



CENTRE DE RECERCA MATEMÀTICA

This is a preprint of: *Data-driven prediction of thresholded time series
of rainfall and self-organized criticality models*
Journal Information: *Physical Review E*,
Author(s): A. Deluca, N. R. Moloney, and A. Corral.
Volume, pages: 91 1-8, DOI:[10.1103/PhysRevE.91.052808]

Data-driven prediction of thresholded time series of rainfall and self-organized criticality modelsAnna Deluca,^{1,*} Nicholas R. Moloney,² and Álvaro Corral^{3,4}¹Max-Planck-Institut für Physik Komplexer Systeme, Nöthnitzer Strasse 38, D-01187 Dresden, Germany²London Mathematical Laboratory, 14 Buckingham Street, London WC2N 6DF, United Kingdom³Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, E-08193 Barcelona, Spain⁴Departament de Matemàtiques, Facultat de Ciències, Universitat Autònoma de Barcelona, E-08193 Barcelona, Spain

(Received 1 October 2014; published 15 May 2015)

We study the occurrence of events, subject to threshold, in a representative self-organized criticality (SOC) sandpile model and in high-resolution rainfall data. The predictability in both systems is analyzed by means of a decision variable sensitive to event clustering, and the quality of the predictions is evaluated by the receiver operating characteristic (ROC) method. In the case of the SOC sandpile model, the scaling of quiet-time distributions with increasing threshold leads to increased predictability of extreme events. A scaling theory allows us to understand all the details of the prediction procedure and to extrapolate the shape of the ROC curves for the most extreme events. For rainfall data, the quiet-time distributions do not scale for high thresholds, which means that the corresponding ROC curves cannot be straightforwardly related to those for lower thresholds. In this way, ROC curves are useful for highlighting differences in predictability of extreme events between toy models and real-world phenomena.

DOI: [10.1103/PhysRevE.91.052808](https://doi.org/10.1103/PhysRevE.91.052808)

PACS number(s): 05.65.+b

I. INTRODUCTION

Many atmospheric processes related to precipitation give rise to structures and correlations across long ranges in space and time, which are the result of the coupling between several nonlinear mechanisms with different spatial and temporal characteristic scales [1–4]. Despite the diversity of individual rain events, an array of statistical measures presents strong statistical regularities [5–10], giving support to the hypothesis that atmospheric convection and precipitation may be a real-world example of self-organized criticality (SOC) [11]. Whereas the usual approach in meteorology and hydrology consists of looking at the occurrence of rain in fixed time intervals (days, months, etc.), “episodic” rain events, similar to avalanches in SOC sandpile models, can be defined by integrating the rain rate over very short time periods. This led to the claim that rain-event sizes are power-law (i.e., scale-free) distributed, at least in the unique site studied [5], in agreement with the SOC hypothesis. Nevertheless, a power-law distribution of this observable is not sufficient evidence for SOC dynamics, as there are many alternative mechanisms that give rise to such behavior (see, for example, Refs. [12,13]).

Further support for the SOC hypothesis was given by Peters and Neelin [6], who found, for rain data over the tropical oceans, (i) a relation between satellite estimates of rain rate and water vapor compatible with a phase transition, with large parts of the troposphere in a convectively active phase, and (ii) that the system was close to the transition point most of the time. This constitutes genuine evidence for SOC. These authors also related this SOC behavior to the concept of atmospheric quasiequilibrium [14], which argues that, since driven processes are generally slow compared to convection, the system should typically be in a far-from-equilibrium statistically stationary state, where driving and dissipation are in balance.

Coming back to local-event-size distributions, recent works have shown that these are indicative of universality as expected in the SOC framework [8–10]. The resulting rain-event-size distributions for several sites distributed worldwide are well approximated by power laws of similar exponents over broad fitting ranges, with differences only in the large-scale cutoffs of the distributions. These differences are attributed to finite-size effects, pointing to distinct system capacities in different places.

This SOC framework raises the question of the possible implications of the critical behavior of atmospheric convection for the prediction of rainfall, which remain unclear, due to the fact that predictability in SOC systems is still not well understood. In particular, prediction studies often focus on the behavior of the most hazardous episodes, i.e., extreme events of rainfall in our case.

In a more general context, Kantz [15] classifies the different scientific approaches to extreme events, such as *extreme value statistics* for the robust estimation of the tails; *data-driven predictions*, which employ conditional probabilities and temporal correlations; *simplistic physical models* to investigate the mechanisms and dynamics from which extreme events emerge; and *detailed disciplinary investigations* of extreme events in a particular system.

Our approach in this paper combines simplistic physical models (sandpile models) and data-driven prediction in order to gain insight into rainfall occurrence, as well as a deeper understanding of SOC phenomena. Because a direct connection between sandpilelike models and rainfall has not yet been established, our purpose is not the modeling of the latter by the former, but rather the comparison of the dynamics of the two systems.

In the next section, we introduce the rain database and the SOC sandpile model used for comparison, and we define the extreme events of interest and present distributions of the quiet times that separate them. In Sec. III, we explain the prediction procedure, which makes use of the hazard function, and evaluate the predictability via the receiver operating characteristics

*adeluca@pks.mpg.de

(ROC) method. Section IV explores the prediction procedure analytically.

II. RAINFALL DATA AND SOC MODELS

We analyze high-resolution local rain intensities across different climates from the Atmospheric Radiation Measurement (ARM) database, and simulated data from the Manna sandpile model. The rain data consist of point-location measurements from sites around the world, with 1-min temporal resolution spanning about 8 months to 7 years, depending on the site. For more details, see Refs. [8,10]. The signal directly measured is the rain rate, giving the “instantaneous” depth of precipitation at a point location every minute, with a resolution of 0.001 mm/h. In fact, anything below 0.2 mm/h should be considered zero (as pointed out in the handbook of the measuring device [16]). This is due to the difficulties in distinguishing rainfall from other phenomena such as mist. A rain “event” is defined as a sequence of rates exceeding a threshold r_c . To be precise, the event starts when the threshold is exceeded for the first time and ends when the rate subsequently falls below threshold. Previous works on rainfall considered only the minimum possible threshold. However, this is not a fundamental physical parameter but rather an unavoidable limitation of the observational procedure. In this paper we are interested in exploring a range of thresholds, starting at the minimum reliable value of 0.2 mm/h and working up to extreme thresholds.

As a prototypical SOC system we investigate the Abelian Manna sandpile model [17] in two dimensions. The model is defined on a two-dimensional square lattice of size L^2 with open boundaries. Each site $i = 1, \dots, L^2$ contains a discrete number of particles z_i . The rules of the dynamics are as follows.

(1) Driving: A particle is added to a randomly chosen site i , $z_i \rightarrow z_i + 1$.

(2) Toppling: If $z_i > 1$ at any site i , the site is relaxed, $z_i \rightarrow z_i - 2$, and the two particles are distributed among randomly and independently chosen nearest neighbors (NN), $z_{NN} \rightarrow z_{NN} + 1$ (with possibly the same site chosen twice). Multiple topplings are performed in parallel. This rule is iterated (if necessary) and each update defines one step in the “fast” avalanche time scale.

(3) Dissipation: Particles that are distributed from edge sites beyond the boundaries of the lattice are removed from the system.

In the following, we drive the system “infinitely slowly” (according to the usual protocol) by adding a new particle to the system only if all sites are relaxed. This effects a time-scale separation between driving and toppling. (For completeness, we also check the robustness of our results when this time-scale separation is broken, as in Refs. [18,19].) Once the system has reached the statistically stationary state (i.e., the number of particles in the pile stabilizes, on average), the toppling activity n , which counts the number of toppling sites at each avalanche time step, is recorded. As with rain data, we make use of an activity threshold n_c and define events as consecutive sequences of $n > n_c$. The events that survive thresholding are broadly distributed over many decades, in the same way as the usual (zero-thresholded) avalanches.

A key observable in our analysis (for both the rain data and the sandpile model) is the quiet time, τ , defined as the time the rate or activity signal spends below threshold. For the Bak-Tang-Wiesenfeld (BTW) model and other SOC models, it is well known that, if the system is slowly driven and the quiet time is measured in the slow time scale, the quiet-time distribution is approximately exponential when $n_c = 0$, and therefore the instantaneous avalanches occur in the manner of a Poisson process [20], which, by definition, has no memory. Paczuski *et al.* [19] showed that this continues to hold even when the time-scale separation is broken (and one measures time on the avalanche time scale), provided $n_c = 0$.

At first sight, this observation appears to rule out the application of SOC models to real-world systems, such as solar flares or earthquakes, for which the quiet-time distributions are not exponential [21–23]. However, as shown in Ref. [19] for the BTW model, when a nonzero threshold is applied ($n_c > 0$), the quiet-time probability density function (PDF) is a decaying power law with exponent $\beta \approx 1.67$, independent of system size. This nonexponential PDF reflects the existence of clustering between events, in agreement with many natural phenomena. In the real world, thresholds are often imposed, e.g., because of the limited resolution of measuring devices, and therefore it is more realistic to consider nonzero thresholds for practical purposes.

Our simulations lead to very similar results for the Manna model with infinitely slow driving. Due to the time-scale separation, there is no well-defined way of measuring intra- and interavalanche quiet times simultaneously. As in model A of Ref. [19], we therefore restrict our attention to intra-avalanche quiet times measured on the fast time scale (i.e., parallel update steps). Figure 1(a) shows, for a system of size $L = 1024$, quiet-time PDFs $P_q(\tau)$, where q denotes the thresholds as quantiles of the activity distribution. For any $n_c > 0$ (i.e., $q > 0$) the PDFs are clearly compatible with a power law with negative exponent $\beta = 1.67$, up until some fast-decaying cutoff. Although the cutoff function moves out with increasing threshold, a convincing data collapse onto a threshold-independent scaling function is possible, as shown in Fig. 1(b). In other words, a scaling law of the form

$$P_q(\tau) \simeq \frac{1}{\tau^\beta} f_0(\tau/a) \quad (1)$$

holds, where a is a scale parameter that depends on q , and f_0 is a scaling function [shown in Fig. 1(c)]. If $1 < \beta < 2$, then the first two moments $\langle \tau \rangle \propto a^{2-\beta}$ and $\langle \tau^2 \rangle \propto a^{3-\beta}$, from which it follows that $a \propto \langle \tau^2 \rangle / \langle \tau \rangle$ and $a^\beta \propto \langle \tau^2 \rangle^2 / \langle \tau \rangle^3$, thereby justifying the rescaling of the axes in Figs. 1(b) and 1(c) [8,24]. This rescaling is more appropriate than that used in Ref. [19]; see also Ref. [25].

Similar behavior is observed when time-scale separation is broken by adding a particle to the system every 100 avalanche steps (whether or not all sites are relaxed). However, the scaling is not so clean: a bump appears between the power law and the subsequent faster decay (not shown). This arises from the mixture of quiet times within avalanches (corresponding to the previous case, model A) and exponential quiet times between avalanches.

Turning to rainfall data, Fig. 2 plots quiet-time PDFs $P_q(\tau)$ for different rate thresholds as measured at the ARM site on

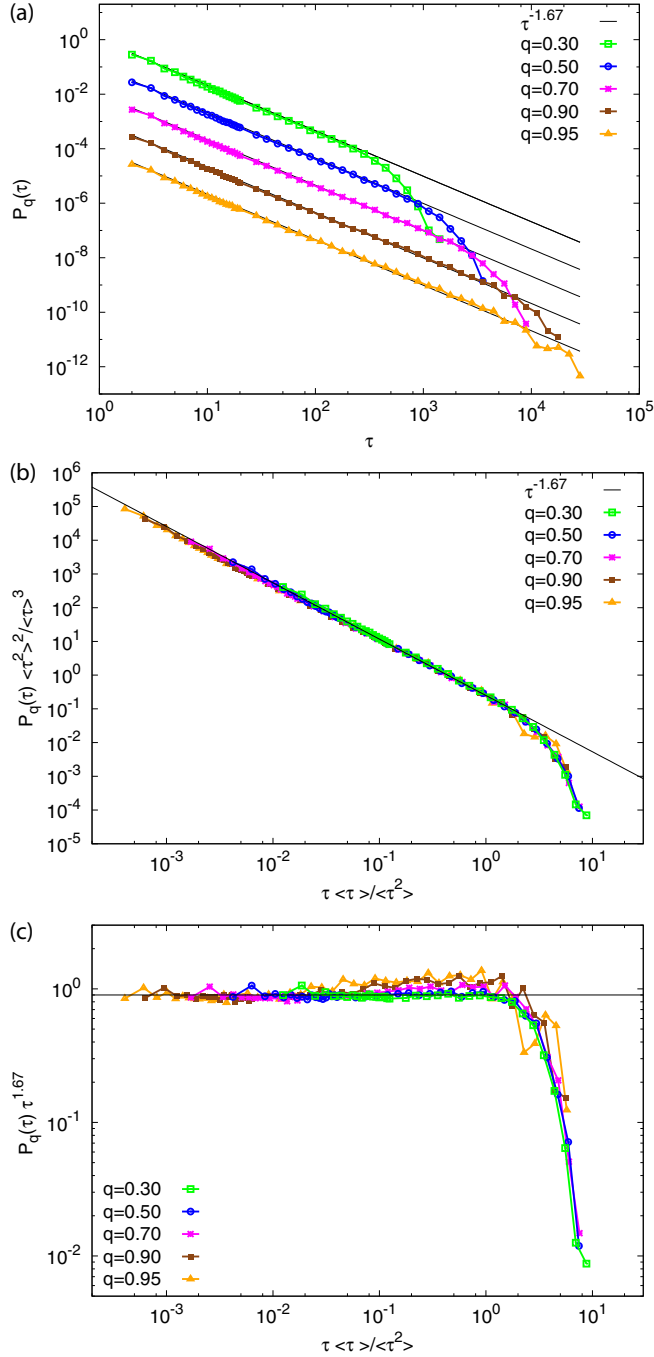


FIG. 1. (Color online) (a) Quiet-time PDFs for the Manna model ($L = 1024$) for different activity thresholds n_c characterized by their quantiles q . Statistics are collected over 1×10^7 particle additions. PDFs for different q are displaced vertically for clarity. The solid black lines have slope -1.67 . (b) Data collapse of the suitably rescaled quiet-time PDFs. (c) Alternative data collapse in order to visualize the scaling function f_0 . In all plots, quantile thresholds q correspond to activity thresholds $n_c = 51, 123, 255, 584$, and 806 .

Manus island. Analyses for rate thresholds below 0.2 mm/h are disregarded owing to the detection limit on measurement devices. Thus, even for $q = 0$ an implicit threshold exists. Contrary to the Manna model, the PDFs are approximately independent of q up until about $q = 0.70$, following a power-

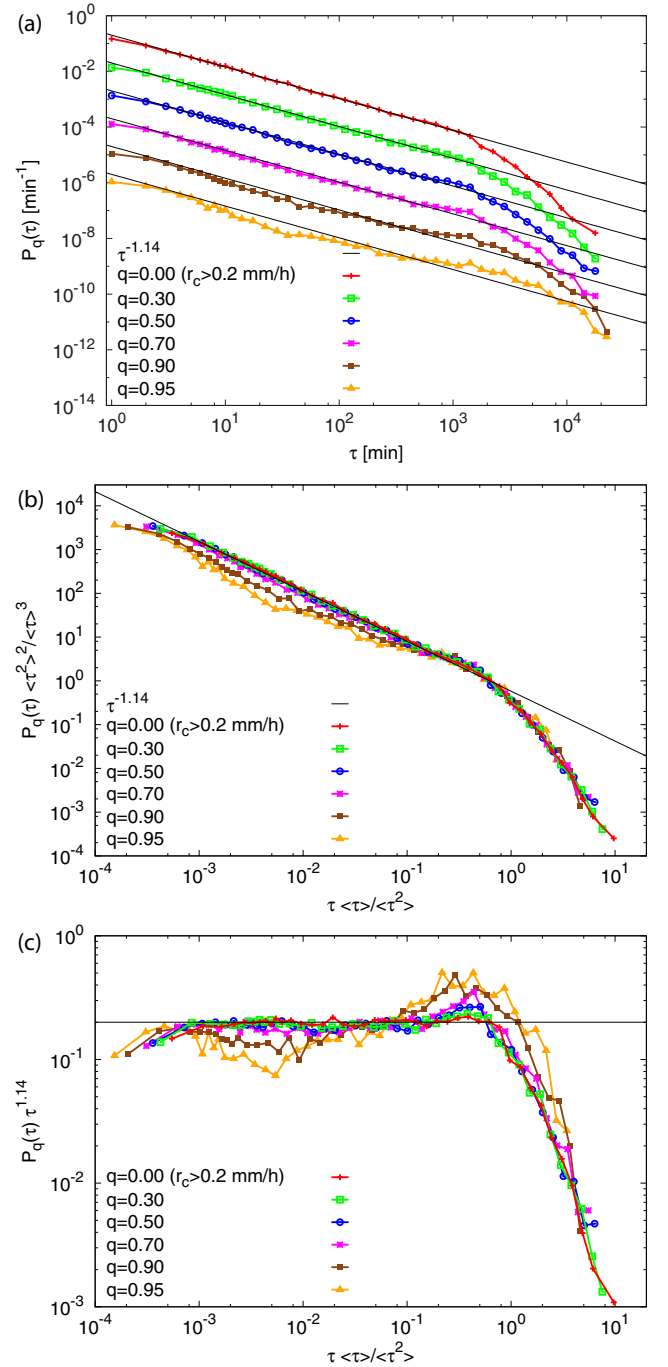


FIG. 2. (Color online) (a) Quiet-time PDFs for different rain rate thresholds, as measured on Manus island for the period 15 February 2005–18 March 2012. PDFs for different q are displaced vertically for clarity. A power-law decay with exponent $\beta = 1.14$ has been fitted to the $q = 0.00$ data to serve as a visual guide. (b) Data collapse of the suitably rescaled quiet-time PDFs. (c) Alternative data collapse in order to visualize the scaling function f_0 . Units in the ordinate are $\text{min}^{0.14}$. Quantile thresholds q correspond to rate thresholds $r_c = 0.20, 0.67, 1.75, 4.74, 22.02$, and 42.76 mm/h.

law decay with an exponent β close to 1.14 and a faster decay at the tail. Beyond $q = 0.70$, differences become more apparent.

Nevertheless, a reasonable data collapse is shown in Fig. 2(b). Other ARM sites show roughly the same behavior.

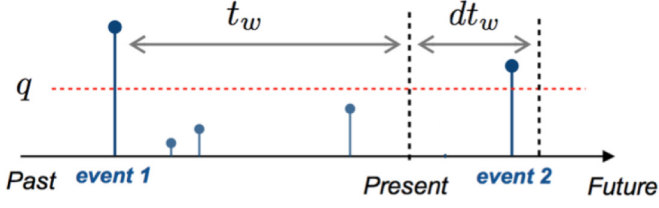


FIG. 3. (Color online) Pictorial definition of the hazard function. Event 2 falls between t_w and $t_w + dt_w$, having exceeded event 1 by t_w .

In the next two sections, we explore how differences in the quiet-time PDFs between the Manna model and rainfall data show up in the predictability of rain events.

III. HAZARD FUNCTION AND ROC CURVES

For the purposes of time-series prediction, we use the hazard function H_q (which is sensitive to both the clustering and repelling of events) as a decision variable. In comparison, the conventional precursory pattern-recognition technique requires a large amount of data, does not capture long-term clustering, and has been found to perform worse in a similar analysis of heartbeat intervals [26]. The hazard function [27] gives the probability per unit time that the quiet time (for events defined by a threshold given by quantile q) terminates between t_w and $t_w + dt_w$, given that it has exceeded t_w , as illustrated in Fig. 3. That is,

$$H_q(t_w)dt_w = \frac{\int_{t_w}^{t_w+dt_w} P_q(\tau)d\tau}{\int_{t_w}^{\infty} P_q(\tau)d\tau} \quad (2)$$

$$= \frac{P_q(t_w)dt_w}{S_q(t_w)}, \quad (3)$$

where $S_q(t_w) = \int_{t_w}^{\infty} P_q(\tau)d\tau$ is the survivor function [27], i.e., the probability that the quiet time is greater than t_w . For future reference it is useful to note that for exponential quiet times (Poisson process), the hazard function is constant. Since the time series under consideration are discrete, dt_w corresponds to one parallel update in the Manna model, and 1 min for the rain data. The hazard function is constructed numerically via the quiet-time PDF and the survivor function. Figure 4 shows results for the Manna model and Manus island rain over the complete temporal record. For the purposes of prediction we only assume access to past information, and the hazard function is therefore constructed solely from past events. We update its estimate every 100 events.

Since the hazard function gives a probabilistic forecast, a deterministic prediction is issued via a discrimination threshold. Specifically, if in a given time step the estimated H_q exceeds the discrimination threshold, an alarm is raised, which is to say that an event is expected in the next dt_w . If H_q does not exceed the threshold, no alarm is raised and no event is expected. The choice of discrimination threshold thus defines time windows for alarms and no alarms, with associated deterministic forecasts [26]. Note also that the prediction protocol does not assume knowledge of the signal when it is below threshold; therefore, the method is useful when this signal is unavailable, unreliable, or noisy. In this

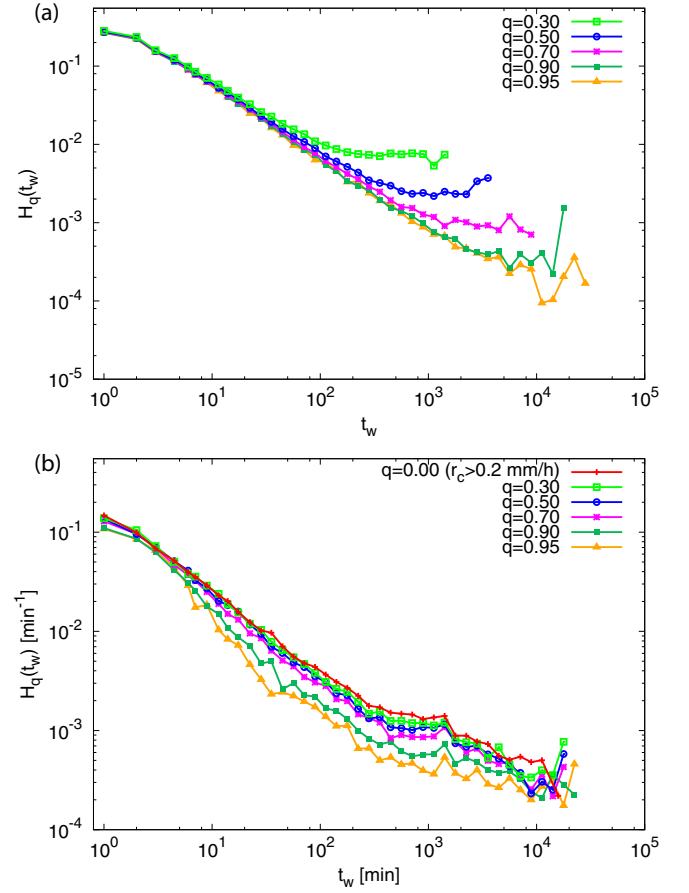


FIG. 4. (Color online) (a) Hazard functions for quiet times in the Manna model for different activity thresholds (same data as in Fig. 1). (b) Hazard functions for quiet times of Manus island rain data for different rain rate thresholds (same data as in Fig. 2).

way, the method is particularly suited to the prediction of extreme events.

The quality of this method of prediction is evaluated using a ROC analysis [28]. For any binary prediction (alarm raised or not) there are four possible outcomes: an alarm is raised and the event does occur [true positive (TP)]; an alarm is raised and the event does not occur [false positive (FP)]; an alarm is not raised and the event does occur [false negative (FN)]; and an alarm is not raised and the event does not occur [true negative (TN)]. The ROC curve summarizes this information by comparing the sensitivity (X_{sens} , the proportion of successfully predicted occurrences) to the specificity (X_{spec} , the proportion of successfully predicted nonoccurrences):

$$X_{\text{sens}} = \frac{R_{\text{TP}}}{R_{\text{TP}} + R_{\text{FN}}}, \quad X_{\text{spec}} = \frac{R_{\text{TN}}}{R_{\text{TN}} + R_{\text{FP}}}, \quad (4)$$

where R_{TP} , R_{FN} , R_{TN} , and R_{FP} refer to the number (or rate) of occurrences or nonoccurrences for each case. Each threshold on the decision variable will give rise to a different point on the ROC curve. For example, in the absence of a threshold an alarm is raised every time step, such that $R_{\text{FN}} = 0$, yielding a sensitivity of 1. Such a protocol will never miss an occurrence, but will also never predict a nonoccurrence, i.e., $R_{\text{TN}} = 0$, and therefore has specificity 0. On the other hand, for an

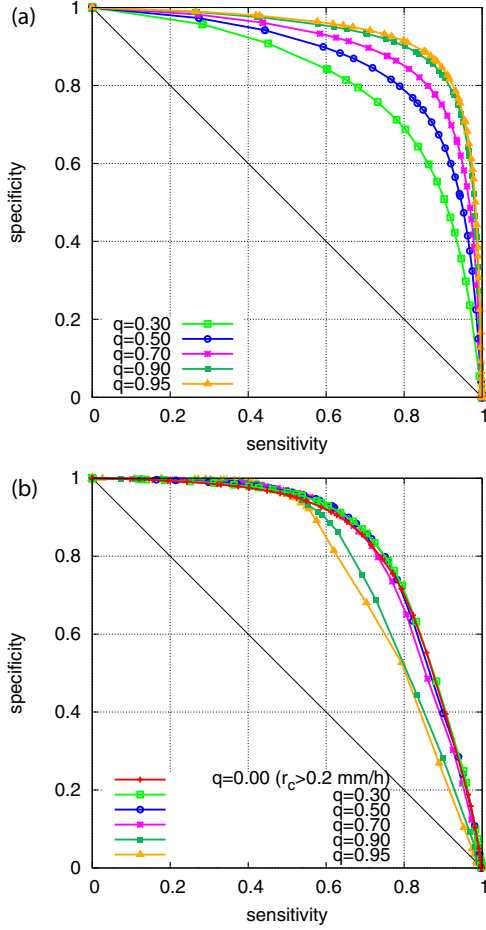


FIG. 5. (Color online) ROC curves for different thresholds in the activity or rate for (a) the Manna model with $L = 1024$ and (b) rainfall data from Manus island. Same data as in Figs. 1 and 2.

unsurpassable threshold an alarm is never raised, such that $R_{FP} = 0$, yielding a specificity of 1. Such a protocol will never miss a nonoccurrence, but will also never predict an occurrence and therefore has $R_{TP} = 0$ and sensitivity 0. The diagonal line in Fig. 5 that joins these two scenarios corresponds to noninformative predictions, i.e., issuing alarms at some fixed rate irrespective of the decision variable. Points above the diagonal represent good predictions and points below poor predictions, in comparison to totally random predictions. The point $(1, 1)$ corresponds to a perfect prediction.

Starting with the Manna model, the sequence of hazard functions in Fig. 4(a) suggests that the “memory” between events persists for increasingly longer quiet times as the activity threshold is increased. This is inferred from the increasing crossover times to the constant “memoryless” portion of the hazard curve. Thus, we anticipate that predictability improves with increasing activity threshold. This is indeed illustrated by the ROC curves in Fig. 5(a), which form an ordered sequence tending towards the corner $(1, 1)$ (perfect predictability). We explore the clustering that gives rise to this increased predictability in the next section.

For rainfall data, the picture is more complicated. One can observe in Fig. 5(b) a bundle of ROC curves which are broadly similar up to thresholds of $q = 0.70$. For higher thresholds we

find that, for sensitivities below 0.4, predictability (specificity) increases with threshold in rain rate (just as in the Manna model), but for higher sensitivities the opposite is true. To interpret what this means for events, recall that the discrimination threshold defines time windows for alarms and no alarms. Therefore, low sensitivities correspond to high thresholds in the hazard function (i.e., only events separated by short quiet times are successfully predicted). Conversely, high sensitivities correspond to low thresholds in the hazard function (i.e., events separated by long quiet times are also successfully predicted). Thus, if we observe increased predictability at low sensitivities, this is attributed to extreme events clustering more and more over short times; however, if we observe decreased predictability at high sensitivities, this is attributed to extreme events correlating less and less over long times. A closer reexamination of the highest quantile curves in Fig. 2(a) bears this out: The power-law exponent β appears to increase in the left part of the distribution, and the range of the exponential tail appears to increase in the right part. These longer times could be affected by seasonal effects which are difficult to resolve owing to the relatively short duration of the time series (7 years). Longer time series would be required in order to make better comparisons with SOC-like dynamics.

IV. ANALYTICAL TREATMENT

For monotonically decaying hazard functions (as is approximately the case for the Manna model and rainfall; see Fig. 4), raising an alarm whenever the function exceeds a threshold value is equivalent to raising an alarm all the while the elapsed time since the last event has not yet exceeded a “threshold” time. This threshold time, Δ , is uniquely determined by the threshold imposed on the hazard function. Under these conditions it is possible to infer some general behavior for the sensitivity and specificity. For ease of notation, we will omit any dependence on activity or rate threshold in the following.

First, consider the sensitivity. Given an event, the subsequent event occurs either before the threshold time, $t_w \leq \Delta$, or after the threshold time, $t_w > \Delta$. In the first case, events are successfully predicted, and therefore the total proportion of successfully predicted events $R_{TP} \propto \int_0^\Delta P(\tau) d\tau = 1 - S(\Delta)$. The remaining proportion $S(\Delta) \propto R_{FN}$ falls to unpredicted events, i.e., false negatives (note the similarity with type I errors in statistics). Thus, the sensitivity X_{sens} is

$$X_{sens} = 1 - S(\Delta),$$

which, if we introduce $\alpha = S(\Delta)$, is independent of the underlying form of the quiet-time distribution (apart from its assumed decaying monotonicity).

Next, consider the specificity. The specificity can also be calculated by instead considering nonevents. Suppose an event takes place after a quiet time τ since the previous event. If $\tau < \Delta$ (with probability $1 - \alpha$), the contribution to false positives is proportional to τ , whereas the contribution to true negatives is zero (since the alarm will be raised all the while). If $\tau > \Delta$ (with probability α), the contribution to false positives is proportional to Δ , whereas the contribution to true negatives is proportional to $\tau - \Delta$ (since the alarm is no longer raised beyond Δ). Taking into account the condition $\tau \leq \Delta$ in the

former case, the mean rate of false positives is thus

$$R_{\text{FP}} \propto (1 - \alpha) \langle \tau | \tau \leq \Delta \rangle + \alpha \Delta,$$

where $\langle \tau | \tau \leq \Delta \rangle = \int_0^\Delta \tau P(\tau | \tau \leq \Delta) d\tau$ is the mean quiet time conditional on $\tau \leq \Delta$ and $P(\tau | \tau \leq \Delta) = P(\tau)/(1 - \alpha)$ is the conditional density for $\tau \leq \Delta$. Likewise, the rate of true negatives is

$$R_{\text{TN}} \propto \alpha \langle \tau | \tau > \Delta \rangle - \alpha \Delta,$$

where $\langle \tau | \tau > \Delta \rangle = \int_\Delta^\infty \tau P(\tau | \tau > \Delta) d\tau$. Therefore, as the proportionality factor is the same in both cases (associated with the elementary time step dt_w), the specificity X_{spec} is

$$\begin{aligned} X_{\text{spec}} &= \frac{\alpha \langle (\tau - \Delta) | \tau > \Delta \rangle}{\alpha \langle \tau | \tau > \Delta \rangle - \alpha \Delta + (1 - \alpha) \langle \tau | \tau \leq \Delta \rangle + \alpha \Delta} \\ &= \frac{\alpha \langle (\tau - \Delta) | \tau > \Delta \rangle}{\langle \tau \rangle}, \end{aligned} \quad (5)$$

where $\langle (\tau - \Delta) | \tau > \Delta \rangle$ is the mean residual quiet time conditional on $\tau > \Delta$, and we have used the fact that the (unconditional) mean quiet time satisfies $\langle \tau \rangle = \alpha \langle \tau | \tau > \Delta \rangle + (1 - \alpha) \langle \tau | \tau \leq \Delta \rangle$. The mean residual quiet time has some counterintuitive properties in the case of a decreasing hazard function [29,30], but it characterizes a random variable in the same way as its probability distribution (if the mean is finite [27]). An additional practical consideration in our current setting is that the finite sample estimates of $\langle \tau \rangle$ and $\langle (\tau - \Delta) | \tau > \Delta \rangle$ may converge very slowly owing to the power-law part of $P(\tau)$. Equation (5) provides the relation between specificity and sensitivity via α (or, equivalently, via Δ , since $\alpha = S(\Delta) = 1 - X_{\text{sens}}$). For example, in the Poisson process the mean residual quiet time is equal to the mean quiet time irrespective of any elapsed time. Thus, $X_{\text{spec}} = \alpha$ and so $X_{\text{spec}} = 1 - X_{\text{sens}}$. In general, in contrast to the sensitivity, the specificity does depend on the underlying form of the quiet-time distribution.

It is instructive to apply the above analysis to a specific form of the quiet-time distribution. For the Manna model, a first approximation is given by a truncated Γ distribution

$$P(\tau) = \frac{1}{a\Gamma(\gamma, m/a)} \left(\frac{a}{\tau}\right)^{1-\gamma} e^{-\tau/a}, \quad (6)$$

where $m \geq 0$ is the lower cutoff of the distribution. The shape parameter $\gamma > 0$ for $m = 0$ and $-\infty < \gamma < \infty$ for $m > 0$, and the scale parameter $a > 0$ [which increases with activity threshold n_c ; see Fig. 1(a)]. The normalizing factor $\Gamma(\gamma, m/a)$ is the (upper) incomplete Γ function, defined by $\Gamma(\gamma, z) = \int_z^\infty x^{\gamma-1} e^{-x} dx$. Note that with this parametrization the power-law exponent is $\beta = 1 - \gamma$. Thus, comparing with the results of Sec. II for the Manna model, we have $\gamma = -0.67$. Also, exponential quiet times (as in a Poisson process) can be recovered as a special case by taking $\gamma = 1$, $m = 0$, and recalling that $\Gamma(1, z) = e^{-z}$.

From Eq. (6), the survivor function is

$$S(\tau) = \int_\tau^\infty P(t_w) dt_w = \frac{\Gamma(\gamma, \tau/a)}{\Gamma(\gamma, m/a)}, \quad (7)$$

and the hazard function is

$$H(\tau) = \frac{P(\tau)}{S(\tau)} = \frac{1}{a\Gamma(\gamma, \tau/a)} \left(\frac{a}{\tau}\right)^{1-\gamma} e^{-\tau/a}. \quad (8)$$

To compute the specificity, we require the mean quiet and mean residual quiet times, which are given by

$$\langle \tau \rangle = a \frac{\Gamma(\gamma + 1, m/a)}{\Gamma(\gamma, m/a)} \quad (9)$$

and

$$\begin{aligned} \langle \tau - \Delta | \tau > \Delta \rangle &= \int_m^\infty (\tau - \Delta) P(\tau | \tau > \Delta) d\tau \\ &= \int_\Delta^\infty (\tau - \Delta) \frac{P(\tau)}{S(\Delta)} d\tau \end{aligned} \quad (10)$$

$$\begin{aligned} &= \frac{a}{\Gamma(\gamma, \Delta/a)} \int_\Delta^\infty \left(\frac{\tau}{a}\right)^\gamma e^{-\tau/a} \frac{d\tau}{a} - \Delta \\ &= a \frac{\Gamma(\gamma + 1, \Delta/a)}{\Gamma(\gamma, \Delta/a)} - \Delta, \end{aligned} \quad (11)$$

for $\Delta > m$. This equation can be recast in a form more suggestive of scaling with the help of the identity $\Gamma(\gamma + 1, z) = z^\gamma e^{-z} + \gamma \Gamma(\gamma, z)$, giving

$$\langle \tau - \Delta | \tau > \Delta \rangle = a \left[\left(\frac{\Delta}{a}\right)^\gamma \frac{e^{-\Delta/a}}{\Gamma(\gamma, \Delta/a)} + \gamma - \frac{\Delta}{a} \right]. \quad (12)$$

As noted previously, high activity thresholds correspond to large values of the scale parameter a , while the lower cutoff m is essentially fixed by the available time resolution, e.g., one time step in the Manna model. Therefore, in order to explore scaling behavior of the sensitivity and specificity, we consider the limit $m/a \rightarrow 0$ (with m constant), in which case

$$\frac{1}{\Gamma(\gamma, m/a)} \simeq \begin{cases} 1/\Gamma(\gamma), & \gamma > 0 \\ -\gamma(m/a)^{-\gamma}, & \gamma < 0, \end{cases} \quad (13)$$

with $\Gamma(\gamma) = \Gamma(\gamma, 0)$ if $\gamma > 0$; see Ref. [31]. Proceeding to the quantities that appear in the ROC curves, we have

$$S(\Delta) \simeq \begin{cases} f_1(\Delta/a), & \gamma > 0 \\ a^\gamma f_1(\Delta/a), & \gamma < 0, \end{cases} \quad (14)$$

$$\langle \tau - \Delta | \tau > \Delta \rangle \simeq a f_2(\Delta/a) \quad \text{for all } \gamma, \quad (15)$$

$$\langle \tau \rangle \simeq \begin{cases} a\gamma, & \gamma > 0 \\ -a^{1+\gamma} m^{-\gamma} \gamma \Gamma(\gamma + 1), & -1 < \gamma < 0 \\ m\gamma/(\gamma + 1), & \gamma < -1, \end{cases} \quad (16)$$

where f_1 and f_2 are different scaling functions, with parameters γ and m . Collecting these results, we have

$$X_{\text{spec}} = \begin{cases} f_3(\Delta/a), & \gamma > -1 \\ a^{\gamma+1} f_3(\Delta/a), & \gamma < -1, \end{cases} \quad (17)$$

where f_3 is another scaling function. The $X_{\text{sens}} = 1 - S(\Delta)$, meanwhile, scales differently for $\gamma > 0$ and $\gamma < 0$. Thus, we

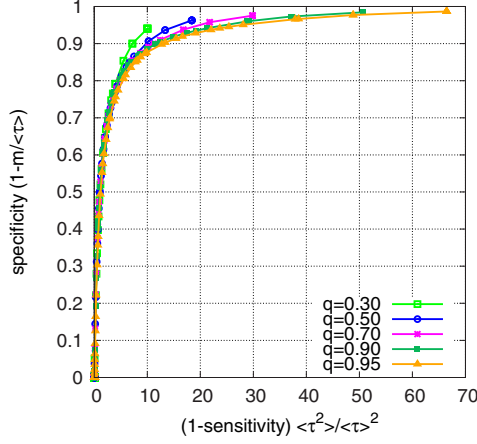


FIG. 6. (Color online) Data collapse after rescaling the ROC curves for different activity thresholds in the Manna model with $L = 1024$. Same data as in Fig. 1.

finally obtain

$$1 - X_{\text{sens}} = \begin{cases} f_4(X_{\text{spec}}), & \gamma > 0 \\ a^\gamma f_4(X_{\text{spec}}), & -1 < \gamma < 0 \\ a^\gamma f_4(X_{\text{spec}}/a^{\gamma+1}), & \gamma < -1, \end{cases} \quad (18)$$

where f_4 is another scaling function. An important consequence of these results is that for $\gamma > 0$ predictability is independent of a , whereas for $-1 < \gamma < 0$ sensitivity increases with a for a fixed specificity. In this sense, extreme events are more predictable, as in Ref. [20], but for different reasons. In fact, the sensitivity tends to 1 as $a \rightarrow \infty$ for any nonzero specificity. This is in agreement with our findings for the Manna model.

A concrete validation of this analysis can be achieved by plotting

$$X_{\text{spec}} \text{ versus } \frac{1 - X_{\text{sens}}}{\langle \tau^2 \rangle^\gamma / \langle \tau \rangle^\gamma},$$

which makes use of the fact that the scale parameter a is proportional to $\langle \tau^2 \rangle / \langle \tau \rangle$ when $-1 < \gamma < 0$, as we have seen in Refs. [8,24]. Noting that $a^\gamma \propto \langle \tau \rangle^2 / \langle \tau^2 \rangle$, a nonparametric version of the scaling law for the ROC curve leads to the plot of

$$X_{\text{spec}} \text{ versus } \frac{1 - X_{\text{sens}}}{\langle \tau \rangle^2 / \langle \tau^2 \rangle},$$

which is also valid for when $-1 < \gamma < 0$ [24]. For a better approximation, the finiteness of the time resolution, i.e., the fact that $m < 0$, should be taken into account in the calculation of the specificity, Eq. (5). If m is not negligible, then the FP rate is reduced by a term proportional to m , which leads to the replacement of $\langle \tau \rangle$ by $\langle \tau \rangle - m$ in the denominator of Eq. (5). This means that all the previous equations remain valid if the specificity is multiplied by a factor $1 - m/\langle \tau \rangle$. Thus, the scaling of the ROC curve is generalized to the plot of

$$X_{\text{spec}} \times \left(1 - \frac{m}{\langle \tau \rangle}\right) \text{ versus } \frac{1 - X_{\text{sens}}}{\langle \tau \rangle^2 / \langle \tau^2 \rangle}.$$

Figure 6 shows the scaled ROC curve for the Manna model. The collapse is reasonable and improves with increasing

threshold q . While the scaling theory works well for the Manna model, the collapse of the ROC curves for rainfall is poor (not shown), because the quiet-time distributions do not scale for high rain-rate thresholds.

As a final comment, although we have carried out this analysis for a truncated Γ distribution, we expect the previous scaling results to hold for any quiet-time distribution of the form

$$P(\tau) = \frac{1}{\tau^{1-\gamma}} f_0(\tau/a),$$

where f_0 is a scaling function taking a constant value for small arguments and decaying fast enough for large arguments. The normalization constant hidden in f_0 may depend on a , m , and γ ; see Ref. [24].

V. CONCLUSIONS

In summary, we have studied the effect of thresholding in a representative SOC model and on actual rainfall data. The predictability of events is studied by means of a decision variable sensitive to the tendency of the events to cluster or repel, and the quality of the predictions is evaluated by the ROC method. Thresholds have been applied to the rate or activity, which, in the context of SOC models, corresponds to observing the process on the fast (avalanche) time scale. In this case, the relative weight of the exponential tail decreases as the threshold increases, leading to higher predictability. For rainfall data, however, the change in the ROC curves with threshold is less clear, since the quiet-time distributions do not seem to scale for high thresholds. These differences in the predictability of extreme events stand out clearly in the ROC methodology: In the Manna model, the ROC curves form a well-ordered sequence of improving predictability with increasing threshold; in the rainfall data, predictability markedly deteriorates above a certain high threshold. This comparison thus yields useful nontrivial insights.

A scaling theory developed for the Manna model, valid for any system in which the quiet-time distribution scales with threshold, helps us understand all the details of the prediction procedure. The philosophy of our paper is similar to that of Ref. [32], but note that in that work the prediction scheme is based on precursory structures rather than the hazard function.

Clearly, our prediction method works best for renewal processes (point processes in which the quiet times are independent of each other). Analogies with other natural hazards [33] suggest that the renewal process is just a first approximation. Extensions to our approach would therefore include previous history, such that the hazard function is a function of previous quiet times. Nevertheless, our analytical approach is valid for any point process to which our prediction procedure is applied (not only for renewal processes). Further refinements would include knowledge of the rain rate or toppling activity below threshold. These extensions are left for future research.

Note added in proof. A relevant precedent to our work, related to Ref. [26], is Ref. [34], where Bogachev and Bunde have compared the performance of the hazard function method versus precursory pattern-recognition techniques on synthetic mono- and multifractal long-range dependent time series.

ACKNOWLEDGMENTS

Rain data were obtained from the Atmospheric Radiation Measurement Program sponsored by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Environmental Sciences Division. The authors were introduced to rain research by Ole Peters.

Research expenses were funded by Project No. FIS2009-09508, from the disappeared Spanish MICINN, Project No. FIS2012-31324, from Spanish MINECO, and Projects No. 2009SGR-164 and No. 2014SGR-1307, from AGAUR. A.D. is grateful for the hospitality of the Max Planck Institute for the Physics of Complex Systems and the Universidade Federal de Minas Gerais.

-
- [1] S. Lovejoy, *Science* **216**, 185 (1982).
 - [2] G. Vattay and A. Harnos, *Phys. Rev. Lett.* **73**, 768 (1994).
 - [3] J. I. Yano, R. Blender, C. Zhang, and K. Fraedrich, *Q. J. R. Meteorol. Soc.* **130**, 1697 (2004); J.-I. Yano, K. Fraedrich, and R. Blender, *J. Climate* **14**, 3608 (2001).
 - [4] E. Bodenschatz, S. P. Malinowski, R. A. Shaw, and F. Stratmann, *Science* **327**, 970 (2010).
 - [5] O. Peters, C. Hertlein, and K. Christensen, *Phys. Rev. Lett.* **88**, 018701 (2002).
 - [6] O. Peters and J. D. Neelin, *Nat. Phys.* **2**, 393 (2006).
 - [7] J. D. Neelin, O. Peters, J. W.-B. Lin, K. Hales, and C. E. Holloway, *Philos. Trans. R. Soc. A* **366**, 2581 (2008).
 - [8] O. Peters, A. Deluca, A. Corral, J. D. Neelin, and C. E. Holloway, *J. Stat. Mech.* (2010) P11030.
 - [9] A. Deluca and A. Corral, *Nonlinear Proc. Geophys.* **21**, 555 (2014).
 - [10] A. Deluca, P. Puig, and A. Corral (unpublished).
 - [11] P. Bak, C. Tang, and K. Wiesenfeld, *Phys. Rev. Lett.* **59**, 381 (1987).
 - [12] R. Dickman, *Phys. Rev. Lett.* **90**, 108701 (2003).
 - [13] M. Mitzenmacher, *Internet Math.* **1**, 226 (2004).
 - [14] A. Arakawa and W. H. Schubert, *J. Atmos. Sci.* **31**, 674 (1974).
 - [15] H. Kantz, in *Network Science*, edited by E. Estrada, M. Fox, D. J. Highman, and G.-L. Oppo (Springer, New York, 2010), pp. 205–216.
 - [16] M. T. Ritsche, U.S. Department of Energy, ARM Technical Report No. DOE/SC-ARM/TR-031, 2008.
 - [17] D. Dhar, *Physica A* **263**, 4 (1999).
 - [18] Á. Corral and M. Paczuski, *Phys. Rev. Lett.* **83**, 572 (1999).
 - [19] M. Paczuski, S. Boettcher, and M. Baiesi, *Phys. Rev. Lett.* **95**, 181102 (2005).
 - [20] A. Garber, S. Hallerberg, and H. Kantz, *Phys. Rev. E* **80**, 026124 (2009).
 - [21] G. Boffetta, V. Carbone, P. Giuliani, P. Veltri, and A. Vulpiani, *Phys. Rev. Lett.* **83**, 4662 (1999).
 - [22] A. Corral, *Phys. Rev. Lett.* **92**, 108501 (2004).
 - [23] A. Corral, *Phys. Rev. Lett.* **95**, 159801 (2005).
 - [24] A. Corral, *Chaos Soliton Fract.* **74**, 99 (2015).
 - [25] G. Pruessner, *Phys. Rev. Lett.* **105**, 029401 (2010).
 - [26] M. I. Bogachev, I. S. Kireenkov, E. M. Nifontov, and A. Bunde, *New J. Phys.* **11**, 063036 (2009).
 - [27] J. D. Kalbfleisch and R. L. Prentice, *The Statistical Analysis of Failure Time Data*, 2nd ed. (Wiley, Hoboken, NJ, 2002).
 - [28] J. P. Egan, *Signal Detection Theory and ROC Analysis* (Academic Press, New York, 1975).
 - [29] P. M. Davis, D. D. Jackson, and Y. Y. Kagan, *Bull. Seismol. Soc. Am.* **79**, 1439 (1989).
 - [30] A. Corral, *Phys. Rev. E* **71**, 017101 (2005).
 - [31] *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (Dover, New York, 1965).
 - [32] S. Hallerberg and H. Kantz, *Phys. Rev. E* **77**, 011108 (2008).
 - [33] A. Corral, *Tectonophysics* **424**, 177 (2006).
 - [34] M. I. Bogachev and A. Bunde, *Physica A* **390**, 2240 (2011).