PHYSICAL REVIEW LETTERS

Making Forecasts for Chaotic Physical Processes

Christopher M. Danforth* and James A. Yorke

University of Maryland, College Park, Maryland 20742-4015, USA
(Received 11 May 2005; published 14 April 2006)

Making a prediction for a chaotic physical process involves specifying the probability associated with each possible outcome. Ensembles of solutions are frequently used to estimate this probability distribution. However, for a typical chaotic physical system H and model L of that system, no solution of L remains close to H for all time. We propose an alternative. This Letter shows how to inflate or systematically perturb the ensemble of solutions of L so that some ensemble member remains close to H for orders of magnitude longer than unperturbed solutions of L. This is true even when the perturbations are significantly smaller than the model error.

Many scientific disciplines require accurate predictions of the future state of chaotic physical systems. Astronomers attempt to predict the trajectories of bodies in the solar system for thousands of years into the future [1], as well as the evolution of galactic clusters [2]. Plasma physicists use nonlinear models to predict magnetic storms and solar wind [3]. Oceanographers forecast sea-surface temperatures in an attempt to predict the likelihood of El Niño Southern Oscillation (ENSO) events in the major oceans up to a year in advance [4]. Meteorologists attempt to predict the path of violent hurricanes with hours of lead time, and larger scale patterns up to a week in advance [5].

Predicting the behavior of a chaotic physical system H using a model L has three obstacles: uncertainty in the initial state, chaos, and model errors, i.e., differences between L and H. Given the initial state $\mathbf{p}_0$ of H, the initial state of L which will yield the trajectory that best matches the physical system is unknown. The accepted procedure is to choose a large collection or ensemble of initial states and follow their L trajectories.

For example, when L is a global weather model and H is the behavior of the atmosphere, an initial state for L is an estimate of the state of the atmosphere over the entire planet. Forecasters view the initial state of the atmosphere as uncertain, but lying within a known ball in state space. The radius σ of this ball corresponds to measurement uncertainties. They choose a finite ensemble of initial states in this ball. Forecasters then take the trajectories of L for each initial state in the ensemble, for example, at time $T = 3$ days later, as predictions. If all such trajectories yield similar behavior at time T, e.g., rain, then the forecaster predicts rain. If the trajectories of L disagree at time T, then the prediction is a nontrivial probability distribution. Even if the model error is small, such a probabilistic forecast will be completely wrong if the H trajectory diverges from the ensemble.

For a chaotic system with imprecise initial state, a perfect L forecast at time T consists of a probability distribution which accurately describes the likelihood of all possible outcomes. Denote the state of H at time T by

$\mathbf{p}_T$. A forecaster hopes that the ensemble is quite close to $\mathbf{p}_T$ at time T. However, only a finite ensemble is followed. Given this limitation, the modeler's goal is that some linear combination of ensemble members remains within the σ-ball around $\mathbf{p}_t$ for the duration of the forecast ($t = 0, 1, \ldots, T$).

The above goal could likely be met if the chaos were of the type called "hyperbolic." Hyperbolicity is not defined here, but hyperbolic systems have the following property [6]. Let $(\mathbf{p}_t)_{t=a}^b$ be a trajectory of a hyperbolic H. Given a $\sigma > 0$, when system L is sufficiently close to H, there exists some trajectory $(\mathbf{y}_t)_{t=a}^b$ of L such that $|\mathbf{y}_t - \mathbf{p}_t| < \sigma$ for all $t \in [a, b]$. In other words, each trajectory of H is σ-shadowed by a trajectory of L. The shadowing property exists for hyperbolic systems in part because the number of expanding (contracting) directions remains constant in such systems. Much of shadowing theory has been developed for hyperbolic systems. Unfortunately, hyperbolic systems are so special that they have been irrelevant to the prediction of virtually all realistic chaotic physical processes. In this Letter, we propose an improved ensemble approach that is more likely to meet the modeler's goal for nonhyperbolic systems.

Do any trajectories of L give accurate predictions?—In Fig. 1, the ensemble of initial states is represented as a disk of radius σ. J_0 is the set of states within σ of the true initial state $\mathbf{p}_0$, $L_t(J_0)$ denotes the trajectories of L at time t of each state in J_0. $H_t(\mathbf{p}_0)$ denotes the trajectory of H at time t. $N_\sigma(\mathbf{p}_t)$ is the set of states within σ of $\mathbf{p}_t$ at time t. As the trajectories of L are followed forward in time, the disk is expected to expand in some directions and contract in others forming a rough ellipsoid. Some ellipsoid axes rapidly become very thin, shrinking to zero thickness exponentially fast. If σ is chosen sufficiently small, after a modest time T no state in $L_T(J_0)$ is within σ of $H_T(\mathbf{p}_0)$. On the other hand, if σ is chosen large enough, the entire attractor will be included and shadowing is trivial. In Fig. 1(b), σ is chosen such that $L_T(J_0)$ has states within σ of $\mathbf{p}_T$. Looking at longer prediction times jT, we ask if any single trajectory of the ensemble remains within σ of
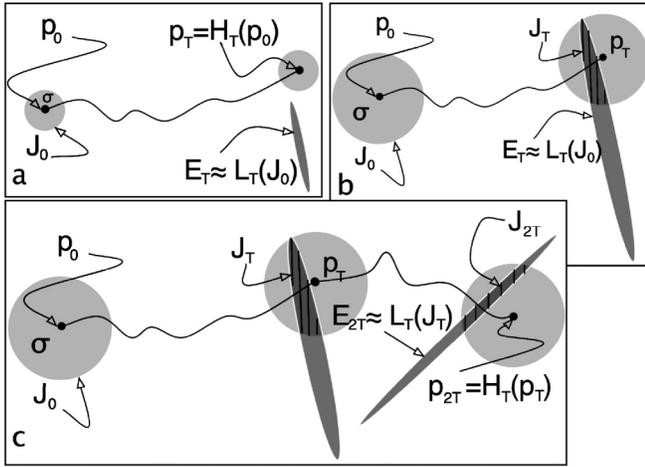
FIG. 1. For a given initial state, models L and H will produce different trajectories. σ balls are shown around states $\mathbf{p}_0$, $\mathbf{p}_T$, $\mathbf{p}_{2T}$ of a trajectory of H. If σ is small (a), shadowing fails in a single step of the process. Increasing σ (b), some trajectories of L remain close to a trajectory of H for time T. These trajectories are given by J_T. For sufficiently close hyperbolic systems L and H, this procedure can be carried out for arbitrarily long times with small σ.

$\mathbf{p}_0, \mathbf{p}_T, \ldots, \mathbf{p}_{jT}$. To answer, at time T all trajectories of L farther than σ from $\mathbf{p}_T$ are ignored. The remaining trajectories lie in J_T, the intersection of $L_T(J_0)$ and $N_\sigma(\mathbf{p}_T)$. Figure 1(c) illustrates that it is possible to continue this procedure, restricting to the trajectories of L that stay within σ of $\mathbf{p}_{jT}$ for each j. As long as this set is nonempty, some members of the ensemble of trajectories of L give accurate predictions.

Unstable dimension variability.—Fig. 1 shows a ball of initial states J_0 contracting in one direction and expanding in another into an ellipse. In Fig. 2(a), the behavior changes locally from one expanding dimension to two and shadowing fails. Increasing σ by a factor of 10 would not prevent such failures. When the dimension is greater than 2 and H is chaotic, it is likely that the number of independent contracting and expanding directions will vary from state to state, see Fig. 3. Such "unstable dimension variability" has been shown to result in shadowing failures, where no trajectory of L stays within σ of an H trajectory [7,11–14].

Outline of our forecast method.—Given a set of states that fill out an ellipsoid E_t^φ and represent a prediction at time t, a prediction at time $t+1$ is produced as follows. (1) Apply L for one time step, yielding $L_1(E_t^\varphi) \approx E_{t+1}$. (2) "Inflate" (see below) the ellipsoid E_{t+1} by φ, yielding E_{t+1}^φ.

Steps 1 and 2 constitute our continually inflated ensemble approach. Strictly for notational simplicity, we inflate only once each time unit. Current ensemble procedures use only step 1. Adding step 2 is our proposed alternative. It makes the procedure more robust in meeting the aforementioned modeler's goal. In practice, a limita-
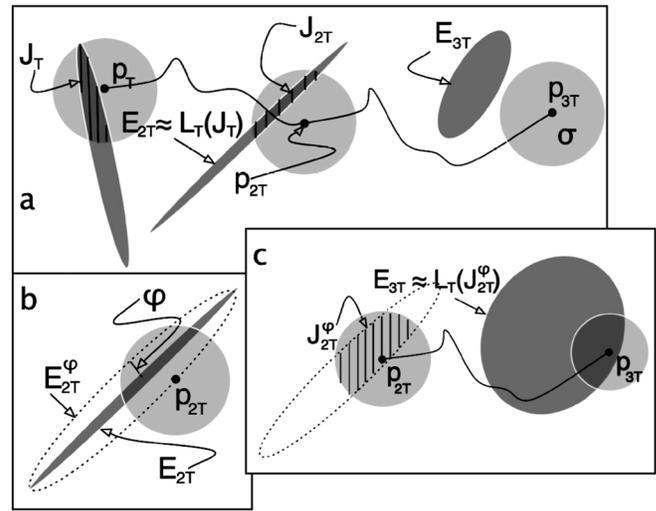
FIG. 2. Most physical systems are nonhyperbolic. In (a), the dynamics contract in one dimension as $\mathbf{p}_T \to H_T(\mathbf{p}_T) \equiv \mathbf{p}_{2T}$. The ellipse $E_{2T} \approx L_T(J_T)$ intersects the σ-ball surrounding $\mathbf{p}_{2T}$, the intersection is denoted J_{2T}. As $\mathbf{p}_{2T} \to H_T(\mathbf{p}_{2T}) \equiv \mathbf{p}_{3T}$, the dynamics expand in both dimensions. The intersection of $L_T(J_{2T})$ and $N_\sigma(\mathbf{p}_{3T})$ is empty and shadowing fails. In (b), E_{2T}^φ is the ellipse E_{2T} inflated by φ. In (c), the intersection of E_{2T}^φ and $N_\sigma(\mathbf{p}_{2T})$ is denoted J_{2T}^φ. Note that J_{2T}^φ contains $\mathbf{p}_{2T}$. Despite expansion in both dimensions, the intersection of $L_T(J_{2T}^\varphi)$ and $N_\sigma(\mathbf{p}_{3T})$ is nonempty. In practice, this procedure is successful at time $T+1$ if J_T^φ contains $\mathbf{p}_T$.

tion of the ensemble method of prediction is that one encounters nonlinearities which distort the ellipsoids. Therefore, the perturbations and time steps in this Letter are chosen sufficiently small for linear approximations to be appropriate. That is, we consider only the case where E_t^φ is a very small ellipsoid.

Step 1: Calculating $L_1(E)$ for an ellipsoid E.—Given an ellipsoid E, $L_1(E)$ is approximated as follows. Choose an ensemble consisting of $\bar{\mathbf{s}}$, the center of E, and states $\mathbf{s}_k$ ($k = 1, 2, \ldots, K$) on the surface of E so that the line from $\bar{\mathbf{s}}$ to $\mathbf{s}_k$ is the kth semiaxis of E. Note that the ensemble is redefined each time step 1 is applied. The image $L_1(E)$ is approxi-
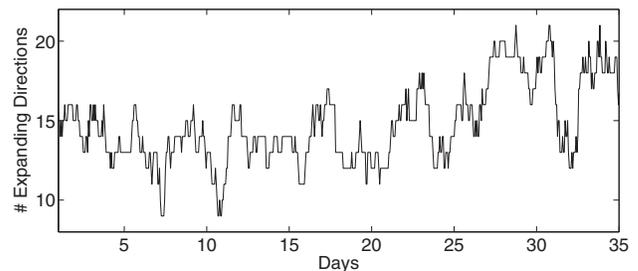
FIG. 3. In the 40 dimensional system discussed later, the number of expanding directions varies from 8 to 23 depending on the state investigated. As a trajectory is followed, the same fluctuations in the local number of expanding directions are observed.

mated by the ellipsoid of linear combinations $L_1(\bar{\mathbf{s}}) + \sum_{k=1}^{K} \beta_k [L_1(\mathbf{s}_k) - L_1(\bar{\mathbf{s}})]$ such that $\sum_{k=1}^{K} \beta_k^2 \leq 1$. Of course, this ellipsoid is not quite $L_1(E)$. Another limitation of the ensemble forecasting method is that many ensemble members may be needed to accurately represent the probability distribution described by the ellipsoid. In this Letter, K is chosen to be equal to the dimension of L. This choice would be computationally prohibitive for systems with millions of dimensions.

Step 2: How to inflate an ellipsoid.—Given an ellipsoid E, write $\mathbf{e}_k$ for the orthonormal basis of unit vectors parallel to the semiaxes. Let $\gamma_k > 0$ be the corresponding semiaxis lengths. $\mathbf{e}_k$ and γ_k can be computed with the singular value decomposition. The thin semiaxes are defined to be those which satisfy $\gamma_k < \sigma$. The ellipsoid E^φ inflated by φ is the ellipsoid with the same center as E and with axes aligned with those of E, but with each thin semiaxis increased by φ. The process of inflation carries any state $\mathbf{u}$ in the ellipsoid E to a state $\mathbf{u}^\varphi$ in E^φ [8].

Testing the effect of inflation.—To test whether the "modelers' goal" is met, after each inflation all states of E_{t+1}^φ that lie more than σ from $\mathbf{p}_{t+1}$ are discarded. If the intersection (denoted J_{t+1}^φ) of E_{t+1}^φ and $N_\sigma(\mathbf{p}_{t+1})$ is empty, shadowing has failed. If J_{t+1}^φ is nonempty, its complicated shape is approximated by an ellipsoid G_{t+1} lying inside J_{t+1}^φ. In doing so, even more states are discarded, but the approximation procedure is computationally tractable. There is no unique choice of G_{t+1}. The center of G_{t+1} is chosen to be the mean of a uniform distribution of states in J_{n+1}^φ. The axes of G_{t+1} are chosen parallel to $\mathbf{e}_k$. These approximations are repeated each time step. As long as the procedure succeeds, the modeler's goal is met.

Since each inflation introduces a small amount of uncertainty into the forecast, as little inflation as possible should be used. However, since the procedure only inflates in thin directions, the additional uncertainty will be damped out if the thin directions continue to contract. Should the dynamics local to the ensemble experience unstable dimension variability and thin directions begin to expand, some ensemble members should remain close to the H trajectory.

Model.—We use a simple model to represent atmospheric behavior, the N-dimensional governing equations, given by [9] are

$$\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F \qquad (1)$$

for $i = 1, 2, \ldots, N$, where the subscripts are treated as periodic with period N. For example, $x_{N+1} \equiv x_1$ so that the variables form a cyclic chain. Each variable represents an unspecified scalar meteorological quantity, such as temperature, at N equally spaced grid sites on a latitude circle. In our experiments, $N = 40$ and $F = 8$ as in [9]. This model shares certain properties with many atmospheric models: a nonlinear advection term, a linear term repre-

senting loss of energy to thermal dissipation, and a constant forcing term F to provide energy. The time unit represents the dissipative decay time of 5 days [9]. There are 13 positive Lyapunov exponents.

Stalking.—Stalking is an aggressive form of shadowing in which the ellipsoids E_t are inflated as described above. Let $(\mathbf{p}_t)_{t=a}^b$ be a sequence representing the true solution (H trajectory). Then given a shadowing distance $\sigma > 0$ and an inflation $\varphi > 0$, $(\mathbf{p}_t)_{t=a}^b$ is φ-σ-stalked so long as J_t^φ is nonempty for all $t \in [a, b]$. The states contained in $(J_t^\varphi)_{t=a}^b$ are called stalking trajectories. If $\varphi = 0$ (no inflation), a stalking trajectory is called a shadowing trajectory. The interval $[a, b]$ is referred to as the stalking time. If no stalking trajectories exist for reasonable σ and φ over an interval of time relevant to prediction, L is an inadequate approximation of H.

An H trajectory.—Equation (1) represents L. A trajectory $\mathbf{p}_t$ representing H is obtained as follows. Given $\mathbf{p}_0$, take a one-fourth order Runge-Kutta time step of size 10^{-2} and denote the result $L_1(\mathbf{p}_0)$. For each t, choose $\mathbf{p}_{t+1}$ randomly from a uniform distribution such that $\mathbf{p}_{t+1}$ is within μ of $L_1(\mathbf{p}_t)$. μ represents model error, the difference between L and H. Fix $\mu = 10^{-6}$ and repeat for $t = 0, 1, \ldots, 10^7$. We say $(\mathbf{p}_t)_{t=a}^b$ is a μ-pseudotrajectory of L because $|\mathbf{p}_t - L_1(\mathbf{p}_{t-1})| \leq \mu$ for all $t \in [a, b]$. We then see how long we can φ-σ-stalk $\mathbf{p}_t$ with an ensemble of φ-pseudotrajectories of (1). If $\varphi \geq \mu$, the H trajectory is itself trivially a φ-σ-stalking trajectory.

Finding the stalking time.—The shadowing distance σ and the inflation φ are fixed throughout each integration and explore the parameter space in φ, recording the average stalking time. In Fig. 4, the stalking time vs the relative inflation φ/μ is plotted for model (1). When $\sigma = 1000\mu$, with no inflation ($\varphi = 0$) the shadowing time is approximately 2 days. Decreasing the shadowing distance by a factor of 10 to $\sigma = 100\mu$ and inflating by $\varphi = 40\%$ of the model error μ gives the same stalking time. When the H trajectory is generated by adding systematic error during integration, slightly more inflation is required to achieve the same results.

Forecasting improvement.—To measure the effect of inflation on forecasts (where J_t is unknown), 5000 independent 25-day H trajectories are calculated. Prediction of an H trajectory is made by following an ellipsoid of trajectories of L, with and without inflation. Figure 5 plots the average distance between the H trajectory and the nearest trajectory of L for ensemble forecasts and continually inflated ensemble forecasts. Ensemble forecasts continually inflated by 50% of the model error produce trajectories of L within σ of an H trajectory for 5 times longer than traditional ensemble forecasts.

Discussion.—We find that modest inflation substantially increases shadowing time. Our "continually inflated ensemble" approach is guaranteed to succeed in the linear regime for inflation in all directions with $\varphi \geq \mu$. We,
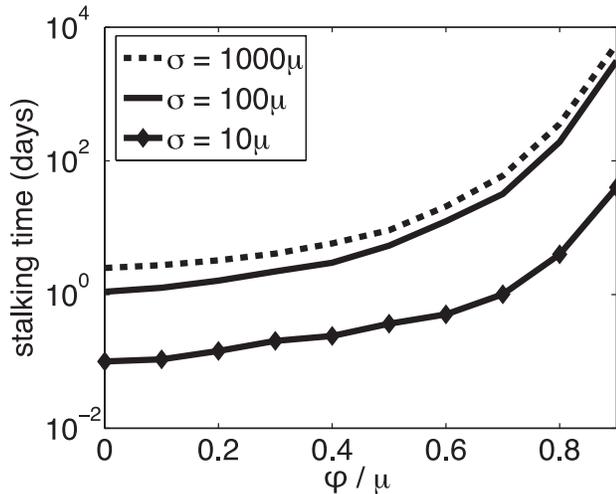
FIG. 4. Stalking time for model (1) measured in days as a function of relative inflation φ/μ, where φ is the inflation, $\mu = 10^{-6}$ is model error ($0 < \varphi < \mu$), and σ is shadowing distance. Trajectories of (1) initially separated by 10^{-16} are uncorrelated after 25 days. If $\varphi/\mu = 0$, the stalking time is the (brief) traditional shadowing time. If $\varphi = \mu$, the stalking time is infinite. The $\sigma = 10\mu$ curve illustrates the phenomenon in Fig. 1(a), where stalking failures occur because the shadowing distance is too small. Increasing σ by a factor of 10, the shadowing time ($\varphi = 0$) increases by a factor of 10.

however, inflate only thin axes and investigate cases where $\varphi < \mu$, so the method can fail (as illustrated in Fig. 4). In practice, the magnitude μ of the model error is unknown, it must be estimated by trial and error. If all directions are inflated, it may be possible to decrease σ. While this Letter deals only with a toy model where all distances are quite small, we hope the approach can be adapted to practical high-dimensional systems.
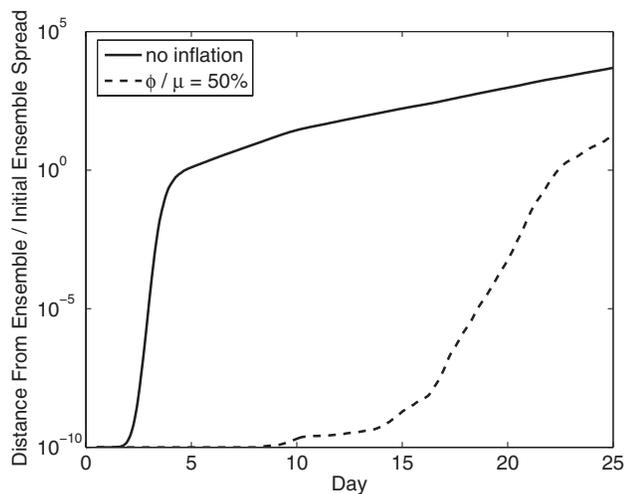
For any moderate shadowing distance σ, no trajectory of L remains within σ of a typical complex high-dimensional physical system H. Orrell *et al.* [10]. estimate that forecasts generated by the European Center for Medium-Range Weather Forecasting operational weather model are dominated by model errors during the first 3 days, and that shadowing the real atmosphere fails after 6 hours, for a reasonable σ. In other words, no trajectories of L initially within observational uncertainty σ remain consistent beyond small T. During the first few days of an operational forecast, inflating the contracting directions of the ensemble of L trajectories every few hours may improve the tracking time. Inflation is not currently used for weather forecasting, but it is used in data assimilation, the practice of combining observations with forecasts to generate the initial set of states for an ensemble.

*Electronic address: danforth@math.umd.edu
[1] G. Sussman and J. Wisdom, Science **257**, 56 (1992).
[2] W. Hayes, Phys. Rev. Lett. **90**, 054104 (2003).
[3] J. Valdivia, A. Sharma, and K. Papadopoulos, Geophys. Res. Lett. **23**, 2899 (1996).
[4] A. Barnston, M. Glantz, and Y. He, Bull. Am. Meteorol. Soc. **80**, 217 (1999).
[5] D. Patil, E. Ott, B. Hunt, E. Kalnay, and J. Yorke, Phys. Rev. Lett. **86**, 5878 (2001).
[6] C. Gregobi, S. M. Hammel, J. Yorke, and T. Sauer, Phys. Rev. Lett. **65**, 1527 (1990).
[7] G.-C. Yuan and J. Yorke, *Proceedings of the American Mathematical Society* (AMS, Providence, 2000), Vol. 128, p. 909.
[8] Let γ_k^φ be the semiaxis lengths after inflation. Given the ensemble $L_1(\mathbf{s}_k), L_1(\bar{\mathbf{s}})$ used to generate E, the ensemble that generates E^φ is created as follows. Choose coefficients ρ_k so that $\mathbf{u} = \sum_{k=1}^K \rho_k \gamma_k \mathbf{e}_k$. Then $\mathbf{u}^\varphi = \sum_{k=1}^K \rho_k \gamma_k^\varphi \mathbf{e}_k$. That is, when $\mathbf{u}$ is expanded in terms of semiaxis vectors $\gamma_k \mathbf{e}_k$, $\mathbf{u}^\varphi$ is the corresponding sum using the inflated semiaxis vectors $\gamma_k^\varphi \mathbf{e}_k$.
[9] E. Lorenz and K. Emanuel, J. Atmos. Sci. **55**, 399 (1998).
[10] D. Orrell, L. Smith, J. Barkmeijer, and T. Palmer, Nonlin. Proc. Geophys. **8**, 357 (2001).
[11] S. Dawson, C. Gregobi, T. Sauer, and J. Yorke, Phys. Rev. Lett. **73**, 1927 (1994).
[12] E. Kostelich, I. Kan, C. Gregobi, E. Ott, and J. Yorke, Physica D (Amsterdam) **109**, 81 (1997).
[13] T. Sauer, Phys. Rev. E **65**, 036220 (2002).
[14] E. Barreto and P. So, Phys. Rev. Lett. **85**, 2490 (2000).

FIG. 5. The distance between an H trajectory and the nearest trajectory of the ensemble ellipsoid is plotted vs time, averaged over 5000 independent 25-day ensemble forecasts (solid) and their corresponding continually inflated ensemble forecasts (dotted). The vertical axis is in units of the initial diameter of the ensemble.