From mechanistic to phenomenological models: integrated forecasting and modeling of earthquake dynamics

Marian Anghel[(1)], Yehuda Ben-Zion[(2)], and Ramiro Rico-Martinez[(3)]

(1) Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM, U. S. A (e-mail: manghel@cnls.lanl.gov). (2) Department of Earth Sciences, University of Southern California, Los Angeles, CA, U. S. A. (e-mail: benzion@usc.edu). (3) Department of Chemical Engineering, Instituto Tecnologico de Celaya, Celaya, Guanajuato, Mexico (e-mail: ramiro@losalamos.princeton.edu).

Abstract

We present a method of constructing low-dimensional nonlinear models describing the main dynamical features of a discrete 2D cellular fault zone, with many degrees of freedom, embedded in a 3D elastic solid. The fault system contains a vertical planar fault with a uniform grid of cells where slip is governed by a static/kinetic friction law surrounded by regions where a uniform slip rate is prescribed to represent the tectonic loading. Quasi-static stress transfer and tectonic loading along the fault are calculated using 3D elastic dislocation theory. A given fault system is characterized by a set of parameters that describe the dynamics, rheology, property disorder and fault geometry. Depending on the location in the system parameter space, we show that the coarse dynamics of the fault is confined to an attractor whose dimension is significantly smaller than the space in which the dynamics takes place. Our strategy of system reduction is to search for a few coherent structures that dominates the dynamics and to capture the interaction between these coherent structures. The identification of the basic interacting structures is obtained by applying the Proper Orthogonal Decomposition (POD) to the surface deformations fields generated during the evolution of the fault. We use a feed-forward artificial neural network (ANN) architecture for the identification of the system dynamics projected onto the subspace (model space) defined by the most energetic coherent structures. The ANN is trained using a standard back-propagation algorithm to predict (map) the values of the observed state at a future time given the observed state at the present time. This ANN then provides an approximate dynamical model for the fault. The map can be evaluated once to provide short term predictions or iterated to obtain prediction for the long term fault dynamics.

Introduction

A first principles approach to modeling and forecasting the dynamics of an earthquake fault is not directly applicable because the governing physical laws and observations of controlling variables are not fully available at present. An alternative

to this modeling approach is to build "phenomenological" models that attempt to estimate the qualitative character of the system's dynamics, and to make short-term predictions based on that understanding, without attempting to provide a detailed explanation of the physical mechanisms that ultimately govern the behavior of the system. The basic idea of this approach is to develop data-driven models based on analysis of the spatio-temporal strain patterns embedded in the observable surface displacements. This approach is based on the observation that the effective dimension of the attractor of the dynamics is often significantly smaller than the original phase space in which the dynamics takes place. In such cases, a few macroscopic observables can approximate very well the present state of the system and predictive models based on the dynamics of a reduced number of macroscopic observables can then be constructed.

In order to identify these low-dimensional dynamical structures we apply the Proper Orthogonal Decomposition (POD) - also known as the Principal Component Analysis (PCA) or Karhunen-Loève expansion - to the ensemble of surface deformation data generated by the system. When dynamic weakening effects are present, we show that a reduced number of deformation modes can explain on average most of the elastic surface deformation. The dynamics of these modes evolves on a low-dimensional attractor in the neighborhood of which the system spends most of its time. The state of the system in this reduced space (model space) is represented by a set of modal coefficients that measures the projection of the ground surface deformation onto each dominant mode. Our next goal is to extract the nonlinear modal dynamics in the model space by constructing a map of the observed state at a future time given the observed an state at the present time. We describe preliminary results in which artificial neural network has been successfully trained to approximate the dynamics of the system projected in the model space.

Our method may be compared to the linear pattern dynamics introduced by Rundle and his coworkers [8]. Their technique is based upon a Karhunen-Loève expansion of the spatio-temporal seismicity data and is used to estimate a linear stochastic model for the evolution of a probability density function for seismic activity. In contrast to this method, which provides a local linear approximation in a probability space, we propose a global nonlinear approximation that describes the effective dynamics in a low-dimensional phase space.

Earthquake Model

The fault system corresponds to a discrete strike-slip fault of length $L = 70$ km and width $W = 17.5$ km embedded in $3D$ elastic continuum [2]. The fault consists of a uniform grid of dynamical cells where slip is governed by static/kinetic friction processes, surrounded by regions with imposed constant slip rate of $V_{pl} = 35$ mm/year, representing the tectonic loading. The brittle deformation at any fault position and time is governed by quasi-static $3D$ elastic dislocation theory and prescribed distributions of stress drops at fault locations where stresses reach the frictional strength thresholds. We assume uniform static brittle strength, set to a value of τ_s =100 bars, everywhere on the fault plane. The stress at any fault position (cell) increases with time due to the gradual tectonic loading and the time-dependent brittle deformation at other fault locations. If the stress τ_{ij} reaches the static strength τ_s, a brittle failure occurs at this location and cell (ij) slips so as to reduce its stress to an

arrest level, $\tau_{a,ij}$. The stress transfered from the failed cell can lead to subsequent brittle failures (i.e., rupture propagation) if the stress anywhere increased to the failure threshold. These failures may, in turn, induce or reinduce more brittle slip events. After an initial slip event the strength drops to a dynamic level, $\tau_{d,ij} < \tau_s$, and reinitiation of brittle slip on an already failed cell occurs when $\tau_{ij} \geq \tau_{d,ij}$ there. The failure iterations end when there are no more brittle instabilities. This marks the end of an earthquake event whose strength is measured by its potency P defined as the integral of slip over the rupture area. At the end of each model earthquake the strength recovers back to τ_s. The static strength, dynamic strength, and arrest stress are related to each other everywhere on the fault as

$$D = \frac{\tau_s - \tau_{a,ij}}{\tau_s - \tau_{d,ij}} \quad , \quad i = 1, \ldots, N_x \;\; j = 1, \ldots, N_y \tag{1}$$

where D is a dynamical overshoot coefficient (see Ben-Zion and Rice [2] for additional details). Its inverse, $1/D$, proportional to $\tau_s - \tau_{d,ij}$, is a measure of the dynamic weakening that characterizes static/kinetic friction models. An important component of the fault rheology is the distribution of arrest stresses along the fault. In the simulations used in this paper we chose random stresses, $\tau_{a,ij} =< \tau_a > +A\xi$, where $< \tau_{a,} >$ represents the fault-averaged arrest stress, ξ is a random number uniformly distributed in the range $[-0.5, 0.5]$, and A is the noise amplitude. The arrest and dynamic stress distributions do not evolve in time (quenched heterogeneities) and the dynamics is deterministic.

Effective dimension of the large scale motions

For the class of systems discussed in this paper the knowledge of the slip deficit, $\phi_{ij} = V_{pl}t - u_{ij}$, fully specifies the state of the fault: the number of degrees of freedom is $M = N_x \times N_z$ and the state of the system is completely represented by a vector in a finite, but high-dimensional vector space R^M, where $M = 4096$ ($N_x = 128, N_y = 32$). We conjecture that a finite dynamic weakening leads to the creation of spatial and temporal correlations that collapse the large scale dynamics of the fault on an attractor of a smaller effective dimension m; $m \ll M$. To prove this assertion we investigate the effective dimension of two fault systems that differ only in the value of the dynamic overshoot coefficient. System A with $D = 1.5$ has a large dynamic weakening, while system B with $D = \infty$ has no dynamic weakening. Both systems have random arrest stress distributions with $< \tau_a >= 80$ bars and $A = 20$ bars amplitude.

Each system observable contains a different amount of information for understanding the underlying dynamics. For prediction purposes we are only interested in the dynamics at the large length and time scales. We have observed that the evolution of the average slip deficit,

$$\phi(t) = \frac{1}{M} \sum_{i,j} (\phi_{ij}(t) - < \phi_{ij} >) \quad , \tag{2}$$

and slip deficit variance,

$$\sigma(t) = \frac{1}{M} \sum_{i,j} (\phi_{ij}(t) - \phi(t))^2 \quad , \tag{3}$$

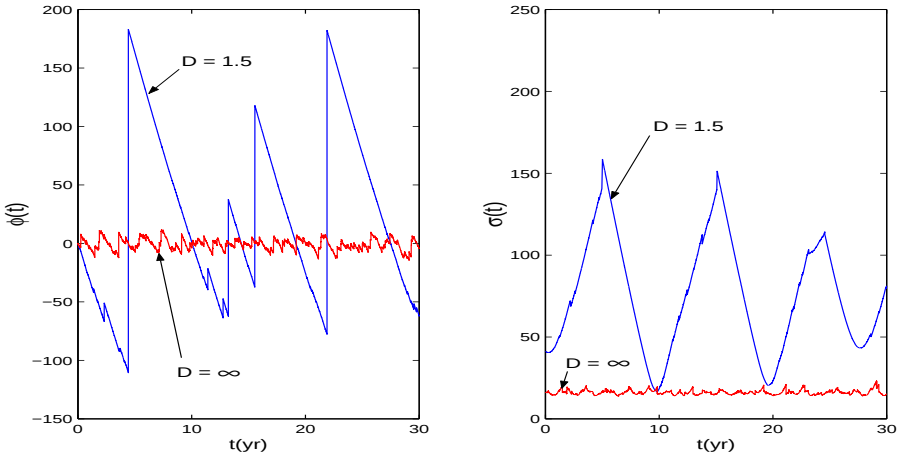

Figure 1: The evolution of the average slip deficit (left) and its variance (right) for the $D = 1.5$ (blue line) and $D = \infty$ (red line) fault systems.

provides a good qualitative measure of the dynamics at these scales. Due to the inhomogeneity of the slip deficit, the slip fluctuations are measured from a local time average $< \phi_{ij} >$. For system A ($D = 1.5$), we observe in Fig. 1 large oscillations and smooth long term behavior between two consecutive large events and a clear separation of time scales between the slow, large scale dynamics and the fast, small scale dynamics. In principle, we can isolate the large scale dynamics which, evolving slowly, will drive the rapidly relaxing small scale dynamics. By contrast, system B ($D = \infty$) has small oscillations and noisy behavior; note the large difference in the scale of the variance evolution for these two cases. For system B there are no large scale events and the signature of small scale events renders the model higher dimensional. A useful upper bound on the dimension of the system that generates a time series is obtained by using a phase space analysis of the dynamics and estimating the optimal value of the embedding dimension m. The embedding concept comes from a phase space analysis of systems with deterministic dynamics. It should be thought [3] as a method for providing a space formally equivalent to the original attractor of the dynamics. In our case the embedding is performed using coordinates made out of an observed slip deficit variance and its time delayed copies,

$$\mathbf{X}_n = [\sigma(n), \sigma(n-d), \ldots, \sigma(n-(m-1)d)] \quad , \tag{4}$$

where $\sigma(n) = \sigma(n\Delta t)$ is the slip deficit variance measured at equal sampling times $\Delta t = 0.1$ years, d is a time lag (an integer multiple of the common lag Δt), and m is an embedding dimension. One way to estimate an optimal value of m is to look for false neighbors in the phase space at a given value of m. To understand this concept, consider the situation that an m dimensional delay reconstruction is an embedding, but an $(m-1)$ dimensional delay reconstruction is not. If the embedding dimension is too small to unfold the attractor, a small R^{m-1} neighborhood will contain points that belong to different parts of the original attractor. Therefore, at a later time, the images of these points under the system's dynamics will split onto different groups, depending on which part of the attractor the points are originally coming

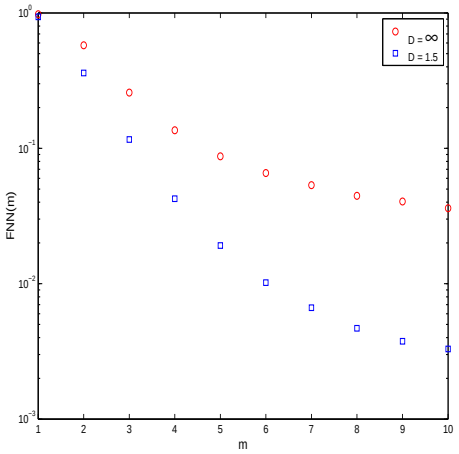

Figure 2: The fraction of false nearest neighbors (FNN) as a function of the embedding dimension $m = 1, \ldots, 10$. Compared to model B ($D = \infty$), there is a dramatic decrease in the number of FNN for model A ($D = 1.5$), as the dimension of the embedding phase space is increased. This suggests that a low-dimensional deterministic dynamics is a good approximation for the large scale dynamics of the fault.

from. This lack of a unique location of all the images in $(m-1)$ dimensions is reflected in finding false neighbors, meaning that determinism is violated. When increasing m, starting with small values, one can detect the minimal embedding dimension by finding no more false neighbors. Then, we can use the fact that under quite general circumstances the attractor formed by $\mathbf{X}_n$ in the embedding space is equivalent to the attractor in the unknown space in which the system is living if m is larger than twice the box counting dimension, d_0, of the attractor. Often, an $m = d_0$ embedding dimension is enough for unfolding the attractor. As shown in Figure 2, for the model with no dynamic weakening the percentage of false neighbors saturates to a low value but it does not drop to zero. In contrast, the behavior of the more realistic $D = 1.5$ model suggests a topological dimension in the range $3 - 6$. Having found an optimal embedding for the dynamics, we can now compute the maximal Lyapunov exponent in order to estimate if the system's dynamics is chaotic. This exponent measures the exponential divergences of nearby trajectories and is an average of these local divergences over the whole data. A positive maximal Lyapunov exponent is a signature of chaos and it also sets a limit for predictability. Its computation is based on the algorithm described by Kantz and Schreiber [3]. We choose a point $\mathbf{X}_n$ of the time series in the embedding space and determine all neighbors $\mathbf{X}_{n'}$ within a neighborhood U_n of radius ϵ. Then we compute the average over the distances of all neighbors to the reference part of the trajectory as a function of the relative time Δn. The logarithm of the average distance at time Δn measures the effective expansion rate over the time span Δn. Repeating this calculation over many values of n, the fluctuations over the effective expansion rates will average out.

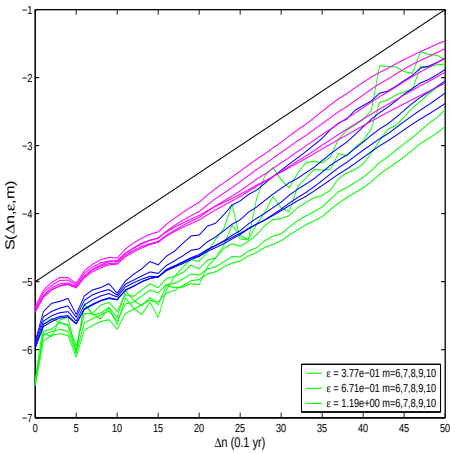

Figure 3: Estimates of the maximal Lyapunov exponent from the slip deficit variance time series data, $\sigma(t)$, for the $D = 1.5$ model. The logarithm of the stretching factor is computed for three different neighborhood sizes, ϵ, and $m = 6, 7, 8, 9, 10$ embedding dimensions. The robust linear behavior of $S(\Delta n)$ reflects the underlying determinism of the data and its slope is an estimate of the maximal Lyapunov exponent, $\lambda_{max} = 0.8$ years^{-1} (the straight line has slope 0.08 and the unit of time is 0.1 years).

Therefore, we compute

$$S(\Delta n, \epsilon, m) = \langle \ln(\frac{1}{|U_n|} \sum_{\mathbf{X}_{n'} \in U_n} |\mathbf{X}_{n+\Delta n} - \mathbf{X}_{n'+\Delta n}|) \rangle_n \quad , \tag{5}$$

where the outer bracket denotes the averaging over the reference point $\mathbf{X}_n$. If for some ranges of Δn, ϵ, and m the function $S(\Delta n, \epsilon, m)$ exhibits a robust linear increase, its slope is an estimate of the maximal Lyapunov exponent λ_{max} per time step. Figure 3 shows three bundles of curves corresponding to three neighborhood sizes, $\epsilon = 0.377$, 0.671, and 1.19, as can be seen for $\Delta n = 0$. Each bundle shows the behavior of the expansion rate for m = 6,7,8,9 and 10 and proves that the result is robust to changes in ϵ and does not depend on the embedding dimension when m is large enough. Our estimate for the maximal Lyapunov exponent is $\lambda_{max} = 0.8$ years^{-1}. We can therefore conclude that the large scale motions have a positive maximal Lyapunov exponent and exhibit sensitive dependence on the initial conditions. This sets a fundamental limitation to long-term earthquake forecasting. The largest Lyapunov exponent, λ_{max}, gives an approximate estimate for the predictability time, $T_p \simeq 1/\lambda_{max} = 1.25$ years, but for extended systems the situation can be very complicated [4]. Generally, the predictability time is scale dependent and can be much longer than the rough estimation $T_p \simeq 1/\lambda$. As we will shortly see, at the large length scale of interest for forecasting, a better estimation of the predictability time is $T_p \simeq 10$ years.

Proper Orthogonal Decomposition

The dynamics of the first and second order moments of the slip and stress fault fields provides an adequate description of the large scale motions. Unfortunately, their measurement poses very difficult problems: the stress field is not directly observable while the slip field requires the solution of an ill-posed inverse problem. Therefore, we propose an alternative analysis of the large scale motions based on the spatio-temporal strain patterns embedded in the surface deformation fields of system A ($D = 1.5$). The space-time signal is obtained by simultaneous measurements of the surface displacements on a 64×32 uniform rectangular grid which covers a 100 km $\times$ 50 km surface area centered around the fault. For each surface deformation we compute an ensemble of N snapshots, $U_\alpha(\mathbf{x}, n) = U_\alpha(\mathbf{x}, t_n)$, $n = 1, \ldots, N$, and $\alpha = x, y, z$, at equally spaced moments, $t_n = n\Delta t$, where $\Delta t = 0.1$ years. The analysis of each deformation direction proceeds identically, therefore we will henceforth drop the deformation index α.

The identification of the active degrees of freedom in the surface deformation fields uses the Principal Orthogonal Decomposition (POD) [1]. In order to perform this analysis it is convenient to separate the flow $U(\mathbf{x}, n) = U(\mathbf{x}, t_n)$ into the time-independent mean flow, $< U(\mathbf{x}) >$, and the fluctuations, $u(\mathbf{x}, n)$, from the ensemble average $< U(\mathbf{x}) >$. Then, we compute the two point correlation matrix of the fluctuations,

$$K(\mathbf{x}, \mathbf{y}) = \frac{1}{N} \sum_{n=1}^{N} u(\mathbf{x}, n) u(\mathbf{y}, n) \quad , \tag{6}$$

whose normalized eigenvalues and eigenfunctions,

$$\int K(\mathbf{x}, \mathbf{y}) \phi_i(\mathbf{y}) d\mathbf{y} = \lambda_i^2 \phi_i(\mathbf{x}) \quad , \qquad \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M \geq 0 \quad , \tag{7}$$

identify and order the spatial modes of deformation. The dominant modes are then used to construct a reduced representation of the model in the low-dimensional subspace spanned by the first m eigenvectors, $\{\phi_i,\, i = 1, \ldots, m; m \ll M\}$:

$$u(\mathbf{x}, n) \simeq \sum_{i=1}^{m} A_i(n) \phi_i(\mathbf{x}) \quad . \tag{8}$$

The eigenvalues λ_i measure the mean square fluctuations of the ensemble in the directions defined by their corresponding eigenfunctions: $\lambda_i =< (u, \phi_i)^2 >$. Therefore, ranked in decreasing order of their eigenvalues, the eigenfunctions (sometimes called empirical eigenfunctions, coherent structures, or dominant modes) will identify the dominant directions in configuration space along which most of the fluctuations take place.

The choice of the embedding dimension m is based on the computation of the cumulative normalized eigenvalue spectrum, Λ_m, defined as:

$$\Lambda_m = \frac{\sum_{j=1}^{m} \lambda_j}{\sum_{j=1}^{M} \lambda_j} \quad . \tag{9}$$

The cumulative spectrum can help us define an effective POD embedding dimension by finding the minimum number of modes needed to capture some specific fraction

$f < 1$ of the total variance of the data:

$$d_{POD} = \arg\min_{m}\{\Lambda_m : \Lambda_m > f\} \quad . \tag{10}$$

For system A, we found that in order to explain $f = 95\%$ of the variance in the u_x, u_y, u_z data sets we need only 1,3, and 5 modes respectively. This is consistent with our earlier estimates of the embedding dimension.

In order to identify the dynamics of the fault in this low dimensional phase space we calculate the projections $A_i(n), i = 1, \ldots, m$, of the surface deformation fields onto the dominant deformation modes, i.e.,

$$A_i(n) = (u(\mathbf{x}, n), \phi_i(\mathbf{x})) \quad . \tag{11}$$

This way, we generate a small number of time series that encapsulate the projection of the system's dynamics onto the m-dimensional *model* space defined by Eq. (8). They encode the evolution, interaction and dynamics of the spatial modes [1]. In Figure 4 we represent the modal time series (red lines) corresponding to the first four u_x spatial modes. The blue lines describe the time evolution of the binned potency released (the total potency released in a time interval of 0.1 years). For the first deformation mode, $A_1(n)$, we notice that the location in time of the amplitude jumps coincides with the time of large events. Moreover, the size of the amplitude jumps is proportional with the size of the event. Therefore, if we can model the evolution of this mode, the time and size of the amplitude jumps will give us accurate information about the time and size of the large earthquake events.

Model Reconstruction and Short-Term Earthquake Forecasting

Using the POD decomposition we have identified a low-dimensional linear space in which system A evolves most of the time and we have reduced its dynamics to a small set of time series, $A_i(n)$, $i = 1, \ldots, m$, describing the evolution of the system in this reduced linear space. We now face the problem of determining the underlying dynamical process from the information available in these time series. They are assumed to be governed by a nonlinear set of ODEs and our modeling approach relies on the ability to identify an approximate m dimensional model,

$$A_i(n+1) = F_i[A_1(n), A_2(n), \ldots, A_m(n)] \quad i = 1, \ldots, m \quad , \tag{12}$$

that describes an explicit Euler approximation to the evolution and interaction of the spatial modes. To identify this nonlinear mapping we employ an artificial neural network (ANN). In this approach the neural network is used as a "black-box" tool in order to develop a process model based only on observations of the system's input-output behavior [5, 6]. In the learning process the network adjusts its internal parameters in such a way as to minimize the squared error between the network output and the desired outputs.

In the present context, we use a feed-forward ANN to learn the dynamics of the reduced model from the time series associated with the evolution of the dominant modes. In order to improve the model forecasting skill it is useful to enlarge the structure of the model to include information about the past history of the modes.

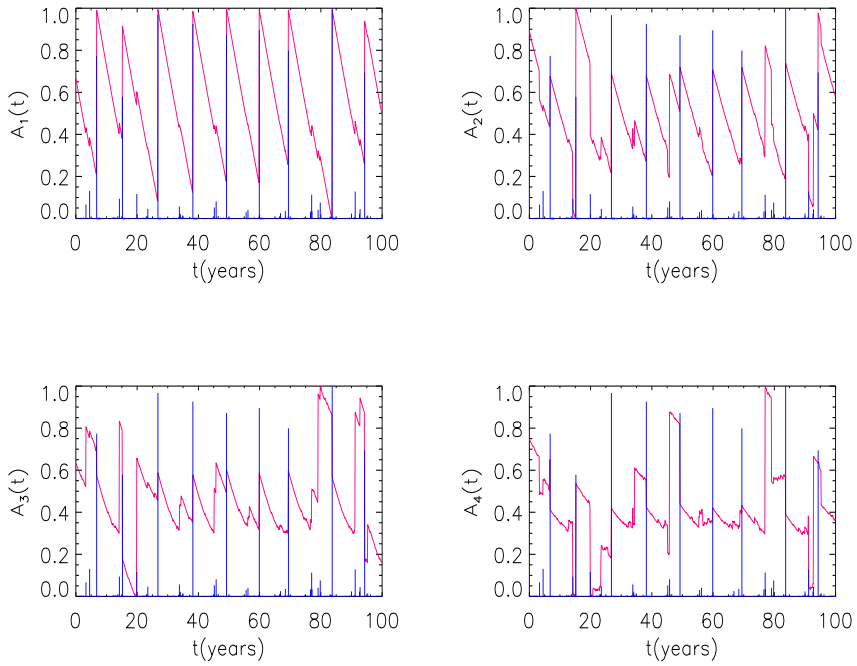

Figure 4: Modal time series (red lines) describing the evolution of the first four u_x spatial modes of system A. This time dependence is determined by projecting the u_x surface deformation onto each POD mode every 0.1 years. The blue lines describe the time evolution of the cumulative potency released over each 0.1 year interval. Note the correlations between the time intervals of high potency released and the discontinuities present in the temporal modes.

Therefore, the input to the network consists of m time-delayed vectors $\mathbf{X}_i$,

$$\mathbf{X}_i = (\underbrace{A_i(n), A_i(n-1), \ldots, A_i(n-d)}_{\text{short-time memory}}, \underbrace{A_i(n-2d), \ldots, A_i(n-(K-1)d)}_{\text{long-time memory}}) \quad , \quad (13)$$

where d is the delay and $(K-1)$ is the number of the time-delay intervals. Due to the presence in the surface dynamics of two different time scales, we have included both short, $A_i(n), A_i(n-1), \ldots, A_i(n-(d-1))$, as well as long time, $A_i(n-d), A_i(n-2d), \ldots, A_i(n-(K-1)d)$, memory information in the time-delayed input vectors. The ANN output then provides a prediction of the temporal mode A_i at moment $(n+1)$,

$$A_i(n+1) = F_i[\mathbf{X}_1(n), \mathbf{X}_2(n), \ldots, \mathbf{X}_m(n)] \quad i = 1, \ldots, m \quad , \quad (14)$$

based on input information describing the past mode histories.

The ANN is trained using a standard back-propagation algorithm [7] to predict (map) the values of the observed state at a future time t_{i+1} given the observed state at the present time t_i. This ANN then provides an approximate dynamical model for the fault. The map can be evaluated once to provide short term predictions or iterated to obtain prediction for the long term fault dynamics.

We have applied this model reconstruction to extract the nonlinear dynamics that controls the evolution of the first two u_x surface deformation modes of the fault system A: $m = 2$. For each mode a time-delayed vector with parameters $K = 6$ and $d = 4$ was used. Because the dimensionality of the input space, $m(K+(d-1)) = 22$, is higher than our estimated embedding dimension, we have decided to perform a POD analysis of the ensemble of input vectors. This time, the POD decomposition performs the analysis of the dominant *temporal* patterns that are created by the modal dynamics. It reduced the 22 dimensional input space to a 6 dimensional representation that contains 99% of the variance of the ensemble of input vectors. The best short-term forecasting performance was obtained for an ANN with two hidden layers of 10 neurons each.

Starting from an initial configuration describing for each input mode the current amplitude and its past K values, we iterate the ANN forward in time for a number of F steps. At each time step the output of the network was used to update and reconstruct the ANN input for the next time step. In Figure 5 we present one of the ANN predictions (blue line) for the $D = 1.5$ model using $F = 200$ steps and we compare it with the true evolution of the most dominant surface mode (red line). The goal is to predict the time and the size of the jumps in the evolution of the first mode amplitude, which, as we have already discussed, corresponds to the time and size (potency) of large seismic events. Due to the time delay involved, the best time resolution of the forecast cannot be in this case less than $T = 0.1$ years. As the fault evolves to the next time step, 0.1 years later, we update the present state of the system and generate a new F step forecast starting from this new state. This procedure is intended to incorporate the information about the system and its current state as is continuously generated by new observations. We observed that depending on the location of the system on its underlying attractor regions of large predictability time coexist with regions of relatively short predictability time [4]. We also noticed a systematic bias in estimating the time to the next large event.

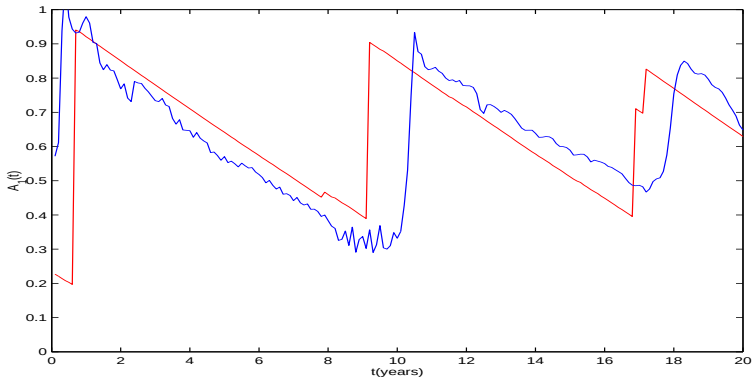

Figure 5: Evolution of the first u_x temporal mode as predicted by the iterated ANN (blue line). The red line represents the projection of the system dynamics onto the first coordinate (mode one) of the model phase space.

Conclusions

We describe a conceptual framework for modeling and forecasting the evolution of a large strike-slip earthquake fault. The approach relies on the detection of spatio-temporal strain patterns embedded in the *observable* surface displacements: no detailed knowledge of the fault geometry, dynamics, or rheology is required. Rather than directly modeling the fault dynamics, we propose instead to model the dynamics of observable surface deformations, which are nonlinearly related to the original dynamics of the fault system. The essence of the method is the discovery that the large length and time scales motions have a strong *low-dimensional, deterministic* component and are therefore amenable to representation by a deterministic model. We have also found that the large scale motions provide reliable forecasting information about the large seismic events. These two fundamental results set the stage for standard data processing and model reconstruction techniques. First, we identify the large scale motions with generalized directions (spatial modes) along which the dynamics has its *largest* fluctuations. Finding these directions is the natural task of the proper orthogonal decomposition applied to the ensemble of surface deformations generated during the evolution of the system. The most dominant spatial modes define the model phase space and provide an optimal embedding for the large scale dynamics of the system. Second, the model reconstruction consists in finding a nonlinear set of ODEs whose trajectory in model phase space approximates the system trajectory projected into the model phase space. This is a *learning* task that can be successfully accomplished by an artificial neural network.

The current paper is concerned primarily with introducing a forecasting methodology. The obtained results are preliminary and a promising framework for getting started. Continuing studies along the directions of this work may have a significant impact on the earthquake predictability problem.

Acknowledgments

We express our gratitude to Yannis Kevrekidis for many useful discussions and insightful suggestions. This research was performed under the auspices of the U. S. Department of Energy at LANL (LA-UR-02-6309) under contract W-7405-ENG-36 and LDRD-DR-2001501 (MA) and the National Earthquake Hazard Reduction Program of the USGS under grant 02HQGR0047 (YBZ).

References

[1] Holmes, P., Lumley, J. L., Berkooz, G., 1996, *Turbulence, Coherent Structures, Dynamical systems and Symmetry*, Cambridge University Press, Cambridge

[2] Ben-Zion, Y. and Rice, J., 1993, J. Geophys. Res., **101**, 14109-14131.

[3] Kantz, H. and Schreiber, T., 1997, *Non-linear time Series Analysis*, Cambridge University Press, Cambridge

[4] Boffetta, G., Cencini, M., Falcioni, M., Vulpiani, A., 2002, Phys. Rep. **356**, 367-474.

[5] Rico-Martinez, R., Krischer, K., Kevrekidis, I. G., 1992, Chem. Eng. Comm. **118**, 25.

[6] Rico-Martinez, R., Kevrekidis, I. G., Krischer, K., 1995, *Neural Networks for chemical Engineers,* ed. Bulsari, A. B., Elsevier Science.

[7] Haykin, S., 1999, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, NJ.

[8] Rundle, J. B., Klein, W., Tiampo, K. F., and Gross, S., 2000, Phys. Rev. E **61**, 2418-2431