Geophys. J. Int. (2001) **144,** 609–624

A simple but effective cellular automaton for earthquakes

Silvia Castellaro and Francesco Mulargia

Settore di Geofisica, Dipartimento di Fisica, Bologna, Italy. E-mails: silvia@ibogeo.df.unibo.it; mulargia@ibogfs.df.unibo.it

Accepted 2000 October 2. Received 2000 September 29; in original form 2000 May 12

SUMMARY

The physics of fractures, which forms the basis of seismic faulting, is not very amenable to simple deterministic differential equations. For this reason a different approach, aimed at reproducing the statistical mechanical properties of earthquakes, has attracted progressively increasing interest. A variety of models have been presented but there seems to be little that can be done to ascertain the merits and defects of each. We set the clock back and attempt to derive a dynamically evolving automaton that is as simple as possible and that incorporates all the basic ingredients and includes strain diffusion, a process often disregarded in simple models in spite of its crucial importance. Our automaton is based on a homogeneous grid of cells and its rupturing is controlled by a generalized local threshold. The automaton also considers local dissipation of energy and time-dependent strain applications. This simple model is capable of reproducing earthquake dynamics, including the effects due to transient loads such as those imposed by elastic waves, with an efficiency superior to that of the most complicated automata and with less stringent assumptions.

Key words: cellular automaton model, earthquake dynamics.

1 INTRODUCTION TO COMPLEX SYSTEMS

While the description of deterministic and stochastic processes is well established, the picture is far less clear for the recently introduced *complex systems*. Generally speaking, a complex system is any system made up of several parts interacting in a strongly non-linear way. This results in behaviour that is difficult to describe with a group of equations.

The classical scientific method looks for

(i) an explanation that fits in with other data and concepts, and
(ii) a testable theory that is as simple as possible.

The second point is one of the paradigms of physics, namely that each model must be written in a form that can be verified by anyone, and the easiest way to do so is to write it in mathematical form. Until a few years ago, differential equations were the most widely used instrument for building mathematical models because of their capability of describing the evolution of a system in time and space. In fact, only some phenomena can be described exhaustively by differential equations, and among the differential equations, only the linear ones, those that can be linearized and a few others have analytical solutions.

The variables in the differential equations are continuous functions. However, such equations are often solved through the application of numerical methods on discretized variables.

Computers have been used mostly as a calculus tool but their capability to model directly complex systems is still on the verge of being fully exploited. In this case, computers can operate according to two paths,

(i) a forward path, through simulation, since they can reproduce and follow the rapid changes of status and interrelations among the parts of a many-body system, and
(ii) an inverse path, since they are able to recognize complex patterns.

We will deal below with the first path. The basis of our computer simulation is a cellular automaton (CA), a conceptual device that can assume certain states and responds to given stimuli (inputs) according to a predefined set of laws. In general, automata can successfully represent many-body interactions and therefore allow one to model complex processes, a task that is beyond the capabilities of classical mathematics. A general reference on the theory and applications of cellular automata is Wolfram (1986).

2 GENERAL TERMS

We know from thermodynamics that certain systems present a discontinuity in some properties such as density, entropy, magnetization, etc. that characterize a phase transition. The point in space where the discontinuity first vanishes, that is, where the first-order transition disappears, is called the critical point, and beyond this point it is possible to move continuously from one phase to the other.

A physical system is said to be in a state of organized criticality (OC) when it approaches a critical point and retreats

from it as a consequence of a certain event, after which it can approach the critical point again and so on, intermittently. A physical system is said to be in a state of self-organized criticality (SOC) when it can organize itself not just into a pattern but into the precise pattern seen at the critical point, and then oscillate around it in a stationary metastable condition. A key property of the SOC state is that the behaviour of the system is relatively insensitive to the details of the dynamics so that it is not necessary to fine-tune the parameters in order to achieve this condition. It has been suggested that earthquakes are similar to critical points and that earthquake distributions have properties strikingly similar to OC and SOC (Main 1996; Sammis & Smith 1999) but it has so far proved impossible to state these issues in definite terms.

Seismic hazard studies assume that the crust undergoes stationary processes. To detect stationarity, one needs to collect data for a time interval longer than the duration of the process. Nobody knows plausible values of the latter for the lithosphere, and evidence in favour of such stationarity is lacking since contrary evidence is also lacking and since it eliminates time dependence. Only an Occam's razor-type reasoning would suggest keeping stationarity.

Although plate motion studies show that the long-term plate tectonic velocities averaged over the last 3 million yr from palaeomagnetic data are virtually identical to those measured in the last 10 yr from satellite data (DeMets 1995), indicating that global plate motion is likely to be stationary, it is hard to assume the same at the smaller scale at which earthquakes occur. The Earth's crust consists of physically and chemically different domains with a decreasing tendency to rupture from the margins inwards. The cumulative distributions of events in the non-stationary intervals may follow locally linear patterns that make it impossible to detect its general state.

The Gutenberg–Richter power law (Gutenberg & Richter 1954) for natural earthquakes holds for events of magnitude greater than a value M_0, below which seismic catalogues are incomplete; similar behaviour occurs at the upper end.

The complexity of lithosphere dynamics is not very amenable to simple deterministic differential equations. In this case, cellular automata can be more efficient.

3 THE MODEL

Following Occam's razor approach, we have attempted to derive a general and as simple as possible automaton capable of reproducing the statistical mechanical properties of earthquakes, incorporating what are assumed to be the basic microphysical aspects of the phenomenon. Our 2-D automaton is based on a square grid of massless equal elements. It does not involve assumptions about the exact physics underlying the process of rupture because it simply considers the stress, strain or energy release as a consequence of loads or rupture events. The way in which this redistribution takes place is not derived from specific physical equations but follows predefined general damping rules.

We decided to avoid the slider block model (Burridge & Knopoff 1967; Ito & Matsuzaki 1990; Narkounskaia *et al.* 1992; Olami *et al.* 1992; Nakanishi 1991; Carlson *et al.* 1994), which is described in Section 7 below, because it is based on planar faults, which are at odds with the unquestionable evidence that faults are fractal, and because of our incomplete knowledge of the physical equations governing the process.

To make the description of our model easier, we define a quantity called *ruptino* (plural *ruptini*), which can be interpreted both as a unit of increase in the ground strain level of the material and as a unit of decrease in the critical value of fracture of the element itself (Fig. 1). Ruptino has this double meaning because rupture is locally controlled by the difference between the rupture threshold and the strain level.

The evolution of the system then travels along a loop of the following transition rules.

(i) A discrete amount of ruptini is added to some elements of the grid.

(ii) A critical value is established for each element of the grid, and when the amount of strain in an element is greater than its threshold, that element becomes unstable and fractures. Its 'deformation' is distributed to its neighbours according to given laws. This redistribution may lead other elements to become unstable, and it is performed for every unstable element until stability of the system is reached again. The number of broken elements gives the size of the fracture event. We call the entire process starting with the addition of one (or more) ruptini and ending with all the elements of the system in a stable condition *iteration*.

(iii) Strain diffusion and the anelastic propagation of deformation are accounted for by further redistributions of strain, starting from the elements that were unstable during the previous step. The number of redistributed ruptini follows a geometric rule decreasing with the distance from the originally unstable cell. At the end of the iteration, new ruptini are added, starting again from the first step.

The specific redistribution laws are as follows. Let $N(i, j)$ be the state of the (i, j) element of the system, τ be the threshold level for rupture and t be the temporal step. At time $t=0$ (Fig. 2),

if $N(i,j) \geq \tau$ then

$$N(i,j) = N(i,j) - \tau, \tag{1}$$
$$N(i \pm 1, j) = N(i \pm 1, j) + \frac{\tau}{4},$$
$$N(i, j \pm 1) = N(i, j \pm 1) + \frac{\tau}{4}.$$

At the successive iteration (after all redistributions have been executed), the algorithm performs the following distributions (Fig. 3):

$$t = 1,\ N(i,j) = N(i,j) - \frac{2}{3}\tau, \tag{2}$$
$$N(i \pm 1, j \pm 1) = N(i \pm 1, j \pm 1) + \frac{\tau}{6},$$
$$t = 2,\ N(i,j) = N(i,j) - \frac{\tau}{3},$$
$$N(i \pm 2, j) = N(i \pm 2, j) + \frac{\tau}{12}, \tag{3}$$
$$N(i, j \pm 2) = N(i, j \pm 2) + \frac{\tau}{12},$$
$$t = 3,\ N(i,j) = N(i,j) - \frac{\tau}{3},$$
$$N(i \pm 1, j \pm 2) = N(i \pm 1, j \pm 2) + \frac{\tau}{12}, \tag{4}$$
$$N(i \pm 2, j \pm 1) = N(i \pm 2, j \pm 1) + \frac{\tau}{12}.$$

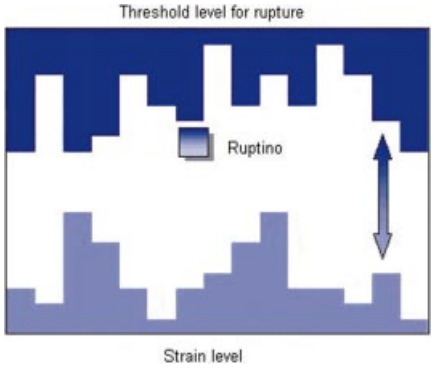

Figure 1. Ruptino is a quantity that represents both an increase in the ground strain level and a decrease in the critical value for rupture of the element considered. When the strain level reaches the threshold, a rupture event occurs.

Ruptini are lost from the grid when redistributions take place near the edges and corners, thus giving a generally dissipative model.

We refer to the action of injecting ruptini into a grid as *loading*. We have explored different ways of loading. In the next sections, we will examine the behaviour of our model for a number of different loading conditions.

3.1 Random load

3.1.1 Load

Consider first of all a grid, the elements of which are at different levels below the rupture threshold. We term this state *heterogeneous*. Starting from an empty grid, we can achieve heterogeneity by adding a certain number of ruptini to random locations. The number of iterations needed (i.e. of ruptini injected) obviously depends on the grid dimensions but we used up to 2×10^6 iterations.

Since most natural materials are heterogeneous and anisotropic, we expect that a heterogeneous (not yet anisotropic, in this study) grid of our CA is the basis for simulating a real material. Heterogeneity has different importance in different simulations. As rock mechanics teaches and engineering tests show, specimens behave more or less like homogeneous bodies only when subjected to loads applied to regions much wider than their largest heterogeneities. Thus, in laboratory experiments a simulation starting from a homogeneous grid of cells may be sufficient in some cases, while in the lithosphere a heterogeneous configuration is far more likely.

The first loading we analysed was a random addition of a single ruptino to a heterogeneous configuration for each step of the running algorithm. This loading condition can be seen as a constant and is accompanied by a lowering of the rupture threshold due to random fluctuations in the mechanical properties of any origin such as fluid injection and local heterogeneities in mechanical behaviour etc. that are represented by ruptini injections.

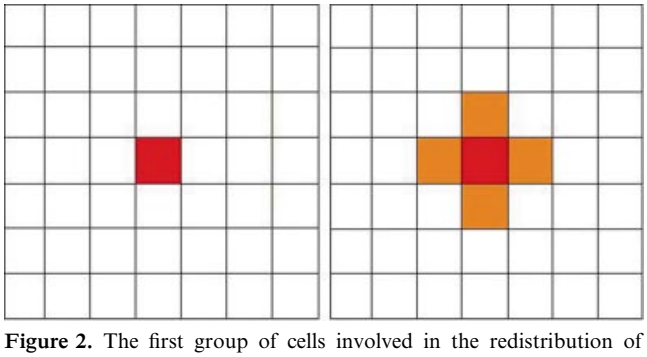

Figure 2. The first group of cells involved in the redistribution of ruptini from an originally unstable element, marked in red.

3.1.2 Results

Defining a cluster as a series of ruptures with a non-event preceding and following it, we observe in 80–90 per cent of cases clustered activity, that is, events occur in groups usually increasing and then decreasing in size according to the classical foreshock–main shock–aftershock pattern of earthquakes (Fig. 4). However, we have verified that the foreshock–main shock–aftershock pattern occurs only in approximately 6–7 per cent of cases if no time-dependent strain diffusion is allowed in accordance with the original Bak & Tang (1989) model.

The percentages of main shocks without foreshocks and aftershocks are shown in Table 1 for grids larger than 20×20 cells, which, due to border effects, is the minimum size for obtaining stable results.

Both cumulative and non-cumulative frequency–size distributions for all shocks and for the three different kinds of shocks (fore-, main and after-) separately follow a gamma-type distribution on a log–log scale (Fig. 5). In more restricted ranges we

Table 1. Percentages of main shocks without fore- and aftershocks.

	Clusters without foreshocks (%)	Clusters without aftershocks (%)
Stationary conditions	40.1 ± 3.3	29.9 ± 2.9

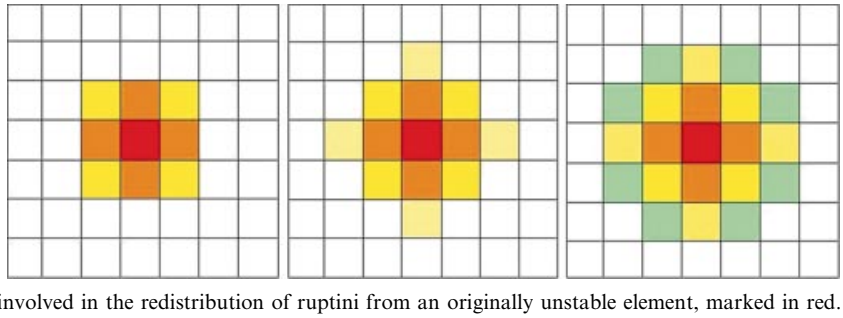

Figure 3. Successive cells involved in the redistribution of ruptini from an originally unstable element, marked in red.

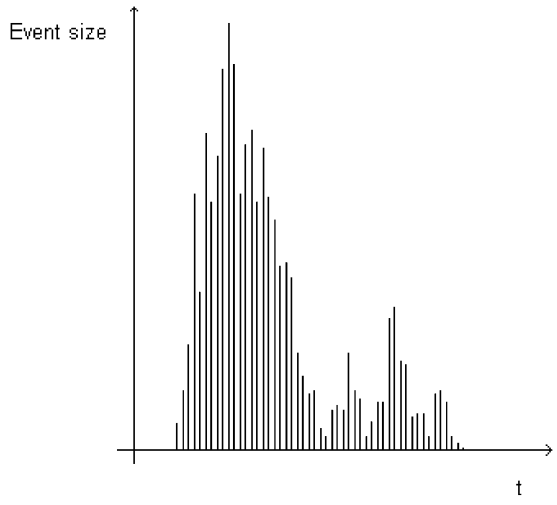

Figure 4. Typical cluster of events composed of foreshocks with an exponential size growth, a main shock and a tail of aftershocks.

observe a generally linear relation between the logarithm of the frequency of events and the logarithm of their size. Obvious deviations from linearity occur for the events with small and large dimensions compared to the grid area.

Changing the number of ruptini randomly dropped at each step gives a single cluster of events. The frequency–size distribution of all the events becomes bell-shaped (see Fig. 6), showing a tendency to cluster around characteristic sizes that increase with the number of particles added. Small-sized events tend progressively to disappear. If one cuts the temporal sequence of events with a threshold, similar to mimicking an instrument with lower sensitivity, several clusters are again obtained and the frequency–size distribution is very similar to those related to the injection of a single ruptino, suggesting that the system dynamic has not changed.

In these experiments we also find a general tendency for the number of foreshocks and aftershocks (when they exist) to increase linearly with the size of the main shock. This relation holds well for big grids, while the scatter in the data relative to lattices with fewer than 30×30 cells shows less clear behaviour due to border effects. As is apparent in Fig. 7, the slope of the curve that relates the number of fore- and aftershocks to the dimension of the main shock increases with grid dimension. This appears intuitively reasonable because the larger the grid, the easier it is to find other smaller events beyond the main shock since more cells lie near the critical value.

We have so far dealt with systems originating from heterogeneous grids. If we start with an empty grid, however, we observe that, apart from very small events, the first rupture

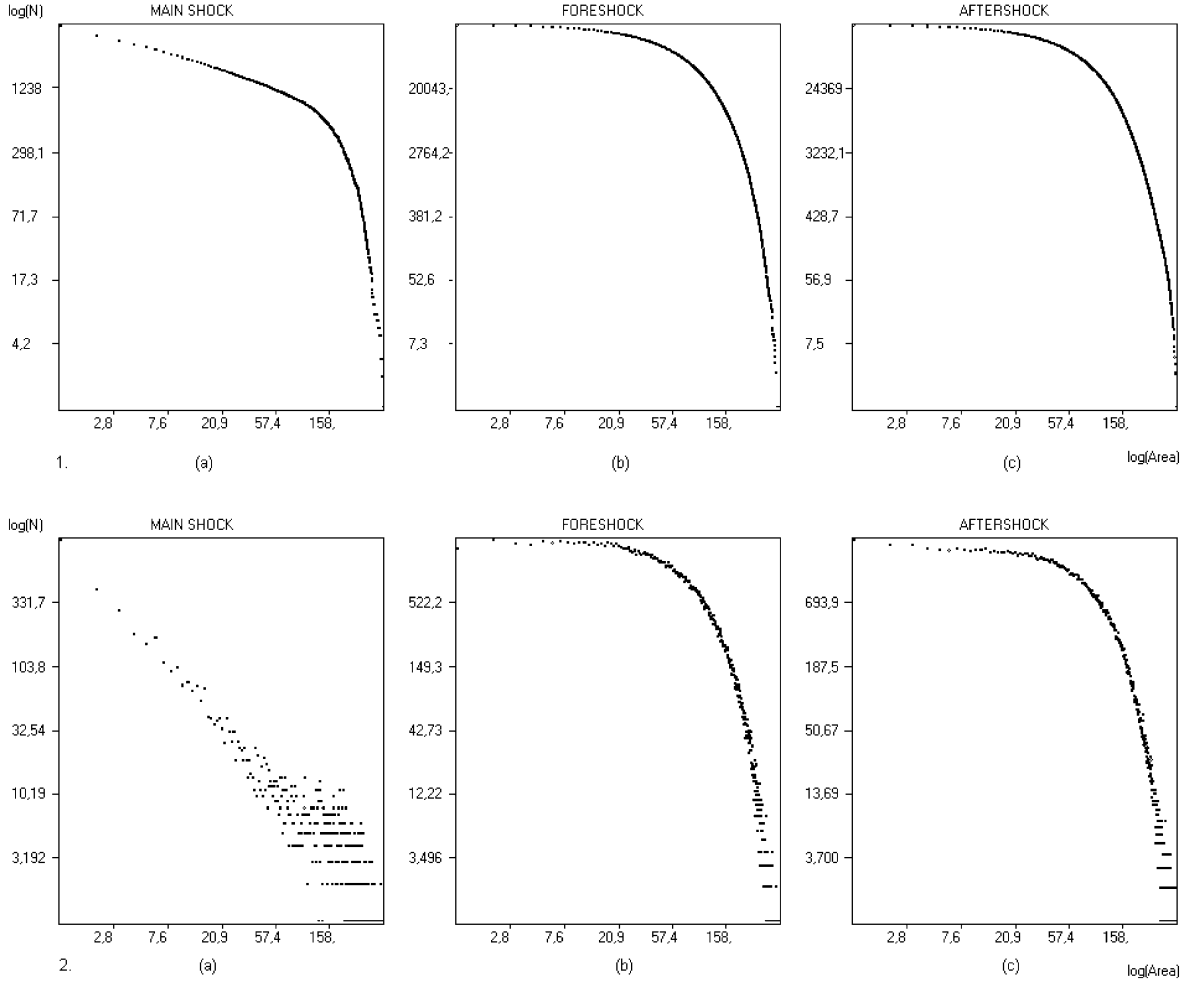

Figure 5. Cumulative (1) and non-cumulative or incremental (2) distributions of (a) foreshocks, (b) main shocks and (c) aftershocks that occurred in a 100×100 grid after 300 000 iterations.

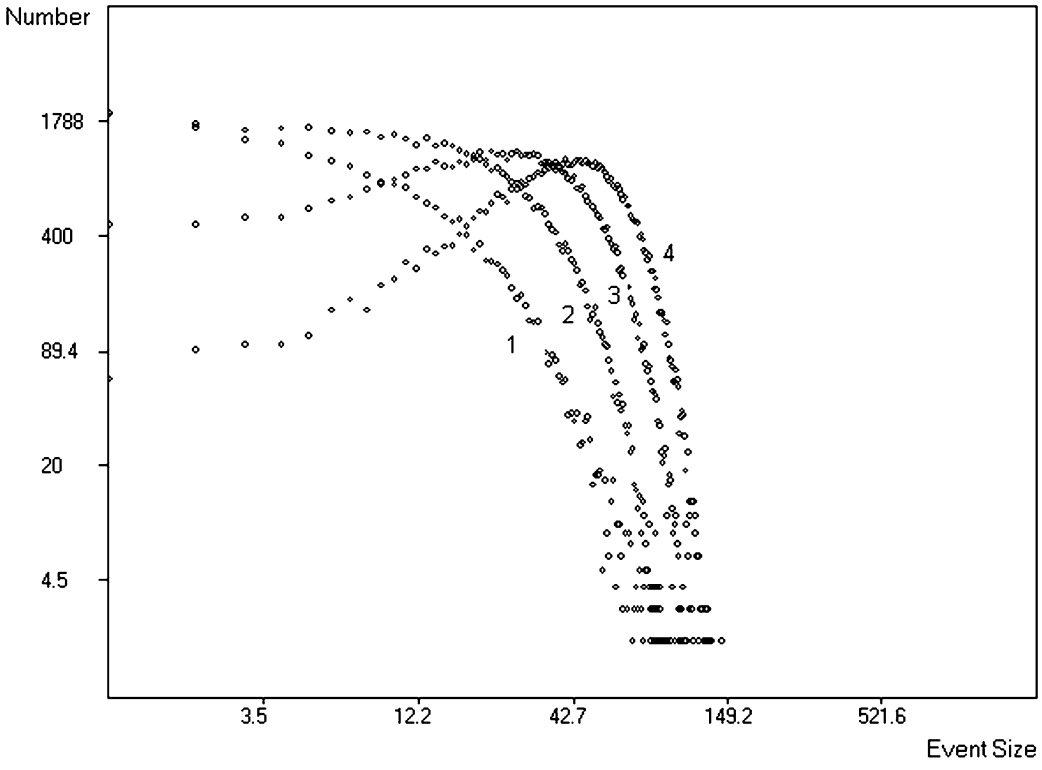

Figure 6. Distributions of events in a 30×30 grid for different amounts of ruptini injected relative to 50 000 iterations. Curve 1 refers to the injection of a single ruptino for each iteration, curve 2 to three ruptini, curve 3 to six ruptini, curve 4 to nine ruptini.

has a very large size compared with the other events under stationary conditions. This effect is due to the fact that, at the beginning, many cells will reach levels close to the critical point simultaneously and the instability reached by one of them, with subsequent redistribution, will trigger instabilities in many of the surroundings cells. After a certain number of iterations, the evolution of our automaton becomes stationary. In particular, we have verified that stationarity is reached when the net input flow of ruptini, namely the amount of ruptini injected into the grid, is equal to the net output flow, namely to the amount of ruptini lost off the grid edges. In Table 2 we report these input and output flows, averaged over 50 000 iterations after the initial transient, for different grid dimensions and for different loads (1 ruptino and 10 ruptini added for each iteration), showing that stationarity and stability are achieved.

Stationarity is approached with different velocities according to the specific transition rules adopted. Non-stationarity, on the other hand, refers both to the non-constant average dimension of the events and to the non-constant shape of the clusters.

Table 2. The typical output flow of ruptini, averaged over 50 000 iterations under stationary conditions, for different grid dimensions and different amounts of load.

Grid side	Input flow	Output flow	Input flow	Output flow
10	1	1.004	10	10.003
20	1	1.010	10	10.005
30	1	1.001	10	9.996
40	1	0.989	10	10.016
50	1	0.996	10	10.005
60	1	0.993	10	10.003

3.2 Time-dependent loads

Dynamic triggering of earthquakes is an important effect that has so far only been 'discovered' empirically in some *in situ* cases (*cf.* Harris 1998) but never studied through cellular automata models. In order to model a wave travelling through a heterogeneous lithosphere we periodically add a sequence of ruptini to successive rows of the specimen. The passing wave can also be interpreted as a periodic lowering in the threshold, induced by a travelling water intrusion, a mechanism often suggested as important in triggering crustal earthquakes.

Note how the systems described previously had borders open only in the outgoing direction, allowing just the loss of particles. The borders we consider here are open in both directions. The passing waves are recognizable from the clusters they produce (Fig. 8), while the distribution function does not change with the number of ruptini injected into the grid (Fig. 9).

3.3 Local dissipation

When a fracture occurs, a certain amount of energy is lost as frictional heat, elastic radiation and plastic deformation. To take into account this local dissipation also, we modify the transition rules of our automaton by letting it redistribute to the neighbours an amount of particles smaller than that lost from the unstable element.

A local dissipation of ruptini, starting from an initially empty grid, produces different patterns of behaviour before stationarity is reached. The temporal evolution of the system shows a first big event (its origin as described in Section 3.1) followed by some smoothing periodic clusters and a final stationarity (Fig. 10).

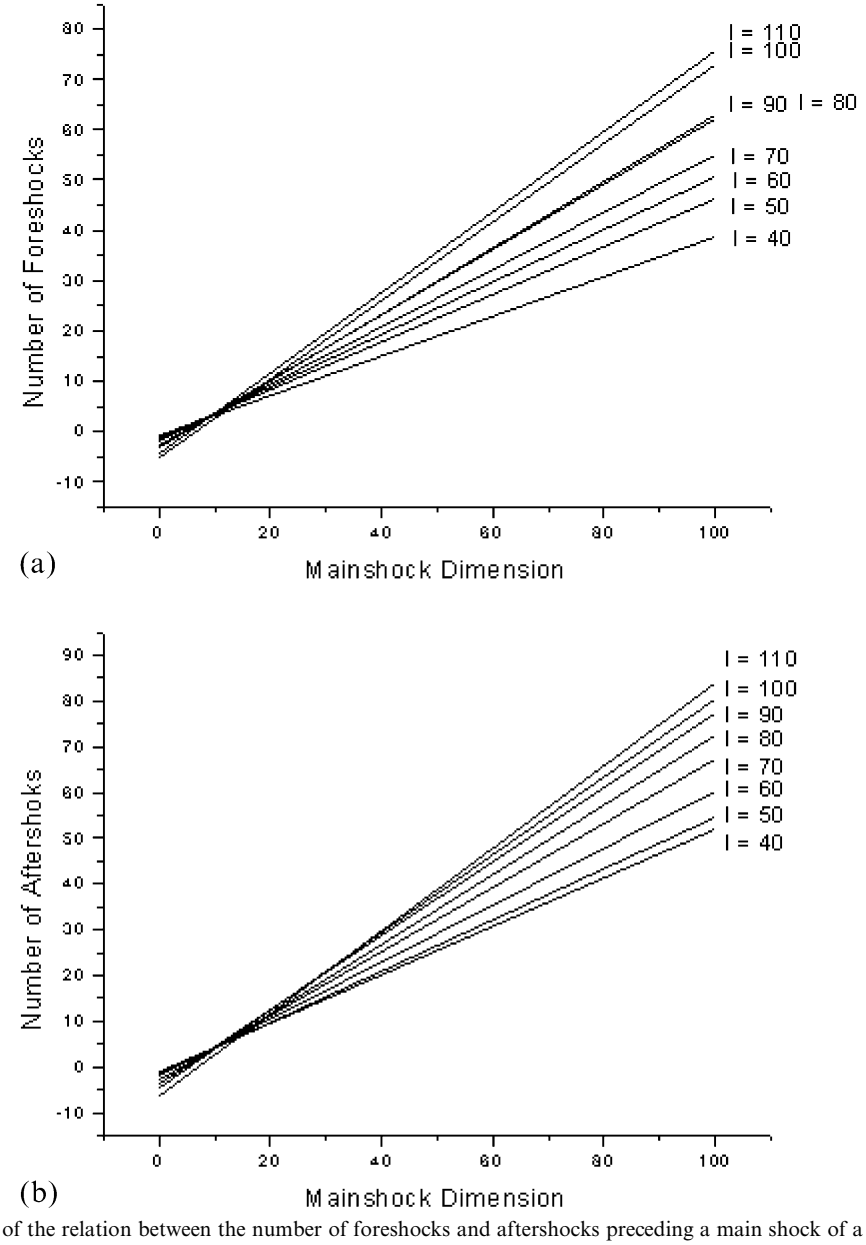

Figure 7. Interpolations of the relation between the number of foreshocks and aftershocks preceding a main shock of a specified area for different lattice dimensions. I is the number of cells in each side of the grid. Errors associated with the quantification of the angular coefficients are lower than 5 per cent and all the statistical tests performed show that the linear fit is a good model for these data.

The average size of the events is obviously smaller than that observed for analogue grids under non-dissipative conditions and the stationary condition is reached after a higher number of iterations. A periodicity is apparent, clearly due to the strain 'shadows' left behind by the large events, similar to those found by Herz & Hopfield (1995) and Main *et al.* (2000).

If we call f the fraction of energy lost by local dissipation, we expect that the period T will increase with decreasing f, following the rule $T=f^{-1}$. In fact, in the case of full strain transfer and SOC (conservative systems), $f=0$ and $T\rightarrow\infty$ so that periodicity disappears. In the case of no strain transfer, $f=1$, the period is $T=1$. The initial random distribution of loads is conserved and leads to a repetition after each full loading period. In the other cases of partial strain transfer, we consider that a cell that nucleates a large cascade will probably cause its neighbours to fail, each of which returns $f/4$ ruptini to the initial cell. This cell will then generate another event after a period proportional to f. We should note that this periodicity is observable only at the beginning, before stationarity is reached. In the other cases, when the grid is highly heterogeneous and there is a strong random load, the memory of strain shadows is destroyed by random fluctuations.

As is seen from Fig. 11, there are two other factors worthy of note:

(i) there exists a direct proportionality between the grid area and the observed periodicity for every amount of dissipation investigated;

(ii) there also exists a direct proportionality between the fraction of energy lost and the related period, so that if the dissipation of a number q of ruptini at each iteration leads to a period T, the dissipation of $2q$ ruptini results in a doubled period, $2T$, and so on.

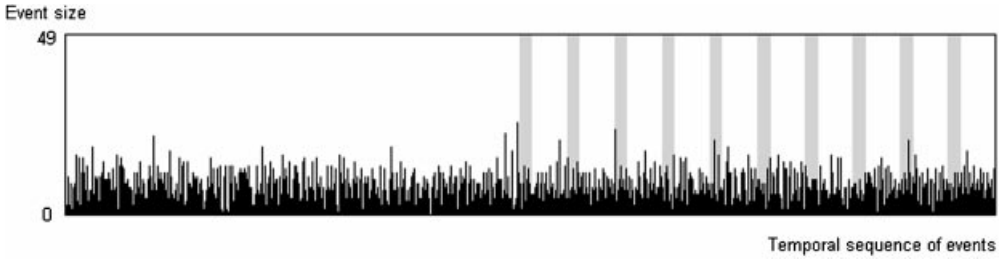

Figure 8. Time-dependent loads: the passing waves of ruptini (occurring during the time intervals shown as grey bands) are recognizable from the cluster they originate.

Let us now compare the sizes of the events, under stationary conditions, in the globally and in the locally dissipative cases. While the size of the largest events in the conservative automaton is limited only by the size of the system (Fig. 12 and Table 3), the size of the largest events in the non-conservative system is limited by the losses during the cascade process and, in particular, it is constant for different grid dimensions. In the non-dissipative case, we also note an almost linear relation between the grid area and the average dimension of the main shocks. The deviation from linearity involves only the smallest grids, with fewer than 40×40 elements.

Table 3. Typical dimensions of the maximum events under stationary conditions for locally dissipative systems with different amounts of dissipation. f is the fraction of energy lost and q is the corresponding amount of ruptini lost from the cell every time it becomes unstable.

Grid side	$f=16\%$ $q=2$	$f=33\%$ $q=4$	$f=66\%$ $q=8$
30	30	21	15
50	30	23	15
60	35	25	16
90	34	25	15

In the case of local dissipation, the number of main shocks without foreshocks and aftershocks changes consistently with respect to the percentages found under the non-dissipative conditions. In Table 4 we report these data relative to different amounts of dissipation under stationary conditions. The number of main shocks without foreshocks and aftershocks increases with the fraction of energy dissipated because dissipation implies a reduction of the strain level of the material around the unstable cells.

Note that the percentage of main shocks without foreshocks is always much higher than that without aftershocks, as is effectively observed in natural occurrences, where the values range between 70 and 80 per cent and 10 and 20 per cent,

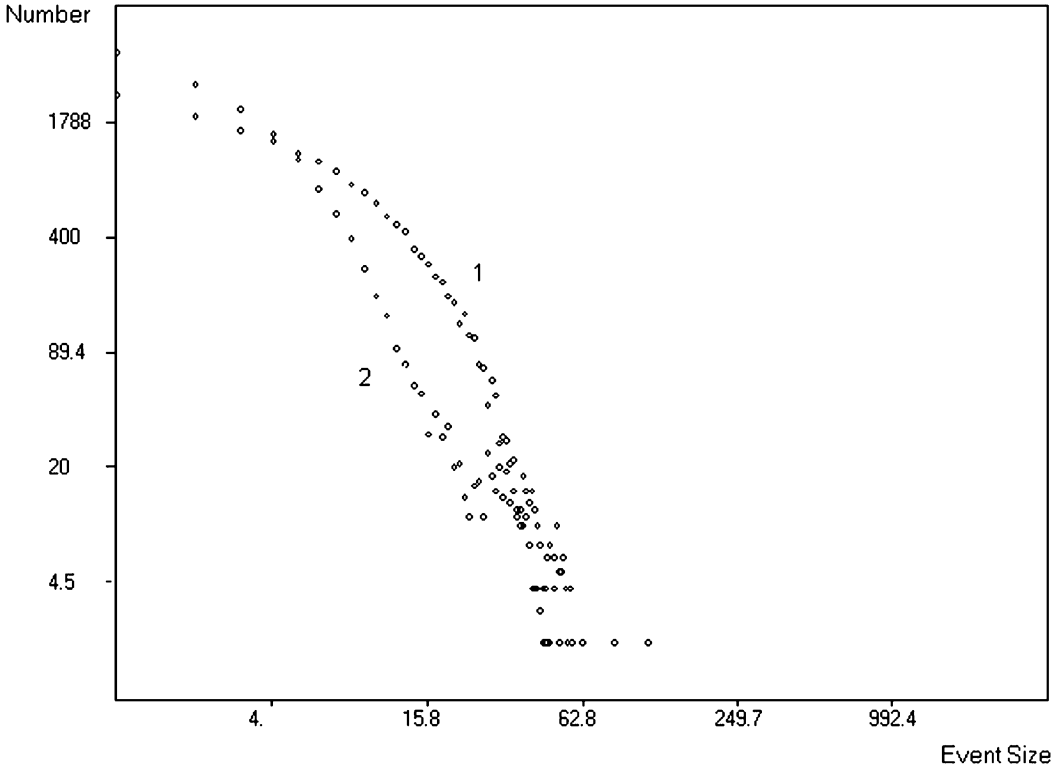

Figure 9. Distributions of events in a 20×20 grid for different amount of ruptini injected in a fixed position relative to 50 000 iterations. Curve 1 refers to a single ruptino injected for each iteration and curve 2 to a passing plane wave with an amplitude equal to 10 ruptini for each element of the grid.

Table 4. Percentages of shocks without fore- and aftershocks under local dissipative conditions.

f (%)	Clusters without foreshocks (%)	Clusters without aftershocks (%)
16	51.2±0.5	27.3±0.3
33	61.2±0.3	33.3±0.1
66	70.2±0.5	41.0±0.5

respectively (Kasahara 1981; Von Seggern *et al.* 1981). We should also emphasize that the actual percentages are strictly connected to the sensitivity of seismometers and to the definition of foreshocks and aftershocks.

3.4 Scaling properties

Recalling that fractals are objects that exhibit similar structures over a range of length scales for which one can define a non-integer dimension, we have investigated the localization of the rupture events in the grid with the aim of finding whether or not its geometry is fractal. There are different procedures to evaluate the fractal dimension. We have applied the box-counting algorithm by dividing the grid into a number of boxes of increasing dimension, ε, and calculating the number of boxes, N, in which at least one event took place (Gonzato *et al.* 1998, 2000). If fractality exists, there is a linear relation between the two quantities on a log–log scale. In fact,

$$N \propto \varepsilon^{-D}, \tag{5}$$

$$\log N \propto -D \log \varepsilon, \tag{6}$$

where D is the fractal dimension of the phenomenon. Here, the scaling range is limited from below by the size of the basic building blocks (the cells) from which the system is composed and from above by the system size. Since the grids we used were seldom larger than 100×100 cells, we could investigate the scaling properties of the system only for two orders of magnitude, both in the case of a random addition of ruptini and in the locally dissipative case. Fig. 13 shows that a linear relation effectively exists between the box dimension and the number of boxes with at least one event on a log–log scale in a scaling range of two orders of magnitude. Although a scaling range over such a small range is unlikely to have much significance *per se* (Malcai *et al.* 1997), the rules are independent of the grid size and this fractal behaviour is likely to be indicative of a real scaling.

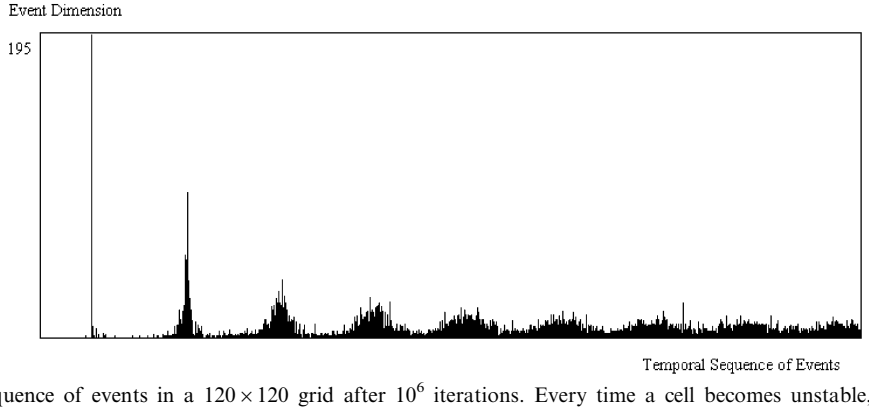

Figure 10. Temporal sequence of events in a 120×120 grid after 10^6 iterations. Every time a cell becomes unstable, eight ruptini are lost as frictional heat.

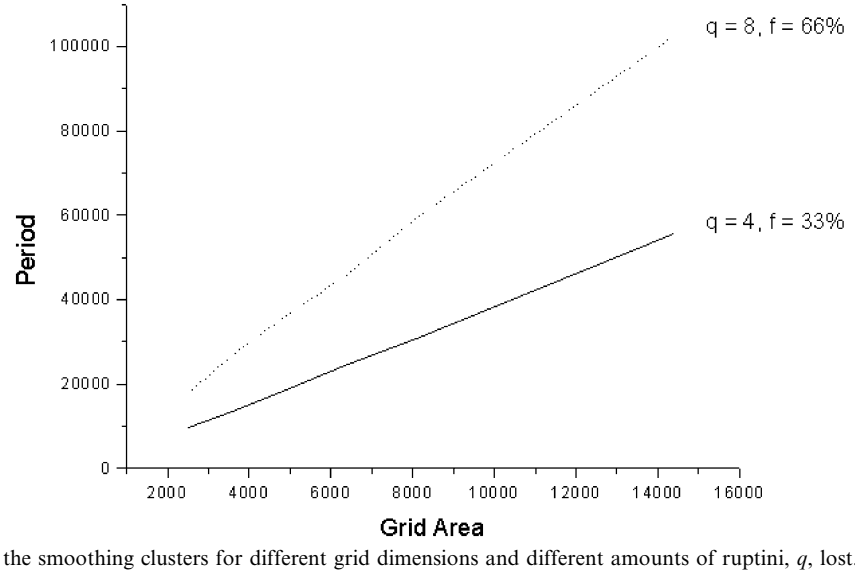

Figure 11. The period of the smoothing clusters for different grid dimensions and different amounts of ruptini, q, lost.

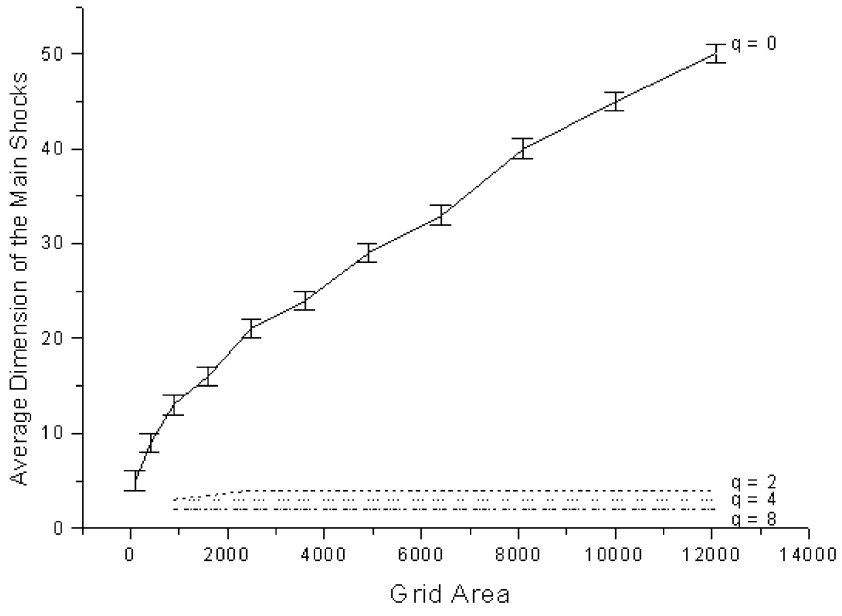

Figure 12. Average dimensions of the main events under stationary conditions versus grid dimension for locally dissipative systems with different amounts of dissipation.

4 CHANGING THRESHOLDS *IN ITINERE*

We know from rock mechanics that if a certain element has a rupture threshold, say R, before the first rupture, after the element has come to fracture this threshold will approach a residual value $r < R$. This is somewhat like saying that instead of nucleating a new fault, reactivating an old one is generally easier, or like saying that after the first rupture has occurred, the fault surface has become smoother and slips more easily. In order to take into account this aspect in our simulations, we have developed an automaton with a memory of the cells in which instabilities did occur by diminishing in these cells the rupture threshold relative to the successive iterations.

As can be seen in Fig. 14, fractures grow around the elements that developed instabilities at the previous stages. A large fracture develops joining smaller weak areas. If these threshold changes have to be taken into account for the lithosphere, its stationarity seems still more unlikely, since it would imply that the whole crust has broken at least once, achieving a new, lower threshold. In Table 5 we show the dynamical properties of this automaton relative to globally dissipative and locally dissipative—with different levels of dissipation—cases.

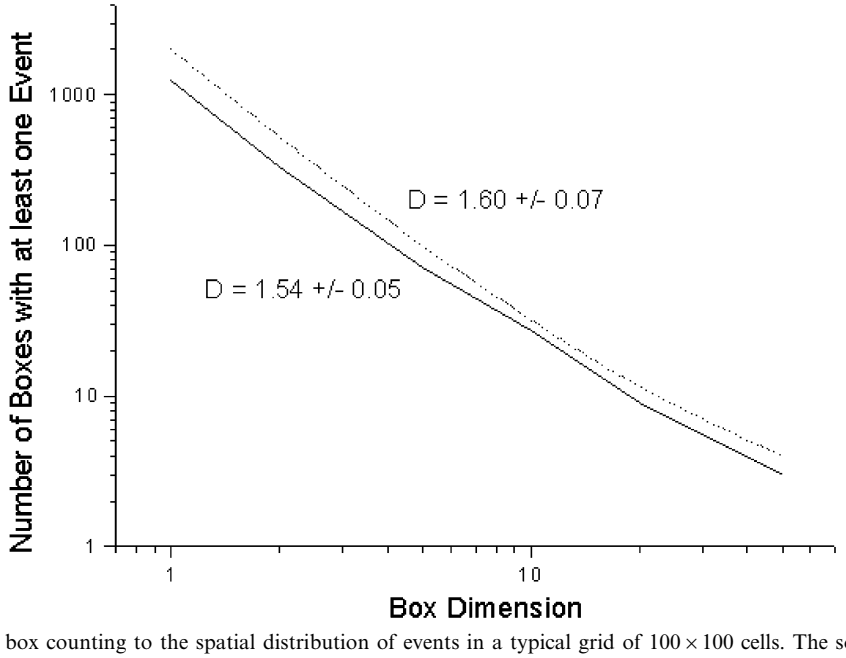

Figure 13. Application of box counting to the spatial distribution of events in a typical grid of 100×100 cells. The solid line (fractal dimension $D = 1.54$) comes from a random load algorithm and the dashed line (fractal dimension $D = 1.60$) from a locally dissipative system with $q = 2$; curves with greater dissipation ($q = 4$ and $q = 8$) are not distinguishable from it in this scaling range.

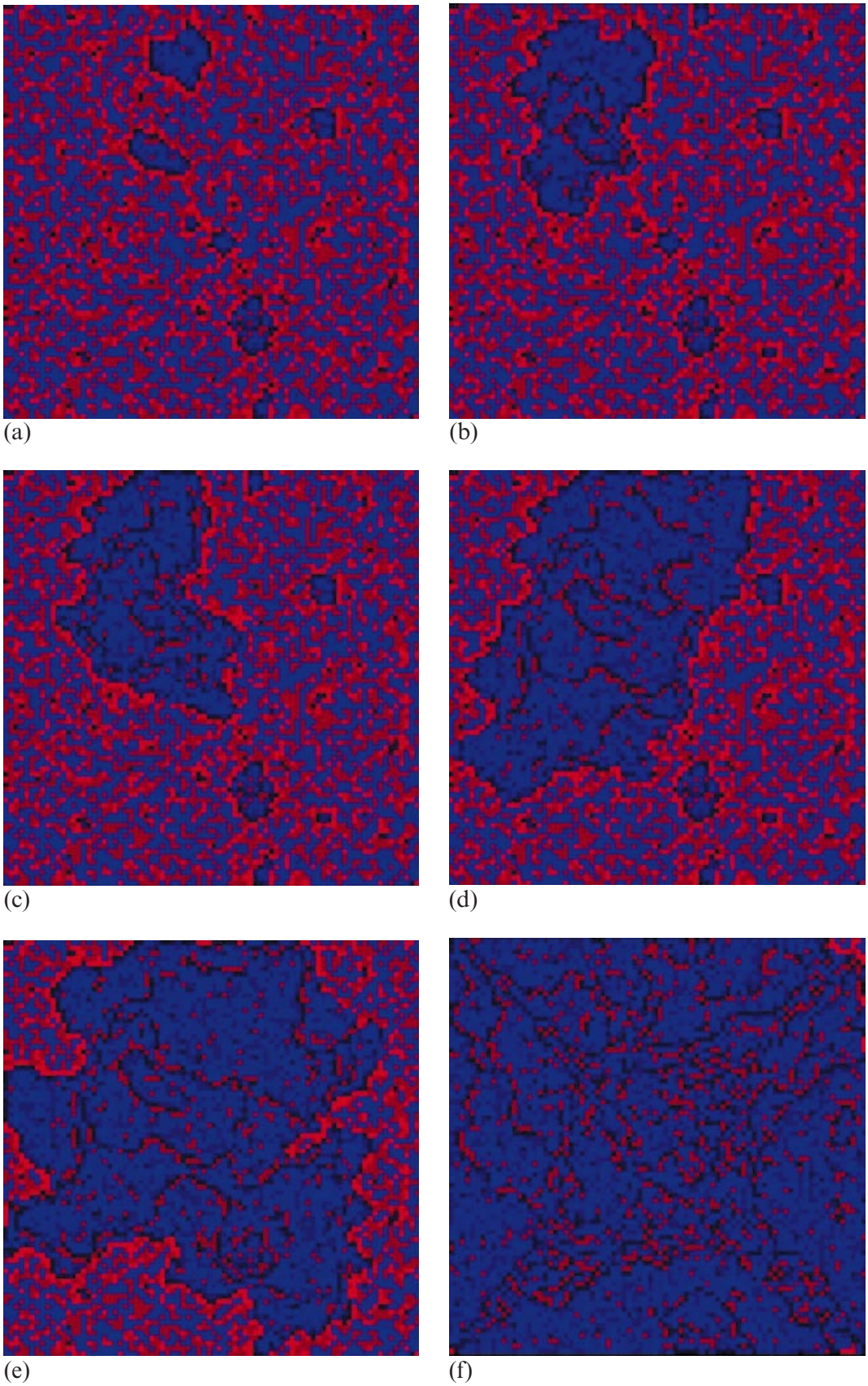

Figure 14. Evolution of a system undergoing loading applied to random locations. After the first rupture, the fracture threshold decreases. The distance of each element from rupture is represented by a scale of colours ranging from blue (stability) to red (instability) after (a) 36 300, (b) 36 350, (c) 36 400, (d) 36 450, (e) 36 500 and (f) 36 600 iterations.

Table 5. Percentages of clusters without foreshocks and without aftershocks for different threshold variations and for different amounts of local dissipation. As the percentages do not change significantly with the degree of dissipation, the complete set of data is given, as an example, only for the case of threshold reduction equal to 60 per cent. In the other cases the variation is within the standard deviation value.

Threshold reduction after the first rupture (%)	Dissipation degree (f)	Clusters without foreshocks (%)	Clusters without aftershocks (%)
20	general	79.0 ± 0.5	40.0 ± 0.5
30	general	77.0 ± 0.5	39.9 ± 0.6
60	0%	66.4 ± 0.2	33.9 ± 0.4
60	16%	66.2 ± 0.2	33.2 ± 0.2
60	33%	66.4 ± 0.2	34.2 ± 0.3
60	66%	66.2 ± 0.1	34.4 ± 0.4

In this automaton, the percentage of events not preceded by foreshocks and/or not followed by aftershocks is not affected by the amount of dissipation but it changes with the degree of threshold reduction, always remaining quite close to that observed in nature in the case of foreshocks. Comparing these results with those in Table 1, where there is no dissipation, one would expect similar percentages, at least in the case in which the threshold reduction factor is low. This is not what happens because the experiments performed in this section differ from those performed in Section 4.1, since the rules for the distribution of ruptini to the neighbours of an unstable cell are different. Referring to the symbols of Section 4.1, here we have adopted

$$\begin{aligned} &\text{if } N(i,j) \geq \tau \quad \text{then} \\ &N(i,j) = N(i,j) - \tau\,, \\ &N(i\pm 1, j) = N(i\pm 1, j) + \frac{\tau}{12}\,, \\ &N(i, j\pm 1) = N(i, j\pm 1) + \frac{\tau}{12}\,, \\ &t = 1,\ N(i\pm 1, j\pm 1) = N(i\pm 1, j\pm 1) + \frac{\tau}{12}\,, \end{aligned} \tag{7}$$

while the other equations do not vary. It has been necessary to smooth the ruptini redistribution law, otherwise redistributing too many ruptini to the nearest cells, together with the threshold reduction, leads to a single superevent that is impossible to stop. It follows that what is important to draw from Tables 1, 4 and 5 is not the absolute values of the percentages because they vary according to the redistribution rules followed. This represents the general behaviour of the system, that is, in every kind of automaton aftershocks occur in 35–40 per cent more cases than foreshocks.

The dimensions of the largest events, once stationarity has been reached, are not controlled by the grid area but, and only very slightly, by the degree of threshold lowering.

5 UNIVERSALITY

When the exponents describing the behaviour of a system do not depend on the parameters of the model, a property of universality is concluded. One can look for this property by comparing the distributions of the events that occur in systems with the same size and the same number of iterations developed according to different rules (Christensen & Olami 1992; Kadanoff *et al.* 1989). In our automata, we observe that the universality of the critical exponents exists only in the conservative cases, while we shift to different classes of universality by changing the non-conservation degree. For a similar discussion relative to the Olami–Feder–Christensen model, see also Carvalho & Prado (2000). Fig. 15(a) shows the distribution for all the events

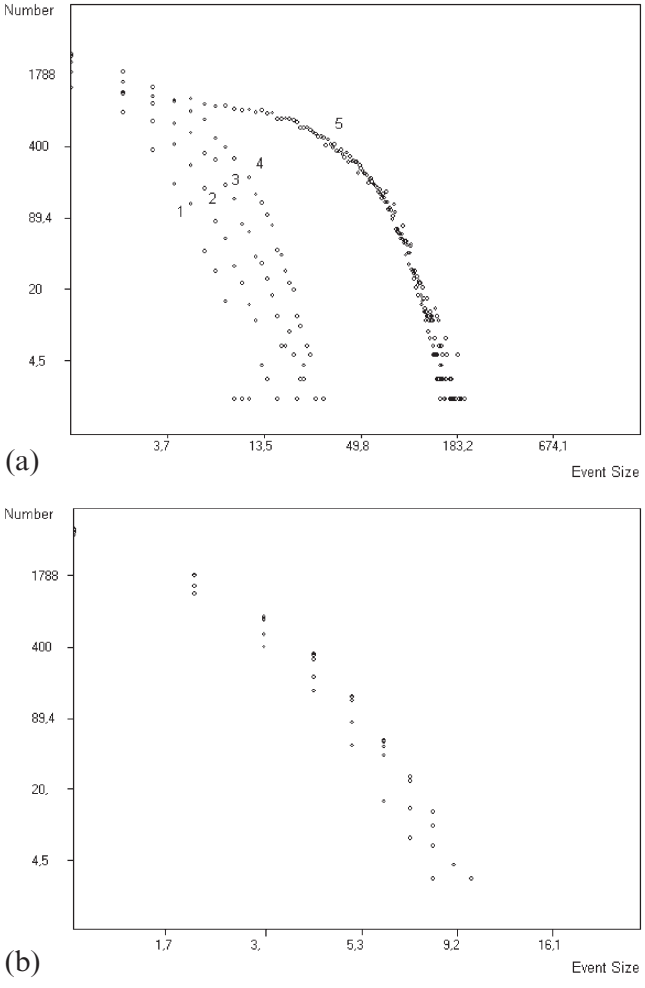

Figure 15. (a) Distributions relative to 50 000 iterations performed on grids of 60×60 particles. Curve 1 refers to the dissipation of $q = 8$ ruptini, curve 2 to $q = 4$, curve 3 to $q = 2$, curve 4 to $q = 1$ and curve 5 to $q = 0$, the conservative case. (b) Distributions relative to the same system as in (a) but with a fracture threshold reduction after the first ruptures. The circles enclose the experiments run with $q = 0$, $q = 1$, $q = 2$, $q = 4$ and $q = 8$ but their distributions are not distinguishable.

that occurred in a 60×60 grid. The system organizes itself into a stationary state after a transient time, dependent on the level of conservation. Here we analyse the statistics only after the system is organized. It is seen that the slopes of the curves become steeper as the dissipation degree increases. We compare this case with the same plot made for a system in which the threshold level changes after the first rupture (Fig. 15b). In this case, the universality of coefficients is preserved. There is a big difference in slope between the conservative and the dissipative cases but, among the latter, the difference is slight.

6 UNSOLVED PROBLEMS

In our automaton we have a generally linear behaviour in the central part of the plot of the frequency–size distribution that is reminiscent of the Gutenberg–Richter power law for real earthquakes cited in Section 2. A basic question is whether any of these driven non-equilibrium threshold systems exhibit similarities to equilibrium systems. In fact, if these models possessed a stable, time-averaged energy distribution function, standard techniques and methods of equilibrium statistical mechanics could then be available for use in the analysis of simulation results and the interpretation of system dynamics (Rundle *et al.* 1997).

As is seen in Fig. 16, our lattice models possess a stable energy distribution. In fact, the average number of ruptini in the grid during the temporal evolution of the system shows a Gaussian distribution as the model approaches stationarity. As Rundle *et al.* (1995) pointed out, this line of reasoning does not depend on the massless nature of the elements considered.

There are a number of questions that our automaton leaves open. The first problem is connected to the temporal meaning of each iteration. Our algorithm works according to two different timescales:

(i) when an element becomes unstable, a sudden redistribution of ruptini takes place and, at the same time, the rupture event increases its area until unstable cells exist;

(ii) the system keeps a memory of the elements that were unstable in the previous steps, and the second timescale involves slow time-dependent strain diffusion, that is, a redistribution of particles from those elements.

Note that the cascades are assumed to be simultaneous since the timescale associated with an earthquake is short compared to the timescale associated with the loading or change in threshold level, represented by ruptini injection. It is difficult to derive a correspondence to these timescales for the Earth, especially the second one. The first should be of the order of a few seconds; the second is much more scattered and should range from a few minutes to many years.

Another limitation of our simulation is that it cannot resolve changes in the rate of occurrence of the events. For instance, it cannot say if there is an acceleration in the succession of occurrences before a main shock simply because we consider foreshocks as events that precede the maximum of each cluster. This means that foreshocks occur at successive iterations in the loop of the programme and the elapsed time between two successive iterations is linked to the rate of particle addition. If we consider the rupture propagation to be instantaneous, we let events take place only at the time step given by the rate of particle injection. The quantification of this rate does not derive from the evolution of the automaton but is imposed from outside.

Another problem, typical of all of these types of models, involves the limited dimension of grids achievable by computers. In principle, it is possible to work with large grids, in three dimensions, modelling realistic portions of the lithosphere, but runs on realistic models of a few hundred thousand cells are demanding even for the most powerful parallel machines.

Finally, as we mentioned in the Introduction, we have little reason to exclude the idea that the actual lithosphere may be subject to continuous activity, that is, ruptures occur all the time, as our automata have shown when loaded with many ruptini at the same time. However, our instruments may only detect a portion of this ground activity due to sensitivity problems and spatial coverage.

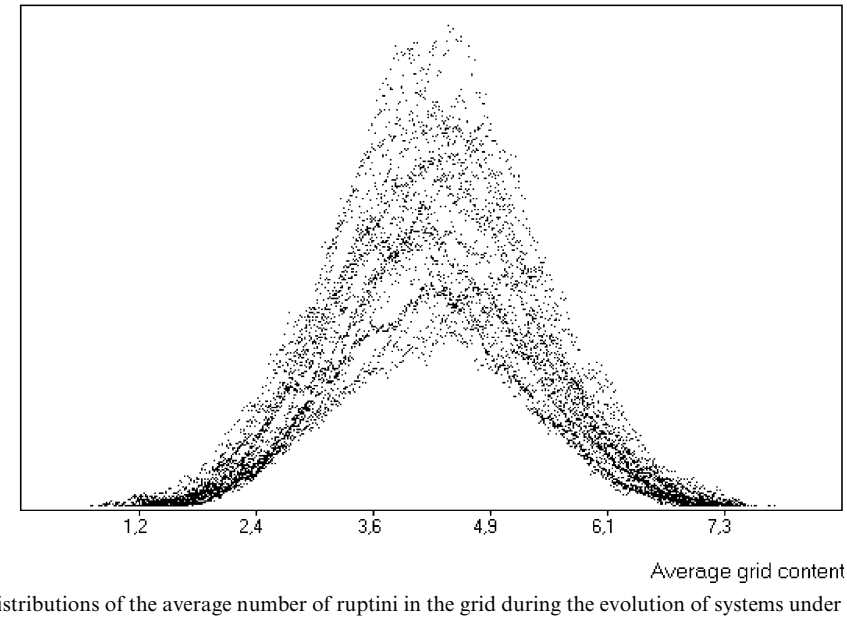

Figure 16. Normalized distributions of the average number of ruptini in the grid during the evolution of systems under stationary conditions. These curves, which fit Gaussian distributions, come from automata of 30×30 cells with different amounts of ruptini (from $q=1$ to $q=8$) locally lost as frictional heat.

7 MEASURING THE COMPLEXITY AND THE PERFORMANCE OF CELLULAR AUTOMATA MODELS

An advantage of the complexity approach we have adopted is that it does not necessarily require huge systems of equations. However, in this case it is difficult to find a criterion to measure the complexity of the model. One can consider the complexity of a cellular automaton model proportional to the number of variables in the algorithm even if this does not help in stating the significance of the parameters. We have called these quantities variables simply because this is their normal name in computer language. Obviously, all the variables that have no effect on the behaviour of the model can be disregarded in the calculation. In this way we can measure the complexity of our model with respect to the three main types of cellular automata proposed so far as analogues of distributed seismicity.

The first slider-block model of any kind was proposed by Burridge & Knopoff (1967). This was a spring-mass oscillator model for which the differential equations were solved. The first massless cellular automaton model was that of Rundle & Jackson (1977). Carlson & Langer (1989) reanalysed the Burridge–Knopoff model by using many more elements. Other variations on the slider-block automaton theme were explored by Ito & Matsuzaki (1990), Nakanishi (1991), Brown *et al.* (1991), Rundle & Brown (1991), Olami *et al.* (1992), Carlson *et al.* (1994), Morein *et al.* (1995) and Narkounskaia *et al.* (1992) with the aim of describing the organization properties of earthquakes. The original Burridge–Knopoff model consists of a lattice of blocks and springs that is pulled across a rough surface, undergoing stick-slip motion. Since each block is connected by springs to the four surrounding blocks, the slippage of a block (i, j) increases the force acting on the four neighbouring blocks $(i \pm 1, j)$ and $(i, j \pm 1)$, and, if the stability criterion is violated, it makes them move. This model would require the simultaneous solution of coupled equations of motion for all the moving blocks. In order to simplify the calculation, only one block at a time is allowed to slip and the system is treated as a cellular automaton. As listed in Tables 6 and 7, this model and its variants need from 13 to 16 variables, seven to eight of which describe the geometrical and physical properties of the elements and the remaining six to nine the transition rules.

The second type of earthquake automaton model was pioneered by Bak & Tang (1989) to establish the conditions that characterize SOC. Bak & Tang's model considered a 2-D array of particles on a square lattice, representing segments of a sliding surface. The particles are subjected to a force from their neighbours that adds to a constantly increasing 'tectonic' driving force. Forces are conserved except at the boundaries. When the total force on a particle exceeds a threshold, the particle slips to a nearby position. The energy release from one particle may lead to instability at a neighbouring position, in which case another redistribution takes place. This cascade process is the earthquake and eventually the average force will reach a statistically stationary value. This model requires only seven variables, only one for its geometrical description and the other six to explain its transition rules. Its main limitation is that it does not allow aftershocks to occur because it does not take into account any sort of time-dependent propagation of energy to the neighbours. After a certain number of iterations, the system reaches a state of SOC and in this state the distribution of events follows a power law, similar to the Gutenberg–Richter law observed in nature.

The third type of model is that of Barriere & Turcotte (1994), who considered a grid of boxes with a fractal distribution of sizes arranged in given patterns of spatial disposition. Particles (representing energy, stress or strain) are randomly added to the boxes until a critical value is reached and a redistribution to the adjacent boxes occurs. The grid loses particles from edges and corners and since there is a continuous input of particles, it is essential for the system to be dissipative, otherwise only a single macro-cluster of avalanches would take place. The authors considered as a foreshock a redistribution from a small box that triggers an instability in a larger box (the main shock). A redistribution from a large box always triggers instabilities in adjacent smaller boxes (aftershocks). Also in this case, the frequency–size statistics for both main shocks and aftershocks satisfies the Gutenberg–Richter relation. Foreshocks occur 28 per cent of the time, in agreement with the measured seismicity, but this result depends on the specific pattern of arrangement of the boxes of different size. With respect to Bak & Tang's model, this one requires two further geometrical variables (Table 8).

Note that while the fractal grid in the model of Barriere & Turcotte was intended to represent a projection of the fault

Table 6. Description and computation of the variables necessary to implement different cellular automata. Bak & Tang's model is the automaton with a homogeneous grid of cells, Barriere & Turcotte's that with a fractal grid, and Kadanoff *et al.*'s is the classical sand-pile model.

Description of the variables	Bak & Tang (1989)	Barriere & Turcotte (1994)	Kadanoff *et al.* (1989)	
			Limited model	Unlimited model
Grid dimension	1	–	1	1
Fractal grid	–	3: side, order of the model (f)	–	–
Number of iterations	1	1	1	1
Rupture thresholds for the cells	1	1	1	1
Amount of stress injected	1	1	1	1
Position for the stress injection	1	1	1	1
Order of the redistribution	1	1	1	1
Law of redistribution	1	f	1	depends on slope
Boundary condition	free	free		1
Total	7	$8+f$	7	7

Table 7. Description and computation of the variables necessary to implement different slider-block cellular automata.

Description of variables	Ito & Matsuzaki (1990)	Narkounskaia *et al.* (1992)			Olami *et al.* (1992)
		General model	Two-block ($n=2$) model	2-D model	
Number of blocks (n)	1	1	1	1	1
Mass of the blocks	–	1	1	1	
Elastic constant of the interconnecting springs	1	1	1	1	2
Number of interconnecting springs	$n-1$	$\binom{n}{2}$	$\binom{n}{2}$	$2\sqrt{n}(\sqrt{n}-1)$	$2\sqrt{n}(\sqrt{n}-1)$
	1	1	1	1	1
Elastic constant of the loader spring	1	1	1	1	1
Number of loading springs	1	1	1	1	1
Velocity of the driving surface	1	1	1	1	1
Frictional coefficients of the sliding blocks	Static 1	*	1	1	1
Number of iterations	1	1	1	1	1
Slip thresholds for the blocks	1	1	1	1	1
Stress release after a slip occurs	1	1	1	1	1
Redistribution rules	1	1	1	1	1
Amount of force to be distributed to the neighbours described in the previous row of this table	1	1	1	1	1
Dissipation	–	–	–	–	1
Probability $1-p$ representing plasticity necessary to stop a fracture	1	–	–	–	–
Memory of the blocks already slipped once	(Takayasu & Matsuzaki) 1	1	–	–	–
Decay of aftershock	2	–	–	–	–
Boundary conditions	periodic	free or periodic	free or periodic	free or periodic	
Total	16	15	13	13	14

* Static, dynamic 2 (the system evolution does not depend on dynamic friction, so that 1 parameter is sufficent).

system onto the lithosphere, and the network of bodies in the model of Burridge & Knopoff was intended to represent a fault surface, the lattice in our model can represent both and can also be seen as any surface, including the topographic surface of the Earth or the surface of a specimen to be fractured in the laboratory. We can consider that the ruptures observed, starting from small, localized cells and coalescing to form bigger areas, occur at different locations on a single fracture surface or that they are the epicentres of the seismic distribution, plotted on other surfaces such as those of the Earth or of a laboratory sample. If we imagine that each element represents a segment of a fracture, we simulate faults that are not planar as the real ones are. The rupture of a fault is in fact not a planar phenomenon but develops through a series of small slips resulting in rough surfaces. If we imagine, instead, that each event is the epicentre of a shock, we obtain the geographic distribution of events projected on the Earth or on a laboratory sample surface.

Our automaton, with its seven to 12 variables (Table 9), one to define the geometry of the grids and the others to set the transition rules, is approximately as simple as that of Bak & Tang, but it satisfies all the main phenomenological features of natural earthquakes, that is, a Gutenberg–Richter distribution, the ratio between the number of foreshocks and aftershocks and the scaling of foreshocks and aftershocks according to power laws. These last two features were not present in Bak & Tang's model. Moreover, it does not constrain the basic geometry to be fractal as Barriere & Turcotte assumed, but it finds this to be the case.

It is immediately clear that the number of variables employed to describe these automata may seem large with respect to classical mathematical models, but none of the latter, even the most complex ones, has ever been able to reproduce the average statistical features of seismicity, not even assuming some of them as tunable parameters (*cf.* Ward 1991).

8 CONCLUSIONS

We have attained our goal of creating a general model that is as simple as possible, albeit accounting for both strain diffusion and local dissipation of energy. Our model outperforms previous models since it is capable of reproducing all the known statistical properties of real seismicity with the smallest number of parameters and with a weak set of assumptions.

Many studies have attempted to describe the lithosphere in terms of a state of SOC. Generally speaking, there seems to exist an independence of the model from the parameters only in the conservative cases, the applicability of which is doubtful. Obtaining power (fractal) laws is a natural consequence of the evolution of this type of automata and we have described different kinds of transition rules that can explain the various forms of power-law relationships known in seismology, although

Table 8. Description and computation of the variables necessary to implement different slider-block cellular automata.

Description of variables	Nakanishi (1991)	Carlson *et al.* (1994)
Number of blocks (n)	1	1
Mass of the blocks	–	1
Elastic constant of the interconnecting springs	1	1
Number of interconnecting springs	one dimension: $n-1$	one dimension: $n-1$
	1	1
Elastic constant of the loader spring	1	1
Number of loading springs	1	1
Position for the stress injection	–	–
Velocity of the driving surface	1	1
Frictional coefficients of the sliding blocks	1	1
Number of iterations	1	1
Number of the iteration the algorithm is performing (necessary in order to evaluate the differential equations)	1	–
Slip thresholds for the blocks	1	1
Force drop after a slip occurs	1	1
Neighbours involved in the redistribution of force after a slip	1	1
Amount of force to be distributed to the neighbours described in the previous row of this table	1	1
Dissipation	–	–
Probability $1-p$ representing plasticity necessary to stop a fracture	–	–
Memory of the blocks already slipped once	–	–
Decay of aftershock	–	–
Boundary conditions	–	–
Total	13	13

they are all isotropic while fracturing is anisotropic. However, they all lead approximately to the same indefinite conclusion, that after a certain number of iterations, varying according to the specific transition rules adopted, the system becomes stationary and there is no way of knowing when and where the next event will take place and how large it will be. Thus, one may know many things about the average properties of the system, but while this leads to a sufficient description for gases, it is not satisfactory for earthquakes.

In our simulation, we have encountered three main obstacles. The first is the non-interpretability of the timescale of the distribution, which directly derives from the way in which the automata are built. The second is ignorance of the object we want to model, that is, the lithosphere and the question whether its dynamics are stationary or not. The third limitation is the grid size (Table 10), even for the most powerful computers, which casts doubts on the applicability of this kind of simulation to real cases.

Table 9. List of the variables needed to run our cellular automaton algorithms. Model A refers to the basic model, without local dissipation of energy and variations in the threshold level for rupture; model B is the model that considers also a local dissipation of energy; model C takes into account a change in the threshold level for rupture after the first fracture has occurred. $\pm$ indicates a parameter that could be omitted in that model. Refer to the text for a detailed explanation of the different models.

Description of variables	Applies to models	Quantity of variables
Grid dimension	A, B, C	1
Number of iterations	A, B, C	1
Rupture thresholds for the cells	A, B, C	1
Variation in the rupture thresholds after the first fracture	C	1
Amount of ruptini injected	A, B, C	1
Position for the ruptini injection	A, B, C	1
Amount of local ruptini dissipation	B, $\pm$C	1
Amount of ruptini distributed randomly after a rupture occurs	$\pm$A, $\pm$B, $\pm$C	*
Order of redistribution	A, B, C	1
Law of redistribution	A, B, C	1
Total	A	7+3
	B	8+3
	C	8+4

* In general, equal to the number of steps in which a complete redistribution takes place; in our model it is typically equal to 3.

Table 10. Properties affected by finite size effects.

Properties affected by size effects	What happens	Size effect disappears for grids larger than
Percentage of main shocks without fore- and aftershocks		20×20
Frequency–size cumulative and non-cumulative distributions of the shocks	deviations from linearity	
Linear increase of the relation number of fore- and aftershocks versus main-shock size	deviations from linearity	30×30
Direct proportionality between the grid area and the periodicity observed under dissipative systems		
Size of the largest events in the conservative automata is limited by the size of the system		in dissipative systems
Linear relation between the grid area and the average dimension of the main shocks	deviations from linearity	40×40

ACKNOWLEDGMENTS

We would like to thank Dr Ian Main and Prof. John Rundle for their review of the manuscript and their constructive suggestions. This work was performed with funding from the Ministero dell'Università e della Ricerca Scientifica e Tecnologica (MURST) 40 and 60 per cent.

REFERENCES

Bak, P. & Tang, C., 1989. Earthquakes as a self-organized critical phenomenon, *J. geophys. Res.,* **94,** 635–637.

Barriere, B. & Turcotte, D., 1994. Seismicity and self-organized criticality, *Phys. Rev.,* **49,** 1151–1160.

Brown, S.R., Scholz, C.H. & Rundle, J.B., 1991. A simplified spring-block model of earthquakes, *Geophys. Res. Lett.,* **18,** 215–222.

Burridge, R. & Knopoff, 1967. Model and theoretical seismicity, *Bull. seism. Soc. Am.,* **57,** 341–371.

Carlson, J.M. & Langer, J.S., 1989. Properties of earthquakes generated by fault dynamics, *Phys. Rev. Lett.,* **22,** 2632–2635.

Carlson, J.M., Langer, J.S. & Shaw, B.E., 1994. Dynamics of earthquake faults, *Rev. Modern Phys.,* **66,** 657–670.

Carvalho, J.X. & Prado, C.P.C., 2000. Self-organized criticality in the Olami-Feder-Christensen model, *Phys. Rev. Lett.,* **84,** 4006–4009.

Christensen, K. & Olami, Z., 1992. Scaling, phase transitions and nonuniversality in a self-organized critical cellular-automaton model, *Phys. Rev. A,* **46,** 1829–1838.

DeMets, C., 1995. Plate motions and crustal deformation, *Rev. Geophys. Suppl.,* **33,** 365–369.

Gonzato, G., Mulargia, F. & Marzocchi, W., 1998. Practical application of fractal analysis: problems and solutions, *Geophys. J. Int.,* **132,** 275–282.

Gonzato, G., Mulargia, F. & Ciccotti, M., 2000. Measuring the fractal dimension of ideal and actual objects: implications for application in geology and geophysics, *Geophys. J. Int.,* **142,** 108–116.

Gutenberg, B. & Richter, C.F., 1954. *Seismicity in the Earth and Related Phenomena,* 2nd edn, Princeton University Press, Princeton, NJ.

Harris, R.A., 1998. Introduction to special section: stress triggers, stress shadows and implication for seismic hazard, *J. geophys. Res.,* **103,** 24 347–24 358.

Herz, A.V.M. & Hopfield, J.J., 1995. Earthquake cycles and neural reverberations: collective oscillations in systems with pulse-coupled threshold elements, *Phys. Rev. Lett.,* **75,** 1222–1225.

Ito, K. & Matsuzaki, M., 1990. Earthquakes as self-organized critical phenomena, *J. geophys. Res.,* **95,** 6853–6860.

Kadanoff, L.P., Nagel, R., Wu, L. & Zhou, S., 1989. Scaling and universality in avalanches, *Phys. Rev. A,* **39,** 6524–6537.

Kasahara, K., 1981. *Earthquake Mechanics,* Cambridge University Press, Cambridge.

Main, I., 1996. Statistical physics, seismogenesis and seismic hazard, *Rev. Geophys.,* **34,** 433–462.

Main, I., O'Brien, G. & Henderson, J.R., 2000. Statistical physics of earthquakes: comparison of distribution exponents for source area and potential energy and the dynamic emergence of log-periodic energy quanta, *J. geophys. Res.,* **105,** 6105–6126.

Malcai, O., Lidar, D.A., Biham, O. & Avirn, D., 1997. Scaling range and cutoffs in empirical fractals, *Phys. Rev. E,* **56,** 2817–2828.

Morein, G., Turcotte, D. & Gabrielov, A., 1995. Statistical mechanics of distributed seismicity, in *Acts of the workshop 'Non-Linear Dynamics and Earthquake Prediction, 6–17 November,* Trieste, Italy.

Nakanishi, H., 1991. Statistical properties of the cellular-automaton model for earthquakes, *Phys. Rev.,* **43,** 6613–6621.

Narkounskaia, G., Huang, J. & Turcotte, D., 1992. Chaotic and self-organized critical behaviour of a generalized slider-block model, *J. Stat. Phys.,* **67,** 1151–1183.

Olami, Z., Feder, H.J.S. & Christensen, K., 1992. Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes, *Phys. Rev. Lett.,* **68,** 1244–1247.

Rundle, J.J. & Brown, S.R., 1991. Origin of rate dependence in frictional sliding, *J. stat. Phys.,* **65,** 403–412.

Rundle, J. & Jackson, 1977. Numerical simulation of earthquake sequences, *Bull. seism. Soc. Am.,* **67,** 1363–1377..

Rundle, J.B., Klein, W., Gross, S. & Turcotte, D., 1995. Boltzmann fluctuations in numerical simulations of nonequilibrium lattice threshold systems, *Phys. Rev. Lett.,* **75,** 1658–1661.

Rundle, J.B., Gross, S., Klein, W., Ferguson, C. & Turcotte, D., 1997. The statistical mechanics of earthquakes, *Tectonophysics,* **277,** 147–164.

Sammis, C.G. & Smith, S.W., 1999. Seismic cycles and the evolution of stress correlation in cellular automaton models of finite fault networks, *Pure appl. Geophys.,* **155,** 307–334.

Von Seggern, D., Alexander, S.S. & Baag, C.E., 1981. Seismicity parameters preceding moderate to major earthquakes, *J. geophys. Res.,* **86,** 9325–9351.

Ward, S.N., 1991. A synthetic seismicity model for the Middle America Trench, *J. geophys. Res.,* **96,** 21 433–21 442.

Wolfram, S.J., 1986. *Theory and Applications of Cellular Automata,* World Scientific, Singapore.