

Predicting Solar Flares by Data Assimilation in Avalanche Models.

I. Model Design and Validation

Eric BÉLANGER, Alain VINCENT, Paul CHARBONNEAU

Département de physique, Université de Montréal, C.P. 6128 Succ. Centre-Ville, Montréal, Qc, H3C-3J7, CANADA.

(e-mails: belanger@astro.umontreal.ca, vincent@astro.umontreal.ca, paulchar@astro.umontreal.ca)

Received ; accepted

Abstract. Data assimilation techniques, developed in the last two decades mainly for weather prediction, produce better forecasts by taking advantage of both theoretical/numerical models and real time observations. In this paper, we explore the possibility of applying the data assimilation techniques known as 4D-VAR to the prediction of solar flares. We do so in the context of a continuous version of the classical cellular automaton-based self-organized critical avalanche models of solar flares introduced by Lu and Hamilton (1991). Such models, although a priori far removed from the physics of magnetic reconnection and magnetohydrodynamical evolution of coronal structures, nonetheless reproduce quite well the observed statistical distribution of flare characteristics. We report here on a large set of data assimilation runs on synthetic energy release time series. Our results indicate that, despite the unpredictable (and unobservable) stochastic nature of the driving/triggering mechanism within the avalanche model, 4D-VAR succeeds in producing optimal initial conditions that reproduce adequately the time series of energy released by avalanches/flares. This is an essential first step towards true forecasting.

Keywords: Sun: flares, data assimilation, self-organized criticality

1. Introduction: Flares and self-organized criticality

Spatially resolved observations of solar flares at short wavelengths have revealed a very broad range of scales in the flaring phenomenon. Probability distributions of global flare characteristics such as peak flux, energy release, duration, are now known to take the form of power laws spanning many decades in size (eight in the case of flare energy; see, e.g., Dennis, 1985; Lu *et al.*, 1993; Aschwanden *et al.*, 2000). This is surprising because the vast majority of flares occur in active regions and activity complexes that have global characteristics (linear size, magnetic flux, peak field strength) that are much more narrowly distributed. This indicates that the flaring phenomenon is intrinsically scale-free, even though its energy reservoir may not be. The relatively slow evolution of active regions is also in stark contrast to the short energy release timescale associated with the flaring phenomenon.

Avalanches are one class of physical phenomena that are characterized by intermittent, scale-free energy release even under conditions of slow, continuous energy loading. In the flare context, the physical picture usually invoked is that of coronal magnetic structures being slowly and stochastically forced by photospheric fluid

motions, leading to the gradual buildup of electrical current sheets in the coronal plasma (Parker, 1983, 1988). Plasma instabilities eventually trigger magnetic reconnection at an unstable site, leading to alterations of the physical conditions at neighbouring current sheets that can then themselves become unstable, and so on in classical avalanching style. Numerous avalanche models inspired by this general scenario have been developed to describe solar flares (see Charbonneau *et al.*, 2001, and references therein). Many of these models manage to produce avalanche size distributions having the form of power-laws, with logarithmic exponent comparing fairly well to observationally-determined values (e.g., Lu *et al.*, 1993; Georgoulis and Vlahos, 1998; Aschwanden and Charbonneau, 2002). Most of these models are based on the idea of self-organized criticality (SOC) (Bak, Tang, and Wiesenfeld, 1987; Jensen, 1998). The “criticality” is akin to phase change in equilibrium thermodynamics, where the effects of a small, localized perturbation can be felt on a dynamical timescale throughout the whole system. The system is said to be “self-organized” when this critical state is a dynamical attractor and is reached in response to external forcing without requiring fine tuning of a control parameter. Generally, SOC is found in slowly-driven open dissipative systems subjected to a self-limiting local threshold instability. The threshold to the instability is crucial, as it allows the system to transit from one metastable state to another while preventing the dynamics to be governed by external forcing (Jensen, 1998). Potential examples in the natural world include various forms of sandpiles, avalanches and landslides, but also earthquakes, forest fires, hydrological networks, traffic jams, magnetospheric substorms, and solar flares (see Bak, 1996, for a spirited exposition).

If flares are truly a manifestation of SOC dynamics, then the outlook for accurate flare forecasting would appear, a priori, pretty grim indeed. Nothing fundamentally distinguishes a large flare from a small one, flare size simply being a matter of the number of current sheets involved in the avalanche of reconnection events. Even worse, the triggers of flares large and small are the same, namely a small (quite possibly unobservable) perturbation affecting the system somewhere locally. However, the occurrence of a large avalanche is only possible if a large, “connected” portion of the system is close to the avalanching threshold. The state of the system, in turn, is a function of its prior history, and in particular of the past occurrence of avalanches, of which the larger ones are (presumably) observable. In other words, past avalanching behavior holds clues to the current state of the system, and therefore to its *potential* avalanching behavior.

The question is then: can this information be retrieved and used to produce reliable avalanche forecasts, despite the stochastic nature of the driving/triggering mechanism? This is the central question we address in this series of papers, using data assimilation techniques. This first paper describes the SOC avalanche model and data assimilation technique we are developing towards forecasting, and demonstrates that the resulting scheme can adequately reproduce the avalanching behavior of the system even in the absence of detailed information on the spa-

tiotemporal behavior of its the stochastic driver. This is a first essential step towards forecasting, which is the topic of the subsequent papers in the series.

This first paper is organized as follows: section 2 gives an overview of a simple, “classical” discrete SOC model based on a cellular automaton, as well as a continuous analog described by a partial differential equation reversed-engineered from the discrete cellular automaton rules. This is the model used in Section 3 in conjunction with the 4D-VAR data assimilation techniques. Section 4 presents results for a wide set of validation experiments demonstrating that 4D-VAR can successfully reproduce the avalanching behavior present in energy release time series, despite the stochastic nature of the external forcing that loads energy into the system and triggers avalanches. We conclude in section 5 by summarizing the main results of this study, and outlining the road lying ahead, towards true forecasting.

2. Avalanche models

2.1. THE LU & HAMILTON MODEL

The sandpile has now become the icon of SOC systems (Bak, Tang, and Wiesenfeld, 1987). As sand grains are dropped one by one on a flat surface, a sandpile will build up, with occasional avalanches of various sizes, until the pile has reached a conical shape with the slope everywhere at or near the angle of repose. Addition of more sand grains can now trigger large avalanches disrupting the whole slope, or the toppling of only a few sand grains, or nothing at all. The system has reached a statistically stationary state where, averaged over a long enough temporal interval, as many sand grains fall off the pile as are dropped on it. Notice that while the loading is slow and gradual, the unloading is strongly intermittent and involves avalanches of all sizes.

The statistical physics of sandpiles has been extensively studied using cellular automata models, where the sandpile is replaced by a lattice of locally interconnected nodes on which a nodal variable related to energy is defined (Kadanoff *et al.*, 1989). In the context of solar flares, the first such model is to be found in the groundbreaking work of Lu and Hamilton (1991) and Lu *et al.* (1993) (but do see also Zirker and Cleveland, 1993). Consider the following two-dimensional scalar version of the Lu and Hamilton (1991) model; a scalar nodal quantity $A_{i,j}^n$ is defined over a $N \times N$ regular cartesian grid with nearest-neighbour connectivity (top+down+right+left; see, e.g., Figure 1 in Charbonneau *et al.*, 2001). Here the superscript n is a discrete time index, and the subscript pair (i, j) identifies a node on the 2D lattice. The cellular automaton is driven by adding small increments in A at randomly selected nodes in the lattice (one per time step), analogous here to dropping sand grains on the pile. A stability criterion is defined in terms of the

local curvature of the field at node (i, j) :

$$\Delta A_{i,j}^n \equiv A_{i,j}^n - \frac{1}{4} \sum_{\text{neighbours}} A_{\text{neighbours}}^n, \quad (1)$$

where the sum runs over the four nearest neighbours at nodes $(i, j \pm 1)$ and $(i \pm 1, j)$. If this quantity exceeds some pre-set threshold A_c (analogous to the slope exceeding the angle of repose on the sandpile), then $A_{i,j}^n$ is redistributed to its four nearest neighbours according to the following rules:

$$A_{i,j}^{n+1} = A_{i,j}^n - \frac{4}{5} \Delta A_{i,j}^n, \quad (2)$$

$$A_{i\pm 1, j\pm 1}^{n+1} = A_{i\pm 1, j\pm 1}^n + \frac{1}{5} \Delta A_{i,j}^n. \quad (3)$$

These redistribution rules are conservative in A ; however, if one identifies A^2 with a measure of energy (more on this below), then it is readily verified that they lead to a decrease of the total lattice energy:

$$E_l^n = \sum_{i,j} (A_{i,j}^n)^2, \quad (4)$$

the excess energy being what is liberated in the “flare”. Lu and Hamilton (1991) and Lu *et al.* (1993) have shown that this driven cellular automaton can produce avalanches with robust power-law exponents resembling those inferred from flare observations (see also Charbonneau *et al.*, 2001, and references therein).

Giving a physical meaning to the scalar A is not trivial. If A is considered as being the magnetic field, then $\nabla \cdot \mathbf{A} \neq 0$. Taking A as the vector potential automatically solves the non-null divergence problem. However, $\sum_{i,j} (A_{i,j}^n)^2$ is then no longer an obvious measure of magnetic energy (Charbonneau *et al.*, 2001). Dimensional analysis does suggest $\nabla \times \mathbf{A} = \mathbf{B}$ or $|\mathbf{A}|^2 \sim L^2 |\mathbf{B}|^2$ so we can at first assume that the vector potential scales with the magnetic field, a conjecture supported by the numerical results of Isliker, Anastasiadis, and Vlahos (2000), using an avalanche model closely related to that described above. Variations in lattice energy E_l from one time step to the next during an avalanche then yields the energy liberated at each time step:

$$E_r^n = \begin{cases} E_l^{n-1} - E_l^n (> 0), & \text{lattice avalanching} \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

The resulting time series of energy release is the target for data assimilation. In the Lu & Hamilton discrete model, E_r can be calculated analytically knowing the redistribution rule and stability measure, but this will no longer be the case with the continuous analog to be introduced shortly; Equation (5) applies equally well to both classes of models.

At this juncture it must be emphasized that the stochasticity in the Lu & Hamilton model (and other published variations thereof) is not a mere noise-like “inconvenience” superimposing itself on an underlying deterministic flaring process. In the Lu & Hamilton SOC avalanche model, stochastic forcing plays the dual role of energy loading *and* avalanche triggering. The model thus has a fundamentally stochastic component.

2.2. A CONTINUOUS SOC MODEL

2.2.1. Reverse engineering

The structure of Equations (2) and (3) suggest that they can be interpreted as the result of applying centered second-order finite difference and a one-step explicit time-stepping algorithm to some partial differential equation (PDE) in two spatial dimension discretized on a regular cartesian grid. Applying this reverse engineering approach (see also Liu *et al.*, 2002) leads to the following PDE describing the evolution of A during an avalanche:

$$\frac{\partial A}{\partial t} = -\frac{\partial^2}{\partial x^2} \left(v(\nabla^2 A) \frac{\partial^2 A}{\partial x^2} \right) - \frac{\partial^2}{\partial y^2} \left(v(\nabla^2 A) \frac{\partial^2 A}{\partial y^2} \right). \quad (6)$$

Likewise, the RHS of Equation (1) is readily interpreted as a second-order centered finite difference representation of the Laplacian operator acting on A , so that the diffusion coefficient $v(\nabla^2 A)$ appearing in Equation (6) is given by:

$$v(\nabla^2 A) = \begin{cases} v_a & \text{if } (\nabla^2 A)^2 > A_c^2 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where A_c is the stability threshold. The numerical value of v_a depends on the redistribution rules that were used in the discrete equations (Equations (2) and (3)) and on the grid size. In the case of the Lu & Hamilton model, in two spatial dimensions, $v_a = \Delta^2/20$ ($\Delta x = \Delta y \equiv \Delta$), with units of *time/length*² implicitly included in the denominator, a value used in all calculations reported upon below. In analogy with the discrete model, we identify the functional $A(x, y, t)$ with a measure of the magnetic vector potential, and assume that A^2 is a measure of energy.

Equation (6) is a fourth-order “hyperdiffusion” equation, albeit a strongly non-linear one since the hyperdiffusion coefficient is only non-zero when the system is avalanching, and even then it remains a discontinuous function of position, being only non-zero at unstable nodes. If all nodes are stable then the quantity A evolves in response to the external forcing only. In the classical cellular automaton, at each (non-avalanching) timestep a small increment of random amplitude in A is added to a single randomly selected node of the lattice. Reverse-engineering of this forcing rule immediately leads to:

$$\frac{\partial A}{\partial t} = F_R(x_0(t), y_0(t)) \quad (8)$$

where the forcing location (x_0, y_0) varies randomly in time. The continuous model then amounts to solving Equation (8) when the system is everywhere stable, and switching to Equation (6) if one or more nodes become unstable. Note that this implies that the forcing F_R is turned off during avalanches. This is the same procedure used in the cellular automaton, and amounts to assuming that there exists a wide separation of timescales between forcing and avalanching, which in fact is well justified in the solar coronal context (see, e.g., Lu *et al.*, 1993).

2.2.2. Sample numerical results

We now proceed to solve numerically the continuous avalanche equation (Equations (6)–(8)), by discretizing it using a centered finite difference scheme with forward Euler differencing for the time derivative. The reader may rightfully wonder why then did we bother reverse-engineering the discrete Lu & Hamilton model in the first place, but the need to proceed in this way will become clear presently. Moreover, in solving Equation (6) numerically as one would any partial differential equations, subtle differences are introduced with respect to the truly discrete model, and these must be understood.

A square domain of linear size $L = 2\pi$ is partitioned using a regular cartesian grid ($\Delta x = \Delta y \equiv \Delta$). Imposing $A = 0$ on domain boundaries, we solve dimensionless forms of the avalanche equations (Equations (6), (7) and (8)), using the hyperdiffusion time scale

$$\tau_v = \frac{L^4}{v_a} \quad (9)$$

as a time unit. This corresponds to the time it would take an avalanche to sweep across the length of the domain. Following a dimensional analysis of the continuous avalanche equation, x and y will now be in L units and t in τ_v units. The diffusion coefficient v , in the non-dimensional equation, is equal to 0 (stable) or to 1 (avalanching). For the scheme to remain numerically stable, a von Neumann stability analysis (Press *et al.*, 1992) indicates that the dimensional time step must be chosen such that

$$\Delta t \leq \frac{\Delta^4}{4v_a} \quad (10)$$

or, in terms of dimensionless time:

$$\frac{\Delta t}{\tau_v} \leq \frac{1}{4} \left(\frac{\Delta}{L} \right)^4. \quad (11)$$

For our working 48×48 mesh, $\frac{1}{4} \left(\frac{\Delta}{L} \right)^4 = 4.7 \times 10^{-8}$, so a non-dimensional time step of size 5.5×10^{-11} safely satisfies that condition, and is used for all simulations unless otherwise specified.

The stability threshold is set at $A_c = 7.0$. Sequences of uniform random deviates are used to define the stochastic forcing. Both the location and magnitude of the perturbation are randomly determined, the former excluding boundary nodes

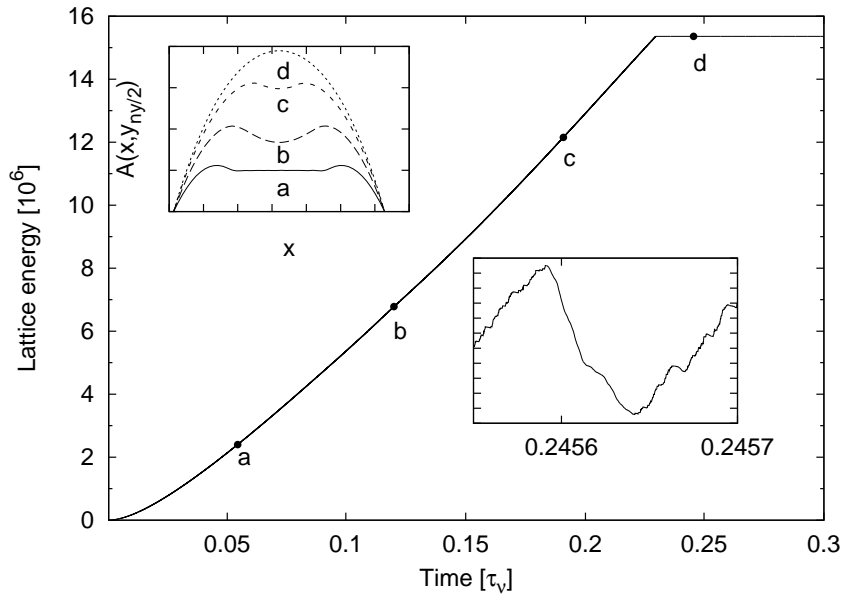


Figure 1. Evolution of the lattice energy as a function of time. The lattice energy increase from the null state ($A = 0$) to the SOC state as the energy input by the perturbations is greater than the energy released in avalanches. At $t = 0.23$, the SOC state is reached: the energy added to the system is released by avalanches, hence the plateau. The top left inset is a cross-section of the lattice variable A along the y -axis, taken at different times. The labels along the lattice energy curve indicates when the cross-sections were taken. The bottom right inset is a zoom of the lattice energy curve about point “d” to show the energy variation in the SOC state, with vertical axis covering 5 energy units.

and the latter constrained to the pre-set interval $[-1 \times 10^{-5} : 4 \times 10^{-5}]$. The perturbations have non-zero mean to ensure buildup of $A(x, y)$ from the initial state $A(x, y) = 0$ throughout. This is again in direct analogy to the Lu & Hamilton model described earlier. Because the continuous model is discretized with centered second-order finite differences, boundary conditions for the two rows of grid points along the boundary are needed. The grid points along the boundary are set to $A = 0$ (like in the discrete model) while the second row is set to $\nabla^2 A = 0$ with the use of the fictional point method.

The first task is to bring the system to the SOC state. Starting with $A = 0$ everywhere, the system is driven to the SOC state by the external forcing, interrupted by avalanching episodes whenever and wherever the stability threshold is exceeded. Figure 1 illustrates the gradual buildup of lattice energy ($\sum A^2$) with time. Initially, the energy gained by the system from the perturbations exceeds the energy released in avalanches. Thus, there is a net increase of the system’s energy. However, at $t \simeq 0.23$ a plateau begins. The system has now reached a statistically-stationary state where, in time-averaged sense, the buildup of A in response to forcing is balanced by A evacuated via the $A = 0$ boundary conditions when avalanches reach the system’s boundaries. Because large avalanches are more likely to reach the

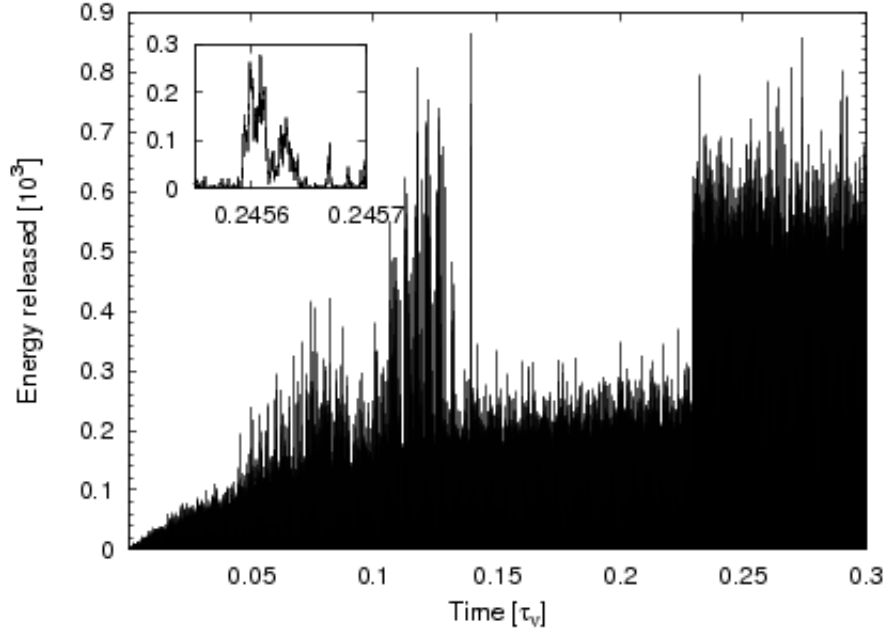


Figure 2. Evolution of the energy released as a function of time. The large avalanches can only be created when the system is in the SOC state. The large avalanches between $t = 0.1$ and $t = 0.14$ are due to a transient SOC state that is created when the stability criterion of the central region take a value of $\sim +A_c$ while the periphery have $\sim -A_c$. This is only a transitional step as the center will fill up to take its final shape of an inverse parabola (“b” \rightarrow “d” on the inset in Figure 1). The inset is a zoom to show that we have individual avalanches separated by calm periods.

boundaries, their frequency of occurrence jumps markedly once the stationary state is reached (cf. Figure 2). This stationary state is the SOC state. The bottom right inset on Figure 1 shows a zooming on the behaviour of the lattice energy during the SOC state in the vicinity of point “d”. Every decrease is the signature of an avalanche and the pronounced drop starting at $t = 0.2456$ corresponds to the large avalanche in the inset of Figure 2. The top left inset shows cross-sections along the y -axis of the evolution of the scalar variable A at different times up to the SOC regime where it approximates the shape of an inverse parabola. A cross-section along the x -axis would yield the same result. All DNS and 4D-VAR runs discussed further below are carried out with the system in the SOC state.

Figure 2 displays the energy released as a function of time. An short segment is shown in inset, where individual avalanches can be distinguished. It is only in the SOC state that avalanche spanning the whole system can be produced with a significant frequency. For this reason, it will be particularly important to ensure that the system is kept in the SOC state throughout the data assimilation process. The large avalanches between $t = 0.1$ and $t = 0.14$ (Figure 2) are due to a transient SOC state that is created when the stability criterion of the central region take a

value of $\sim +A_c$ while the periphery have $\sim -A_c$ (such as for the curve “b” on the top left inset of Figure 1). This is only a transitional step as the center will fill up to eventually take its final shape of an inverse parabola and the system will then be in its true SOC state.

Even though the avalanche model is a continuous one, it retains some discrete features. Usually, in the numerical solution of partial differential equations, reducing the time steps by a factor of two but using twice as many of these smaller time steps, should not affect the outcome significantly. However, this is not the case here, because of the the spatiotemporally discrete nature of the hyperdiffusion coefficient, as the latter is only turned on when the system is unstable. When this happens, the run with the smaller time step Δt will redistribute more finely, as seen in Equation (6). When stability is recovered, the formerly avalanching portion of the lattice will find itself closer to the stability threshold, thus the lattice energy will increase significantly. The redistribution is also affected by the value of the diffusion coefficient (with a dependency on the distance between the grid points) which must be lower for higher resolution grids otherwise too much energy is redistributed and the system restabilizes too far below the stability threshold, thus leading to loss of SOC at the expense of large, quasiperiodic avalanches.

The last difference between the Lu & Hamilton and continuous models pertains to the effect of an increase in the number of grid points. In the Lu & Hamilton model, the distance between each grid point is constant ($\Delta = 1$) which leads to an increase of the size of the domain when the grid point number is increased. As for the continuous model, even though the domain size can be increased, we will keep it constant in this paper so an increase in the number of grid points will increase the resolution. The system will then be more effective in reproducing fine structures as seen in Figure 3, which shows the normalized frequency distributions for the stability criterion (Equation (1)) evaluated at each grid point at ten different times for grid resolution of 24×24 , 36×36 and 48×48 . As with the discrete model, at any given time only a small fraction of nodes are very close to the instability threshold. The distribution is broader and centered further below the stability threshold for the higher resolution runs. The reason behind this behaviour is that the larger number of degrees of freedom in the high resolution run allows for the presence of finer structures, which make it easier for the lattice to exceed stability, in the sense that fluctuations of $\Delta A_{i,j}^n$ about its lattice mean become larger as the spatial resolution is increased.

2.2.3. Characteristics of avalanches in the SOC state

In the SOC state, several characteristics of avalanches have probability distribution functions (hereafter PDF) that behave as power laws. This is the case for the avalanche energy (E), namely the total energy released by the lattice over the duration of the avalanche; the peak energy release (P), corresponding to maximum energy released by the avalanche in a single time step; and the duration (T), which is simply the time elapsed from the beginning of an avalanche to its end. Figure 4

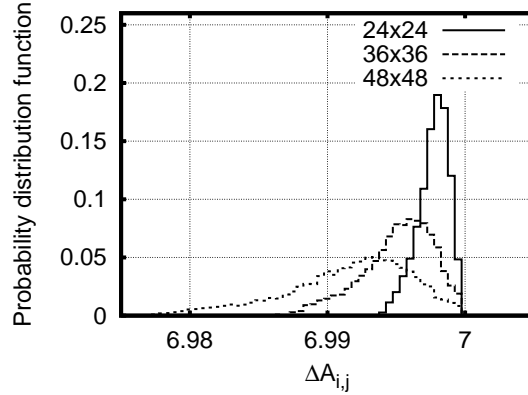


Figure 3. Normalized frequency distributions for the stability criterion evaluated at each grid point for runs of 24×24 , 36×36 and 48×48 . Data taken at ten different times, when the system was stable, were combined to generate the PDF. As expected, the peaks of the PDF lies close but below the stability threshold of $A_c = 7.0$. As the resolution increases, the width of the distribution widens and move further from the threshold. This is due to the increase of fine structures which permits more degrees of freedom to the large resolution systems, thus making it easier to stray away from the threshold.

(panels A–C) show probability distribution functions for E , P and T , constructed from runs of different resolutions (24×24 , 36×36 and 48×48) of 10^7 time steps, spanning 1000 hyperdiffusion times, and each including up to 6×10^5 avalanches. Least-squares fit of the form:

$$f(x) = f_0 x^{-\alpha} \quad (12)$$

are overplotted on the distributions. The fits used the data from the large resolution runs (48×48). Because the mesh spacing affects the statistics for small avalanches, the first few bins are excluded from these fits. The power law indices obtained were $\alpha_E = 1.407 \pm 0.02$, $\alpha_P = 1.800 \pm 0.04$ and $\alpha_T = 2.067 \pm 0.03$. These first two values are in good agreement with those obtained with the discrete formulation of this avalanches model (see Table II in Charbonneau *et al.* (2001)). The α_T value is a bit higher, but is in fact more comparable to the values obtained by Lu *et al.* (1993) (Table 1). Figure 4 (panel D) show the frequency distribution of the waiting time (WTD; now on semi-log scale). The waiting time (ΔT) is the time interval during which the system is in a stable state between two consecutive avalanches. The WTD is well-fitted with exponential of the form:

$$f(\Delta t) = f_0 e^{-\beta \Delta t}, \quad (13)$$

here with an inverse e -folding timescale $\beta = 2158 \pm 4$. The exponential form of the WTD comes from the statistically uniformity of the external forcing, here a stationary random process (see Wheatland, 2000). Other types of forcings can be introduced in SOC avalanche models for solar flares, producing WTD distributions in better agreement with observations (see, e.g., Norman *et al.*, 2001), but for the

purposes of the forthcoming validation exercise we retain the uniform driver of the original Lu & Hamilton model.

To sum up, and notwithstanding some subtleties related to mesh refinement, numerical solutions of the reverse-engineered avalanche equation leads to avalanching behavior essentially identical to that of the original, discrete model of Lu and Hamilton (1991). This “validation” of the continuous model may well appear tautological, but it represents an essential starting point to our forecasting scheme because recasting the model in continuous form is required by the 4D-VAR data assimilation formalism, the topic to which we now turn.

3. Data assimilation (4D-VAR)

The underlying idea of data assimilation is to use observed data to introduce corrections to a model—often a numerical simulation—of a physical phenomenon, typically with the aim of correcting for missing data or to produce forecasts. Generally speaking, data assimilation methods can be subdivided into three categories: successive correction, sequential (Kalman filter) and variational methods (Le Dimet and Talagrand, 1986; Daley, 1991; Kalnay, 2003). The successive correction method, also known as Cressman method, consist in the correction of a background field, previously obtained through a previous forecast or a trivial state due to physical constraints, until it includes the given observations (Cressman, 1959). Most sequential methods are based on the Kalman filter, which uses the model error and the observational error statistics to find the optimal combination of the model and observational data (Kantha and Clayson, 2000). The variational methods consist in finding the space-time trajectory of the state variables that will minimize a cost function measuring the discrepancy between the forecast and the observations (Talagrand and Courtier, 1987; Courtier and Talagrand, 1990). The 4D-VAR method belongs to this third class.

3.1. 4D-VAR: AN OVERVIEW

Four-dimensional variational data assimilation (hereafter 4D-VAR) is an efficient technique for incorporating observations in numerical forecasting models (Talagrand and Courtier, 1987; Courtier and Talagrand, 1990). The 4D-VAR method consists in minimizing a scalar cost function measuring the deviation between the forecast and the observations. The physical fields produced by data assimilation must correspond to the observations, while abiding to the known physical laws and/or statistical relations characterizing the system being treated (Le Dimet and Talagrand, 1986).

Figure 5 shows an overview of the 4D-VAR method as applied to a classical forecasting problem, namely using the known state of a variable ψ_0 at some initial instant of time, to forecast its value ψ_T at a later time T . This forecast is to be

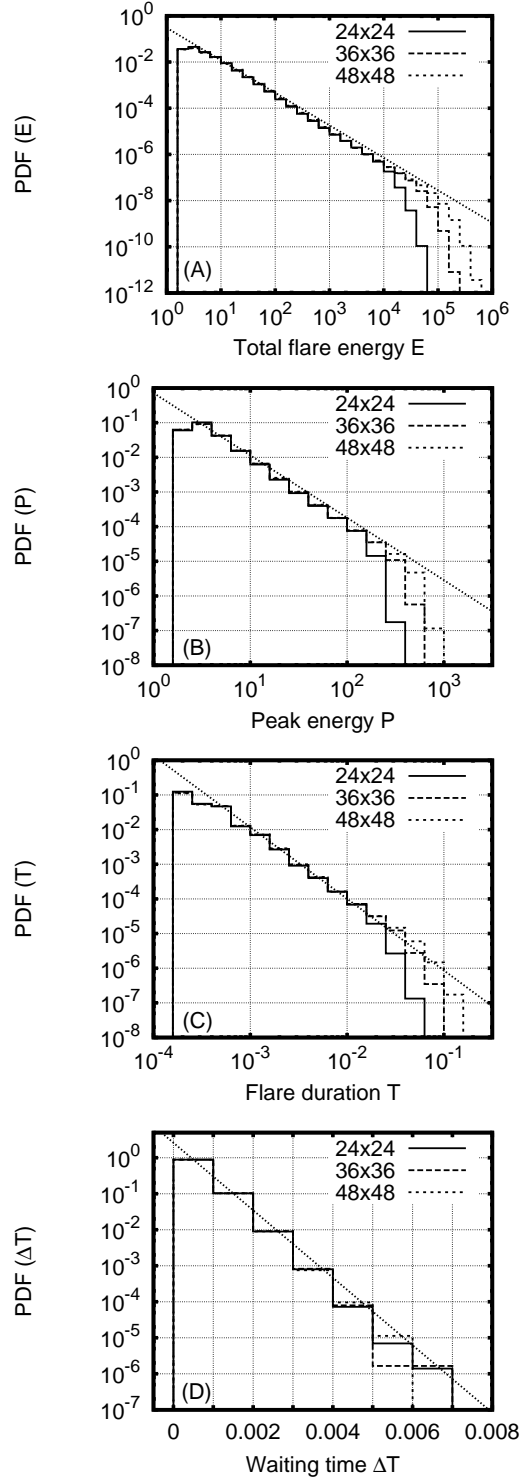


Figure 4. Normalized frequency distributions for the total energy (E), the peak energy (P), the duration (T) and the waiting time (ΔT) for runs of different spatial resolutions (24×24 , 36×36 and 48×48) of $t = 10^3$ containing up to 6×10^5 avalanches. Bins of a constant logarithmic width $\Delta \log x = 0.2$ were used to construct the PDFs. Power law indices of $\alpha_E = 1.407 \pm 0.02$, $\alpha_P = 1.800 \pm 0.04$ and $\alpha_T = 2.067 \pm 0.03$ were obtained by fitting data for the 48×48 run. An exponential was fitted to the waiting time distribution, yielding an inverse e-folding time $\beta = 2158 \pm 4$.

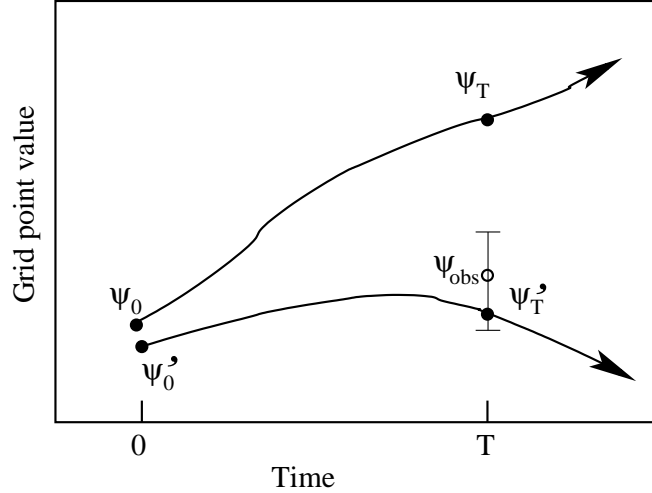


Figure 5. Overview of variational data assimilation. Point ψ_0 is an estimate of the initial condition at time 0. Using this initial condition, a forecast ψ_T is obtained at time T . The 4D-VAR method uses the difference between the forecast ψ_T and the observation ψ_{obs} to generate a new initial condition ψ'_0 which will produce a better forecast (ψ'_T) at time T (Errico, 1997).

compared to an actual observation ψ_{obs} . Here this first forecast falls outside of the observation's error bar. The 4D-VAR method uses the difference between the forecast ψ_T and the observation ψ_{obs} to generate a new initial condition ψ'_0 that now produces an improved forecast ψ'_T at time T which is closer to the observation (Errico, 1997). The procedure can be repeated until some pre-set goodness of fit criterion between ψ'_T and ψ_{obs} is reached. In classical data assimilation applications (for example numerical weather forecasting), the point ψ_0 would correspond to a field variable discretized on a single node (i, j) of a $N \times N$ spatial mesh; 4D-VAR must then solve concurrently as many variational problems as there are variables, times the number of spatial mesh points on which the numerical simulation used to advance the field from 0 to T is performed. Data assimilation in numerical simulations is a computationally intensive undertaking!

3.2. THE COST FUNCTION

Generally, in variational problems, we want to minimize the cost function \mathcal{J} :

$$\mathcal{J} = \int_0^T \int_{\Omega} f(\boldsymbol{\psi}, \mathbf{x}, t) \, d\mathbf{x} \, dt, \quad (14)$$

where $f(\boldsymbol{\psi}, \mathbf{x}, t)$ is a scalar function, defined over a domain Ω and a time interval $[0, T]$, of the state variable $\boldsymbol{\psi}$ (Sanders and Katopodes, 2000). More precisely, in data assimilation, we want to minimize the error between the forecast and the

observations $(\boldsymbol{\psi}_T - \boldsymbol{\psi}_{\text{obs}})$ (cf. Figure 5):

$$\mathcal{J} = \frac{1}{2} \int_0^T \int_{\Omega} (\boldsymbol{\psi}_T - \boldsymbol{\psi}_{\text{obs}})^{\top} \mathbf{W} (\boldsymbol{\psi}_T - \boldsymbol{\psi}_{\text{obs}}) \, \mathrm{d}\mathbf{x} \, \mathrm{d}t, \quad (15)$$

where \mathbf{W} is a matrix of statistical weights given by the instrumental errors in the observations and \top represent the transpose. The squared residual $(\boldsymbol{\psi}_T - \boldsymbol{\psi}_{\text{obs}})^2$ is used instead of absolute values, to avoid introducing discontinuities when the cost function will be differentiated (next section). The physical equations (i.e. the continuous avalanche equation (Equations (6)-(8))), which can be schematically written as:

$$\mathcal{E}(\boldsymbol{\psi}, \mathbf{x}, t) = 0, \quad (16)$$

are acting as constraints during the minimization (Talagrand and Courtier, 1987).

In this paper, the cost function will use the time series of released energy as defined via Equation (5) as the state variable $\boldsymbol{\psi}$. Via the use of Equation (4), we get:

$$\begin{aligned} E_r &= \frac{\mathrm{d}E_l}{\mathrm{d}t} \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \sum_{i,j} A_{i,j}^2 \\ &= \sum_{i,j} \frac{\mathrm{d}A_{i,j}^2}{\mathrm{d}t}. \end{aligned} \quad (17)$$

If we bin the time series of the released energy E_r (§ 4.1.2), we have:

$$\bar{E}_r = \frac{\sum^{\text{bin}} E_r}{\Delta b}, \quad (18)$$

where Δb is the number of elements per bin. The cost function is then:

$$\mathcal{J} = \frac{1}{2} \int_0^T (\bar{E}_r - \bar{E}_r^{\text{obs}})^2 \, \mathrm{d}t. \quad (19)$$

The identity matrix is used for the observation errors matrix \mathbf{W} because synthetically generated observations will be used (§ 4) in what follows. Covariance error matrices are certainly an important and delicate point when dealing with real data.

3.3. THE LAGRANGIAN FORMULATION

We want to minimize the cost function \mathcal{J} given the constraint $\mathcal{E}(\boldsymbol{\psi}, \mathbf{x}, t) = 0$. Since this is a problem of minimization with constraints, a Lagrangian formulation is used:

$$\mathcal{L}(\boldsymbol{\psi}, \boldsymbol{\lambda}) = \mathcal{J}(\boldsymbol{\psi}) + \int_0^T \int_{\Omega} \boldsymbol{\lambda}(\mathbf{x}, t) \cdot \mathcal{E}(\boldsymbol{\psi}, \mathbf{x}, t) \, \mathrm{d}\mathbf{x} \, \mathrm{d}t, \quad (20)$$

where $\boldsymbol{\lambda}(\mathbf{x}, t)$ are the Lagrange undetermined multipliers, also called adjoint variables (Sanders and Katopodes, 1999). The variational operator δ is then applied on the Lagrangian to find its stationary points:

$$\begin{aligned}\delta\mathcal{L} &= \nabla_{\boldsymbol{\psi}}\mathcal{L} \cdot \delta\boldsymbol{\psi} + \nabla_{\boldsymbol{\lambda}}\mathcal{L} \cdot \delta\boldsymbol{\lambda} \\ &= \frac{\partial\mathcal{L}}{\partial\boldsymbol{\psi}}\delta\boldsymbol{\psi} + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\lambda}}\delta\boldsymbol{\lambda} .\end{aligned}\quad (21)$$

One can use integration by parts to transfer the differential operators from the state variable $\boldsymbol{\psi}$ to the adjoint variable $\boldsymbol{\lambda}$. For an arbitrary displacement $(\delta\boldsymbol{\psi}, \delta\boldsymbol{\lambda})$, the minimum is reached only when $\delta\mathcal{L} = 0$ (Daley, 1991). This indicates that the derivative of the Lagrangian with respect to each direction must be zero:

$$\frac{\partial\mathcal{L}}{\partial\boldsymbol{\lambda}} = \mathcal{E}(\boldsymbol{\psi}, \mathbf{x}, t) = 0 , \quad (22)$$

and

$$\frac{\partial\mathcal{L}}{\partial\boldsymbol{\psi}} = \text{Adj}(\boldsymbol{\lambda}) + \frac{\partial\mathcal{J}}{\partial\boldsymbol{\psi}} = 0 , \quad (23)$$

where $\text{Adj}(\boldsymbol{\lambda})$ represents the adjoint equations (Schröter, Seiler, and Wenzel, 1993). As noted by Le Dimet and Talagrand (1986), this set of equations (Equations (22) and (23)) are the Euler-Lagrange equations.

For the 2D avalanche model the direct and adjoint equations are:

$$\begin{aligned}\frac{\partial A}{\partial t} &= -\frac{\partial^2}{\partial x^2} \left(v(\nabla^2 A) \frac{\partial^2 A}{\partial x^2} \right) - \frac{\partial^2}{\partial y^2} \left(v(\nabla^2 A) \frac{\partial^2 A}{\partial y^2} \right) && \text{avalanching} \\ \frac{\partial A}{\partial t} &= F_R(x_0(t), y_0(t)) && \text{stable}\end{aligned}\quad (24)$$

and

$$\frac{\partial A^*}{\partial \tau} = -\frac{\partial^2}{\partial x^2} \left(v(\nabla^2 A) \frac{\partial^2 A^*}{\partial x^2} \right) - \frac{\partial^2}{\partial y^2} \left(v(\nabla^2 A) \frac{\partial^2 A^*}{\partial y^2} \right) - \frac{\partial\mathcal{J}}{\partial A} \quad (25)$$

respectively, where the generic variable $\boldsymbol{\psi}$ have been replaced by the variable A defined over the lattice, the adjoint variable $\boldsymbol{\lambda}$ has been renamed A^* as it is the adjoint variable associated with A , and τ is a reverse time ($\tau = T - t$).

Evaluating the term $\frac{\partial\mathcal{J}}{\partial A}$ is particularly delicate because the cost function \mathcal{J} is not directly given in terms of a spatial quantity related to A , but as a time series of a nonlinear, non-local function of that quantity:

$$\begin{aligned}\frac{\partial\mathcal{J}}{\partial A} &= \int_0^T \frac{1}{2} \frac{\partial(\bar{E}_r - \bar{E}_r^{\text{obs}})^2}{\partial A} dt \\ &= \int_0^T (\bar{E}_r - \bar{E}_r^{\text{obs}}) \frac{\partial\bar{E}_r}{\partial A} dt \\ &= \int_0^T \frac{(\bar{E}_r - \bar{E}_r^{\text{obs}})}{\Delta b} \frac{\partial \sum_{\text{bin}} E_r}{\partial A} dt .\end{aligned}\quad (26)$$

Using the theorem of implicit functions to rewrite the derivative gives:

$$\frac{\partial \sum_{\text{bin}} E_r}{\partial A} = - \frac{\partial \sum_{\text{bin}} E_r}{\partial t} \bigg/ \frac{\partial A}{\partial t} \quad (27)$$

provided $\frac{\partial A}{\partial t} \neq 0$ so the $\frac{\partial \mathcal{J}}{\partial A}$ term is only to be evaluated when the system is avalanching. Substituting Equation (27) in Equation (26):

$$\begin{aligned} \frac{\partial \mathcal{J}}{\partial A} &= - \int_0^T \left[\frac{(\bar{E}_r - \bar{E}_r^{\text{obs}})}{\Delta b} \frac{\partial \sum_{\text{bin}} E_r}{\partial t} \bigg/ \frac{\partial A}{\partial t} \right] dt \\ &= - \int_0^T \left[\frac{(\bar{E}_r - \bar{E}_r^{\text{obs}})}{\Delta b} \left(\sum_{\text{bin}} \frac{\partial E_r}{\partial t} \right) \bigg/ \frac{\partial A}{\partial t} \right] dt . \end{aligned} \quad (28)$$

If we use the definition of E_r (Equation (17)), we get:

$$\frac{\partial \mathcal{J}}{\partial A} = - \int_0^T \left[\frac{(\bar{E}_r - \bar{E}_r^{\text{obs}})}{\Delta b} \left(\sum_{\text{bin}} \left(\sum_{i,j} \frac{\partial^2 A_{i,j}^2}{\partial t^2} \right) \right) \bigg/ \frac{\partial A}{\partial t} \right] dt . \quad (29)$$

The initial and boundary conditions for the adjoint equation arise from the integration by parts, namely the terms evaluated at the limits of the integrals. The initial conditions are $A^*|_{\tau=0} = 0$ and the boundary conditions are:

$$\begin{aligned} A^*(0, y, \tau) &= 0 & A^*(L_x, y, \tau) &= 0 \\ A^*(x, 0, \tau) &= 0 & A^*(x, L_y, \tau) &= 0 \\ \frac{\partial A^*}{\partial x} \bigg|_{x=0} &= 0 & \frac{\partial A^*}{\partial y} \bigg|_{y=0} &= 0 \\ \frac{\partial^2 A^*}{\partial x^2} \bigg|_{x=0} &= 0 & \frac{\partial^2 A^*}{\partial y^2} \bigg|_{y=0} &= 0 \\ \frac{\partial A^*}{\partial x} \bigg|_{L_x=0} &= 0 & \frac{\partial A^*}{\partial y} \bigg|_{L_y=0} &= 0 \\ \frac{\partial^2 A^*}{\partial x^2} \bigg|_{L_x=0} &= 0 & \frac{\partial^2 A^*}{\partial y^2} \bigg|_{L_y=0} &= 0 \end{aligned} \quad (30)$$

Unfortunately, there is no efficient method to directly solve the Euler-Lagrange equations (Equations (22) and (23)); therefore, we must formulate the problem as an unconstrained problem (Talagrand and Courtier, 1987). Although we have a random parameter in our model, the model can still be regarded as being deterministic because the initial conditions, more precisely the ‘‘connected’’ portion of the system close to the avalanching threshold, will dictate the evolution of the system with time. This highlight the fact that the cost function is an implicit function of the initial conditions: it is by varying the initial conditions that we will find the solution

of the physical equations which minimize the cost function (Ehrendorfer, 1992). In the language of optimal control theory, the initial conditions are the control variable in this problem (Lions, 1968).

As most minimization algorithm requires the gradient of the function to be minimized, we need the gradient of the cost function with respect to the initial conditions. However, it is not possible to calculate this gradient analytically as the cost function is an explicit function of the final conditions (i.e. forecast). It turns out that the more efficient way to calculate the gradient of the cost function with respect to the initial conditions is to use the adjoint equations evaluated at $\tau = T$ (Courtier and Talagrand, 1990):

$$\nabla \mathcal{J}_{A^0} = A^*(x, y, \tau = T) \quad (31)$$

which then requires numerical integration of the adjoint equations from $\tau = 0$ to $\tau = T$, i.e., backwards in time from $t = T$ to $t = 0$. This is the reason behind the interest and use of the adjoint equations.

3.4. MINIMIZATION ALGORITHM

The minimization of the cost function is usually carried out via a minimization algorithm such as steepest descent, conjugate gradient or quasi-Newton methods. The steepest descent is a simple method but it converges linearly. In this study, the conjugate gradient is used because of its quadratic convergence. The quasi-Newton method also converges quadratically and is popular among meteorologists. However, it requires the computation of the Hessian matrix. Even if an approximation of the Hessian is normally used, convergence problems may arise if it becomes nearly singular (Press *et al.*, 1992). One can solve these kind of problems but this leads to algorithm of a greater complexity, and usually more computationally intensive.

The algorithm implementation of 4D-VAR data assimilation runs as shown on Figure 6. Starting from initial conditions obtained by current experimental observations or a previous numerical simulation, a direct simulation generates a traditional (DNS) forecast. After reading the observations taken at the end of the forecast period, the cost function is evaluated, followed by the evaluation of its gradient. If the gradient is smaller than a chosen tolerance which accounts for error due to numerical precision, then the minimum of the cost function has been reached and the optimal 4D-VAR forecast is obtained. In the case that the cost function is not minimized, we must iterate. A new set of initial conditions are generated and will be used as the starting point of a new forecast. The cost function and its gradient are reevaluated and checked again against the termination criterion. This procedure is repeated until the latter is met. The iteration loop in Figure 6 takes place within the conjugate gradient minimization algorithm. It is also the conjugate gradient which modifies the initial conditions $A_{i,j}^0$ at each iteration:

$$[A_{i,j}^0]^{k+1} = [A_{i,j}^0]^k + \alpha^k p_{i,j}^k, \quad (32)$$

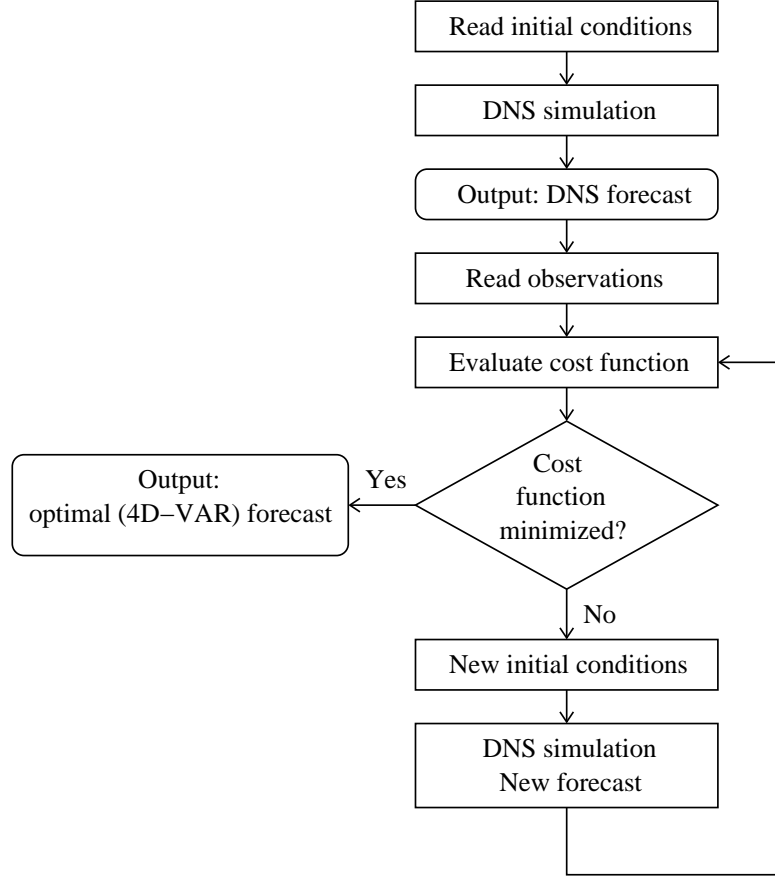


Figure 6. Algorithm of 4D-VAR data assimilation. From initial conditions, a traditional (DNS) forecast is made. Then the observations are read. The value of the cost function's gradient indicates if minimization was achieved. If not, a new set of initial conditions are produced by backwards integration of the adjoint equations, a new forecast is produced, and the verification is repeated. This procedure is iterated until minimization of the cost function is achieved. The output is the optimal 4D-VAR forecast and associated initial condition.

where, for the k th iteration of the conjugate gradient, $p_{i,j}^k$ is the conjugate direction vector multiplied by an amplitude α^k . These conjugate direction vectors are obtained with the use of the gradient:

$$p_{i,j}^{k+1} = \nabla \mathcal{J}_{i,j}^{k+1} + \gamma^k p_{i,j}^k \quad (33)$$

where

$$\gamma^k = \frac{(\nabla \mathcal{J}_{i,j}^{k+1} - \nabla \mathcal{J}_{i,j}^k) \cdot \nabla \mathcal{J}_{i,j}^{k+1}}{\nabla \mathcal{J}_{i,j}^k \cdot \nabla \mathcal{J}_{i,j}^k}, \quad (34)$$

and the gradients are evaluated using Equation (31). The amplitude α^k is calculated with:

$$\alpha^k = b - \frac{1}{2} \frac{(b-a)^2[\mathcal{J}(b) - \mathcal{J}(c)] - (b-c)^2[\mathcal{J}(b) - \mathcal{J}(a)]}{(b-a)[\mathcal{J}(b) - \mathcal{J}(c)] - (b-c)[\mathcal{J}(b) - \mathcal{J}(a)]} \quad (35)$$

where a , b and c are modified initial conditions that were obtained through a line minimization method known as inverse parabolic interpolation, which iteratively finds a triplet of points such that the minimum of a parabola passing through these three points will be as close as possible to the minimum of the function in that given interval (Press *et al.*, 1992).

3.5. BEYOND CLASSICAL 4D-VAR

There are many ways in which our use of 4D-VAR goes beyond the "classical" formulation of 4D-VAR, as given on Figure 5. We are using 4D-VAR to assimilate data into a finite grid sized cellular automaton made continuous only for the purpose of writing adjoint equations. Grid and time steps are fractions of the characteristic scales of the hyperdiffusive processes involved and are not infinitesimals. However, as pointed by Isliker, Anastasiadis, and Vlahos (2000), it is still possible to compute derivatives and thus operators. To the best of our knowledge, at the present time the only other area in which data assimilation techniques are being developed in conjunction with a cellular automaton is in seismic data assimilation of a stochastic random fault model (Rundle *et al.*, 2003; González *et al.*, 2006).

We are also using a time series of a global, model-produced variable to define the error, as opposed to the spatial state of the system measured at some time intervals beyond the initial condition in which data is being assimilated. Assimilating time series of a global variable instead of (or together with) spatial states of a system at non-zero times t is truly compatible with the 4D-VAR approach, as opposed to 3D-VAR for instance. Models in environmental sciences have been and will increasingly be assimilating time-series of data (see for instance Carton, Chepurin, and Cao (2000) in Oceanology or Bertino, Evensen, and Wackernagel (2002) in Estuary modeling or Eymin and Fournier (2005) in Geomagnetism).

Finally, our model system includes an essential stochastic component, namely the driving. Ocean and atmosphere general circulation or solid Earth dynamics are all random driven (noisy) systems. Data assimilation, including 4D-VAR, in principle can be applied to any such system (e.g., Nichols, 2003; Moore *et al.*, 2004). Random driving, inaccessible to observations, cannot be assimilated, but is not seen as a systematic bias in the model (see, e.g., Anderson, 2003).

4. Validation Experiments

4.1. EXPERIMENTAL DESIGN

4.1.1. Synthetic data

The validation experiments use synthetic observations generated from the same SOC avalanche model used to carry out data assimilation (certainly an optimal situation from the data assimilation point of view). Of course, entirely independent realizations of the stochastic driving are used to generate the synthetic observations, and these realizations are *not* made available to 4D-VAR, as our challenge is precisely to see whether 4D-VAR can still adequately assimilate the observations without “knowing” about the stochastic driving.

The synthetic energy release time series cover one hyperdiffusion time, and were produced on a 48×48 grid, with time step $\Delta t / \tau_v = 5.5 \times 10^{-11}$, and forcing parameters as given in §2.2.2.

4.1.2. Thresholding and binning

Figure 7 illustrates a segment of representative simulation run. The gray line in the bottom panel of Figure 7 is a typical time series for the energy released by the avalanches. Time is measured in units of the magnetic hyperdiffusion time scale, and energy in some arbitrary units. This arbitrary energy scale can be mapped to the standard flare classification by dividing the peak energy that covers three decades (panel B of Figure 4), in four ranges equal in logarithmic size. Thus, the following association can be made: $P < 8$ are B-class flares, $8 < P < 40$ are C-class flares, $40 < P < 200$ are M-class flares and $P > 200$ are X-class flares.

Data assimilation is carried out on a binned version of the energy time series, shown as a thick solid line in Figure 7. Bin width of $\Delta b = 100$ (i.e. 100 time steps) were chosen, as they large enough to remove the small structures but small enough to keep the general features of the avalanche This binning facilitates the minimization of the cost function by eliminating the fine structure details, which are unnecessary as we are mostly interested in forecasting the flaring time, peak flux, and total released energy.

Two parameters are used to build statistics. First, there is an energy threshold (horizontal dashed lines on Figure 7). Only avalanches with an energy above the threshold are considered in the statistics. The signals below the threshold are treated as being noise or low energy avalanches that are unimportant. Here the two thresholds that will be used in this paper are shown: a threshold at 90 and a larger one at a value of 200 energy units. The lower threshold retains the model’s equivalents of upper half of the medium size (M-class) and large size (X-class) flares while the larger one will only take the high energy X-flares into account. These classes of flares are the ones having the most important effects on Earth. The second parameter is the forecast window. A forecast window of a range of $\delta t = 0.55 \times 10^{-7}$ is depicted in the upper left corner of Figure 7. The forecast

window is the maximal time interval between an observed peak and a forecasted peak to have a match. In the case of a forecast window of $\delta t = 0.55 \times 10^{-7}$, the predicted peak can either be $\delta t = 0.275 \times 10^{-7}$ before or after the observed peak. Hence, a smaller forecast window implies a more precise forecast.

The top four panels of Figure 7 show the evolution of the avalanching regions at different stages throughout the spatiotemporal evolution of a large avalanche. Starting from an initial perturbation at a given random point, the ‘‘preflare’’ stage begins with an avalanching region occupying a small region (panel A, Figure 7) of the domain that will gradually increase (panel B, Figure 7) to reach a large size, occupying an important portion of the domain, at the ‘‘impulsive stage’’ at the end of which the peak is reached (panel C, Figure 7). At this point, the fragmented hyperdiffusion front has reached the sides of the domain. In the ‘‘decay’’ stage, the avalanching regions start to decrease in size (panel D, Figure 7). Eventually, the system will return to a stable state, and driving resumes until the triggering of the next flare/avalanche.

4.1.3. *Maintaining the SOC state*

To prevent the system from leaving the SOC state due to the 4D-VAR correction made to the initial conditions, the lattice energy of the corrected initial conditions is compared to the average lattice energy, \bar{E}_l , of the SOC state. If the lattice energy of the corrected initial conditions is not found within the variance σ , the 4D-VAR correction is adjusted by a factor ε defined as:

$$\varepsilon = \begin{cases} -\frac{|E_l - \bar{E}_l|}{\sigma} & \text{if } E_l > \bar{E}_l + \sigma \\ 1 & \text{if } \bar{E}_l - \sigma < E_l < \bar{E}_l + \sigma \\ \frac{|E_l - \bar{E}_l|}{\sigma} & \text{if } E_l < \bar{E}_l - \sigma \end{cases} \quad (36)$$

which brings the initial conditions’ lattice energy within the variance.

4.1.4. *Random number sequences*

For the minimization procedure within the 4D-VAR algorithm, the same seed is used to initialize the random number generator, so that during each iteration within 4D-VAR (cf. Figure 6), the same sequence of perturbations are added in the same order to the same mesh points.

4.1.5. *Running 4D-VAR*

With a realization of stochastic driving different from that having been used to produce the synthetic observations (top panel Figure 8), a DNS ‘‘forecast’’ is produced (middle panel Figure 8). The same initial conditions were used in both cases, but the different realizations of the driving have led, perhaps not surprisingly, to very different time series of energy release. Although the DNS run has reproduced the small avalanche at $t = 1.1 \times 10^{-7}$, it missed the large one at $t = 3.5 \times 10^{-7}$. The 4D-VAR run (bottom panel Figure 8), using the same driving realization as the

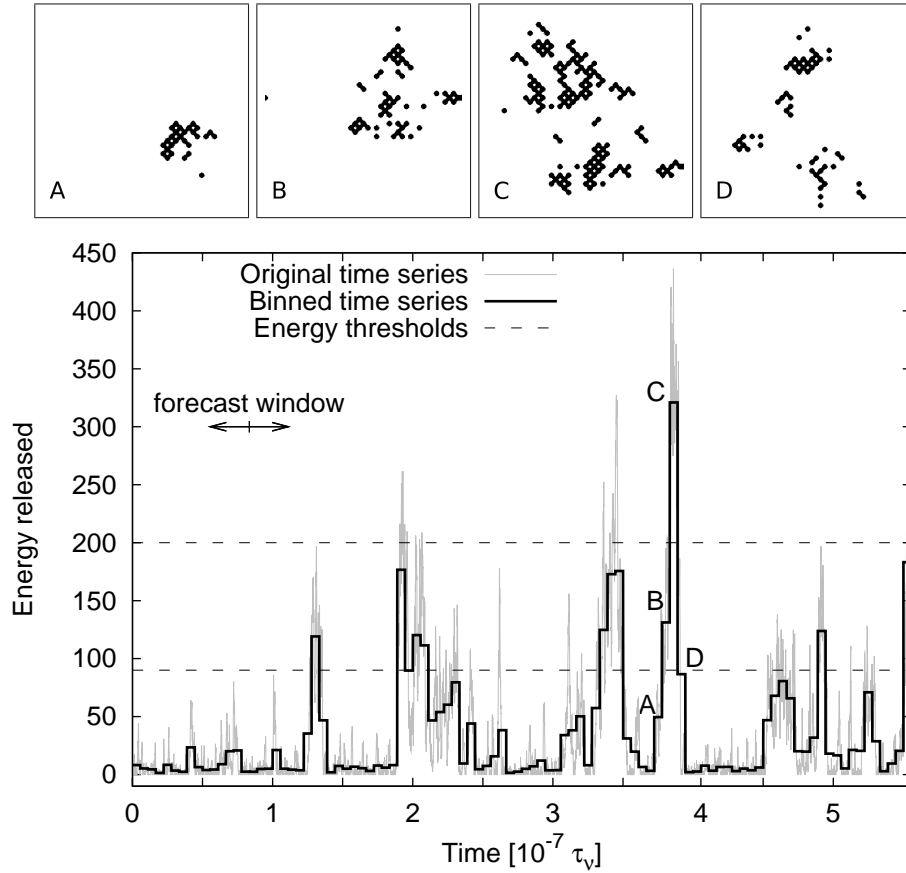


Figure 7. Several aspects of the experimental design are displayed in the bottom panel. The gray line is an original time series of released energy generated by the avalanche model (Equation (6)). The thick solid line is the same series which has been binned by using bins of width of $\Delta b = 100$. Only the avalanches above the energy threshold (horizontal dashed lines) will be considered. In the top left corner, a typical forecast window is shown. It determines the maximum time interval between an observed and forecasted avalanche in order to have a match. The top four panels show the avalanching regions evolving in space and time. From the initial perturbation at a random point, the avalanching region increases from a small region (A) to a large region occupying an important portion of the domain where the hyperdiffusion has reached the sides of the domain (B & C). The avalanche is typically fragmented in its decay phase (D).

DNS forecast but with corrected initial conditions, has correctly reproduced the large avalanche. Figure 8 is a typical case when the 4D-VAR method performs well. There are cases where the DNS is already quite good, and 4D-VAR cannot produce significant improvement; this is in fact expected, and moreover is the reason why true forecasting may be possible despite the stochastic nature of the forcing (more on this in the concluding section). In this representative sample run, minimization

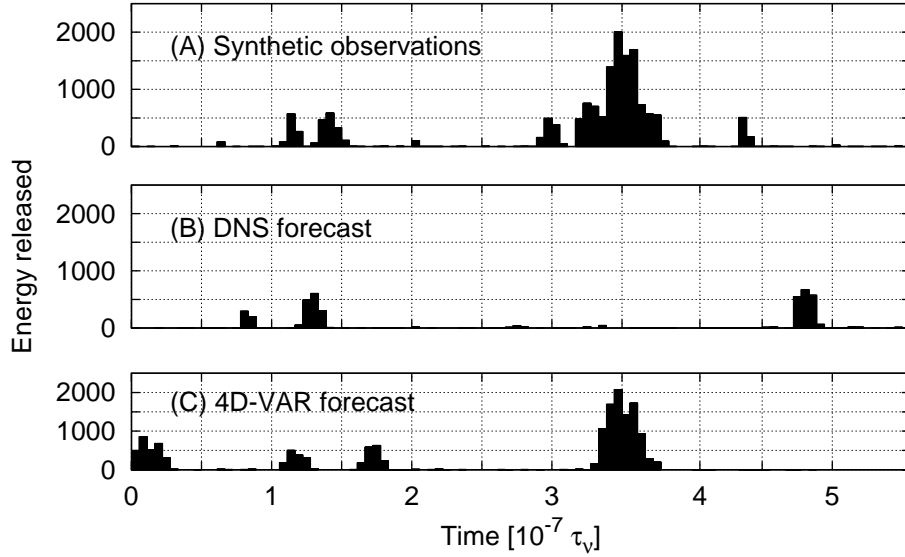


Figure 8. This figure displays a sample run of 4D-VAR data assimilation. The synthetic observations being assimilated are in the top panel. Even using the same initial condition $A(x, y, t = 0)$, a single DNS “forecast” (middle panel) often produces poor results, a direct consequence of the stochastic nature of the driving process. Retaining the same driving but allowing 4D-VAR to alter the initial condition (bottom panel) results in a much better representation of the observations.

of the cost function by 4D-VAR was achieved in a mere 5 iterations of the conjugate gradient.

4.2. PERFORMANCE

4.2.1. Performance measurements

In anticipation of true forecasts, it is instructive to analyze the performance of the 4D-VAR runs in terms of matches, misses and false alarms. Only the avalanches with a peak above the energy threshold are considered. With the forecast window centered at the peak of each observed avalanche, the forecast is examined to see if one of its avalanche take place inside the window. If it is the case, we have a match. If we use an energy threshold of 200 and a forecast window of $\delta t = 0.55 \times 10^{-7}$, Figure 8 has 2 matches: the large avalanche at $t = 3.4 \times 10^{-7}$ and a smaller one at $t = 1.1 \times 10^{-7}$. The avalanche at $t = 1.1 \times 10^{-7}$ is a possible match for either one of the avalanches at $t = 1.1 \times 10^{-7}$ and $t = 1.38 \times 10^{-7}$. Such situations are treated as a single match. Hence, a match can be considered as an event happening in a time interval, determined by the forecast window, regardless if there is a single avalanche or multiple consecutive avalanches. A miss happens when the observed

avalanche does not have a counterpart in the 4D-VAR run. Figure 8C has 2 misses: at $t = 2.9 \times 10^{-7}$ and $t = 4.3 \times 10^{-7}$. False alarms are avalanches appearing in the 4D-VAR run which do not have counterparts in the observations. Figure 8C has one false alarm at $t = 1.7 \times 10^{-7}$. The avalanche at $t = 0.1 \times 10^{-7}$ is not considered as a false alarm as it is an artifact of the 4D-VAR method. The correction to the initial conditions has left the system in an unstable state so the forecast time series began with an avalanche. Finally, the two small avalanches at $t = 0.6 \times 10^{-7}$ and $t = 2 \times 10^{-7}$ in the observations are not included in the analysis as they fall below the energy threshold.

The final value of the cost function is not necessarily an optimal measure of a successful run, as it simply measures the mean quadratic difference between the observation and model time series over the whole duration of the assimilation interval. In the flare forecasting context, one is primarily interested in predicting the timing of discrete events, namely the largest flares/avalanches, and ideally also a measure of their peak flux and/or total released energy. Consequently, we define the following quality factor Q to assess numerically whether a given run was successful or not in “catching” avalanches in the observations:

$$Q = \alpha \sum_{\text{match}} \left| \frac{E_o}{E_o - E_f} \right| - \beta \sum_{\text{miss}} \left(\frac{E_o}{E_{tot}} \right) - \gamma \sum_{\text{false alarm}} \left(\frac{E_f}{E_{tot}} \right). \quad (37)$$

In Equation (37), the first term is the sum, for all pair of matching flares, of the inverse of the error difference between the total energy of the observed (E_o) and modeled (E_f) avalanches multiplied by a factor α . The two next terms are penalties due to the misses and false alarms. The penalty are defined as the energy of the missed (E_o) and false alarm (E_f) avalanches normalized by the total energy released by all the observed avalanches (E_{tot}). These three individual contributions to Q are then each assigned a distinct weighting factor, chosen here as $\alpha = 4$, $\beta = 2$, and $\gamma = 1$ where the largest weight is for the matches. Note also that under these definitions, missing a large avalanche incurs a larger penalty than missing a small one. The misses have a larger weight than the false alarms because a miss leaves us unprepared to handle the consequence of the flare. On the other hand, in a false alarm we may incur additional costs even though no flare is triggered.

4.2.2. Performance statistics

Statistics of hits, misses and false alarms has been gathered for 100 4D-VAR runs. These runs have been realized with combinations of ten sets of observations and ten sets of distinct random number sequences each setting a distinct realization of stochastic driving within the SOC model. For all runs included in the statistical analysis to follow, an energy threshold of 90 and a forecast window of $\delta t = 0.55 \times 10^{-7}$ were used.

The idea here is to investigate the performance of the system *over the assimilation interval* (we are still not forecasting at this stage!), in terms of the number of consecutive matches that could be obtained before the first miss. The results are

Table I. Performance of the system in matching n consecutive avalanches before the first miss. The results from the regular DNS runs (numbers in square brackets) were added for comparison purpose. The second and third columns lists the number of runs that had a match or a miss for the n^{th} avalanche, respectively. Fourth and fifth columns displays the number of false alarms between the n^{th} and $(n - 1)^{\text{th}}$ avalanches and the corresponding numbers of runs which had these false alarms. The sixth column keeps track of the runs that have neither a match nor a miss because there are no longer any avalanches above the threshold.

Number of consecutive avalanches	Number of matches	Number of misses	Number of false alarms since last match		Number of avalanches below threshold
			False alarms	Number of runs	
1	65 [41]	35 [59]	0	44 [26]	0 [0]
			1	12 [9]	
			2	8 [3]	
			3	0 [2]	
			4	1 [1]	
2	27 [22]	28 [14]	0	22 [15]	10 [5]
			1	4 [3]	
			2	1 [4]	
3	11 [13]	7 [6]	0	10 [8]	9 [3]
			1	1 [4]	
			2	0 [1]	
4	8 [1]	1 [5]	0	8 [1]	2 [7]
5	2 [0]	0 [0]	0	2 [0]	6 [1]
6	2 [0]	0 [0]	0	2 [0]	0 [0]
7	0 [0]	0 [0]	0	0 [0]	2 [0]

tabulated in Table I for each n^{th} avalanche (first column). The second and third columns lists the number of runs that had a match or a miss for the n^{th} avalanche, respectively. The fourth and fifth columns displays the number of false alarms between the n^{th} and $(n - 1)^{\text{th}}$ avalanches and the number of runs which had these false alarms. The last column accounts for the fact that the runs do not have the same number of avalanches above the threshold. The results of applying the same analysis to regular DNS runs are added for comparison purpose (numbers in square brackets).

Examining the first row, we see that, of the 100 runs, 65 4D-VAR runs succeeded in reproducing the first avalanche (above threshold) while 35 runs missed it. Of the 65 runs with the match, 44 had no false alarms between the beginning of the

run to the first avalanche while 21 of them had between 1 and 4 false alarms in this time interval. The DNS runs were less successful in reproducing the first avalanche as 41 of them were able to do so. The proportion of false alarms before the first avalanche is similar for both the 4D-VAR and the DNS runs. Every run, 4D-VAR and DNS, had at least one avalanche above the threshold. This first result is quite interesting: 44% of the 4D-VAR runs successfully reproduced the first observed avalanche without making a false alarm; this is a 60% increase in performance compared to the DNS runs. If we move on to the second row, of the 65 4D-VAR runs which reproduced the first avalanche, 27 of them were also able to reproduce the second avalanche. The remaining 38 runs either missed the second avalanche (28 runs) or did not had a second avalanche higher than the threshold (10 runs). The number of false alarms between the first and second avalanches follows the same trends as the ones before the first avalanche. The difference between the 4D-VAR and DNS runs is less pronounced as 22 DNS runs matched the second avalanche. This implies that the 4D-VAR method is very good at reproducing the first avalanches but afterward the performance degrades to become equivalent to the DNS method. However, the 4D-VAR method did reproduce runs of 5 and 6 consecutive avalanche when the DNS runs no longer produced avalanches above the energy threshold. Finally, the number of consecutive matches continue to decrease until either an avalanche is missed or all observed avalanches are reproduced.

The reason behind this somewhat sudden decrease in performance come from the fact that avalanches, especially large ones, have a deep impact on the energy distribution on the lattice. Thus, even with synthetic data produced by the same model used for data assimilation, it become more difficult to reproduce the next avalanche, which explains the constant decrease in the number of matches. This is a direct consequence of the stochastic nature of the driving process, and already heralds the finite forecasting window that can be expected when operating in true forecasting mode. Nevertheless, one tenth of the total 100 runs could still reproduce the first three avalanches. Only 2 runs reproduced the first six avalanches above threshold (although 61 runs did not have a sixth avalanche to reproduce). However, at this point, the size of the time interval for the run ($t = 5.5 \times 10^{-7}$) is felt as 9 runs that matches the first two runs did not have a third avalanche above the threshold. The runs have an average of 4 avalanches above the energy threshold.

4.2.3. *Performance of the code*

At a spatial resolution of 48×48 , it takes about 5 minutes of wallclock time to complete a data assimilation run over 10000 time steps on an Intel Itanium 2 processor. Note that no particular efforts were made towards code optimization. Although the writing of the adjoint equations in the 4D-VAR implementation can be difficult, the resulting data assimilation scheme is quite fast. The hope is that this performance will not degrade too much once real data, including observational errors, will be used for true flare forecasting.

5. Conclusion: towards forecasting

What have we actually achieved with this whole data assimilation procedure? Let's go back to the idea of predicting flare occurrence and energy release via a direct numerical simulation based on a SOC avalanche model. To run such a model forward in time, two things must be specified: (1) the current state of the lattice at time t_0 , to be used as initial condition for the forecast, and (2) the spatiotemporal sequence of perturbations throughout the forecasting interval. If the latter are truly stochastic in nature, they remain completely unknown at t_0 . In the context of Parker's nanoflare hypothesis, which provides the physical underpinning of avalanche models in the present situation, these perturbations amount to small kinks between adjacent magnetic fieldlines, building up in response to slow forcing of the structure's photospheric magnetic footpoints. Not only do these kinks develop in response to stochastic forcing, but they also occurs on spatial scales inaccessible to direct observation. This means that flare forecasting using an avalanche model will always retain a stochastic component.

What we have shown in this paper is that past avalanching behavior can be reproduced reasonably well using data assimilation, *even without detailed knowledge of the stochastic forcing*. At the end of the assimilation interval, the lattice is in a state that is compatible with (and determined by) past flaring behavior. This then represents the optimal initial condition from which to carry out a DNS forecast. This, of course, does not guarantee that any given DNS forecast will be accurate, but that an ensemble of DNS forecasts will show avalanching patterns that reflect, at least in part, the state of the lattice at t_0 . In particular, if this initial condition is characterized by a large, connected portion of the lattice close to the instability threshold, then one would expect that a large avalanche is likely in the near future, irrespective of the spatiotemporal details of the forcing. It should then be possible to forecast with some accuracy the largest upcoming avalanches using statistical ensembles of DNS runs. Small avalanches, on the other hand, will depend more sensitively on details of the (stochastic) forcing. In such cases, even ensemble of DNS runs are less likely to produce useful forecasts. In the space weather context, this not too problematic, since it is precisely the largest flares/avalanches for which one is seeking accurate forecasts. These expectation are examined in detail in the following paper in this series (Bélanger, Vincent, and Charbonneau, 2007).

SOC avalanche models are certainly not the only modeling framework for solar flares within which data assimilation can be carried out. A good case in point is the CISM project (see Solomon, 2005; Siscoe and Solomon, 2006), an ambitious sun-to-ionosphere data assimilation framework based on a suite of coupled 3D MHD models. The attractive feature of SOC models—arguably their single most attractive feature—is that, by all appearance, they correctly capture the global statistical behavior of energy release by solar flares, including in particular its power-law form and associated exponent. This makes such models ideal candidates for data assimilation-based forecasting, despite their extreme physical simplicity and the

inevitable stochastic effects associated with the driving mechanism responsible for energy injection into the overlying coronal magnetic structures.

We note, in closing, that within Parker's physical picture of coronal structures being forced by photospheric fluid motions, such stochastic effects would also need to be incorporated into any full-scale MHD models of coronal structures to be used for flare forecasting. Data assimilation could help here as well (see, e.g., Schrijver and DeRosa, 2003), but spatially and/or temporally unresolved fluid motions would again introduce a form of stochastic "noise" in the MHD simulations, with inevitable degradation of forecasting performance even if such models would be based on true physical equations rather than some largely ad hoc cellular automaton. The latter, however, is such a simpler model to simulate that it becomes possible in practice to carry out ensemble DNS forecasting in reasonable wallclock time even on mid-range computational platforms. This is an essential requirement of operational forecasting.

Acknowledgements

The authors would like to acknowledge financial support of the FQRNT-NATEQ (Projet de recherche en équipe: Avalanches et prédiction des éruptions solaires) and NSERC Discovery Grant grants to A. V. and P. C. The authors are also grateful to RQCHP for providing computational support.

References

- Anderson, J. L. 2003, in *Predictability of a Data Assimilation / Prediction System*, NCAR Data Assimilation Initiative
- Aschwanden, M. J. and Charbonneau, P. 2002, *Astrophys. J.*, 566, L59
- Aschwanden, M. J., Tarbell, T. D., Nightingale, R. W., *et al.* 2000, *Astrophys. J.*, 535, 1047
- Bak, P. 1996, *How nature works: The science of self-organized criticality* (New York: Springer-Verlag)
- Bak, P., Tang, C., and Wiesenfeld, K. 1987, *Phys. Rev. Lett.*, 59, 381
- Bertino, L., Evensen, G., and Wackernagel, H. 2002, *Inverse Problems*, 18, 1
- Bélangier, E., Vincent, A., and Charbonneau, P. 2007, in preparation.
- Carton, J. A., Chepurin, G., and Cao, X. 2000, *J. Phys. Oceanogr.*, 30, 311

- Charbonneau, P., McIntosh, S. W., Liu, H.-L., and Bogdan, T. J. 2001, *Solar Phys.*, 203, 321
- Courtier, P. and Talagrand, O. 1990, *Tellus*, 42A, 531
- Cressman, G. P. 1959, *Mon. Wea. Rev.*, 87, 367
- Daley, R. 1991, *Atmospheric Data Analysis*, Cambridge atmospheric and space science series (Cambridge: Cambridge University Press)
- Dennis, B. R. 1985, *Solar Phys.*, 100, 465
- Ehrendorfer, M. 1992, *Quart. J. Roy. Meteor. Soc.*, 118, 673
- Errico, R. M. 1997, *Bull. Amer. Meteor. Soc.*, 78, 2577
- Eymin, C. and Fournier, A. 2005, IAGA2005
- Georgoulis, M. and Vlahos, L. 1998, *Astron. Astrophys.*, 336, 721
- González, Á., Vázquez-Prada, M., Gómez, J. B., and Pacheco, A. F. 2006, *Tectonophysics*, 424, 319
- Isliker, H., Anastasiadis, A., and Vlahos, L. 2000, *Astron. Astrophys.*, 363, 1134
- Jensen, H. J. 1998, *Self-Organized Criticality* (Cambridge: Cambridge University Press)
- Kadanoff, L. P., Nagel, S. R., Wu, L., and Zhou, S. 1989, *Phys. Rev. A*, 39, 6524
- Kalnay, E. 2003, *Atmospheric modeling, data assimilation and predictability* (Cambridge: Cambridge University Press)
- Kantha, L. H. and Clayson, C. A. 2000, *International Geophysics Series*, Vol. 66, *Numerical Models of Oceans and Oceanic Processes* (Academic Press)
- Le Dimet, F.-X. and Talagrand, O. 1986, *Tellus*, 38A, 97
- Lions, J. L. 1968, *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles* (Paris: Dunod)
- Liu, H.-L., Charbonneau, P., Pouquet, A., Bogdan, T., and McIntosh, S. 2002, *Phys. Rev. E*, 66, 056111
- Lu, E. T. and Hamilton, R. J. 1991, *Astrophys. J. Lett.*, 380, L89
- Lu, E. T., Hamilton, R. J., McTiernan, J. M., and Bromund, K. R. 1993, *Astrophys. J.*, 412, 841
- Moore, A., Arango, H., Lorenzo, E., *et al.* 2004, *Ocean Modelling*, 7, 227

- Nichols, N. 2003, in Proceedings of the NATO Advanced Study Institute on Data Assimilation for the Earth System
- Norman, J. P., Charbonneau, P., McIntosh, S. W., and Liu, H.-L. 2001, *Astrophys. J.*, 557, 891
- Parker, E. N. 1983, *Astrophys. J.*, 264, 635
- Parker, E. N. 1988, *Astrophys. J.*, 330, 474
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. 1992, *Numerical Recipes in FORTRAN : The Art of Scientific Computing* (New York: Cambridge University Press)
- Rundle, J., Klein, W., Tiampo, K., Donnellan, A., and Fox, G. 2003, *Lecture Notes in Computer Science*, Vol. 2659, *Computational Science - ICCS 2003* (Springer Berlin/Heidelberg), 827–836
- Sanders, B. F. and Katopodes, N. D. 1999, *Journal of Irrigation and Drainage Engineering*, 125, 287
- Sanders, B. F. and Katopodes, N. D. 2000, *Journal of Engineering Mechanics*, 126, 909
- Schrijver, C. J. and DeRosa, M. L. 2003, *Solar Phys.*, 212, 165
- Schröter, J., Seiler, U., and Wenzel, M. 1993, *J. Phys. Oceanogr.*, 23, 925
- Siscoe, G. and Solomon, S. 2006, *Space Weather*, 4, S04002
- Solomon, S. 2005, CEDAR/GEM Meeting, Santa Fe, New Mexico
- Talagrand, O. and Courtier, P. 1987, *Quart. J. Roy. Meteor. Soc.*, 113, 1311
- Wheatland, M. S. 2000, *Astrophys. J.*, 536, L109
- Zirker, J. B. and Cleveland, F. M. 1993, *Solar Phys.*, 145, 119