Physics – Uspekhi **57** (3) 280 – 291 (2014)

100th ANNIVERSARY OF THE BIRTH OF Ya B ZELDOVICH

PACS numbers: **02.50.–r**, **05.45.–a**, 92.05.Df

Climate as a game of chance

A Ruzmaikin

DOI: 10.3367/UFNe.0184.201403f.0297

Contents

Abstract. We use general concepts and simple models to examine the role of randomness in chaotic systems, like Earth's climate, in response to external forcing. The response of a simple homogeneous system is determined by its correlation function in accordance with the fluctuation–dissipation theorem. A structured (patterned) system responds in a more complicated way. Whereas its mean state (for example, Earth's global temperature) is changing only slightly, extreme events (such as floods and droughts) are increasing more dramatically in number. The statistics of extremes reveals remarkable properties, in particular, clustering (troubles never come alone, the saying goes) and are here illustrated by precipitation and space climate processes.

1. Introduction

In 1987, the last year of Zeldoich's life, Yakob Borisovich, Dmitry Sokoloff, and I wrote the book *Almighty Chance* [1]. It was a time when scientists were rediscovering randomness in physics. There were very active discussions on whether turbulence could be described as a mix of a large number of modes with irrational ratios of frequencies [2] or as a simple chaotic system with a few degrees of freedom [3]. Mathematicians brought out Lapunov's exponents, phase space trajectories, and strange attractors. Our book exposed that fascinating time by discussing the role of randomness in the origin of magnetic fields of planets, stars, and galaxies, our main interests at that time. Zeldovich was pleased to dust off the classical work of Einstein and Smoluchowsky on Brownian motion, which we, together with Stanislav Molchanov, a talented mathematician from Kolmogorov's school, extended to the transport of scalars and vectors using modern mathematical techniques [4, 5].

One of the great features of Zeldovich's personality was keeping the pulse of time. He instantly and actively responded to new paths in science and in many cases greatly contributed to their progress. In the spirit of his style, I believe it would be appropriate to discuss the role of randomness at the current level of scientific and public interests. And the most exciting issue, at least from the author's viewpoint, is the role of randomness in our Earth's very chaotic climate system.

Today, nobody denies that the climate of Earth is changing. Earth's global temperature is rising on a long time scale, and the number of extreme weather events such as floods and droughts is increasing in magnitude and frequency. The current news is full of reports of devastating consequences of these extreme events. We can easily pick up a few major events from the Internet: The heat wave in Europe in the summer of 2003 cost almost 40,000 lives. The July 2010 Pakistan floods resulting from heavy monsoon rains covered about one-fifth of Pakistan's total land area and affected about 20 million people. The great Russian heat wave of July 2010, with nothing similar of that magnitude over the last one thousand years, ended with about 15,000 dead. Sandy, the deadliest and most destructive hurricane of the 2012 Atlantic hurricane season, and the second costliest hurricane (after Katrina) in the United States history, cost over 68 billion US dollars. Not to mention the August 2013 painful, unprecedented flood in the Amur region, the birth place of the author of this review.

The debates are now centering around the question of whether current climate change is caused by humans, mostly due to the release of carbon dioxide (CO_2) and other greenhouse gases into Earth's atmosphere. There are skeptical scientists and non-scientists who believe that the climate is changing mainly due to natural processes intrinsic to Earth's ocean–atmosphere system and that the anthropogenic (human) effect is small. Answering this question requires a careful investigation of Earth's climate system responses to external forcing and finding whether the response to anthropogenic forcing, such as release of CO_2, is different from the response to natural forcing, such as

A Ruzmaikin Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA
E-mail: alexander.ruzmaikin@jpl.nasa.gov

Received 21 October 2013, revised 9 December 2013
Uspekhi Fizicheskikh Nauk **184** (3) 297 – 311 (2014)
DOI: 10.3367/UFNr.0184.201403f.0297
Translated by A Ruzmaikin; edited by A M Semikhatov

volcanos or the sun. An extensive account of the current state of research and analyses of observations related to this question can be found in the recently released 5th IPCC Report [6]. In *Physics–Uspekhi*, Byalko [7] gave a nice introduction to climate change over an extended time period using a simple relaxation approach.

Discussions of statistical issues close to those to be touched on below can be found in books [8] and [9]. Numeric methods of modeling atmospheric circulation and climate changes are being rapidly developed, but we do not discuss them here. This modest review, which is written in the spirit of book [1], is biased towards the author's own research and focused on only one aspect, the role of randomness in the climate response to external forcing.

The review is organized as follows. In Section 2, we start with the introduction to well-known studies of the response of purely random or chaotic homogeneous systems to weak external forcing. We then discuss the role of structures embedded into the system, by first using the examples of the attractors of the Lorentz dynamical system and of a simple dynamical system based on the double-well potential, and then considering Earth's real climate structures, called 'climate patterns'. A key conclusion of these discussions is that forcing only weakly affects the mean states of the system but strongly increases the probability of large deviations (extreme events). Section 3 is devoted to discussing the basics of statistics of extremes and presents examples of the application of the extreme value statistics to space weather and precipitation extremes.

2. Response to forcing

Linear and nonlinear systems respond differently to external forcing. A classic example of a linear system response is Hooke's law of elasticity, which states that the amount by which a material body is deformed is linearly proportional to the force causing the deformation. This linear approximation is still widely used in climate studies [6], for example, in evaluating the sensitivity s of the change in Earth's global temperature T caused by a change in radiative forcing F:

$$\Delta T = s\Delta F .$$

The radiative forcing is defined as the difference between the solar energy received by Earth and the energy radiated back to space. It is usually quantified as a change in the heat flux ΔF at the atmospheric tropopause in units of [W m^{-2}]. A typical value $s = 0.8$ K W^{-1} m^2 allows making a quick estimate of global warming caused by any heat imbalance. For example, for the evaluation of global warming due to the increase in CO_2, the formula $\Delta F = 5.4 \ln(CO_2/C_0)$, where C_0 is a reference concentration of CO_2, is often used (cf. wikipedia.org/wiki/Radiative_forcing).

The response of nonlinear systems to external forcing is conceptually different. The issue is not the magnitude (sensitivity) of the response but the greater involvement of the system's rich characteristics of randomness and the different ways these characteristics are treated. For example, it was suggested that averaged atmospheric variability, after subtracting clearly observed periodicities such as seasonal cycles, can be completely characterized as a fluctuating system that displays a different behavior on shorter (< 10 days typical for weather) and longer (years, typical for climate) time scales [10]. Such a 'soup' type treatment of Earth's climate system interprets the fluctuations as combinations of external forcings (anthropogenic, solar, volcanic, etc.) and internal feedbacks caused by processes such as deep-ocean or land-ice dynamics. We here adopt a different, more traditional view of separating the system fluctuations from the external forcing.

The classical approach to studying the response of a random system to forcing is based on considering an isotropic, uniform medium that is fully defined by the second-order correlations (covariances) of its fluctuations. The response of this type of system is determined by the general fluctuation–dissipation theorem (FDT), which can be applied to a wide range of physical systems [11–15].

2.1 Fluctuation–dissipation theorem

According to the FDT, the response of a random system to weak external forcing is determined by covariances and lagged covariances of fluctuations of the undisturbed system. This theorem was applied in [16, 17], among others, to Earth's climate system, and in [18–20], to barotropic and two-level baroclinic atmospheric models in a multivariate setting to find the sensitivity of the atmospheric dynamics to weak forcing.

The equations for unperturbed and perturbed systems have the form

$$\frac{\mathrm{d}\mathbf{u}_0}{\mathrm{d}t} = \mathbf{N}(\mathbf{u}_0, \lambda) , \tag{1}$$

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \mathbf{N}(\mathbf{u}, \lambda) + F , \tag{2}$$

where $\mathbf{u}_0$ and $\mathbf{u}$ are the unperturbed and perturbed climate variables, λ is a set of parameters, and $\mathbf{N}$ is a nonlinear operator. Under the assumption that the probability distribution function of solutions of Eqns (1) and (2) is the Gaussian $\exp[-C^{-1}(0)(\mathbf{u}\,\mathbf{u})/2]$, it follows that the difference between the mean values of $\mathbf{u}_0$ and $\mathbf{u}$ is

$$\langle \mathbf{u_0} \rangle - \langle \mathbf{u} \rangle = \int_0^t C(\tau)\, C^{-1}(0)\, F \mathrm{d}\tau , \tag{3}$$

where $C(\tau)$ is the lag-τ covariance matrix of $\mathbf{u}$.

This is the main result provided by the FDT. It can be easily understood if we treat the system as a well-mixed random system similar to classical, uniform, isotropic turbulence. In this approximation, the system is fully described by its second-order correlations, and therefore its response to external forcing is determined by the correlation (covariance) function. In fact, all proofs of the FDT are based on assumptions that reduce the system to a uniformly mixed, Gaussian, stationary medium.

The classic FDT characterizes the response of a fluctuating, dissipative system to small, stationary, external forcing. In an excellent review in *Physics–Uspekhi*, devoted to the legacy of V L Ginzburg, Pitaevskii discussed the general situation where the external force is not small and not stationary [21]. He showed what can be done in this highly nonlinear situation using the approach developed by Bochkov and Kuzovlev [22] and the Jarzynski equality [23]

$$\left\langle \exp\left(-\beta W\right) \right\rangle = \exp\left(\beta \Delta \mathcal{F}\right) ,$$

where W is the work on the system done by the external force, $1/\beta$ is the temperature, and $\Delta\mathcal{F}$ is the difference between the

equilibrium free energy and its value before the force application. A non-evident fact is that extreme values of W, when it takes rare, very small values, substantially contribute to $\langle \exp(-\beta W) \rangle$ due to the exponential dependence. Pitaevskii described mechanical and biological experiments confirming this extension of the FDT and emphasizing the benefits of using traditionally unrelated branches of science. We here show that there is another potential domain of application of these fundamental ideas: climate.

The FDT is focused on changes in the mean value of the random system variables in response to an external forcing. But real systems are usually highly structured and the standard statistical mean does not provide a sufficient measure of the response. Mathematically, nonlinear system structures are characterized by preferred states defined by their internal processes and called 'attractors'. The dynamics are defined by residence in the states and transitions between them. The question is how an external forcing changes the states, residence times, and other characteristics of the system. The answer to this question is critical to our understanding of climate change, because Earth's real climate is a good example of a very structured system. In the following sections, we discuss a possible approach to understanding the effects of external forcing of structured systems. We start with the simplest dynamical system, which has two basic attractors in its phase space. We then discuss how Earth's climate structures fit into this category.

2.2 Forcing a two-state dynamical system

Insight into a structured system response to a weak forcing can be gained by considering the forcing of a dynamical system with known states (attractors). The simplest and best known is the Lorenz system [24]

$$\dot{x} = -\sigma(x - y) + F_0 \cos\theta\,, \tag{4}$$

$$\dot{y} = -xz + rx - y + F_0 \sin\theta\,, \tag{5}$$

$$\dot{z} = xy - bz\,, \tag{6}$$

which (for the parameters chosen as $r = 28$, $\sigma = 10$, $b = 8/3$, and $F_0 = 0$) has two attractors (Fig. 1a). What happens when this system is perturbed by a force of a constant amplitude F_0 [see the right-hand sides of Eqns (4) and (5)] applied at an angle θ was investigated in [25]. It was shown there that the probability density functions of the residence time in these two attractors are exponential, and that the strongest response to forcing is a change in the frequency of the occurrence of extremely persistent events, rather than a much weaker change in the mean residence time; in other words, the tails of the density have enhanced sensitivity to forcing.

To understand this interesting result, in the spirit of the approach in [26], we consider an even simpler system than the Lorenz one, a mechanical system

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{\mathrm{d}U}{\mathrm{d}x}\,, \tag{7}$$

where U is a double-well potential complemented with a constant external force F_0:

$$U = -\frac{x^2}{2} + \frac{x^4}{4} - F_0\, x\,. \tag{8}$$

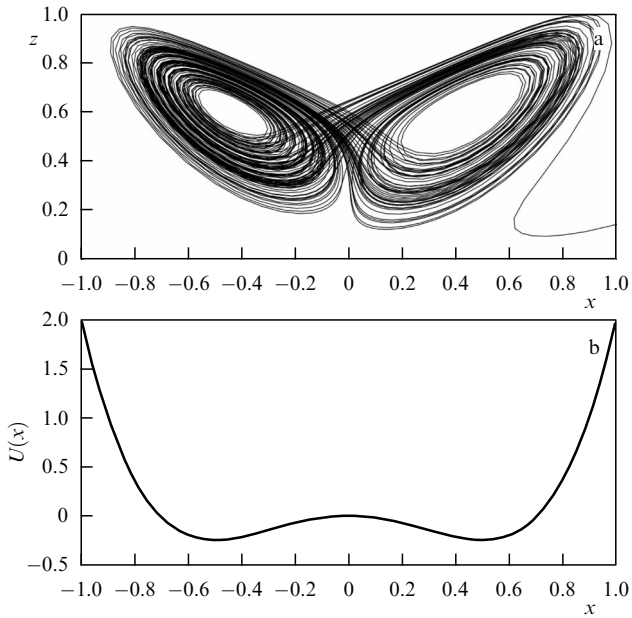

Figure 1. (a) Two attractors of the Lorenz system seen in the (x, z) plane. (a) Two states in a double-well potential dynamical system.

The potential wells are similar to the two attractors of the Lorenz system (Fig. 1b). We now generate random transitions from one potential well to another. This can be done by applying a random forcing to the right-hand side of Eqn (7).

In the absence of the force, $F_0 = 0$, the mean residence times in the positive (right) and negative (left) wells are equal: $\tau_\mathrm{p} = \tau_\mathrm{n}$. But when $F_0 \neq 0$, one well becomes deeper than the other, and the residence times become different.

Figure 2 shows the difference found by numerical simulation for $F_0 = -0.05$. In this case, τ_p becomes smaller than τ_n, i.e., the system spends more time in the negative well. It is also true that the duration of random times spent in the negative well becomes larger. This less trivial fact can be discerned from the difference in the tail of probability distributions of the well duration times shown in Fig. 2c.

These results suggest that the most significant effect of forcing can be a change in the frequency of the occurrence of negative events (or positive events if the positive well is made deeper), rather than a change in the mean state. In the next section, we return to this issue and consider an application of the concept developed here to real structures of Earth's climate system, called 'climate patterns' here.

2.3 Climate patterns

The concept of climate patterns was introduced by the celebrated British scientist Sir Gilbert Walker in his address to the Royal Society, and was published in [27]. Walker looked into the surprising anecdotal reports of connections between the weather in distant parts of Earth. By analyzing the available observations, he recognized that the correlations are real and introduced three large-scale coherent oscillating patterns of the atmosphere–ocean variability, which he called 'swayings' and which are now known as the North Atlantic Oscillation, North Pacific Oscillation, and Southern (South Pacific–Indian Ocean) Oscillation. Subsequently, numerous studies have demonstrated that the atmospheric and oceanic anomalies (deviations from the mean) are associated with spatially coherent patterns. From

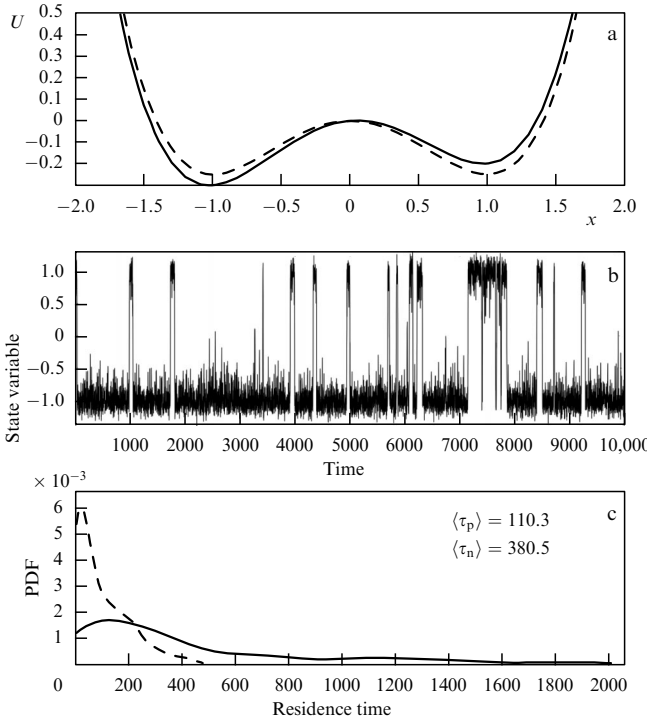

Figure 2. (a) Double-well potential perturbed by a force lowering its negative (left) well. (b) Time series of the state variable in the perturbed potential. (c) Distribution functions of the residence times in the two potential wells. The distribution function for the negative well has a higher mean value (shown on the plot) and a much longer tail. The times in (b) and (c) are shown in arbitrary units.

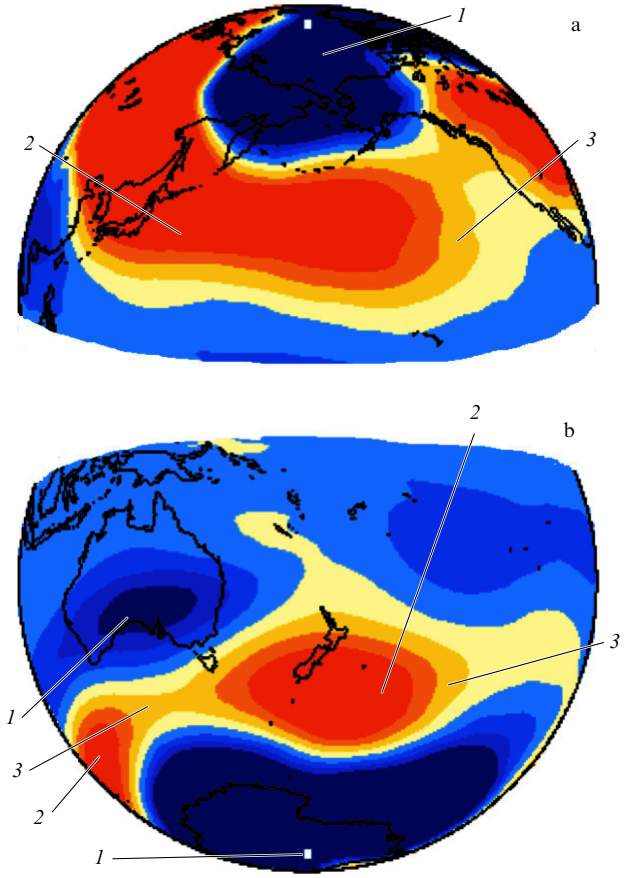

Figure 3. Illustration of the (a) NAM and (b) SAM patterns of the tropospheric temperature. Low-temperature anomalies (deviations from the mean over about a 50-year period) over the poles (area *1*, dark grey) are surrounded by a warm ring (area *2* and *3*, light grey), which includes Eurasia and USA. The cold poles are accompanied by a cooling of the tropics. For more details, see www.jisao.washington.edu/wallace/ncar-notes.

a mathematical standpoint, the patterns can be effectively represented by the empirical orthogonal functions (EOFs) introduced into atmospheric science in [28]. They can be calculated from monthly grids of the sea level pressure (or geopotential heights), temperature, and zonal wind anomalies (cf. [29]).

The time series of principal components (PCs) of these EOFs can be used as empirical indices of patterns. One of the most explored patterns is the North Atlantic Oscillation (NAO), which was originally defined as a December-to-March mean of the sea level pressure difference between Iceland and Portugal [30]. It is closely related to the more spatially extended North Annular Mode (NAM) [31]. The counterpart of the NAM in the southern hemisphere is called the SAM (Fig. 3). These modes are influenced by solar variability [43, 76, 77]. The weather conditions in the northern (southern) hemisphere substantially depend on the state of the NAM (SAM), as is illustrated by the data on weather extremes compiled from [32] (Table 1). The first column defines the type of weather event and its location. The numbers in the second and third columns show those events falling on positive and negative states of the NAM that exceed one standard deviation in absolute value. Many more cold temperature and frozen precipitation anomalies (extreme events) occurred during the negative state of the NAM when the polar regions were warmer. "Blocking Days" refers to days when the normal eastward progression of synoptic disturbances is obstructed, leading to episodes of prolonged extreme weather conditions. The results, which are based on daily data from 1958–1997 winters, are adapted from [32]. The negative NAM state means a warm Arctic and cold Eurasia. In fact, observational evidence shows that significant cold anomalies over the Far East in early winter and zonally elongated cold anomalies from Europe to the Far East in late winter are associated with a decrease in the Arctic sea-ice cover in the preceding summer-to-autumn seasons [33]. Results from numerical experiments using an atmospheric general circulation model support these notions. An anomalous decrease in the wintertime sea ice concentration in the Barents–Kara Seas was associated with extreme cold events like the 2005–2006 winter in Europe [34].

Other examples of climate patterns are the Pacific–North America (PNA) pattern, the Cold Ocean–Warm Land (COWL) pattern [35], the Pacific Decadal Oscillation (PDO) [36], the well-known El Niño–Southern Oscillation (ENSO) in the Pacific [37], and the Quasi-Biennial Oscillation (QBO) in the stratosphere [38]. As shown in [39], the QBO during mid-winter has an impact on the northern hemisphere weather similar to the NAM but somewhat weaker. This impact operated through the change in the strength and stability of the stratospheric polar vortex. The easterly (westerly) phase of the QBO favors an increased incidence of extreme cold (warm) events. The signature of the QBO in NH wintertime temperatures is roughly comparable in amplitude to that observed in relation to the El Niño–Southern Oscillation phenomenon.

Table 1. Extreme weather events associated with the state of the NAM. The first column defines the event type.

Events and location	NAM +	NAM–
Daily min temperature		
< −15 °C Chicago	29	84
< 3 °C Paris	23	97
< −29 °C Novosibirsk	21	85
< −19 °C Bejing	21	55
< −1 °C Tokyo	20	95
Frozen precipitation		
Trace snow, Dallas, TX	1	17
> 5 cm snow, Baltimore	11	63
> 0.5 cm snow, Paris	11	25
> −0 cm snow, Tokyo	8	25
Winds		
> 25 knots, Seattle	78	27
> 35 knots, Astoria, OR	55	20
> 30 knots, Boston	22	45
> 50 knots, Keflavik, Iceland	81	19
Blocking days		
Alaska (170° E–150° W, 60° – 75° N)	53	98
North Atlantic (50° – 0° W, 60° – 75° N)	1	225
Russia (40° – 70° E, 60° – 75° N)	29	82

2.4 Origin of climate patterns

What is the origin of climate patterns? It is generally agreed that climate patterns are naturally excited by the atmosphere–ocean dynamical system. However, physical mechanisms of the excitation are still in an early state of investigation and vary from pattern to pattern. Carl Gustaf Rossby [40, pp. 656–661] was the first to emphasize the importance of two main ingredients of atmospheric dynamics: the zonal–mean zonal wind and nonzonally symmetric deviations of pressure. He described the nonzonally symmetric deviations as waves, which are now known as Rossby or planetary waves.

Rossby waves are driven by the Coriolis force. Their lowest mode changes sign once over the whole longitudinal circle. These large-scale waves are most effectively generated by winter flow over mountains and by sea–land temperature contrasts. The waves propagate in horizontal and vertical directions [41]. The vertical propagation of the waves into the stratosphere with decreasing the air density dramatically increases their amplitude. This increase often leads to nonlinear wave breaking accompanied by energy release that produces temperature anomalies and sometimes reverses the direction of the zonal wind. The zonal wind, in turn, affects the wave propagation by modifying the refractive index. It was suggested and demonstrated in numerical simulations that the excitation of the NAM, which characterizes the zonally symmetric anomalies of atmospheric circulation, involves interaction between planetary waves and the zonal mean flow in the atmosphere [42, 43, 45]. The same interaction produces an orthogonal mode, the PNA climate pattern [29]. This nonlinear wave–zonal flow interaction can be envisioned as a dynamical system with two basic states in its phase space corresponding to positive (negative) NAM states [43–45], thus justifying the simple double-well potential model discussed in Section 2.2.

2.5 Forcing the climate patterns

Introducing an external forcing can change either the states themselves or the residence times spent in the states. In [40], it is hypothesized that forcing does not change the states and only affects the mean residence times (occupation frequencies) of the states.

The Rossby conjecture was further developed in [46] and [47]. For visual illustration, Palmer presented a picture with two tea cups representing the states, a ball randomly thrown from above to simulate occupation of the states, and a fan imitating the external force (Fig. 4a). However, the analysis presented in Section 2.2 shows that the change in the mean residence times is a small effect compared to a stronger change in the tail of the probability distribution of the residence times, i.e., the increase in the frequency of occurrence of extreme events.

To understand why the system responds strongly via extremes, we recall [77] that the energy barrier ΔU separating the two wells impedes the system from transition to the other state (Fig. 4b). This critical feature is missing in the Palmer figure of two equal solid tea cups (Fig. 4a) and in the original Rossby hypothesis. In fact, the external force affects the state by making one potential well deeper than the other. When a

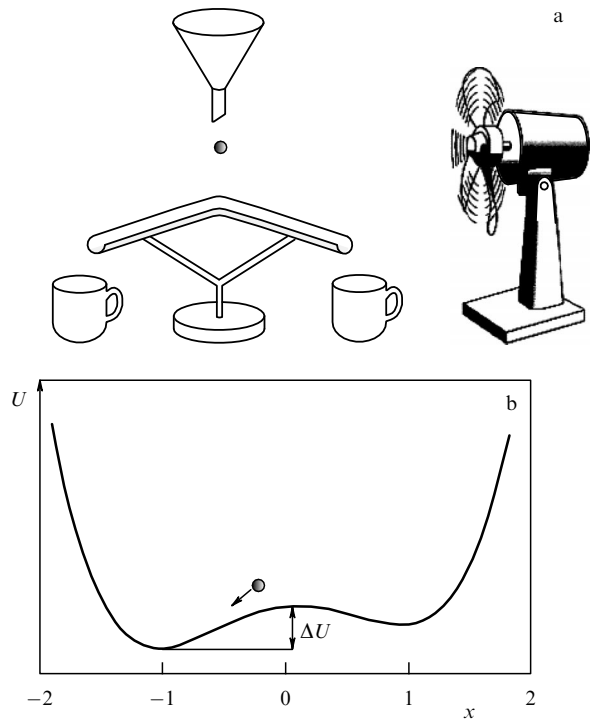

Figure 4. (a) Illustration of forcing a two-state system suggested in [47]. Note that the tea cups are solidly fixed. (b) Illustration of forcing a two-state system suggested in [77] using a two-well potential. The forcing affects the depth of the potential well, making the transition from one state to another biased toward one of the states.

change in the depth is significant, the system more often stays longer in the deeper well, which explains the prolonged persistence of this state (pattern). An example of a real two-state climate system is the NAM. The long residence time in one of its states may be relevant to the predominance of the negative NAM pattern during the prolonged period of reduced solar forcing of climate at the Maunder Minimum of solar activity [48] and the recent Arctic warming accompanied by the cooling of Eurasia [33].

This reasoning is supported by simple estimations based on the two-well model described above. As is well known from 20th-century studies, the probability p of having a residence time τ is proportional to $\exp(\tau/\tau_K)$, where $\tau_K \propto \exp(\Delta U/\sigma)$ is the mean residence time, with σ describing the rate of stochastic forcing that provides transitions from one state to the other [26]. An external forcing that affects the depth of one of the states effectively increases (or decreases) the barrier. Due to the exponential sensitivity of the mean residence time τ_K to the height of the barrier, even a small change in the barrier can induce a noticeable change in the mean residence time. But the change in the probability of having a very long residence time $\tau_c \gg \tau_K$ is larger by the factor τ_c/τ_K because $dp/p = (\tau_c/\tau_K)\, d\tau_K/\tau_K$! Numerical simulations of a model double-well potential system with stochastic transitions between the wells described in Section 2.2 support this rough estimate. When one of the wells is made deeper by changing the forcing in potential (8), the probability distribution of residence times in this well displays a longer tail (Fig. 2c).

Does external forcing also change the spatial structure of the patterns? Model studies of the simple Lorenz and double-well systems discussed above show that the positions of the attractors or wells change only slightly. Climate variations during periods over the past 165,000 yr, when the ice sheet distribution resembled that of the present day, are accumulated in [49]. This condition is met during the Holocene period, extending approximately 11,000 yr before the present. These numerical experiments revealed that the orbital forcing led to variations in the NAM but preserved its spatial structure. Nonetheless, it remains unclear whether changing the boundary conditions, for example due to large-scale melting of ice in Greenland or Antarctica, could change the NAM pattern.

Based on the model estimations discussed, the following conjecture of the climate system response to external forcing emerges: external forcing (such as solar or anthropogenic) only weakly affects the climate patterns and their mean residence times but increases the probability of the occurrence of long residence times. In other words, under external influences, the changes in mean climate values, such as the global temperature, are less pronounced than the increased duration and intensity of certain climate patterns that can be associated with cold, wet conditions in some regions (floods) and warm, dry conditions (droughts) in other regions (see Table 1). Hence, studying extremes has a high priority. In the next section, we outline some ideas related to the statistics of extremes.

3. Extremes, extremes

The history of scientific studies of extremes goes back to the early 18th century. In 1709, the Swiss mathematician Nicolaus Bernoulli formulated the following problem: if n men of equal age die within t years, what is the life expectancy for the last survivor? He also suggested a method for solving this problem: multiple times, place n points at random on a line of length t and calculate the mean largest distance from the origin. This was the commencement of statistics of extremes. Nicolaus Bernoulli and Daniel Bernoulli, who is well known to physicists for his pioneering work in hydrodynamics, were among the first foreign scientists to have been invited to join the newly opened Russian Academy of Sciences in Saint Petersburg in 1725.

3.1 Stable probability distributions

The statistics of extremes appears to be completely different from the standard statistics of large numbers familiar to physicists. The familiar statistics is based on the central limit theorem, which states that a sample of a large number of random numbers has a nonrandom mean value, and the deviation from this value obeys the Gaussian distribution. But the number of extremes is usually very small, and therefore the central limit theorem is not applicable. Amazingly enough, 20th-century statisticians figured out what to do with the statistics of small numbers. There were earlier anecdotal attempts, such as the work of Polish statistician Ladislav von Bortkiewich (1868–1931), who was born in St. Petersburg and graduated from St. Petersburg University. In his paper "Klienen Zahlen" published in Leipzig in 1898, Bortkiewich tried to solve the so-called Prussian army horse-kick problem of the probability for an army officer to be killed by a horse.

The fundamentals of the statistics of small numbers were later laid out by three outstanding statisticians: Ronald Fisher (1890–1962), Leonard Tippett (1902–1985) [50], and Boris Vladimirovich Gnedenko (1912–1995) [51]. The result of their work is formulated in the Fisher–Tippett–Gnedenko (FTG) theorem, which states that if $e_1, e_2, .., e_n, ...$ are independent, identically distributed random events, then the cumulative probability of the maxima of n events $M_n = \max(e_1, e_2, ..., e_n)$ when $n \to \infty$ must obey one of the three cumulative probability distributions:

$$P_G = \exp\left[-\exp\left(-\frac{x-\mu}{\sigma}\right)\right], \tag{9}$$

$$P_F = \exp\left[-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}\right], \quad x-\mu > 0, \tag{10}$$

$$P_W = \exp\left[-\left(\frac{-x+\mu}{\sigma}\right)^{\alpha}\right], \quad x-\mu < 0, \tag{11}$$

where μ, σ, and α are parameters defining the centroids, scaling, and the rate of convergence to unity. For simplicity, we keep the same notation of the parameters, although they are different for each of the three distributions.

The first distribution had been discovered previously by the German statistician Emil Gumbel (1891–1966) in his studies of floods, the second by the brilliant French statistician Maurice Fréchet (1878–1973), and the third by the Swedish engineer Waloddi Weibull (1887–1979).

The three distributions in Eqns (9)–(11) can be combined into the so-called generalized extreme value distribution

$$G(x) = \exp\left\{-\left(1+\frac{\gamma(x-\mu)}{\sigma}\right)^{-1/\gamma}\right\}, \quad 1+\frac{\gamma(x-\mu)}{\sigma} > 0, \tag{12}$$

where $\sigma > 0$ is the scale, μ is the location, and γ is the shape parameter of the distribution.

When $\gamma \to 0$, we obtain the Gumbel cumulative distribution $\exp\{-\exp[-(x-\mu)/\sigma]\}$. If $\gamma < 0$, the right-hand tail is bounded and G becomes the reversed Weibull law. Finally, if $\gamma > 0$, then the right-hand tail decays as a power law, and G reduces to the Fréchet distribution with $\alpha = 1/\gamma$.

The base on which the proof of the nonevident FTG theorem stands is the so-called 'stability postulate' introduced by Fréchet in 1927. In short, it condenses to the following: if the distribution of random events is $P(x)$, i.e., $\text{Prob}\,(e_i < x) = P(x)$, then, due to the independence of the events,

$$\text{Prob}\,(M_n < x) = \text{Prob}\,(e_1 < x) \times \text{Prob}\,(e_2 < x) \ldots \text{Prob}\,(e_n < x) = \big[P(x)\big]^n .$$

The stability postulate states that there is a set of distributions for which the complicated function $[P(x)]^n$ reduces to a rescaled original distribution $P(x)$:

$$\big[P(x)\big]^n = P(a_n x + b_n) . \tag{13}$$

The FTG theorem identifies Eqns (9)–(11) as the only distributions satisfying condition (13). In particular, it was shown in [51] that the constants a_n and b_n are only weakly dependent on n and are equal to $a_n = 1$, $b_n = -\sigma \log(n)$ for the Gumbel distribution and to $a_n = n^{-1/\alpha}$, $b_n = \mu(1 - n^{-1/\alpha})$ for the Fréchet distribution.

The FTG theorem provides a statistical foundation for the characterization of extremes. In principle, we should just take a sample of extremes and fit it to one of the distributions described by Eqns (9)–(11) or to the generalized extreme-value distribution function (12). However, in practice, the quality of the fit is difficult to evaluate because of the scarcity of extreme events and a lack of precise mathematical techniques to do the fitting. The fitting curve depends on the selected samples of data, adjustable parameters, and the skill of the researcher. A simpler approach is fitting only the tail of the distribution. With the full distribution remaining unknown, this allows using only the data above the selected threshold. The weakest point of the tail fitting is the arbitrary selection of the threshold.

There are other practical limitations on the direct use of the FTG theorem. One of them is the severe restriction on independent extreme events. Observed extremes often arrive in clusters, i.e., close to each other (see Section 3.2), in accordance with the folk wisdom that troubles never come alone. Another limitation is that the FTG theorem applies to only the distribution of their magnitude and does not directly answer the question of how often extremes happen, i.e., what is the rate of the occurrence of extremes. For practical purposes, we need to know the distribution of time between extreme events. If timing of the events were similar to the arrival of random telephone calls, i.e., controlled by the well-known Poisson process, then the times between the arrival of subsequent events $\tau = t(i+1) - t(i)$ would be distributed exponentially as $\exp(-\tau/\tau_0)$, where τ_0 is the mean time between the events. But in general, the times of arrival of extremes do not follow this classic distribution.

Fortunately, there are simple, practical methods to determine the extreme tail of the distribution and the frequency of the occurrence of independent and correlated (clustered) extreme events. One method is called the max-spectrum [52], which was applied in [53] to the study of extreme solar events.

The method uses the scaling of the data maxima observed at different time scales and does not involve a fit to an empirically determined distribution function. The scaling approach is familiar to physicists and widely used in many physical applications, e.g., in turbulence studies (recall the Kolmogorov–Obukhov law for velocity fluctuations). The simplicity and usefulness of scaling methods is nicely illustrated in [54, 55]. Scaling can naturally be extended when new data become available, which is useful in studying extreme events due to the scarcity of data samples at the time of data analysis. It also allows interpolating the behavior of a variable beyond the limits of a given data set if there is no indication of any preferred value that could break the scaling.

The max-spectrum method uses a range of data to effectively estimate the threshold separating extreme values from typical values without any additional assumptions. It produces two exponents: one defines the high tail of the distribution function and the other characterizes the clustering of extremes in time. This method is briefly described below in application to data samples presented by time series. For technical details and mathematical proofs, the reader is referred to [52] and [56].

We consider the data time series $X(i)$ of length N, $1 \leqslant i \leqslant N$ and form nonoverlapping time blocks of length 2^j for each time scale index $j = 1, 2, 3, \ldots, [\log_2 N]$ (brackets denote the integer value), i.e., we progressively double the time scale. At each fixed scale, we calculate the data maximum within each block:

$$D(j,k) = \max_{1 \leqslant i \leqslant 2^j} X\big(2^j(k-1) + i\big) , \quad k = 1, 2, \ldots, b_j ,$$

where $b_j = [N/2^j]$ is the number of blocks (of the length 2^j) and i labels the data points within the kth block. We note that j, the log-block-size, plays the role of a time-scale parameter. We observe that the blocks of scale j are naturally nested in the blocks of the larger scale $(j+1)$.

We now average the logarithms of the maxima $D(j,k)$ over all blocks having the scale j:

$$Y(j) = \frac{1}{b_j} \sum_{k=1}^{b_j} \log_2 D(j,k) .$$

The function $Y(j)$, i.e., a set of $[\log_2 N]$ numbers, is called the 'max-spectrum' of the data. An important result, established in [52], is that if the max-spectrum is linear for sufficiently large j,

$$Y(j) \simeq \frac{j}{\alpha} + C , \tag{14}$$

where C and $\alpha > 0$ are constants, then the tail of the data distribution follows a power law with the exponent α. If the tail is not a power law, e.g., exponential, Gaussian, or lognormal, then the max-spectrum levels off at large scales. It is proved in [52] that the exponent α is the same for both statistically independent and dependent (correlated) data if the time series are stationary and have the same distribution function. The dependence, which means the clustering of the times of extreme events, affects only the intercept in Eqn (14):

$$Y(j) \simeq \frac{j}{\alpha} + C + \frac{\log_2 \theta}{\alpha} , \tag{15}$$

where the quantity θ $(0 < \theta \leqslant 1)$ is called the extremal index.

The extremal index is used in statistical studies to characterize the temporal clustering of extreme events [57, 58].

It allows representing the distribution function of maxima of n dependent events as a distribution function of the maxima of roughly $n\theta$ independent events, i.e., grouping the n dependent events into $n\theta$ independent clusters. It also allows generalizing the FTG theorem to the case of dependent (clustered) random events:

$$\mathcal{P}\left(\frac{M_n - d_n}{c_n} \leqslant x\right) = G^{\theta}(x), \ \mathcal{P}\left(\frac{M_n^* - d_n}{c_n} \leqslant x\right) = G(x)\,,$$

where M_n and M_n^* are the maxima of n dependent and independent random events deduced from the same probability distribution as the data time series X_t, $c_n > 0$ and d_n are normalization constants, and $G(x)$ is the cumulative extreme value distribution function (12). The constants $c_n > 0$, and d_n depend on the distribution function of the data. For example, if the X_t are distributed in accordance with the power law with an exponent $\alpha > 0$, then $c_n = n^{1/\alpha}$ and $d_n = 0$. We note that the constants $c_n > 0$ and d_n are the same for both dependent and independent events.

The extremal index refers only to the temporal dependence between extreme events but not between all events. The smaller the index is, the stronger the extreme events are interdependent, as is exhibited by clustering of time intervals between events. In the limit case $\theta = 1$ (independent events), we consider the onset times t_i of events exceeding a specified threshold $X_t = u$, which can be chosen, e.g., as the 90th or 95th percentile of the data distribution, or from physical considerations. Then the distribution of times between two consecutive events $\tau_i = t_i - t_{i-1}$, $i = 1, 2, \ldots$ is simply $\mathcal{P}(\tau = k) = (1-p)^{k-1}p$, where $k = 1, 2, 3, \ldots$ labels the time steps and $p = p(u)$ is the probability of the occurrence of one event in unit time. For large thresholds, p is small and this distribution is essentially exponential with the expectation value $1/p = 1/\mathcal{P}(X_t > u)$.

Equations (14) and (15) suggest a method for estimating both α and θ [52, 56]. The inverse exponent $1/\alpha$ is obtained as the slope of the line fitted to the max-spectrum of the data. The best linear fit outlines the self-similar part of the max-spectrum. It must be taken into account that in practice, as the scale j increases, fewer block-maxima $D(j,k)$ (indexed by k) are available and the variability of the max-spectrum $Y(j)$ increases. The best way to deal with this problem is to apply the method of generalized least squares, which accounts for the bias–variance trade-off [52].

Taking Eqns (14) and (15) into account, we can easily obtain estimates of the extremal index. By permuting the data with a substitute of data points (bootstrap) or by simply randomly permuting the original data time series, we can obtain a time series X_i^*, $1 \leqslant i \leqslant N$, that has the same distribution function as the original data set, but with the dependence (i.e., correlations between data points) destroyed. Repeating this operation creates a large set of pseudo time series in which the original data dependence is destroyed and the events can be viewed as nearly independent in time. For each such time series, we compute the max-spectrum $Y^*(j)$, $1 \leqslant j \leqslant [\log_2 N]$, that satisfies Eqn (14). The max-spectrum of the original data $Y(j)$ satisfies Eqn (15) with the same constant C; thus, the difference between the two spectra yields an estimate of θ:

$$\widehat{\theta}(j) = 2^{-\hat{\alpha}\left(Y^*(j) - Y(j)\right)}\,, \tag{16}$$

where $\widehat{\alpha}$ is the estimate of the tail exponent α obtained from the slope of the max-spectrum. Because there is a large sample of pseudo-independent time series, we obtain many realizations of $\widehat{\theta}(j)$ at each scale j. The median or the mean of these estimates can be taken as a point-estimator of θ on the scale j. The whole sample of estimates can be used to quantify the estimation error on each scale.

3.2 Clustering of extremes

When the time intervals between extreme events do not follow the standard exponential law for independent events, the extreme events are correlated, i.e., occur in clusters. The standard tools of time series analysis, such as the autocorrelation function of a process under consideration, cannot provide information about clustering of extremes. Two basic ideas have been introduced to tackle the problem: the cluster Poisson process [58] and an asymptotic covariance function called an 'extremogram' [59].

The mean time interval between events within a cluster depends on the threshold defining the extremes. It was shown in [58] that, with the asymptotically larger and larger thresholds, the time intervals τ_i between the extremes converge (under time rescaling) to a cluster (compound) Poisson process, which is similar to the standard Poisson process, but with a random number of events arriving clustered in time. A cluster Poisson process can be distinguished from the standard Poisson process by the extremal index $0 < \theta < 1$, the reciprocal of which defines the mean size of clusters [57].

To obtain more detailed information about the clusters, we can apply the statistical methodology called 'declustering'. The methodology employs a 'declustering threshold time' τ_c defined by the extremal index [60]. If the time interval between two extreme events is less than τ_c, these events can be grouped into a cluster, i.e., τ_c separates intra-cluster time intervals from inter-cluster time intervals.

To estimate the declustering threshold time, we consider an ordered collection of all times between consecutive extreme events,

$$\tau_1 \geqslant \tau_2 \geqslant \ldots \geqslant \tau_{m-1}\,, \tag{17}$$

and take τ_c as the $\theta \times m$th largest among them [60]. This choice of the declustering time is justified by the fact that the extremal index taken as the inverse mean of the cluster size allows estimating the number of clusters. Indeed, if our time series consists of m extreme events (i.e., m events exceed a threshold u), they are on average grouped into $\theta \times m$ clusters. We now consider the time intervals τ_i between the extreme events as they occur in a real time sequence, $i = 1, 2, 3, \ldots$. If m_1 subsequent time intervals have $\tau_i < \tau_c$, the associated extreme events constitute a cluster of size m_1. If the time τ_k exceeds τ_c, the extreme events occurring in the time step k and $k+1$ belong to different clusters, which can, in particular, be size-one clusters, i.e., single extreme events.

The declustering threshold time can be estimated in other ways. For example, one can use the observed distribution of time intervals between extremes in comparison with the exponential distribution expected for independent events.

The extremogram method used for clustering extremes generalizes the concept of the tail dependence coefficient,

$$\Lambda(\tau) = \lim_{x \to \infty} P(X(t) > x | X(t+\tau) > x)\,, \tag{18}$$

which describes the correlation between a pair of extreme data points shifted by a lag equal to τ.

The collection of values of Λ contains information about the serial dependence between extremes in the time series X. The following limit is called the extremogram [59]:

$$\rho_{AB}(\tau) = \lim_{n\to\infty} \frac{P\big(a_n^{-1}X(t) \in A, a_n^{-1}X(t+\tau) \in B\big)}{P\big(a_n^{-1}X(t) > A\big)}\,, \qquad (19)$$

where A and B are selected sets and a_n^{-1} is an increasing sequence of numbers such that $P(|X| > a_n) \propto n^{-1}$. If we choose $A = B = (1, \infty)$, the extremogram $\rho_{AB}(\tau)$ simply becomes the probability λ_k of observing another extreme event after the time k an extreme event has already been observed, i.e.,

$$\lambda_k = P\big(X_k > u | X_0 > u\big)\,, \quad k = 1, 2, \ldots\,, \qquad (20)$$

for a threshold u. If the X_k were time independent, this conditional probability would equal the unconditional probability $P(X_k > u)$. Therefore, the X_k are statistically dependent on time if λ_k is significantly different from $P(X_k > u)$. The parameter λ_k can be estimated with the following empirical statistics:

$$\widehat{\lambda}_k = \frac{\big[\sum_{j=1}^{n-k} I(X_{j+k} > u, X_j > u)\big]/(n-k)}{\big[\sum_{j=1}^{n} I(X_j > u)\big]/n}\,, \qquad (21)$$

where $I(A)$ equals unity if the event A occurs and zero otherwise.

In the next two subsections, we illustrate the application of the described statistical methods to real data.

3.3 Precipitation extremes

We first consider the application of the max-spectrum method to precipitation extremes over the Pacific Ocean. Rainfall over the oceans determines surface layer stratification, the ocean freshwater balance, and ocean mixing. Its variations are often determined by El Niño. Here, we consider the rainfall over the so-called Niño 3.4 region located in the middle of the Pacific Ocean (5°S–5°N; 170°W–120°W) using satellite data obtained from the Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E) sensor on NASA's Aqua satellite. The data can be downloaded from http://nsidc.org/data/amsre/. The AMSR-E instrument provides measurements of land, ocean, and atmospheric parameters for the investigation of global water and energy cycles, including precipitation rates, sea surface temperatures, sea ice concentration, snow water equivalent, soil moisture, surface wetness, wind speeds, atmospheric cloud water, and water vapor.

The rainfall data used were generated daily in the period from June 2002 to June 2010 and averaged over the Niño 3.4 region (Fig. 5). There is a slight influence of the annual cycle seen in Fig. 6. It was filtered out in the search for the extremes, although it does not produce a substantial effect on the max-spectrum. The cumulative distribution function of this time series and the max-spectrum are shown in Fig. 6. The spectral index is found to be $\alpha = 2.2 \pm 0.1$. The extremal index is quite small (< 0.2), indicating substantial clustering (more than 5 events in a cluster on average) of extreme rain events in this region.

3.4 Space climate extremes

Another example is the application of the max-spectrum method to the study of extreme space climate events [53].

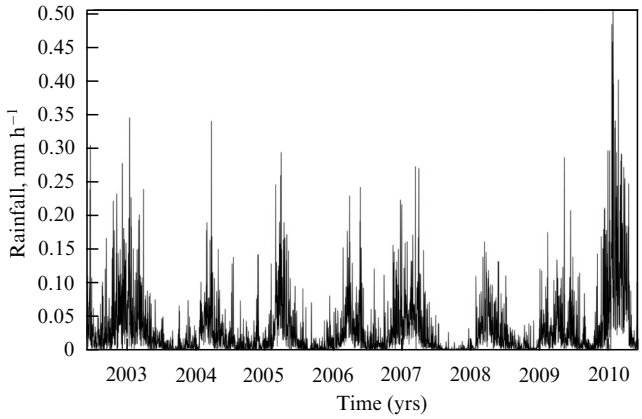

Figure 5. Rainfall record over the Niño 3.4 region in the Pacific Ocean.

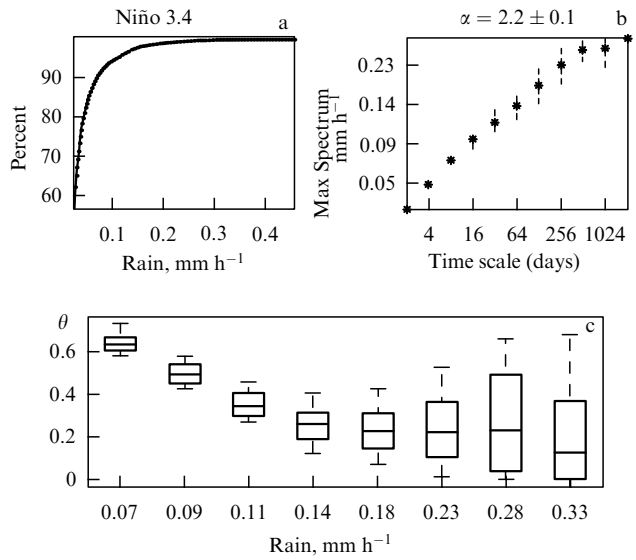

Figure 6. (a) Cumulative distribution function of the rainfall data shown in Fig. 5. (b) Max-spectrum of the rainfall in the Niño 3.4 region. (c) Extremal index as a function of rainfall percentiles.

The currently popular term "space weather" refers to severe disturbances in Earth's upper atmosphere and in the near-Earth space environment that are driven by solar activity [61]. Similarly to the meteorological terminology, external disturbances in Earth's atmosphere and its space environment on a long time scale (e.g., a year or the 11-year solar cycle) are referred to as space climate [62]. The main space disturbances come from the solar plasma and particle ejection and galactic cosmic rays [63]. Space climate estimates are required for the design of space missions. Manifestations of these disturbances are seen in beautiful natural extreme events, the auroras. Whether the aurora is seen as moving white and green bands or a diffuse red light depends on observer's position on Earth. The aurora takes place in an annulus around the geomagnetic poles [64]. Within 20 degrees of latitude of the equator, the sky is red during an aurora and seeing it would be a once in a lifetime experience. In the far north, the light is most likely green, and can be seen during most dark nights. At mid-latitudes, the phenomenon might be seen once or twice a year and the moving light would be either whitish or greenish or a vivid red. Documentation of auroras in Europe and the Orient cover the time period as far back as

the seventh century BC. These observations have been gathered in a catalog [65] that has been used in scientific research on solar variability and its influence on Earth's climate.

During strong auroras, ionospheric currents can be hazardous by inducing damaging electrical currents in electric power grids, as happened during the great geomagnetic storm of March 1989, when the collapse of northeastern Canada's Hydro-Quebec power grid left millions of people without electricity for up to 9 hours. Such currents also contribute to the corrosion of oil and gas pipelines [66]. Space-driven ionospheric density disturbances interfere with high-frequency radio communications along airplane routes, requiring aircraft to be diverted to lower latitudes, and affecting navigation signals from Global Positioning System (GPS) satellites [61]. Bursts of energetic particles and radiation belt enhancements during strong space weather events can cause operational anomalies and damage the electronic equipment on spacecraft [67].

The main solar plasma events are initiated by disturbances due to the sudden release of large amounts ($> 10^{16}$ g) of solar plasma from the solar corona into the solar wind. These events, which are called coronal mass ejections (CMEs), are generated by dynamo activity in the solar interior that produces magnetic fields permeating into the solar corona.

A part of this field erupts as a result of an instability or loss-of-equilibrium process. Once a CME is underway, a whole host of additional processes are triggered, including magnetic reconnection, shock formation, and particle acceleration [68, 69]. Coronal mass ejections propagate from the Sun through the interplanetary space, some of them toward Earth. They vary widely in their speeds. When viewed near the Sun, some are relatively slow (< 200 km s^{-1}) and others have very high speeds exceeding 2,500 km s^{-1} [70]. These high-speed (fast) CMEs are most interesting in the context of space climate extremes. The fast CMEs and the shocks they generate in the solar wind are directly responsible for solar energetic particle (SEP) events [71]. The interaction of a strong southward magnetic field associated with fast CMEs with Earth's magnetic field causes major geomagnetic storms [72, 73]. The fast CMEs are especially geoeffective, because this interaction is governed by the induced electric field, which is a product of the speed and magnetic field of the CME propagating in the solar wind.

Coronal mass ejections are associated with active (sunspot) regions, which appear on the surface of the Sun. The frequency of occurrence of active regions is regulated by the solar cycle. Observations have shown that active regions have a tendency to cluster, i.e., new active regions preferably emerge in the vicinity of old ones [74]. The clusters can live for as long as six months, and there are indications that the fastest CMEs originate mainly from them [75].

The plane of the sky speed of a CME propagating through the solar corona has been measured by coronagraphs carried on spacecraft, such as the CMEs measured by the Large Angle and Spectrometric Coronagraph Experiment on board the Solar and Heliospheric Observatory (LASCO SOHO). They are listed in a catalog developed in cooperation with the Naval Research Laboratory and the Solar Data Analysis Center at the Goddard Space Flight Center and at the Center for Solar Physics and Space Weather at the Catholic University of America. The entries begin in January 1996. This data record was used in [53]. To avoid the obvious large nonstationarity due to the solar cycle dependence, the data set

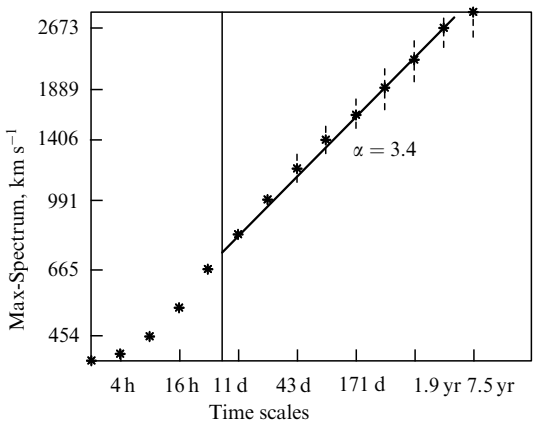

Figure 7. Max-spectrum of CME speeds at progressively increasing time scales. The error bars are estimated using the generalized regression [52] and correspond to 95% confidence intervals. The $\log_2$ units for $Y(j)$ are converted into km s^{-1} and the scales j into time units according to the formula 2^j. The vertical line segment indicates the starting scale selected for the evaluation of α. The speed on this scale can be interpreted as the beginning of the distribution function tail, which defines fast CMEs, shown on the right of the vertical line.

was limited to the high-activity part of solar cycle 23 (from January 1999 to December 2006), resulting in 9.408 CMEs. The CME speeds used in this paper are given in the catalog as obtained by a 2nd-order polynomial fit to the time–height measurements during the CME propagation through the solar corona. The data input to the max-spectrum method includes all CMEs without preselection of those with high speeds. Notably, the threshold speed at which the tail of the distribution function begins was not preselected but estimated from the onset of the power tail.

The resulting max-spectrum of the CME speeds is shown in Fig. 7. The best fit to the slope gives evidence that the cumulative distribution function of the CME speeds has a Fréchet-type power-law tail with the exponent $\alpha = 3.4$. The lower boundary of the linear portion of the max-spectrum identifies the onset of the power-law tail, i.e., the corresponding speed threshold, and the onset of the self-similar range. This gives a meaningful definition of the 'fast' CMEs loosely used by solar scientists previously. Bearing in mind the analogy with the standard, self-similar cascade process in turbulence, which is fully defined by a Kolmogorov-type spectral index, it is reasonable to conjecture that the physical process leading to fast CME production is the same from about 700 km s^{-1} to the highest velocities in the data set.

Figure 8 shows the extremal index estimated by the max-spectrum method. It also provides 'confidence intervals' as illustrated with the histogram on the lower panel in Fig. 6 plotted for the thresholds 1,000–2,300 km s^{-1}. The resulting empirical 95% confidence interval for θ is from 0.33 to 0.60 with the mid-point $\theta = 0.49 \approx 0.5$. The θ in the range 0.3–0.6 with the mean 0.5 can be taken as an estimate of the extremal index. The inverse value of the index gives the estimate of the average cluster size equal to 2–3; on average, therefore, the appearance of a fast CME is followed by one or two other fast CMEs.

Application of the declustering methodology allows finding the number and content of clusters of fast CMEs. As an example, we consider the threshold $U = 1000$ km s^{-1} and $\theta = 0.5$. With this threshold, we have $n = 586$ fast CMEs with the 'declustering time' $\tau_c = 42$ h. The maximum number of

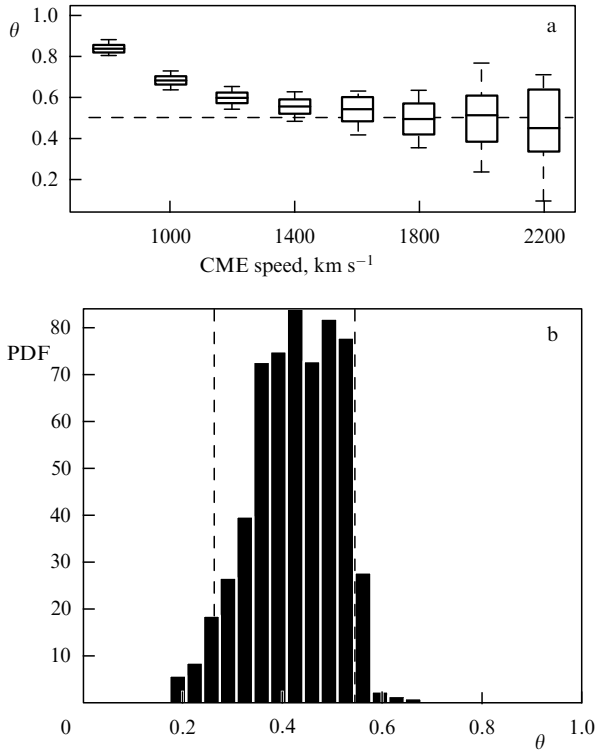

Figure 8. (a) Extremal index obtained by the max-spectrum method. The box plots for each time scale are obtained from 100 independent realizations of the randomized $\widehat{\theta}$. The central mark in a box is the median, the box edges are the 10th and 90th percentiles, and whiskers extend to the most extreme data points. (b) Histogram of the $\widehat{\theta}$ for speed thresholds from 1,000 to 2,300 km s^{-1}.

fast CMEs in a cluster found from the LASCO catalog used here is 9. The average duration time between CMEs within a cluster is 18 hours, with the standard error of 2 hours. Table 2 provides more detailed information about the probability and the corresponding duration of clusters as a function of their size (the number of CMEs in a cluster). The first column (size) lists the number of CMEs in the cluster. The second and the third columns give the number of clusters of this size and the total number of CMEs in these clusters. The fourth column provides estimates and the standard error (in parenthesis) of the probabilities that a cluster of the corresponding size is recorded. The last column lists the expected mean durations of the clusters (with the standard error in parentheses). We see that about 30% of the fast CMEs are single; the rest appear in clusters of different sizes. There is a statistically significant portion of clusters (about 35%) with five or more members, with the average duration of about 110 hours. Similar estimates can be made using different threshold speeds.

Table 2. Example of predictive statistics for clusters of CMEs with a speed exceeding 1000 km s^{-1}.

Size	No. of clusters	No. of CMEs in clusters	Probability	Duration, hr
1	177	177	0.61 (0.03)	—
2	53	106	0.18 (0.02)	20.1 (1.7)
3	18	54	0.06 (0.01)	39.7 (3.8)
4	20	80	0.07 (0.01)	56.8 (4.5)
5	7	35	0.02 (0.01)	70 (7.2)
> 5	17	169	0.06 (0.01)	107.7 (10.6)

4. Summary and reminiscences

The main conclusion of this review is that climate changes caused by humans do not so much affect the mean state of Earth's climate, such as its global temperature, but greatly increase the number of extreme events, such as floods and droughts. This may add to our understanding of climate change. The application of advanced statistical methods to space climate may be of interest to scientists and engineers involved in space exploration.

Reflecting on the quote ascribed to Mark Twain, "Go to heaven for the climate and hell for the company," [78] I also hope that the review exposes us to another side of contemporary scientific knowledge. In old times, physicists had no need for most statistics or details of noise. They lived in the realm of the law of large numbers; knowing how to calculate means and standard deviations was sufficient to take care of experimental errors. Nowadays, they have discovered that rare random events can play a critical role in physical processes and they appreciate the importance of the law of small numbers, which was actually formulated by outstanding mathematicians and statisticians long ago.

This review, written by a student of Zeldovich's school of physics, tries to reflect on this striking revolution in our minds by discussing the role played by randomness in chaotic systems, such as Earth and space climates, in response to external forcing. In the style of Zeldovich's school, it uses general concepts and simple models. But being his undergraduate student and PhD student and having worked with Zel'dovich for 20 years, I noticed that he often based his simple, toy model presentations on deep mathematical theorems. Once, I witnessed how he surprised a mathematician at the Keldysh Institute of Applied Mathematics, where we worked, with exact knowledge of some topological invariants. Another time, I saw him proudly exiting his office after giving a lesson on the correct use of statistics in a nontrivial case to a famous experimental physicist. I believe that other students and colleagues of Zeldovich may have other examples of the deep mathematical knowledge of this great physicist.

Acknowledgments

I thank the referee for helpful, critical comments and M A Shelyakhovskaya and M S Aksent'eva for editing and providing useful references to papers related to this review. I would also like to thank Stilian Stoev, who introduced me to statistics of extremes, and Joan Feynman, who indicated to me the best ways of handling data. The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

1. Zeldovich Ya B, Ruzmaikin A A, Sokoloff D D *The Almighty Chance* (Lecture Notes in Physics, Vol. 20) (Singapore: World Scientific, 1990)
2. Landau L D, Lifshitz E M *Fluid Mechanics* (London: Pergamon Press, 1959); Translated from Russian: *Mekhanika Sploshnykh Sred* (Moscow: GITTL, 1953)

3. Ruelle D, Takens F *Commun. Math. Phys.* **20** 167 (1971)
4. Zel'dovich Ya B, Molchanov S A, Ruzmaikin A A, Sokolov D D *Sov. Phys. JETP* **62** 1188 (1985); *Zh. Eksp. Teor. Fiz.* **89** 2061 (1985)
5. Zeldovich Ya B, Molchanov S A, Ruzmaikin A A, Sokoloff D D *Sov. Sci. Rev. C Math. Phys. Rev.* **7** 1 (1988)
6. IPCC — Intergovernmental Panel on Climate Change, http://www.ipcc.ch/ (2013), publications
7. Byalko A V *Phys. Usp.* **55** 103 (2012); *Usp. Fiz. Nauk* **182** 111 (2012)
8. Demchenko P F, Kislov A V *Stokhasticheskaya Dinamika Prirodnykh Ob'ektov: Brounovskoe Dvizhenie i Geofizicheskie Prilozheniya* (Stochastic Dynamics of Natural Objects: Brownian Motion and Geophysical Applications) (Moscow: GEOS, 2010)
9. Golitsyn G S *Statistika i Dinamika Prirodnykh Protsessov i Yavlenii: Metody, Instrumentarii, Rezul'taty* (Statistics and Dynamics of Natural Processes and Phenomena: Methods, Tools, Results) (Moscow: KRASAND, 2013)
10. Lovejoy S "What is climate?" *EOS Trans. Am. Geophys. Union* **94** (1) 1 (2013)
11. Nyquist H *Phys. Rev.* **32** 110 (1928)
12. Callen H B, Welton T A *Phys. Rev.* **83** 34 (1951)
13. Kubo R *J. Phys. Soc. Jpn.* **12** 570 (1957)
14. Kraichnan R H *Phys. Rev.* **113** 1181 (1959)
15. Landau L D, Lifshitz E M, Pitaevskii L P *Electrodynamics of Continuous Media* 2nd ed. (Oxford: Pergamon Press, 1984); Translated from Russian: *Elektrodinamika Sploshnykh Sred* 2nd ed. (Moscow: Nauka, 1982)
16. Leith C E *J. Atmos. Sci.* **32** 2022 (1975)
17. Gritsun A, Branstator G *J. Atmos. Sci.* **64** 2558 (2007)
18. Dymnikov V P, Gritsun A S *Nonlin. Processes Geophys.* **8** 201 (2001)
19. Dymnikov V P, Gritsun A S *Izv. Atmos. Ocean. Phys.* **41** 266 (2005); *Izv. Akad. Nauk SSSR. Fiz. Atm. Okeana* **41** 294 (2005)
20. Gritsun A S *Russ. J. Numer. Anal. Math. Model.* **16** 115 (2001)
21. Pitaevskii L P *Phys. Usp.* **54** 625 (2011); *Usp. Fiz. Nauk* **181** 647 (2011)
22. Bochkov G N, Kuzovlev Yu E *Sov. Phys. JETP* **49** 543 (1979); *Zh. Eksp. Teor. Fiz.* **76** 1071 (1979)
23. Jarzynski C *Phys. Rev. Lett.* **78** 2690 (1997)
24. Lorenz E N *J. Atmos. Sci.* **20** 130 (1963)
25. Khatiwala S, Shaw B E, Cane M A *Geophys. Res. Lett.* **28** 2633 (2001)
26. Kramers H A *Physica* **7** 284 (1940)
27. Walker G *Quart. J. R. Meteorol. Soc.* **54** 79 (1928)
28. Obukhov A M *Usp. Mat. Nauk* **2** (2) 196 (1947)
29. Quadrelli R, Wallace J M *J. Climate* **17** 3728 (2004)
30. Hurrell J W *Science* **269** 676 (1995)
31. Thompson D W J, Wallace J M *Geophys. Res. Lett.* **25** 1297 (1998)
32. Thompson D W J, Wallace J M *Science* **293** 85 (2001)
33. Honda M, Inoue J, Yamane S *Geophys. Res. Lett.* **36** L08707 (2009)
34. Petoukhov V, Semenov V A *J. Geophys. Res. Atmos.* **115** (D21) (2010), DOI: 10.1029/2009JD013568
35. Wallace J M, Zhang Y, Renwick J A *Science* **270** 780 (1995)
36. Mantua N J et al. *Bull. Amer. Meteor. Soc.* **78** 1069 (1997)
37. Philander S G *El Niño, La Niña, and the Southern Oscillation* (San Diego: Academic Press, 1990)
38. Baldwin M P et al. *Rev. Geophys.* **39** 179 (2001)
39. Thompson D W J, Baldwin M P, Wallace J M *J. Climate* **15** 1421 (2002)
40. Rossby C G "The scientific basis of modern meteorology", in *Climate and Man. Yearbook of Agriculture* (Washington: United States Department of Agriculture, 1941) p. 599
41. Charney J G, Drazin P G *J. Geophys. Res.* **66** 83 (1961)
42. Limpasuvan V, Hartmann D L *J. Climate* **13** 4414 (2000)
43. Ruzmaikin A, Lawrence J, Cadavid C *J. Climate* **16** 1593 (2003)
44. Ruzmaikin A, Feynman J, Yung Y L *J. Geophys. Res. Atmos.* **111** (D21) (2006), DOI: 10.1029/2006JD007462
45. Ruzmaikin A, Cadavid A C, Lawrence J K *J. Atmos. Solar-Terr. Phys.* **68** 1311 (2006)
46. Corti S, Molteni F, Palmer T N *Nature* **398** 799 (1999)
47. Palmer T N *J. Climate* **12** 575 (1999)
48. Ruzmaikin A et al. *Geophys. Res. Lett.* **31** L12201 (2004)
49. Hall A et al. *J. Climate* **18** 1315 (2005)
50. Fisher R A, Tippett L H C *Proc. Camb. Philos. Soc.* **24** 180 (1928)
51. Gnedenko B *Ann. Math.* **44** 423 (1943)
52. Stoev S A, Michailidis G, Taqqu M S *IEEE Trans. Inform. Theory* **57** 1615 (2011)
53. Ruzmaikin A, Feynman J, Stoev S A *J. Geophys. Res. Space Phys.* **116** (A4) (2011), DOI: 10.1029/2010JA016247
54. Golitsyn G S *Phys. Usp.* **51** 723 (2008); *Usp. Fiz. Nauk* **178** 753 (2008)
55. Golitsyn G S *Dokl. Phys.* **49** 501 (2004) *Dokl. Ross. Akad. Nauk* **398** 177 (2004)
56. Hamidieh K, Stoev S, Michailidis G *J. Comput. Graph. Stat.* **18** 731 (2009)
57. Leadbetter M R, Lindgren G, Rootzén H *Extremes and Related Properties of Random Sequences and Processes* (New York: Springer-Verlag, 1983)
58. Hsing T, Hüsler J, Leadbetter M R *Probab. Theory Related Fields* **78** 97 (1988)
59. Davis R A, Mikosch T *Bernoulli* **15** 977 (2009)
60. Ferro C A T, Segers J *J.R. Statist. Soc. B* **65** 545 (2003)
61. *Severe Space Weather Events – Understanding Societal and Economic Impacts* (Workshop Report) (Washington, D.C.: The National Academies Press, 2008); http://www.nap.edu/catalog/12507.html
62. Mursula K, Usoskin I G, Maris G *Adv. Space Res.* **40** 885 (2007)
63. Dorman L I *Phys. Usp.* **53** 496 (2010); *Usp. Fiz. Nauk* **180** 519 (2010)
64. Feldstein Y I *J. Geophys. Res.* **78** 1210 (1973)
65. Siscoe G L *Rev. Geophys. Space Phys.* **18** 647 (1980)
66. Boteler D, Marti L "Space weather situational awareness for power systems", in *39th COSPAR Scientific Assembly, 14 – 22 July 2012, Mysore, India*, Abstract E2.1-4-12, p. 225
67. Feynman J, Gabriel S B *J. Geophys. Res. Space Phys.* **105** 10543 (2000)
68. Forbes T G et al. *Space Sci. Rev.* **123** 251 (2006)
69. Amari T et al. *Astrophys. J.* **742** L27 (2011)
70. Kahler S *Rev. Geophys.* **25** 663 (1987)
71. Reames D V *Space Sci. Rev.* **90** 413 (1999)
72. Hirshberg J, Colburn D S *Planet. Space Sci.* **17** 1183 (1969)
73. Gopalswamy N *J. Atmos. Solar-Terr. Phys.* **70** 2078 (2008)
74. Gaizauskas V et al. *Astrophys. J.* **265** 1056 (1983)
75. Ruzmaikin A, Feynman J, in *Multiscale Phenomena in Space Plasmas. Proc. of the 1998 Cambridge Workshop* (Physics of Space Plasmas, Eds T Chang, J Jasperse) (Cambridge, Mass.: MIT Center for Theor. Geo/Plasma Physics, 1998) p. 295
76. Ruzmaikin A, Feynman J *J. Geophys. Res. Atmos.* **107** (D14) ACL7-1 (2002)
77. Ruzmaikin A "Effects of solar variability on the Earth's climate patterns" *Adv. Space Res.* **40** 1146 (2007)
78. Mark Twain Quotations, Newspaper Collections, and Related Resources (Heaven), http://www.twainquotes.com/Heaven.html