# Prediction of stasis and crisis in the Bak-Sneppen model.

Alejandro Tejedor[a], Javier B. Gómez[b], Amalio F. Pacheco[a,c]

[a]Department of Theoretical Physics, University of Zaragoza, 50009 Zaragoza, Spain.
[b]Department of Earth Sciences, University of Zaragoza, 50009 Zaragoza, Spain.
[c]Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, 50009 Zaragoza, Spain.

**Abstract**

Several recurrences in the dynamics of an individual species in the one-dimensional Bak-Sneppen model are analysed. The distributions of the time intervals for stasis and crisis are separately calculated together with the respective hazard functions for the transition between them. The predictabilities of when a crisis will start and when it will conclude are evaluated by using one- and two-parameter strategies and the information is represented in standard error diagrams.

*Key words:* Bak-Sneppen Model, Self-organized criticality, Predictability
*PACS:* 89.75.Fb

## 1. Introduction

The idea of self-organized criticality (SOC) introduced in [1] has had a great impact in recent decades in many different areas such as physics, biology, geosciences and social sciences [2]. Perhaps the most simple and elegant model of this type is the so called Bak-Sneppen model (BSM) [3][4][5][6].

The one-dimensional (1-d) BSM is a linear array of $N$ sites. Each site represents a species, and is assigned initially a random number uniformly distributed between 0 and 1 called fitness, $f_i$ ($1 \le i \le N$). At each time step, $n$, the site with lowest fitness is identified and the species at that site is mutated, that is, a new random number is assigned to this site. The interaction is introduced by also assigning new random numbers to the two nearest neighboring sites of the mutated site. Iterating this process, after many updates the set of $f_i$ in the array approaches a stationary distribution where most of them are bigger than a self-organized critical value $f_c = 0.66702$. In this stationary state an avalanche starts when $\forall i$, $f_i > f_c$ and for a while there appear sites where $f_i < f_c$. The avalanche ends when all the sites in the system again fulfill $f_i > f_c$, $\forall i$.

This model was initially designed to describe the coevolution of natural species, but it has also been used to model earthquakes [7] and the effect on commercial companies of introducing market regulations [8]. In this paper, we examine two new recurrence properties of the 1-dim BSM that can be interpreted within the original spirit of the evolution of species but which also offer an appealing perspective when observed from the point of view of the evolution of a set of companies interacting in the market.

Let us consider any particular site, $i$, in the 1-d BSM. If the element in question has a fitness barrier, $f_i$, under the critical value $f_c$, it belongs to an avalanche and the element is said to be in crisis (here we prefer the word crisis instead of *activity*). And if its fitness is higher than the critical value, the element is said to be in stasis. This is illustrated in Fig. 1.

The time intervals separating subsequent returns to crisis are called $T_{first}$.The distribution of these waiting time intervals is a power-law function, denoted by $P_{first}(n)$, whose exponent is $\tau_{first} = -1.58$ [6]. However, any interval $T_{first}$ is formed by two consecutive subintervals, the first when the element is in crisis, $T_c$, and the second when the element is in stasis, $T_s$ (Fig. 1).

$$T_{first} = T_c + T_s \quad (1)$$

Each subinterval has its own density distribution, denoted here $P_c(n)$ and $P_s(n)$. We propose that the information carried by these two probability distributions is much more interesting than that carried by the combined $P(n)$.

From the point of view of companies interacting in

*Email addresses:* atejedor@unizar.es (Alejandro Tejedor), jgomez@unizar.es (Javier B. Gómez), amalio@unizar.es (Amalio F. Pacheco)

a market, it is more important to know how long a crisis will last than the time at which the next one will occur for the obvious reason that the duration of the present crisis will determine their fate, independently of the time span until the next one. An interval of crisis of a company can be induced by its own weakness, small $f$, or by the weakness of its nearest neighbours in the context of an economic network. The crisis ends at the moment when $f$ is bigger than $f_c$, and a stasis interval starts for that particular company.

In Section 2 we perform simulations for sets of different size, $N$, to evaluate the mean, standard deviation and variation coefficient (aperiodicity) of the distributions of the crisis- and stasis- intervals of a particular company. Then, for the size $N = 256$ these two distributions will be plotted to appreciate their power-law nature and the corresponding critical exponents will be identified. The hazard rates for the transition from crisis to stasis and vice versa will also be calculated. In Section 3, the possibility of predicting these two types of transitions will be evaluated by means of simple one- and two-parameter strategies. These results will be graphically displayed in the so-called error diagrams [9]. Finally, in Section 4 these results are discussed.

## 2. Results for the waiting time statistics of stasis and crisis intervals

In Fig. 2 we have plotted the mean, $\mu$, and the standard deviation, $\sigma$, of the crisis and stasis intervals as a function of the size of the system, $N$. For stasis intervals, $\mu$ and $\sigma$ increase with $N$ as a power law with exponents $\approx 1.1$ and $\approx 1.7$ respectively. In contrast, in the crisis intervals, $\mu$ and $\sigma$ are almost constant with values of around 15 and 180 time units respectively.

More important than $\mu$ and $\sigma$ separately, the ratio $\frac{\sigma}{\mu}$ known as the coefficient of variation or aperiodicity, $\alpha$, is a key parameter in the predictability of the system. Fig. 3 plots the aperiodicity, $\alpha$, of the crisis and stasis intervals as a function of $N$. Note that whilst $\alpha$ in the stasis intervals is a strictly increasing function, in the crisis intervals it is almost constant $\alpha \approx 11$ (for $N > 256$).

Fig. 4 shows the density distribution function of stasis intervals, $P_s(n)$, for a system $N = 256$, that is the probability that a stasis interval ends after $n$ time steps. This distribution is a power law with an exponent $\tau_s = -1.477 \pm 0.004$. This graph has been built juxtaposing a first part (up to $n < 50$) and a second part ($n > 50$). In the first part we have used a linear binning of the data. In the second, where the amount of data is scarce, a logarithmic binning has been used. It is manifest that there is no kink in the curve, as would be expected, given that the same density of probability is plotted to the left and to the right of $n = 50$.

Fig. 4 also shows the density distribution of crisis intervals, $P_c(n)$ ($N = 256$), i.e. the probability that an interval of crisis finishes after $n$ time steps. Note that in this case the initial points (low $n$) do not follow a power-law. The slope of the straight part is $\tau_c = -2.060 \pm 0.011$. In this case a linear binning has been also done for $n < 50$ and a logarithmic binning for $n > 50$.

In Fig. 5 we have plotted the hazard rates for entering into a crisis, $H_s$, and the recovery rate for leaving the crisis, $H_c$, for a $N = 256$ system. Both rates refer to the function

$$H_k(n) = \frac{P_k(n)}{\sum_{i=n}^{\infty} P_k(i)} \tag{2}$$

where $k$ refers to crisis ($k = c$) or stasis ($k = s$) intervals.

According to Eq. 2, $H_k(n)$ is the conditional probability that if a crisis/stasis has not started for $i < n$, it starts just in the $n^{th}$ time interval. As $P_k(n)$ is a discrete probability distribution, $H_k(n)$ is dimensionless. Note that for the crisis→stasis transition ($H_c$) we prefer the term recovery instead of hazard for obvious reasons.

Because of the power law behaviour of the distributions of both crisis and stasis intervals $P_k(n) \propto n^{\tau_k}$ ($\tau_k < 0$), a simple calculation can be made for $|\tau_k| > 1$ and for sufficiently large $n$:

$$\begin{aligned} H_k(n) &= \frac{n^{-|\tau_k|}}{\sum_{i=n}^{\infty} i^{-|\tau_k|}} \approx \\ &\approx \frac{n^{-|\tau_k|}}{\frac{i^{-|\tau_k|+1}}{-|\tau_k|+1}\Big|_n^{\infty}} \propto n^{-1} \end{aligned} \tag{3}$$

which agrees with the numerical results shown in Fig. 5.

## 3. Using error diagrams to evaluate several strategies for predicting the beginning and end of crisis intervals

A convenient way to asses the predictability of the occurrence of certain *target* events in the context of a temporal series is to declare alarms at particular times, maintain them during a certain interval, and then disconnect them. The aim of any strategy is that the alarm is ON when the target events occur in order not to miss an event, and simultaneously to minimize the connection time of the alarm. A success in the prediction corresponds to a target event that occurs while the alarm is ON, and an error in the prediction corresponds to a target event that occurs while the alarm is OFF. Thus, the fraction of errors, $f_e$, is the number of missed target events divided by the total number of target events that have occurred during a long interval of time, $T$, while the fraction of alarm time, $f_a$, is the total time that the alarm was ON divided by $T$. A so-called loss function, $L$, is then defined to take into account the relative importance, in terms of cost, of $f_e$ and $f_a$. Here we will use $L = f_e + f_a$, where failure to predict and alarm time are equally penalized. Our strategies will be parametric and thus we will explore the value of the parameters that minimize $L$. As each type of strategy produces a minimum for $L$, denoted by $L^*$, this method allows them to be easily compared.

In the Random Guessing Strategy (RGS), the decision of putting the alarm ON or OFF is made randomly in each temporal step. RGS leads to $L^* = 1$, and therefore any strategy for which $L^* > 1$ is considered useless. In some cases, if a specific strategy leads to $L^* > 1$, it is possible to find a complementary strategy for which $L^* < 1$.

A Reference Strategy (RS) [10] consists in waiting $n$ time units after the occurrence of each target event, setting the alarm, and maintaining it until the occurrence of the following target event. The RS is simple and does not require knowledge of the internal mechanisms of the system. The complementary option to RS, $\overline{\mathrm{RS}}$, is also easily implemented: the alarm is connected just after the occurrence of a target event, and is maintained for an n-step interval. After these steps, if the target event has not happened the alarm is removed. In both RS and $\overline{\mathrm{RS}}$,the aim is finding the $n$ that minimizes $L$, that is, they are one-parameter strategies. In this paper, we have also studied a two-parameter strategy, denoted by bi$\overline{\mathrm{RS}}$. It has been specifically designed to try to predict the beginning of a crisis interval. In this new strategy, bi$\overline{\mathrm{RS}}$, the rules of $\overline{\mathrm{RS}}$ hold, but a new parameter is added: the alarm will also be connected for a period of time, $n'$, if any of the second neighbours of the company of reference enters into a crisis while our company is in stasis.

A convenient way to graphically display the results of the application of a strategy is by means of an Error Diagram [9], where the fraction of errors $f_e$ runs along the horizontal axis and the fraction of alarm $f_a$ runs along the vertical axis.

In Fig.6 we show the results of applying the above-mentioned strategies to predict the transition stasis→crisis for $N = 256$. As is apparent, the RS is useless here, whilst $\overline{\mathrm{RS}}$ provides an optimal $L^* = 0.14$ result for $n = 179$. The biparametric strategy bi$\overline{\mathrm{RS}}$ is even better providing $L^* = 0.10$ for $n = 81$ and $n' = 5$.

Fig.7 shows the results of predicting the crisis→stasis transition (i.e, the recovery from a crisis), for $N = 256$. Here RS is again useless, while $\overline{\mathrm{RS}}$ provides a significant improvement $L^* = 0.47$ for $n = 15$. The results for bi$\overline{\mathrm{RS}}$ are worse than those of $\overline{\mathrm{RS}}$, $L^* = 0.55$ for $n = 9$ and $n' = 1$. (For $n' = 0$, the bi$\overline{\mathrm{RS}}$ coincides with the $\overline{\mathrm{RS}}$).

## 4. Discussion

We have studied the recurrences of the crisis and stasis intervals in the dynamics of a particular element (a company), in the 1-dim Bak-Sneppen Model, for different array sizes. We have then made specific simulations for an array of $N = 256$ sites, and two new critical exponents, $\tau_s$ and $\tau_c$, have been identified. The considerable difference between $\tau_s$ and $\tau_c$ implies that in comparison, the occurrence of long intervals is much more difficult in crisis than in stasis. This is confirmed in Fig. 2 where it is apparent that the crisis intervals are much shorter than the stasis intervals. This behaviour is reinforced as $N$ becomes larger.

This is understood recalling the dynamics of the BS model. The fact that a definite site, at a definite time, is in crisis implies that it is an element of an avalanche existing in the whole system. During the time that an avalanche is spatially located near the element of reference, this element undergoes fluctuations between the two states (crisis and stasis) and when the avalanche finally leaves that area, or when it ends, a long stasis interval starts.

With respect to Fig. 5, both rates show a power law behaviour with an exponent of about $-1$.This implies that as the time elapsed since the last crisis (stasis) increases, the probability of occurrence of the next one decreases.

Regarding the predictability of the recurrences as shown in Figs. 6 and 7, several facts are clear. The values $\alpha > 1$ of the aperiodicity as shown in Fig.3 warn that for the two types of intervals, the phenomenon of clustering is present. This is produced by the rapid fluctuations borne by an element when it is affected by a low-fitness group of an avalanche and explains the poor performance of the RS and the goodness of the $\overline{\mathrm{RS}}$. The positive result of the bi$\overline{\mathrm{RS}}$ in predicting stasis→crisis transitions is due to the fact that it gives (proxy) information about the approach of the low-fitness group. This type of information is not available in the simple $\overline{\mathrm{RS}}$ method. In the crisis→stasis transition, on the other hand, the bi$\overline{\mathrm{RS}}$ is not good because it should not be expected that when a second neighbour recovers the stasis state, our company will do the same immediately afterwards.

We conclude that the model presented here obviously does not attempt to describe, for example, a global crisis in the world economy but only the habitual behaviour of one company interacting with others in the same sector: good steady intervals (here denoted by stasis intervals) followed by turbulent intervals that require the redefining of goals and making in-depth changes (here denoted by crisis intervals). The dissection made here to the intervals $T_{first}$, can likewise be performed if the Bak-Sneppen model is placed not in a regular 1-d lattice, as assumed here, but for instance in 2-d or in a scale-free network [11] and also, of course, if the rules of the model are changed as described in [8] .


## ACKNOWLEDGEMENTS

This work was supported by a Project of the Spanish Ministry of Education and Science. Several long numerical simulations were carried out using the computing facilities of BIFI.


## References


[1] Bak, P. and Tang, C. and Wiesenfeld K. (1987), Self-organized criticality: An explanation of the 1/f noise, *Phys. Rev. Lett.*, *59*(4), 381–384.

[2] Jensen, H. J. (1998), Self-Organized Criticality : Emergent Complex Behavior in Physical and Biological Systems, *Cambridge University Press*.

[3] Bak, P. and Sneppen,K. (1993),Punctuated equilibrium and criticality in a simple model of evolution, *Phys. Rev. Lett.*, *71*(24), 4083–4086.

[4] Maslov, S. and Paczuski, M. and Bak P. (1994), Avalanches and $1/f$ Noise in Evolution and Growth Models, *Phys. Rev. Lett.*, *73*(16), 2162–2165.

[5] Grassberger, P. (1995), The Bak-Sneppen model for punctuated evolution, *Physics Letters A*, *200*(3-4), 277 - 282.

[6] Paczuski, M. and Maslov, S. and Bak P. (1996), Avalanche dynamics in evolution, growth, and depinning models, *Phys. Rev. E*, *53*(1), 414–443.

[7] Ito, K. (1995), Punctuated-equilibrium model of biological evolution is also a self-organized-criticality model of earthquakes, *Phys. Rev. E*, *52*(3), 3232–3233.

[8] Cuniberti, G. and Vallerani, A. and Vega J. L. (1999), Effects of regulation on a self-organized market, *Quantitative Finance*, *1*,332.

[9] Molchan, G. M. (1997), Earthquake prediction as a decision-making problem, *Pure Appl. Geophys.*, *149*, 233–247.

[10] Newman, W. I. and Turcotte,D. L (1993),A simple model for the earthquake cycle combining self-organized complexity with critical point behavior, *Nonlin. Proc. Geophys.*, *9*, 453-461.

[11] Moreno, Y. and Vazquez,A. (2002),The Bak-Sneppen model on scale-free networks, *Europhys. Lett.*, *57*(5), 765771.


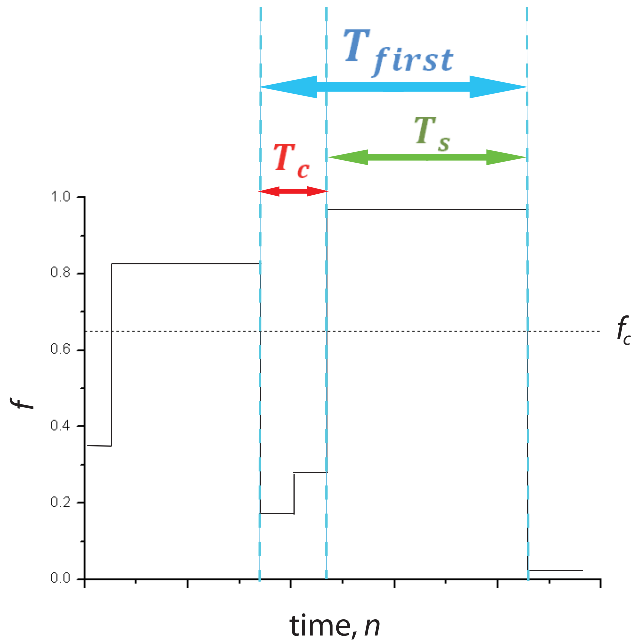


Figure 1: Illustration of the different time intervals: $T_{first}$,$T_s$ and $T_c$. $f$ is the fitness of an individual element, and $n$ is the discrete time of the model.

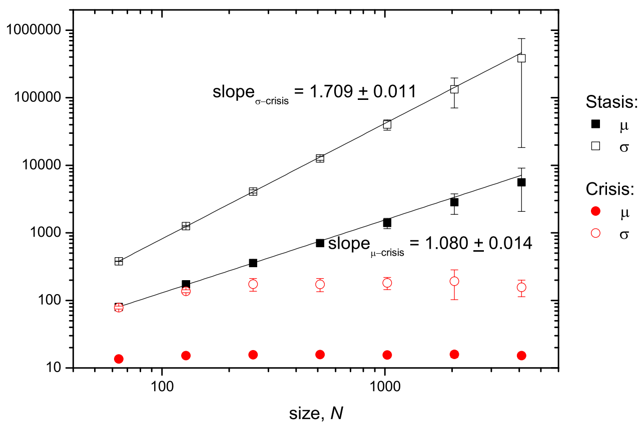


Figure 2: Mean and standard deviation of the crisis and stasis intervals as a function of the size of the system, $N$.

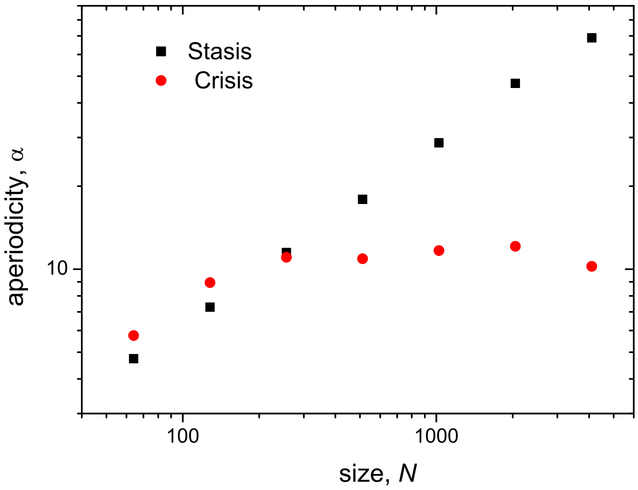


Figure 3: Aperiodicity for stasis and crisis intervals as a function of $N$.

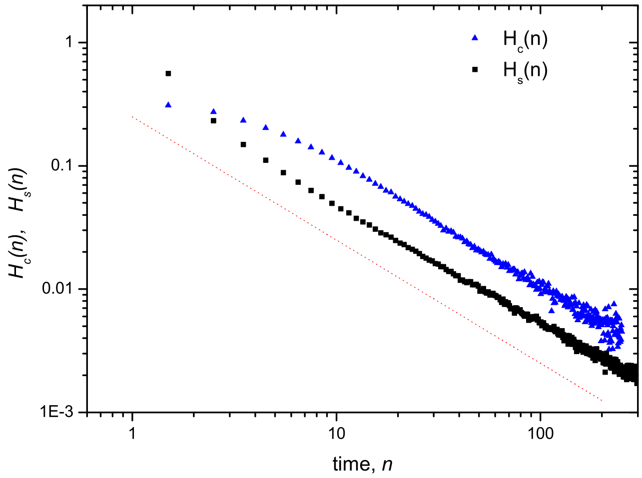


Figure 5: Hazard rate for entering into a crisis, $H_s(n)$, (black squares) and Recovery rate for leaving the crisis, $H_c(n)$ (blue triangles).($N$ = 256) A dotted line with slope -1 is plotted for reference.

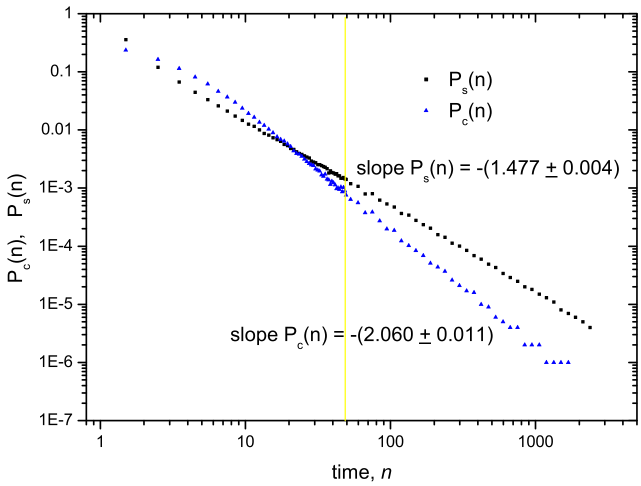


Figure 4: Distribution function of stasis (black squares) and crisis (blue triangles) intervals.($N$ = 256)

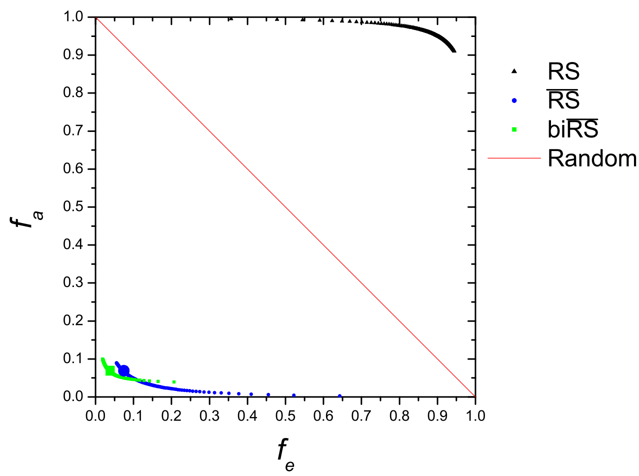


Figure 6: Error diagram to evaluate the predictability of the transition stasis→crisis $N$ = 256)

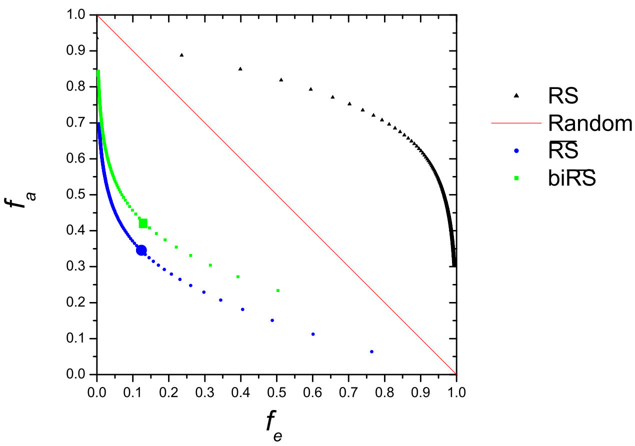


Figure 7: Error diagram to evaluate the predictability of the transition crisis→stasis $N$ = 256)