dynamical systems perspective on atmospheric predictability: Eastern mediterranean weather regimes as a case study, *Science advances* **5**, eaau0936 (2019).
- [139] G. Messori, R. Caballero, and M. Gaetani, On cold spells in north america and storminess in western europe, *Geophysical Research Letters* **43**, 6620 (2016).
- [140] E. M. Fischer, S. I. Seneviratne, P. L. Vidale, D. Lüthi, and C. Schär, Soil moisture–atmosphere interactions during the 2003 european summer heat wave, *Journal of Climate* **20**, 5081 (2007).
- [141] E. Mills, Insurance in a Climate of Change, *Science* **309**, 1040 (2005), publisher: American Association for the Advancement of Science.
- [142] K. Kornhuber, D. Coumou, E. Vogel, C. Lesk, J. F. Donges, J. Lehmann, and R. M. Horton, Amplified rossby waves enhance risk of concurrent heatwaves in major breadbasket regions, *Nature Climate Change* **10**, 48 (2020).
- [143] R. Leeding, J. Riboldi, and G. Messori, On pan-atlantic cold, wet and windy compound extremes, *Weather and Climate Extremes* **39**, 100524 (2023).
- [144] G. Messori and D. Faranda, On the systematic occurrence of compound cold spells in north america and wet or windy extremes in europe, *Geophysical Research Letters* **50**, e2022GL101008 (2023).
- [145] J. Riboldi, R. Leeding, A. Segalini, and G. Messori, Multiple large-scale dynamical pathways for pan–atlantic compound cold and windy extremes, *Geophysical Research Letters* **50**, e2022GL102528 (2023).
- [146] D. Faranda, G. Messori, and P. Yiou, Diagnosing concurrent drivers of weather extremes: application to warm and cold days in north america, *Climate Dynamics* **54**, 2187 (2020).
- [147] M. P. Cellitti, J. E. Walsh, R. M. Rauber, and D. H. Portis, Extreme cold air outbreaks over the united states, the polar vortex, and the large-scale circulation, *Journal of Geophysical Research: Atmospheres* **111** (2006).
- [148] R. Grotjahn, R. Black, R. Leung, M. F. Wehner, M. Barlow, M. Bosilovich, A. Gershunov, W. J. Gutowski, J. R. Gyakum, R. W. Katz, *et al.*, North american extreme temperature events and related large

- scale meteorological patterns: a review of statistical methods, dynamics, modeling, and trends, *Climate Dynamics* **46**, 1151 (2016).
- [149] E. National Academies of Sciences, Medicine, *et al.*, *Attribution of extreme weather events in the context of climate change* (National Academies Press, 2016).
- [150] A. T. Rädler, P. H. Groenemeijer, E. Faust, R. Sausen, and T. Púčík, Frequency of severe thunderstorms across europe expected to increase in the 21st century due to rising instability, *npj Climate and Atmospheric Science* **2**, 30 (2019).
- [151] M. K. Tippett, A. H. Sobel, and S. J. Camargo, Association of us tornado occurrence with monthly environmental parameters, *Geophysical research letters* **39** (2012).
- [152] S. J. Camargo and A. A. Wing, Tropical cyclones in climate models, *Wiley Interdisciplinary Reviews: Climate Change* **7**, 211 (2016).
- [153] D. Faranda, S. Bourdin, M. Ginesta, M. Krouma, G. Messori, R. Noyelle, F. Pons, and P. Yiou, A climate-change attribution retrospective of some impactful weather extremes of 2021, *Weather and Climate Dynamics Discussions*, 1 (2022).
- [154] D. Faranda, V. Lembo, M. Iyer, D. Kuzzay, S. Chibbaro, F. Daviaud, and B. Dubrulle, Computation and characterization of local subfilter-scale energy transfers in atmospheric flows, *Journal of the Atmospheric Sciences* **75**, 2175 (2018).
- [155] A. Leonard, Energy cascade in large-eddy simulations of turbulent fluid flows, in *Advances in geophysics*, Vol. 18 (Elsevier, 1975) pp. 237–248.
- [156] M. Germano, Turbulence: the filtering approach, *Journal of Fluid Mechanics* **238**, 325–336 (1992).
- [157] C. Meneveau, Statistics of Turbulence Subgrid-Scale Stresses - Necessary Conditions and Experimental Tests, in *Physics of Fluids* (1994) pp. 815–833.
- [158] C. Tong, J. C. Wyngaard, and J. Bresseur, Experimental study of the subgrid-scale stresses in the atmospheric surface layer, *Journal of Atmospheric Sciences* **56**, 2277 (1999).
- [159] N. Vercauteren, E. Bou-Zeid, M. Parlange, U. Lemmin, H. Huwald, J. S. Selker, and C. Meneveau, Subgrid-scale dynamics of water vapour, heat, and momentum over a lake, *Boundary-Layer Meteorology* **128**, 205 (2008), publisher: Springer.
- [160] S. Vannitsem and M. Ghil, Evidence of coupling in ocean-atmosphere dynamics over the north atlantic, *Geophysical Research Letters* **44**, 2016 (2017).
- [161] S. Sroka and S. Guimond, Organized kinetic energy backscatter in the hurricane boundary layer from radar measurements, *Journal of Fluid Mechanics* **924**, A21 (2021).
- [162] J. Duchon and R. Robert, Inertial energy dissipation for weak solutions of incompressible euler and navier-stokes equations, *Nonlinearity* **13**, 249 (2000).
- [163] D. Kuzzay, E.-W. Saw, F. J. Martins, D. Faranda, J.-M. Foucaut, F. Daviaud, and B. Dubrulle, New method for detecting singularities in experimental incompressible flows, *Nonlinearity* **30**, 2381 (2017).
- [164] A. H. Sobel, A. A. Wing, S. J. Camargo, C. M. Patricola, G. A. Vecchi, C.-Y. Lee, and M. K. Tippett, Tropical cyclone frequency, *Earth’s Future* **9**, e2021EF002275 (2021).
- [165] K. A. Emanuel, An air-sea interaction theory for tropical cyclones. part i: Steady-state maintenance, *Journal of Atmospheric Sciences* **43**, 585 (1986).
- [166] S. Camargo, A. H. Sobel, A. G. Barnston, and K. A. Emanuel, Tropical cyclone genesis potential index in climate models, *Tellus A: Dynamic Meteorology and Oceanography* **59**, 428 (2007).
- [167] K. Emanuel and D. S. Nolan, Tropical cyclone activity and the global climate system, in *26th conference on hurricanes and tropical meteorology* (2004).
- [168] M. K. Tippett, S. J. Camargo, and A. H. Sobel, A poisson regression index for tropical cyclone genesis and the role of large-scale vorticity in genesis, *Journal of Climate* **24**, 2335 (2011).
- [169] B. Wang and H. Murakami, Dynamic genesis potential index for diagnosing present-day and future global tropical cyclone genesis, *Environmental Research Letters* **15**, 114008 (2020).
- [170] C. E. Menkes, M. Lengaigne, P. Marchesiello, N. C. Jourdain, E. M. Vincent, J. Lefèvre, F. Chauvin, and J.-F. Royer, Comparison of tropical cyclogenesis indices on seasonal to interannual timescales, *Climate dynamics* **38**, 301 (2012).
- [171] L. Cavicchia, E. Scoccimarro, G. Ascenso, A. Castelletti, M. Giuliani, and S. Gualdi, Tropical cyclone genesis potential indices in the era of high-resolution climate models, *Geophysical research letters* (submitted).
- [172] K. Emanuel, Tropical cyclone seeds, transition probabilities, and genesis, *Journal of Climate* **35**, 3557 (2022).
- [173] L. Cavicchia, E. Scoccimarro, G. Ascenso, A. Castelletti, M. Giuliani, and S. Gualdi, Tropical cyclone genesis potential indices in a new high-resolution climate models ensemble: Limitations and way forward, *Geophysical Research Letters* **50**, e2023GL103001 (2023).
- [174] K. A. Emanuel, S. Ravela, E. Vivant, and C. Risi, Application of stochastic and deterministic modeling to hurricane wind risk assessment, *Massachusetts Institute of Technology, Cambridge, Massachusetts* (2006).
- [175] K. Emmanuel, Using physics to help assess tropical cyclone risk, (2017).
- [176] S. Hallegatte, The use of synthetic hurricane tracks in risk analysis and climate change damage assessment, *Journal of applied meteorology and climatology* **46**, 1956 (2007).
- [177] T. W. B. Leijnse, A. Giardino, K. Nederhoff, and S. Caires, Generating reliable estimates of tropical-cyclone-induced coastal hazards along the bay of bengal for current and future climates using synthetic tracks, *Natural Hazards and Earth System Sciences* **22**, 1863 (2022).
- [178] J. Rumpf, H. Weindl, P. Höppe, E. Rauch, and V. Schmidt, Stochastic modelling of tropical cyclone tracks, *Mathematical Methods of Operations Research* **66**, 475 (2007).
- [179] N. Bloemendaal, I. D. Haigh, H. de Moel, S. Muis, R. J. Haarsma, and J. C. Aerts, Generation of a global synthetic tropical cyclone hazard dataset using storm, *Scientific data* **7**, 40 (2020).
- [180] K. J. Walsh, J. L. McBride, P. J. Klotzbach, S. Balachandran, S. J. Camargo, G. Holland, T. R. Knutson, J. P. Kossin, T.-c. Lee, A. Sobel, *et al.*, Tropical cyclones and climate change, *Wiley Interdisciplinary Reviews: Climate Change* **7**, 65 (2016).
- [181] C.-Y. Lee, S. J. Camargo, A. H. Sobel, and M. K. Tippett, Statistical-dynamical downscaling projections of

- tropical cyclone activity in a warming climate: Two diverging genesis scenarios, *Journal of Climate* **33**, 4815 (2020).
- [182] Y. Frenkel, A. J. Majda, and B. Khouider, Using the stochastic multicloud model to improve tropical convective parameterization: A paradigm example, *Journal of the Atmospheric Sciences* **69**, 1080 (2012).
- [183] J. T. Allen, M. K. Tippett, and A. H. Sobel, An empirical model relating us monthly hail occurrence to large-scale meteorological environment, *Journal of Advances in Modeling Earth Systems* **7**, 226 (2015).
- [184] B. Khouider, J. Biello, and A. J. Majda, A stochastic multicloud model for tropical convection, (2010).
- [185] S. Hottovy and S. N. Stechmann, A spatiotemporal stochastic model for tropical precipitation and water vapor dynamics, *Journal of the Atmospheric Sciences* **72**, 4721 (2015).
- [186] N. Ban, J. Schmidli, and C. Schär, Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations, *Journal of Geophysical Research: Atmospheres* **119**, 7889 (2014).
- [187] E. Coppola, S. Sobolowski, E. Pichelli, F. Raffaele, B. Ahrens, I. Anders, N. Ban, S. Bastin, M. Belda, D. Belusic, *et al.*, A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over europe and the mediterranean, *Climate Dynamics* **55**, 3 (2020).
- [188] G. Fossler, E. J. Kendon, D. Stephenson, and S. Tucker, Convection-permitting models offer promise of more certain extreme rainfall projections, *Geophysical Research Letters* **47**, e2020GL088151 (2020).
- [189] E. Pichelli, E. Coppola, S. Sobolowski, N. Ban, F. Giorgi, P. Stocchi, A. Alias, D. Belušić, S. Berthou, C. Caillaud, *et al.*, The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation, *Climate Dynamics* **56**, 3581 (2021).
- [190] M. Ginesta, P. Yiou, G. Messori, and D. Faranda, A methodology for attributing severe extratropical cyclones to climate change based on reanalysis data: the case study of storm alex 2020, *Climate Dynamics* **61**, 229 (2023).
- [191] C. Cadiou, R. Noyelle, N. Malhomme, and D. Faranda, Challenges in attributing the 2022 australian rain bomb to climate change, *Asia-Pacific Journal of Atmospheric Sciences* **59**, 83 (2023).
- [192] D. Faranda, S. Pascale, and B. Bulut, Persistent anti-cyclonic conditions and climate change exacerbated the exceptional 2022 european-mediterranean drought, *Environmental Research Letters* **18**, 034030 (2023).
- [193] W. Hazeleger, B. J. van den Hurk, E. Min, G. J. van Oldenborgh, A. C. Petersen, D. A. Stainforth, E. Vasileiadou, and L. A. Smith, Tales of future weather, *Nature Climate Change* **5**, 107 (2015).
- [194] T. G. Shepherd, A common framework for approaches to extreme event attribution, *Current Climate Change Reports* **2**, 28 (2016).
- [195] E. A. Lloyd and T. G. Shepherd, Climate change attribution and legal contexts: evidence and the role of storylines, *Climatic Change* **167**, 28 (2021).
- [196] P. Yiou, A. Jézéquel, P. Naveau, F. E. Otto, R. Vautard, and M. Vrac, A statistical framework for conditional extreme event attribution, *Advances in Statistical Climatology, Meteorology and Oceanography* **3**, 17 (2017).
- [197] A. C. M. Freitas, J. M. Freitas, M. Magalhães, and S. Vaienti, Point processes of non stationary sequences generated by sequential and random dynamical systems, *Journal of Statistical Physics* **181**, 1365 (2020).