Forecasting Confined Spatiotemporal Chaos with Genetic Algorithms

Cristóbal López,[1] Alberto Álvarez,[2] and Emilio Hernández-García[1]

[1]*Instituto Mediterráneo de Estudios Avanzados, IMEDEA (CSIC-Universitat de les Illes Balears), 07071 Palma de Mallorca, Spain*
[2]*SACLANT Undersea Research Centre, 19138 San Bartolomeo, La Spezia, Italy*
(Received 29 March 2000)

A technique to forecast spatiotemporal time series is presented. It uses a proper orthogonal or Karhunen-Loève decomposition to encode large spatiotemporal data sets in a few time series, and genetic algorithms to efficiently extract dynamical rules from the data. The method works very well for confined systems displaying spatiotemporal chaos, as exemplified here by forecasting the evolution of the one-dimensional complex Ginzburg-Landau equation in a finite domain.

Nonlinear time-series analysis provides tools to identify dynamical systems from measured data [1]. The approach has been greatly developed in the last few years and provides, under the assumption of deterministic behavior, useful recipes for system control, noise reduction, and forecasting. Applications of these techniques to situations of spatiotemporal chaos, however, is still in its beginnings [2,3]. There are two main reasons for this: (a) The large attractor dimensions of spatiotemporally chaotic systems, increasing with system size, pose serious difficulties to the standard methods of delay embedding and attractor reconstruction. (b) The right choice of variables is far from obvious: whereas the time evolution of an observable at a particular space point could be enough in some particular situations, decaying space correlations and propagation phenomena would turn this to be a poorly performing choice in most cases.

A very efficient method for time-series prediction using genetic algorithms (GA) has been recently proposed in [4] for nonextended systems. Comparatively small data sets are enough to use this technique, which makes it competitive in facing difficulty (a), i.e., prediction in the presence of attractors of a large dimension. In this Letter we extend the GA approach to the forecasting of *confined spatiotemporal chaos*. By this we mean the situation in which chaotic dynamics in an extended system is strongly affected by the presence of boundaries. Our interest in this situation, somehow intermediate between low-dimensional chaos and homogeneous extensive chaos, arises from its relevance to real experimental situations [5,6], and from recent work [7] leading to theoretical understanding: the boundaries break translational symmetry and the resulting phase rigidity restricts the shape of the chaotic fluctuations allowed. This manifests, for example, in the appearance of nontrivial average patterns [5,7] and in inhomogeneities in other statistical characteristics [6,8]. Under these circumstances the empirical orthogonal functions (EOF's) [9,10] obtained from a proper orthogonal decomposition (POD, also known as Karhunen-Loève decomposition) provide an excellent basis for describing the system dynamics. They are different from simple Fourier modes and contain information (optimal in a precise sense) on the broken trans-

lational symmetry. The amplitudes of the most important EOF's will be the variables chosen in response to difficulty (b). By increasing system size, the dynamics would leave the regime of confined spatiotemporal chaos and at some point extracting and using information on local structures as in the methods of [2,3] would become advantageous over the POD representation. The GA's, however, could still be used as predictors for the new representation.

We now describe in more detail our method for spatiotemporal forecasting, in which the POD is used to encode the large spatiotemporal data set in a few time series, and the GA approach is used to obtain the corresponding forecasts. Given a time series of spatial patterns $U(\mathbf{x}, n)$, where $n = 1, \ldots, N$ labels the temporal sequence and $\mathbf{x}$ labels the M spatial points in a d-dimensional mesh, the POD decomposes the fluctuations around the temporal mean $u(\mathbf{x}, n) \equiv U(\mathbf{x}, n) - \langle U(\mathbf{x}, n) \rangle_n$ into modes ranked by their temporal variance. As a result, a set of spatial EOF's and associated temporal amplitude functions are obtained. The EOF's $\phi_i(\mathbf{x})$ ($i = 1, \ldots, M$) are the (orthogonal) eigenfunctions of the covariance matrix of the data $C(\mathbf{x}, \mathbf{x}') = \langle u(\mathbf{x}, n) u(\mathbf{x}', n) \rangle_n$ and are the spatial structures statistically more representative of the fluctuations in the data set. Temporal amplitude functions $a_i(n)$, describing the dynamics of the system, are obtained from the modal decomposition $u(\mathbf{x}, n) = \sum_{i=1}^{M} a_i(n) \phi_i(\mathbf{x})$. If only $K < M$ of the EOF's (the ones containing the highest temporal variance as measured by the corresponding eigenvalues) are used in the reconstruction process, the set of reconstructed patterns

$$u^K(\mathbf{x}, n) = \sum_{i=1}^{K} a_i(n) \phi_i(\mathbf{x}) \qquad (1)$$

is still the best approximation one can obtain by linearly combining K arbitrary spatial patterns multiplied by K arbitrary amplitude functions [9]. Even more, it has been shown for several chaotic and even turbulent confined systems [9,10] that taking a few dominating modes $K \ll M$ provides a good approximation to the complete data set.

Forecasting of the amplitude functions is performed with a genetic algorithm. In general, GA's are computational methods to solve optimization problems in which the

optimal solution is searched iteratively with steps inspired in the Darwinian processes of natural selection and survival of the fittest [11]. Here the optimization problem to be solved is finding the empirical model best describing the data, that is, finding the optimum function F_i that minimizes the difference $E_i^2 \equiv \sum_{n=1}^{N}[a_i(n) - \tilde{a}_i(n)]^2$ between the values $a_i(n)$ of each time series and the corresponding estimator given by

$$\tilde{a}_i(n) = F_i[a_i(n-1), a_i(n-2), \ldots, a_i(n-D)],$$
(2)

with $D + 1 \leq n \leq N$. Finding $F_i, i = 1, \ldots, M$ amounts in identifying the dynamical system behind the data set. Once found, Eq. (2) can be used to predict the future evolution of the system. If D is large enough, the existence of the exact F_i's is guaranteed by Takens theorem and its extensions [1], but a smaller D can give approximate dynamics F_i with already a reasonably low error E_i. In addition, we are not looking for all the M estimators but only for the K associated to the dominant EOF's. In our approach, the time series associated to each EOF are modeled independently. More general multivariate estimators, with each $\tilde{a}_i$ possibly dependent on different a_j's, may in principle be used, but we restrict ourselves to the choice (2) for algorithmic simplicity.

The power of the GA resides in that a huge functional space is explored in order to find an optimal F_i. Each possible F_i is a formula consisting in a combination of numerical constants, variables, and arithmetic operators. This combination is stored in the computer as a symbolic string. The only limitation to the allowed functional forms (besides the limitation to arithmetic operations) is the maximum allowed length of the symbolic string. The search procedure begins by randomly generating an initial population of potential estimators F_i that will be subjected to the evolutionary process. The evolution is carried out by selecting from the initial population the strongest individuals, i.e., the functions that best fit the data, giving a smaller E_i. In practice, only a temporal part of the data set is used in this step (the *training set*), whereas the rest of the data are used later for validating the efficiency of the prediction method (*validating set*). The strongest strings choose a mate for reproduction while the weaker strings disappear. "Reproduction" consists in interchanging parts of the symbolic strings (the "genetic material") between the two mating individuals. As a result, a new generation of individuals (which includes the original "parent" string) is generated. The new population is then subjected to mutation processes that change, with low probability, small parts of the symbolic strings. The evolutionary steps are repeated with the new generation, and the process is iterated until an optimum individual is finally found or after a fixed number of generations. Further details about the implementation of the algorithm can be consulted in [12].

The formulas F_i are optimized only for predicting the value of $a_i(n)$ in terms of the D amplitudes immediately before in time. We call this "one-step-ahead forecast."

One can in principle iterate the formulas to obtain successively predictions for $a_i(n + 1)$, $a_i(n + 2)$, etc. But this will normally lead to results rapidly diverging with respect to the correct values because of error accumulation and amplification [4].

However, GA's can be designed specifically to forecast values of the time series not necessarily in the immediate future. For example, finding the function F_i^T minimizing the error between the actual series and the estimator

$$\tilde{a}_i^T(n) = F_i^T[a_i(n-T), a_i(n-T-1), \ldots, a_i(n-D)],$$
(3)

with $D + 1 \leq n \leq N$, allows direct prediction of $a_i(N + T)$, that is, prediction of T steps ahead, without iteration.

Numerical results. — To illustrate the forecasting method we generate a data set from the numerical simulation of a well-studied model equation displaying spatiotemporal chaos, the one-dimensional complex Ginzburg-Landau equation (CGLE), supplemented with Dirichlet boundary conditions at the ends of a finite interval [10]. We write it as

$$\partial_t A(x,t) = q^2(1 + \alpha)\partial_x^2 A + A - (1 + i\beta)A|A|^2,$$
(4)

where q, α, and β are real and positive and $A(x,t)$ is a complex-valued field. We solve it in the interval $[0, \pi]$ so that the boundary conditions read $A(0) = A(\pi) = 0$. By simple scaling of the spatial coordinate one sees that this is equivalent to rewriting the equation with $q = 1$, but solving it in a domain of size $L = \pi/q$. Thus the parameter q is equivalent to an inverse system size, and decreasing it is equivalent to increasing system size. Following [10] we fix $\alpha = 4$ and $\beta = -4$ [13]. For $q < 0.2$ the system displays spatiotemporal chaos for most of the initial conditions. Decreasing q one encounters the regime of confined spatiotemporal chaos that we are interested in before approaching homogeneous extensive chaos at large system sizes ($q \rightarrow 0$) [14]. According to [10], the correlation dimension of the dynamical attractor for $q = 0.14$ is 9.08. We sample our simulation every $\tau = 0.1$ time units and at spatial locations separated $\Delta = \pi/100$ space units, and we follow it for 80 time units (800 samples) after discarding the initial transient starting from random initial conditions (this sampling leads to $N = 800$ and $M = 100$). This will be our "training set" to be fed into the GA. The simulation is then continued for a few more time units to provide the "validation set" which is hidden to the GA. It is used later to check the accuracy of the predictions.

We choose as the basic field to be forecasted the *modulus* of the complex field $U(x,n) = |A(x, t = n\tau)|$. The algorithm seems to perform slightly better in forecasting the real or the imaginary parts of A, but we use U to show that the algorithm works well with nonlinear combinations of the basic dynamical quantities. In Fig. 1 we show parts of typical spatiotemporal evolutions for $q = 0.12$ and $q = 0.16$. Clearly, reducing q decreases the spatial
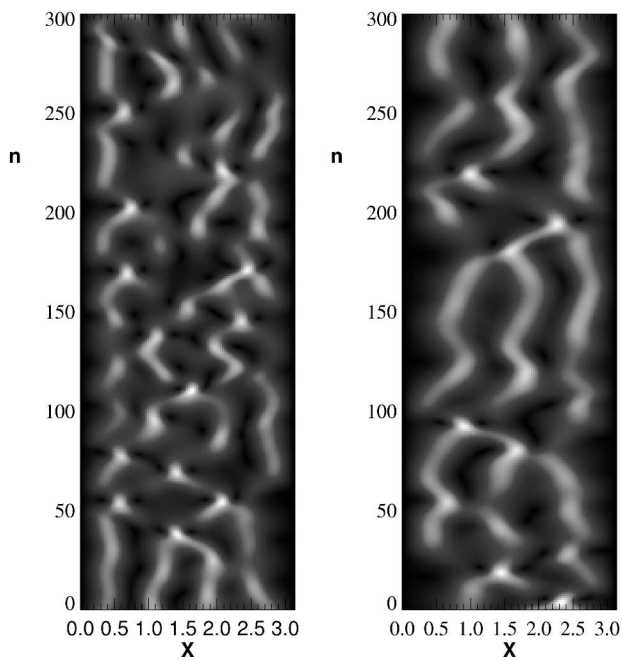
FIG. 1.   Spatiotemporal evolutions of $U(x, n)$, as given by the CGLE for $q = 0.12$ (left) and $q = 0.16$ (right). Black corresponds to $U = 0$, and lighter gray to high values of U.

scales, as corresponding to an effectively larger system size, but also the complexity of the evolution is increased. In both cases it is clear that the motion of the dynamical structures is constrained by the presence of the walls, as corresponding to *confined* spatiotemporal chaos.

We solve Eq. (4) for $q = 0.18, 0.16, 0.14, 0.12$ and perform the POD on the fluctuations $u(x, n)$ of the modulus around its temporal mean value in the resulting data sets. The number of relevant EOF's (which we define to be those accounting for at least 99% of the data variance [10]) are, respectively, 9, 11, 13, and 15. We note that this confirms the expected approximate linear scaling of the number of EOF's with increasing system size L ($\propto q^{-1}$) [15]. It is somehow surprising that this extensive scaling appears even when chaos is not homogeneous, but is still influenced by the boundaries. This fact has been observed in other systems before [6,8]. For illustrative purposes, we show in Fig. 2 the two most relevant EOF's from our training set at $q = 0.16$, and the corresponding temporal amplitude functions. The chaotic character of these series is evident.

We next apply the GA to each of the amplitude functions of the relevant EOF's. We use the following parameters for all the values of q: number of generations in the evolutionary process, 2000; number of individuals in each generation, 120; maximum number of symbols allowed for each symbolic string, 20; maximum delay in (2) or (3), $D = 18$ [16]. Tuning of these parameters for each particular value of q would improve forecasting, but it would make comparisons more difficult. Predictions for the field $u(x, t)$ are then built up by reconstruction according to (1) with K the number of relevant EOF's defined above. In Fig. 3
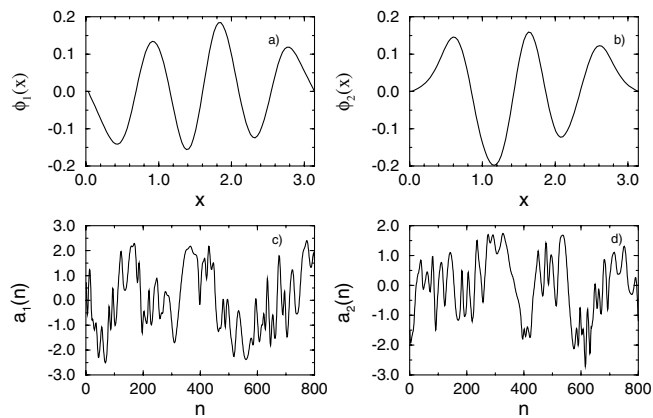

FIG. 2.   The first two EOF's (a),(b) and the corresponding time amplitude functions (c),(d) from the training set at $q = 0.16$.

we show the one-step-ahead forecasted fields, more concretely the prediction for the first step beyond the training set, $n = 801$. It is compared with the actual numerical pattern in the validation set, for $q = 0.12, 0.14$ and $q = 0.16$, displaying an excellent performance.

We quantify the quality of the prediction in terms of the mean square error $\epsilon_q(n)$:

$$\epsilon_q^2(n) \equiv \frac{1}{M} \sum_{j=1}^{M} [\tilde{u}^K(x = j\Delta, n) - u(x = j\Delta, n)]^2, \quad (5)$$

where $\tilde{u}^K(x, n)$ is the predicted pattern reconstructed from Eq. (1) and $u(x, n)$ is the actual pattern from the validation set. As stated before, GA's can be used to predict future values some time steps ahead, without the need of iterating the one-step-ahead predictor (which early becomes useless because of the expected exponential growth of errors). Figure 4 shows $\epsilon_q(n)$ as a function of n for $q = 0.16$ calculated from (a) the one-step-ahead prediction formulas obtained from the training set, but applied
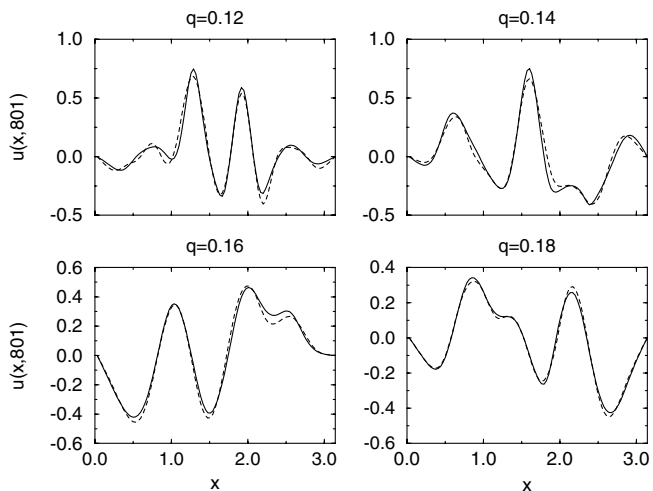

FIG. 3.   The forecasted moduli fields (dashed line) as compared to the real ones (solid line) for the one-step-ahead prediction for several values of q.
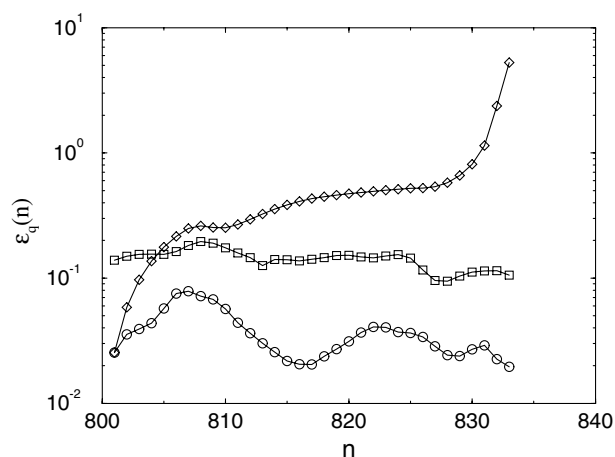
FIG. 4.   Errors as a function of n in the validation set, for $q =$ 0.16. Circles: One-step-ahead prediction. Diamonds: Iteration of the one-step-ahead formulas starting from the training set ($n \leq 800$). Squares: Five-step-ahead prediction.

to obtain the pattern at step n from the previous D values in the validation set; (b) iteration of the one-step-ahead formulas starting from the last D data in the training set; (c) five-step-ahead prediction from a formula of type (3) with $T = 5$, obtained by the GA in the training set, and used into the validation set. We see that the improvement in accuracy is notorious when iteration is avoided. We note that the errors in methods (a) and (c) remain bounded even when n is far from the values from which the prediction formulas were estimated (i.e., the training set $n < 800$). This confirms that the method is not simply fitting data, but rather it has really found approximate dynamical rules within the deterministic spatiotemporal series.

Figure 5 displays the average error $\langle \epsilon_q \rangle$, which is the temporal average of $\epsilon_q(n)$ with n in the validation range displayed in Fig. 4, as a function of q (for the one-step-ahead prediction). Despite including more EOF's in the reconstruction for decreasing q, the prediction error shows a tendency to increase. This is a consequence of the increase in complexity (and in attractor dimension) of the
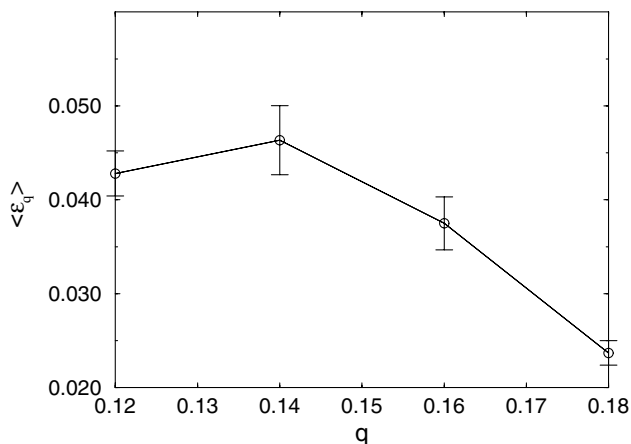

FIG. 5.   Mean error for the one-step-ahead prediction in the validation set as a function of q.

dynamics by the effective increase in system size ($\approx q^{-1}$). Since we keep the maximum delay D fixed, the embedding of the data set becomes more incomplete at smaller q and the prediction deteriorates. In addition, for smaller q the *confined* or *boundary influenced* character of the spatiotemporal chaos in the system is lost and a description in terms of local structures will be certainly more efficient [2].

In summary, we have presented a method to forecast the evolution of spatially extended systems based in the combination of POD and GA's. The method performs very well in situations of confined spatiotemporal chaos as exemplified by the CGLE in a finite interval. We are testing the method for prediction from noisy natural data sets. Results obtained in forecasting sea surface temperature patterns [17] are encouraging.

[1] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, England, 1997).
[2] U. Parlitz and G. Mayer-Kress, Phys. Rev. E **51**, R2709 (1995); U. Parlitz and C. Merkwirth, Phys. Rev. Lett. **84**, 1890 (2000).
[3] S. Ørstavik and J. Stark, Phys. Lett. A **247**, 145 (1998).
[4] G. G. Szpiro, Phys. Rev. E **55**, 2557 (1997).
[5] B. J. Gluckman, C. B. Arnold, and J. P. Gollub, Phys. Rev. E **51**, 1128 (1995); L. Ning *et al.,* Phys. Rev. Lett. **71**, 2216 (1993).
[6] S. M. Zoldi *et al.,* Phys. Rev. E **58**, R6903 (1998).
[7] V. M. Eguíluz *et al.,* Phys. Rev. E **59**, 2822 (1999).
[8] M. Meixner *et al.,* Phys. Rev. E **61**, 1382 (2000).
[9] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry* (Cambridge University Press, Cambridge, England, 1996).
[10] L. Sirovich, Physica (Amsterdam) **37D**, 126 (1989); J. D. Rodriguez and L. Sirovich, Physica (Amsterdam) **43D**, 77 (1990); L. Sirovich, J. D. Rodriguez, and B. Knight, Physica (Amsterdam) **43D**, 63 (1990).
[11] J. H. Holland, *Adaptation in Natural and Artificial Systems* (University of Michigan Press, Ann Arbor, 1992).
[12] A. Álvarez, A. Orfila, and J. Tintoré, "DARWIN–An Evolutionary Program for Nonlinear Modelling of Chaotic Time Series" (unpublished).
[13] The parameters c_0 and ρ of the notation of Sirovich *et al.* [10] are related to ours by $c_0 = 1/\alpha$ and $\rho = -1/\beta$. Their time and space variables need to be multiplied by factors ρ and $\sqrt{\rho/c_0}$, respectively, to obtain our time and space variables.
[14] Our parameter regime leads to the so-called "defect chaos" in the large-size limit. See B. I. Shraiman *et al.,* Physica (Amsterdam) **57D**, 241 (1992).
[15] S. Zoldi and H. Greenside, Phys. Rev. Lett. **78**, 1687 (1997).
[16] To find each of the temporal amplitude predictors takes about 400 seconds of CPU time in a AlphaStation 600s/333.
[17] A. Álvarez *et al.,* Geophys. Res. Lett. (to be published); e-print chao-dyn/9911012.