Q. J. R. Meteorol. Soc. (1999), **125**, pp. 2855–2886

Uncertainty dynamics and predictability in chaotic systems

By L. A. Smith[1]*, C. Ziehmann and K. Fraedrich[2]

[1]*Mathematical Institute, University of Oxford, UK*
[2]*Meteorologisches Institut der Universität Hamburg, Germany*

(Received April 1997; revised December 1998)

Summary

An initial uncertainty in the state of a chaotic system is expected to grow even under a perfect model; the dynamics of this uncertainty during the early stages of its evolution are investigated. A variety of "error growth" statistics are contrasted, illustrating their relative strengths when applied to chaotic systems, all within a perfect model scenario. A procedure is introduced which can establish the existence of regions of a strange attractor within which *all* infinitesimal uncertainties *decrease* with time. It is proven that such regions exist in the Lorenz attractor, and a number of previous numerical observations are interpreted in light of this result; similar regions of decreasing uncertainty exist in the Ikeda attractor. It is proven that no such regions exist in either the Rössler system or the Moore-Spiegel system. Numerically, strange attractors in each of these systems are observed to sample regions of state space where the Jacobians have eigenvalues with negative real parts, yet when the Jacobians are not normal matrices this does not guarantee that uncertainties will decrease. Discussions of predictability often focus on the evolution of infinitesimal uncertainties; clearly, as long as an uncertainty remains infinitesimal it cannot pose a limit to predictability. To reflect realistic boundaries, any proposed "limit of predictability" must be defined with respect to the nonlinear behaviour of perfect ensembles. Such limits may vary significantly with the initial state of the system, the accuracy of the observations, and the aim of the forecaster. Perfect model analogues of operational weather forecasting ensemble schemes with finite initial uncertainties are contrasted with both perfect ensembles and uncertainty statistics based upon the dynamics infinitesimal uncertainties.

Keywords: Predictability Ensemble Prediction Nonlinear Systems Lyapunov Exponents Error Growth Singular Vectors Uncertainty Perfect Model

1. Introduction

While there is widespread agreement as to the best approach for predicting large dynamical systems like the Earth's atmosphere, optimal methods for quantifying predictability itself remain the subject of debate. It has been 40 years since Thompson focussed attention on quantifying the role which uncertainty in the initial condition played in limiting the predictability of the atmosphere (Thompson 1957); during this time, operational weather forecasting has progressed to the point of explicit attempts to quantify the evolution of this uncertainty during each forecast (Toth and Kalnay (1993), Palmer *et al.* (1992) and references thereof). The aims of early work by Lorenz, Epstein, Leith and others was summarised by Tennekes (1991), who proclaimed that no forecast was complete without an estimate of the forecast error. Over this same period, the proliferation of chaotic dynamical systems with a few, or a few dozen, degrees of freedom involved a wide community in the issues of uncertainty growth in general, and the long time behaviour of infinitesimal uncertainties in particular.

The present paper critically examines the spectrum of statistics which seek to quantify the dynamics of uncertainty, a spectrum which ranges from the instantaneous growth rate of infinitesimal uncertainties to the expected increase in finite uncertainties after a finite period of time; we explore the extent to which each is (or is not) relevant to establishing a global "limit of predictability" or to the lesser goal of estimating the likely time over which observations of a particular quantity may be gainfully predicted. Among other things, we demonstrate that positive Lyapunov exponents place no finite limits

* Corresponding author: Mathematical Institute, University of Oxford, Oxford, OX1 3LB, U.K.

on prediction *per se*, and prove that several common chaotic systems have finite regions within which *all* uncertainties *decrease* with time.

Predictability analysis based on forecast experiments requires a careful distinction between error and uncertainty. Theoretical (internal) predictability experiments are performed by identical dynamical systems leading to the so-called 'perfect model scenario.' Alternatively, practical (external) or operational experiments are made with an imperfect model forecasting observational data; model error exists in the particular model employed, and measurement errors in the data. The computations presented below assume a perfect model scenario with uncertain initial conditions, and then quantify the growth of uncertainty in situations without forecast model error; that is, we only consider the effect of (observational) uncertainties in the initial condition. Hence there is no error *per se*; in this case predictability describes the evolution of an initial uncertainty due to internally occurring instabilities.

The loss of superposition and the extreme inhomogeneity common in nonlinear chaotic systems require local measures of predictability; this has resulted in the introduction of a plethora of such local measures (see Lorenz (1965), Nicolis *et al.* (1983), Benzi and Carnevale (1989) and references thereof). Yet the fact that this well-known non-uniformity limits the application of global measures of predictability has been widely overlooked. In addition to contrasting local and global measures, a variety of distinct local measures of predictability are discussed in Section 2. Here rate-like statistics and time-like statistics are contrasted in an inhomogeneous system, confusing the two is to make the elementary error of assuming that the inverse of the average reflects the average of the inverse. This need not be the case.

Section 3 illustrates the fact that the evolution of an initial uncertainty need not be uniform even in a uniform, linear system; following Farrell (1994), we stress the role of non-orthogonal eigenbases on uncertainty growth. These exact results are then used in Section 4 to interpret the numerics of common nonlinear systems and to clarify the confusion in the literature regarding "oscillating error growth" and "super-exponential" error growth. Considering the dynamics of uncertainty on strange attractors requires the prescription of suitable initial conditions and distributions of uncertainties; as noted by Haubs and Haken (1985) and quantified by Nese (1989) and others (Abarbanel *et al.* (1991), Kostelich and Yorke (1990), Mukougawa *et al.* 1991, Doerner *et al.* (1991), Smith (1992,1994a)) there is structure in the variations of predictability in state space. Following Ziehmann-Schlumbohm (1994), an analytic approach for establishing regions of *decreasing* uncertainty in chaotic systems is presented in Section 5 and applied to three common strange attractors as illustrated in figure 9. Here we are again restricted to infinitesimal uncertainties, but the results apply for evolution over a finite time. We also consider the implications of the various orientations of likely initial uncertainty, and show that the most relevant initial uncertainties need not be related to the most analytically tractable; indeed the results are not only system specific, but also depend on the details of the observation and aim of the particular forecast problem at hand.

In Section 6 we examine the efficacy of infinitesimal measures of predictability in reflecting the dynamics of finite uncertainty. Examining the behaviour of perfect ensembles (Smith 1996), we quantify the role of magnitude as well as the orientation in ensemble formation for two simple dynamical systems; recently, Buizza (1996) has considered analogous effects in an operational atmospheric model. Unconstrained and constrained ensembles are contrasted in this section.

What exactly is meant by "Predictability"? We shall use the word in two distinct contexts. In the perfect model scenario, predictability is lost either (i) when an initial uncertainty increases by a factor of q or (ii) when the forecast adds no new informa-

tion to the climatology. As is common in meteorology (see, for example, Toth (1991)), predictability is quantified by a time scale; this time scale indicates when the forecast uncertainty will exceed some bound in (i), while in (ii) it indicates when information of the initial condition is lost. It is widely held that both of these time-scales are related to the inverse of the largest Lyapunov exponent (defined in section 2b below). Evidence is provided to illustrate that generally this is not the case; an explicit counterexample is given in Smith (1994a).

The systems analysed in this paper fall within the lower levels of a hierarchy of models of increasing complexity, the higher levels of which include operational numerical weather prediction (NWP) and climate models. While generalization between different nonlinear systems is always hazardous, the results presented here, and in particular the shortcomings of measures of predictability based on infinitesimal uncertainties, should be of use in the design and analysis of predictability experiments on more complex members of the hierarchy.

2. Local dynamics of infinitesimal uncertainties and predictability

Practical operational forecasting is concerned with the evolution of finite uncertainties over some interval of time in a particular dynamical system. In this section we discuss the dynamics of *infinitesimal* uncertainties and motivate several natural definitions of different *types* of initial uncertainty, returning to finite uncertainties in section 6. Rates and time scales provide two distinct classes of statistic employed as predictability measures. To define a rate some duration must be chosen *a priori*; this duration may be infinitesimal to yield an instantaneous growth rate, or finite to yield an "effective" growth rate, or infinite again yielding "effective" growth rates and their kin (*e.g.* global Lyapunov exponents). Alternatively we may compute time scales directly by considering uncertainty doubling (quadrupling, etc) times, and *then* take averages if single value is required. Each of these rates and times provide different information about the predictability of the system.

In general both rates and times are local: each statistic will depend both on the initial condition and on the orientation of the uncertainty. We will contrast different distributions of orientations: if the orientation is distributed isotropically we have an *unconstrained* distribution; alternatively we may consider a subset of orientations, for example only those in locally expanding directions, and thereby restrict attention to a *constrained* distribution. While the isotropic distribution is often the most analytically tractable, it may be the least physically meaningful.

Global measures of predictability are often defined by averaging local measures, weighted, of course, by the likelihood of each locality as defined by the climatology (or natural measure, see Ott (1994)). When estimating such global measures, we aim to sample initial conditions uniformly with respect to this natural measure. For systems which evolve on a strange attractor, sampling with respect to the natural measure is often referred to as sampling "uniformly on the attractor." The interpretation of such averages, (*e.g.* contrasting geometric and arithmetic means) is discussed in section 4. Again, we stress the limited scope of infinitesimal uncertainties; the consideration of finite amplitude ensembles in section 6, exposes the limited utility of infinitesimal measures.

(*a*) *Uncertainty dynamics*

We formulate our dynamical system as a set of m autonomous nonlinear ordinary differential equations

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{1}$$

where $\mathbf{x} \in R^m$, $\mathbf{f} : R^m \to R^m$. The Lorenz system (Lorenz 1963, Sparrow, 1982) provides a concrete example obtained by retaining the nontrivial dynamics from Saltzman's expansion of thermally driven convection in a 2-D fluid. The Lorenz system is:

$$\begin{aligned} \dot{x} &= -\sigma x + \sigma y \\ \dot{y} &= -xz + rx - y \\ \dot{z} &= xy - bz \end{aligned} \tag{2}$$

where x describes the intensity of the convective motion, y characterizes the temperature difference between ascending and descending fluid elements, and z is proportional to the deviations of the vertical temperature profile from its equilibrium value. The parameters σ, r and b are the Prandtl-number, Rayleigh-number and a geometrical parameter reflecting the shape of the convective cell. With $\sigma = 10$, $r = 28$ and $b = \frac{8}{3}$ one obtains the standard Lorenz attractor with the well-known butterfly shape. The three fixed points of the system are $\mathbf{x}_{f,0} = (0, 0, 0)$, $\mathbf{x}_{f,\pm} = (\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$.

Given $\mathbf{f}$ and an initial condition $\mathbf{x_0} \equiv \mathbf{x}(t = 0)$, the trajectory $\mathbf{x}(t)$ is uniquely determined. But suppose there is an uncertainty, $\boldsymbol{\epsilon}_0 \in \mathcal{R}^m$, in the initial condition $\mathbf{x}_0$; it is the growth of this uncertainty we wish to quantify. The dynamics of an infinitesimally small uncertainty is governed by the linearization of the flow, that is

$$\dot{\boldsymbol{\epsilon}} = J(\mathbf{x})\boldsymbol{\epsilon} \tag{3}$$

where $J(\mathbf{x})$ is the Jacobian of $\mathbf{f}$ at $\mathbf{x}$. In the Lorenz case

$$J(\mathbf{x}) = \begin{pmatrix} -\sigma & \sigma & 0 \\ (r-z) & -1 & -x \\ y & x & -b \end{pmatrix}. \tag{4}$$

Given a single forecast trajectory in the perfect model scenario, uncertainty is defined in terms of the distance between the (inexact) forecast and the true state as a function of time, that is the magnitude of the uncertainty at time t is simply $\epsilon(t) = |\boldsymbol{\epsilon}(t)|$. Equations 1 and 3 provide a dynamical system for infinitesimal uncertainties, from which we proceed to define effective rates in 2(b) and time scales in 2(c).

(*b*) *Growth rates: Effective and Instantaneous*

An uncertainty of size ϵ_0 which after a time Δt has grown to $\epsilon(\Delta t)$ has an *effective* growth rate

$$\overline{r}_{\Delta t} = \frac{1}{\Delta t} \log \frac{\epsilon(\Delta t)}{\epsilon_0}. \tag{5}$$

The independent, and rarely satisfied, additional assumption that the growth of the uncertainties is uniform in time allows the relation $\epsilon(t) \simeq \epsilon_0 e^{\overline{r}_{\Delta t} t}$. The failure to distinguish "effective" growth from a uniform growth has resulted in much confusion.

In the limit $\Delta t \to 0$, the instantaneous growth rate, r, is simply $\frac{1}{\epsilon}\frac{d\epsilon}{dt}|_t$, while the instantaneous growth rate at time t of an initial uncertainty which was ϵ_0 at $t = 0$ is

$$\frac{1}{\epsilon}\frac{d\epsilon}{dt}\bigg|_t = \frac{1}{\epsilon}\frac{d}{dt}\sqrt{\boldsymbol{\epsilon}(t)^T\boldsymbol{\epsilon}(t)} = \frac{1}{2\epsilon^2}\left(\dot{\boldsymbol{\epsilon}}(t)^T\boldsymbol{\epsilon}(t) + \boldsymbol{\epsilon}(t)^T\dot{\boldsymbol{\epsilon}}(t)\right) \tag{6}$$

where the superscript T is the transpose operator. For simplicity, we shall assume throughout this paper that the matrices of interest are full rank with distinct eigenvalues, and write the eigen-decomposition of J as $J = S\Lambda S^{-1}$, where the diagonal matrix Λ contains the eigenvalues, λ_i, of J and S its eigenvectors. Whether ϵ is increasing or decreasing (equivalently whether r is greater than or less than zero), will depend on (i) the eigenvalues of $J(\mathbf{x})$, (ii) the projections of the uncertainty onto the eigenvectors, $\boldsymbol{\epsilon}^T S$, and (iii) its decomposition into the eigenvectors, $S^{-1}\boldsymbol{\epsilon}$. The eigenvalues of J alone do not supply sufficient conditions to determine the sign of r since, in general, the eigen-basis is not orthogonal (equivalently, J is non-normal).

Combining equations (3) and (6), and dropping the explicit time dependence of $\mathbf{x}$ and $\boldsymbol{\epsilon}$ for clarity, we have

$$r(\boldsymbol{\epsilon}_0, \mathbf{x}_0, t) = \frac{\boldsymbol{\epsilon}^T\left(\frac{J(\mathbf{x})+J(\mathbf{x})^T}{2}\right)\boldsymbol{\epsilon}}{\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}} \tag{7}$$

As $\frac{J(\mathbf{x})+J(\mathbf{x})^T}{2}$ is symmetric, its eigenvalues are real and the eigenvectors are orthogonal. In addition to the instantaneous growth rate for a particular $\boldsymbol{\epsilon}$, equation (7) provides bounds on the growth rate at $\mathbf{x}$ for any $\boldsymbol{\epsilon}$. Interpreting $r(\boldsymbol{\epsilon}_0, \mathbf{x}_0, t)$ as the Rayleigh-quotient of $\boldsymbol{\epsilon}_t$ (see Strang 1988) the largest eigenvalue of $\frac{J(\mathbf{x})+J(\mathbf{x})^T}{2}$ provides an upper bound on $r(\boldsymbol{\epsilon}_0, \mathbf{x}_0, t)$ for any $\boldsymbol{\epsilon}$ at $\mathbf{x}$; similarly its smallest eigenvalue is a lower bound.

In particular, if *all* eigenvalues of $\frac{J(\mathbf{x}_0)+J(\mathbf{x}_0)^T}{2}$ are positive (negative), $r(\boldsymbol{\epsilon}_0, \mathbf{x}_0, 0)$ will be positive (negative) *independent* of the orientation of the uncertainty. Thus if this symmetric matrix is definite, the sign of $r(\boldsymbol{\epsilon}_0, \mathbf{x}_0, 0)$ is determined; only its magnitude depends on the orientation of the uncertainty.

In the Lorenz system

$$\frac{J(\mathbf{x}) + J(\mathbf{x})^T}{2} = \begin{pmatrix} -\sigma & \frac{\sigma+r-z}{2} & \frac{y}{2} \\ \frac{\sigma+r-z}{2} & -1 & 0 \\ \frac{y}{2} & 0 & -b \end{pmatrix}. \tag{8}$$

It follows immediately that there is no state $\mathbf{x}$ in state space where this matrix is positive definite. Yet it can be negative definite, as shown in section 5. The ensemble average of the instantaneous growth rates for *isotropically distributed* uncertainties is given by the arithmetic mean, $\langle r(|\boldsymbol{\epsilon}_0|, \mathbf{x}_0, 0)\rangle_{\boldsymbol{\epsilon}_0} = \frac{1}{m}\sum_{i=1}^{m}\lambda_i$, where the λ_i are the eigenvalues of $J(\mathbf{x}_0)$ (or equivalently, those of $\frac{J(\mathbf{x}_0)+J(\mathbf{x}_0)^T}{2}$).

The *effective* growth rates over *finite* Δt depend on an integral (or solution) of equation (3) along a trajectory $\mathbf{x}(t)$ given $\mathbf{x}_0$ and $\boldsymbol{\epsilon}_0$. This requires the solution of the first variational equation (see Parker and Chua (1989)) and leads to a linear resolvent operator, $M(\mathbf{x}_0, \Delta t)$, (also called "tangent propagator") which maps $\boldsymbol{\epsilon}_0$ at $(\mathbf{x}_0, t_0)$ to $\boldsymbol{\epsilon}(t)$ observed at $(\mathbf{x}(t), t)$, that is

$$\boldsymbol{\epsilon}(t_0 + \Delta t) = M(\mathbf{x}_0, \Delta t)\boldsymbol{\epsilon}_0. \tag{9}$$

Taking $t_0 = 0$, the magnitude of $\boldsymbol{\epsilon}(t)$ is

$$\epsilon(t) = \|\boldsymbol{\epsilon}(t)\| = \sqrt{\boldsymbol{\epsilon}(t)^T \boldsymbol{\epsilon}(t)} = \sqrt{\boldsymbol{\epsilon}_0 M(\mathbf{x}_0, t)^T M(\mathbf{x}_0, t) \boldsymbol{\epsilon}_0}. \tag{10}$$

The effective growth rate $\overline{r}_t$ is then defined as

$$\overline{r}_t(\boldsymbol{\epsilon}_0, \mathbf{x}_0) = \frac{1}{t} \log \sqrt{\frac{\boldsymbol{\epsilon}_0 M(\mathbf{x}_0, t)^T M(\mathbf{x}_0, t) \boldsymbol{\epsilon}_0}{\boldsymbol{\epsilon}_0{}^T \boldsymbol{\epsilon}_0}}. \tag{11}$$

As the term under the root is a quadratic form, $\overline{r}_t(\boldsymbol{\epsilon}_0, \mathbf{x}_0)$ is bound by the eigenvalues of $M^T M$. The singular value decomposition (SVD) of M is $M = U\Sigma V^T$ with orthogonal matrices V (U) containing the right (left) singular vectors as columns and Σ the diagonal matrix of the singular values σ_i with $\sigma_i \geq \sigma_j$ for $i < j$. The σ_i^2 are the eigenvalues of $M^T M$. Under M, each right singular vector, $\mathbf{v}_i$, is rotated into the corresponding left singular vector $\mathbf{u}_i$, and stretched (shrunk) by the factor σ_i, for σ_i greater than (less than) one; that is $M\mathbf{v}_i = \sigma_i \mathbf{u}_i$.

For a given $\mathbf{x}_0$ and t, the σ_i define the finite-time Lyapunov exponents (Lorenz (1965), Abarbanel *et al.* (1991), Ziehmann *et al.* (1999)), and the maximum $\overline{r}_t$ is observed when the uncertainty is aligned with the first right singular vector $\mathbf{v}_1$. In the limit $t \to \infty$, $\overline{r}_t(\boldsymbol{\epsilon}_0, \mathbf{x}_0)$ approaches the largest global Lyapunov-exponent, Λ_1, for almost all pairs $(\mathbf{x}_0, \boldsymbol{\epsilon}_0)$ while the (global) Lyapunov-spectrum $\Lambda_i, i = 1, \ldots, m$ of the system corresponds to the logarithms of the singular values of $M(\mathbf{x}_0, t)$, assuming these limits exist (Oseledec 1968).

As $t \to \infty$, the local orientation of the first global *Lyapunov vector* at each point $\mathbf{x}(t)$ on the attractor is defined by $\mathbf{u}_1(t)$; defining this local direction at a fixed location $\mathbf{x}_0$ assumes the singular value decomposition of $\lim_{\Delta t \to -\infty} M(\mathbf{x}_0, \Delta t)$ is evaluated over a bounded trajectory, and thus this direction is not defined for almost every $\mathbf{x}_0$ not on the attractor. While the singular vectors can be determined from local dynamics over finite time, the orientation of the Lyapunov vectors cannot, simply because the true Lyapunov vectors depend on an integral which extends into the infinite past*. The orientations of Lyapunov and the singular vectors in the 2-D Ikeda map (Ikeda 1979) are illustrated in figure 1. The left panel shows numerical estimates of the Lyapunov vector and the first right ($\Delta t = 4$) singular vector at a point $\mathbf{x}(t)$, while in the right panel their images at $\mathbf{x}(t + \Delta t)$ are shown. Not only are the global Lyapunov vectors and exponents defined only for orbits that are recurrent in state space (as opposed to on a transient), but these quantities are defined for a specific dynamical system and trajectory; in the meteorological context this corresponds to the perfect model scenario with a noise free trajectory. Even with a perfect model, a time series of points obtained from observations under variational assimilation (see Pires *et al.* 1996) need not correspond to a trajectory which meets the above requirements. Under other methods of assimilation, the series need not yield a model trajectory at all.

For simplicity, we have employed the Euclidean norm throughout this paper; generalized stability analysis and the extension to other metrics is discussed in Orr (1907),

* Our numerical approximation of these local directions satisfies the obvious necessary conditions (*e.g.* each integration is of a duration such that large variations in the initial orientation have negligible effects, say less than one part in 10^6, in the final orientation. In the Ikeda Map, this usually requires only 10 to 50 pre-iterates.). Nevertheless this is not a sufficient condition. The same restriction, of course, applies to estimating Lyapunov exponents; for this reason there are very few systems which can be proven to be chaotic, we rely only on numerical evidence. All Lyapunov vectors and exponents presented in the current paper are subject to this caveat.

Farrell (1990, 1994), Molteni and Palmer (1993), and Buizza and Palmer (1995) and references therein. While the singular values and both the finite time and the finite sample Lyapunov exponents (see Ziehmann *et al.* 1999) depend on the choice of norm, the global Lyapunov exponents do not.

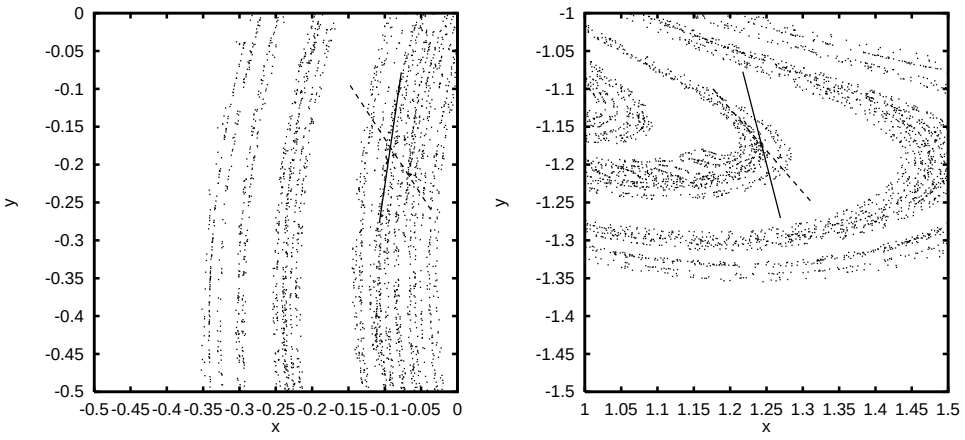

Figure 1. Two details of the Ikeda attractor showing estimated Lyapunov vectors (solid) and Singular vectors (dashed) at two points along a trajectory; the two points are separated by 4 iterations. Singular vectors are defined over these 4 iterations. The left panel shows the earlier point and the right singular vectors at initial time while the right panel shows the later point and the left singular vectors. These estimates are, of course, numerical approximations as noted in the text.

(*c*) *Time scales*

Given an uncertainty $\boldsymbol{\epsilon}_0$ at $\mathbf{x}_0$ and $t = 0$, a q-pling time $\tau_q(\boldsymbol{\epsilon}_0, \mathbf{x}_0)$ is defined by the smallest time t for which $\epsilon(t) \geq q\epsilon_0$. In symbols

$$\tau_q(\mathbf{x}_0, \boldsymbol{\epsilon}_0) = \min_{t>0}\{t \,\Big|\, \|F_t(\mathbf{x} + \boldsymbol{\epsilon}_\mathbf{0}) - F_t(\mathbf{x})\| \geq q\|\boldsymbol{\epsilon}_0\|\}. \tag{12}$$

The choice $q = 2$ leads to the doubling time τ_2, while $q = e$ yields the so-called e-folding time. In the limit $\|\boldsymbol{\epsilon}_0\| \to 0$, the τ_q are given by simultaneous integration of equations (1) and (3), but in this case they reflect only infinitesimal uncertainties.

The τ_q directly quantify the time at which a certain threshold is reached: this information cannot be derived from the effective rates. Further, the arithmetic average τ_q need bear no relation to the statistics of $\frac{1}{r_{\Delta t}}$ for any fixed Δt: the inverse of the average of the inverse of X is almost invariably a poor estimate of the (arithmetic) average of X. The potentially extreme effects of using such an estimate are illustrated by considering a random variable x, drawn uniformly from the interval $(0, 1)$. The average value of x converges to $\frac{1}{2}$ as expected; the average value of $\frac{1}{x}$ does not converge to 2. Similarly, $\frac{1}{\Lambda_1}$ (measured in time/bits) is not simply related to the mean (or median) doubling time. Lyapunov exponents need not reflect predictability. This comes about first, because a duration *must* be chosen *a priori* as part of the definition of an "effective rate," and

second, because the definition of Lyapunov exponents assumes the uncertainty remains infinitesimal. If the growth rate is not constant, then neither global Lyapunov exponents nor their finite-time counterparts provide a time scale for predictability. In this sense, the τ_q better quantify the growth of uncertainties than either $\overline{r}_{\Delta t}$ or the Λ_i even for infinitesimal uncertainties. While the τ_q do provide a specific time-scale for predictability, the variation of τ_q with $\mathbf{x}$ constrains the interpretation of a single τ_q as defining a limit of predictability whenever $\tau_{q^2} \neq 2\tau_q$ for some q sufficiently small as to be of practical interest.

3. Non-Uniform uncertainty evolution in uniform linear systems

Before considering nonlinear systems, we examine the complexity already inherent in uniform linear systems with constant, if non-normal, Jacobian. Here the asymptotic growth rate is determined by the eigenvalues of J, and the transient dynamics can be solved analytically, the evolution of an initial uncertainty $\boldsymbol{\epsilon}_0$ is

$$\dot{\boldsymbol{\epsilon}} = J\boldsymbol{\epsilon} = S\Lambda S^{-1}\boldsymbol{\epsilon} \tag{13}$$

solving for $\epsilon(t)$ yields

$$\boldsymbol{\epsilon}(t) = Se^{\boldsymbol{\Lambda} t}S^{-1}\boldsymbol{\epsilon}_0 = \sum_{i=1}^{m} c_i e^{\lambda_i t}\boldsymbol{\eta}_i \tag{14}$$

where $\mathbf{c} = S^{-1}\boldsymbol{\epsilon}_0$ is the decomposition of the initial uncertainty in the eigenvector basis $\boldsymbol{\eta}_i$. For a linear 2-D system the magnitude $\epsilon(t)$ grows as

$$\epsilon(t) = \sqrt{c_1^2 e^{2\lambda_1 t} + c_2^2 e^{2\lambda_2 t} + 2c_1 c_2 e^{(\lambda_1+\lambda_2)t}\boldsymbol{\eta}_1 \cdot \boldsymbol{\eta}_2}. \tag{15}$$

Asymptotically, $\epsilon(t)$ grows exponentially as $e^{\lambda_1 t}$ but *for finite time* it does not. In this uniform case there is no dependence on $\mathbf{x_0}$ but only on initial orientation; it follows from (15) that the instantaneous growth rate at time t is simply

$$r(\boldsymbol{\epsilon}_0, t) = \frac{\lambda_1 c_1^2 e^{2\lambda_1 t} + \lambda_2 c_2^2 e^{2\lambda_2 t} + (\lambda_1+\lambda_2)c_1 c_2 e^{(\lambda_1+\lambda_2)t}\boldsymbol{\eta}_1 \cdot \boldsymbol{\eta}_2}{c_1^2 e^{2\lambda_1 t} + c_2^2 e^{2\lambda_2 t} + 2c_1 c_2 e^{(\lambda_1+\lambda_2)t}\boldsymbol{\eta}_1 \cdot \boldsymbol{\eta}_2} \tag{16}$$

and the effective growth rate is given by

$$\overline{r}_t(\boldsymbol{\epsilon}_0) = \frac{1}{2t}[\log(c_1^2 e^{2\lambda_1 t} + c_2^2 e^{2\lambda_2 t} + 2c_1 c_2 e^{(\lambda_1+\lambda_2)t}\boldsymbol{\eta}_1 \cdot \boldsymbol{\eta}_2) - \log(c_1^2 + c_2^2 + 2c_1 c_2 \boldsymbol{\eta}_1 \cdot \boldsymbol{\eta}_2)]. \tag{17}$$

Since $2\lambda_1 \geq \lambda_1 + \lambda_2 \geq 2\lambda_2$ both the instantaneous and the effective growth rates will approach λ_1 for almost every $\boldsymbol{\epsilon}_0$ (*i.e.* as long as $c_1 \neq 0$). To disentangle effects due to magnitude of the eigenvalues and the non-orthogonality of the eigen-basis, we construct J by defining an eigen-basis of real eigenvectors separated by angle δ through

$$S = \begin{pmatrix} \cos\alpha & \cos(\alpha+\delta) \\ \sin\alpha & \sin(\alpha+\delta) \end{pmatrix}. \tag{18}$$

Thus $\boldsymbol{\eta}_1 \cdot \boldsymbol{\eta}_2 = \cos\delta$ and the inverse of S is

$$S^{-1} = \frac{1}{\sin\delta} \begin{pmatrix} \sin(\alpha+\delta) & -\cos(\alpha+\delta) \\ -\sin\alpha & \cos\alpha \end{pmatrix}. \tag{19}$$

Hence we may consider any 2-D constant Jacobian system with real eigenvalues λ_1 and λ_2 and real eigenvectors

$$J = \frac{1}{\sin\delta} \begin{pmatrix} \cos\alpha & \cos(\alpha+\delta) \\ \sin\alpha & \sin(\alpha+\delta) \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \sin(\alpha+\delta) & -\cos(\alpha+\delta) \\ -\sin\alpha & \cos\alpha \end{pmatrix}. \tag{20}$$

Any initial uncertainty $\boldsymbol{\epsilon}_0 = (\cos\gamma, \sin\gamma)$ of unit length decomposes into the eigenvector basis (18) with $c_1 = \frac{1}{\sin\delta}(\sin(\alpha+\delta)\cos\gamma - \cos(\alpha+\delta)\sin\gamma)$ and $c_2 = \frac{1}{\sin\delta}(\cos\alpha\sin\gamma - \sin\alpha\cos\gamma)$.

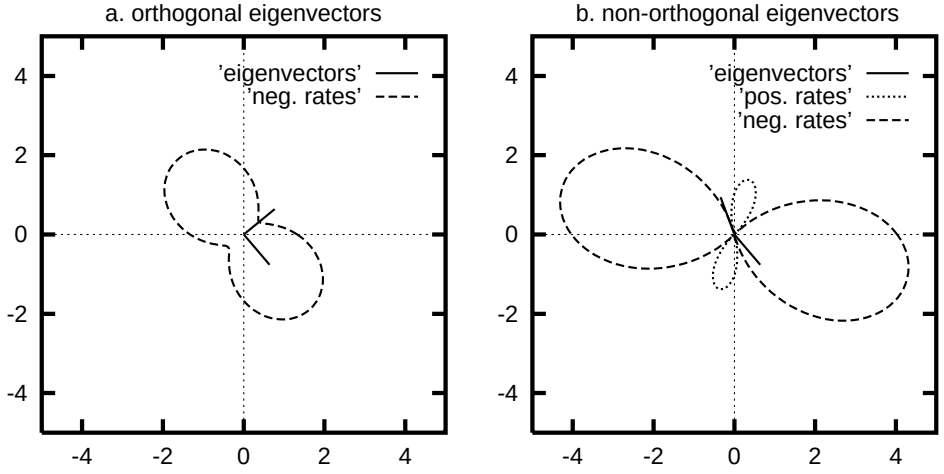

Figure 2. Initial instantaneous growth rates in a linear 2-dimensional system with both negative eigenvalues $\lambda_1 = -0.5$, $\lambda_2 = -2.5$ for (a) $\delta = 90^0$ and (b) $\delta = 160^0$. The magnitude of the growth rates in a specific direction is reflected by the radius, negative rates are long-dashed lines, positive rates dotted, the eigenvectors are also shown.

Figure 2 shows the effect of non-orthogonality on the instantaneous growth rates at time $t = 0$ when *both* eigenvalues are negative. In the orthogonal case (panel a) all rates are negative and the maxima and minima correspond to the eigenvalues. Yet a corresponding non-orthogonal case (panel b) includes large positive growth rates. In the orthogonal case the maximum growth rate is, of course, aligned with eigenvector $\boldsymbol{\eta}_1$, while in the non-orthogonal case it is oriented in the direction $\gamma_{max} = \alpha - \frac{\arctan(\cot(\delta))}{2}$.

Graphs showing curves of the instantaneous growth rate as a function of time for various initial directions are often cited in the literature. It is important to remember that even when the instantaneous growth rate in different directions rapidly equilibrates, the difference in magnitude between perturbations initially in these directions may remain substantially different. This difference is more easily seen in graphs of the effective growth rate as a function of time, although it is least ambiguously reflected by simply plotting the magnitudes as a functions of time. If the initial uncertainty is isotropic, γ is uniformly distributed and integration of (16) and (17) over angle yields the ensemble-averaged instantaneous growth rate $\langle r(\boldsymbol{\epsilon}_0, t)\rangle_\gamma$ and the ensemble-averaged effective growth rate $\langle \overline{r}_t(\boldsymbol{\epsilon}_0)\rangle_\gamma$, the behaviour of which are shown in figure 3. Both $\langle r(\boldsymbol{\epsilon}_0, t)\rangle_\gamma$ and $\langle \overline{r}_t(\boldsymbol{\epsilon}_0)\rangle_\gamma$ evolve from an initial value of $\frac{\lambda_1+\lambda_2}{2}$ toward λ_1 as $t \to \infty$. Clearly the short term be-

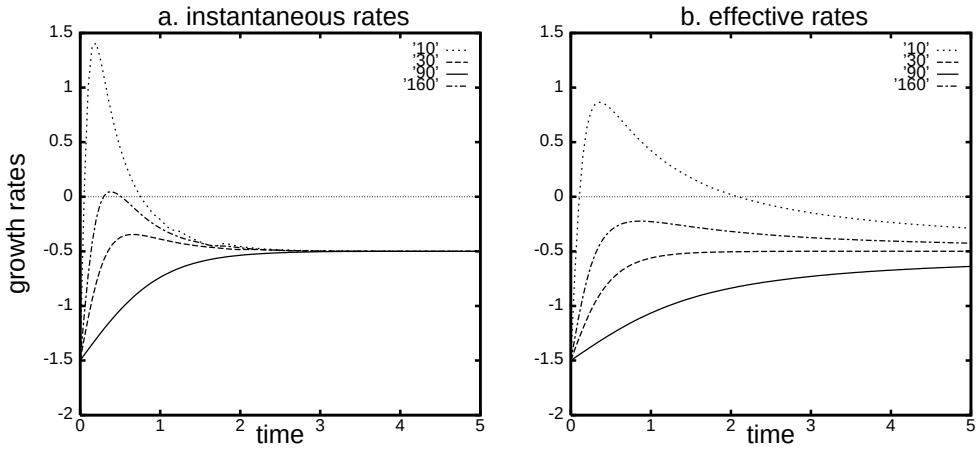

Figure 3. Ensemble-averages of the instantaneous (a) and the effective (b) growth rates of an ensemble of uniformly distributed uncertainties as a function of time in a linear 2-dimensional system with $\alpha = 30^0$, $\lambda_1 = -0.5$, $\lambda_2 = -2.5$ for $\delta = 10^0, 30^0, 90^0, 160^0$.

haviour is *not* represented by λ_1 even in a uniform (linear) system if $m > 1$. In the orthogonal case, λ_1 and λ_m determine the maximum and minimum growth rates; this is not the case in a uniform non-orthogonal system where even the ensemble-averaged growth rates may significantly exceed the largest eigenvalue.

This apparently anomalous growth simply results from non-orthogonality: the difference between the ensemble average of the squared uncertainties in an arbitrary non-orthogonal system and the corresponding orthogonal system (with the same eigenvalues) is simply:

$$\langle \epsilon(t)^2\rangle_{\gamma,non-orthogonal} - \langle \epsilon(t)^2\rangle_{\gamma,orthogonal} = \frac{1}{2}\cot^2(\delta)[e^{\lambda_1 t} - e^{\lambda_2 t}]^2 \qquad (21)$$

which is nonnegative for all times $t > 0$ regardless of δ. When δ is small (*i.e.* the eigen-directions are similar), this difference becomes very large and the growth rates may exceed λ_1. This has been termed "super-exponential" growth (or "sub-exponential" if the rate is smaller than λ_1) by Nicolis *et al.* (1995). We note that there is no faster than exponential

growth; rather $\epsilon(t)$ is simply a sum of exponentials. Indeed, in non-normal systems it is often the rapid exponential *decay* of one component of $\boldsymbol{\epsilon}$ that results in the growth of the $|\boldsymbol{\epsilon}|$. From equation (21) we see that, for linear systems with identical eigenvalues but different eigenbases, the expected growth rate is smallest in the orthogonal case. We also note that the effective rates converge to λ_1 much more slowly than the instantaneous rates.

These results for linear systems hold important implications for nonlinear systems as well. For instance, common algorithms for estimating the largest Lyapunov exponent from time series data (Wolf *et al.* (1985), Kurths and Herzel (1987)) approximate the linear evolution about a fiducial trajectory by monitoring its distance from a reference trajectory. When this distance grows too large, the reference trajectory is replaced by a closer reference trajectory with a "similar" orientation. The results of this section demonstrate that small errors in angle may significantly effect the estimated exponent.

In terms of predictability, we have shown how non-normality may yield growth rates larger than the largest eigenvalue; Nicolis et al. (1995) the "non-uniformity" as another mechanism for this behaviour. Growth rates exceeding either the maximum eigenvalue or Lyapunov-exponent are also discussed in Lacarra and Talagrand (1988) and Trevisan and Legnani (1995).

4. Uncertainty dynamics in nonlinear systems - Numerical results

Growth rates and q-pling times cannot be obtained analytically in a nonlinear system where trajectories must be approximated numerically; in this section we present results for the Lorenz (Lorenz 1963), Moore-Spiegel (Moore and Spiegel 1966) and Rössler (Rössler 1976) systems. Our first task is to obtain a set of initial conditions and local orientations. For each system, these are obtained from a very long, post-transient integration of the augmented dynamical system consisting of equations 1 and 3. N points are chosen at random* from this long (but finite) trajectory subject to the requirement that they are well separated in time (*i.e.* a minimum separation of $\geq 10^3$ characteristic time). This yields both the initial conditions $\mathbf{x}$ and local orientation corresponding to the first *global* Lyapunov exponent Λ_1, which we will refer to as the Lyapunov direction, $\mathbf{l}_1$, at $\mathbf{x}$. As noted in Section 2(b), these orientations are only approximate as they are not determined from an infinitely long trajectory. The fixed time singular direction, $\mathbf{v_1}$, corresponds to the first right singular vector of $M(\mathbf{x}, \Delta t_{opt})$. The value of N and the optimization time, Δt_{opt}, vary with the system and experiment. In addition to these well-defined orientations, we shall also consider (at each $\mathbf{x}$) one random orientation, $\mathbf{r}_1$, drawn from an isotropic distribution.

At each $\mathbf{x}$, uncertainty growth rates are obtained by integrating the augmented equations (for each initial orientation) for a fixed time. In this case, if the distribution of initial values of $\mathbf{x}$ reflect the natural measure, then the distribution of final images will also. To determine the q-pling times, the experiment is repeated but each initial condition is integrated *not* for a fixed time, but rather until the uncertainty has increased by a factor of q. For each q, $\mathbf{x}$, and orientation Θ we denote this time by $\tau_q(\mathbf{x}, \Theta)$. In this case, the distribution of final points need not reflect the natural measure; figure 5 shows that the distribution of final points for the doubling time is indeed very different from the natural measure in the Lorenz case: doubling only occurs in regions of the state space with small z-values suggesting the hypothesis that the "missing" portion of the attractor lies in a region which is characterized by enhanced predictability.

* The distribution formed by taking points at random from along a trajectory approaches the natural measure as the length of the trajectory increases toward infinity.

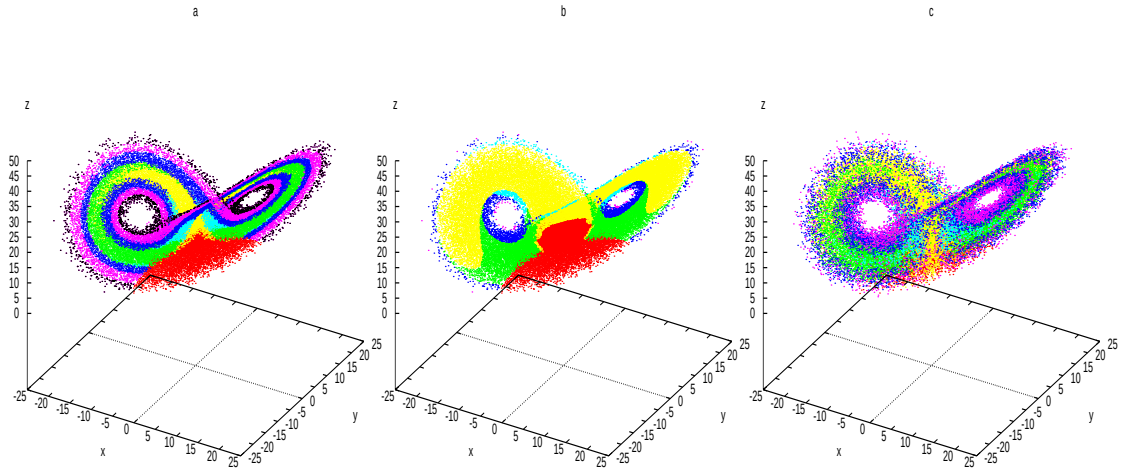

Figure 4. Views of the Lorenz attractor showing the doubling times for 2^{16} initial conditions for the Lyapunov direction (a), the maximum singular vector optimized for the doubling time obtained in the Lyapunov direction (b), and in random direction (c). Red indicates $\tau_2 < 0.15$, yellow $0.15 \leq \tau_2 < 0.5$, green $0.5 \leq \tau_2 < 0.8$, turquoise $0.8 \leq \tau_2 < 1$, blue $1 \leq \tau_2 < 2$, magenta $2 \leq \tau_2 < 4$, and black for $\tau_2 \geq 4$.

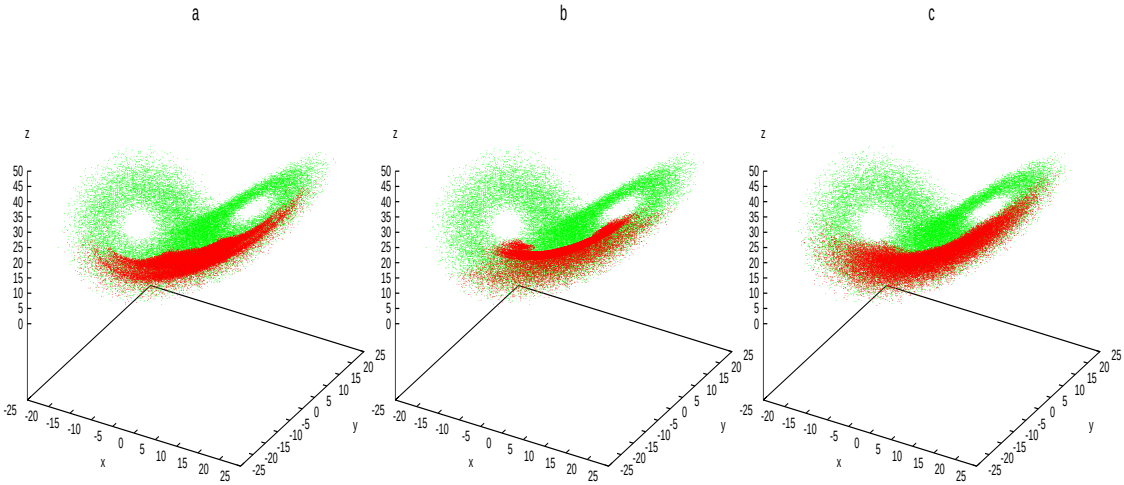

Figure 5. Three views showing where uncertainty doubling occurs in the 3 uncertainty directions described in figure 4. The starting points (green) reflect the natural measure on the attractor while the points where doubling occurs (red) do not. Each panel contains $N = 2^{16}$ initial conditions and their images.

In forecasting, one would like a local estimate of predictability for a particular state: in figure 4 each initial condition is colored by its doubling time; the three panels correspond to the three different initial orientations. A point is colored red if the magnitude of the uncertainty doubled within less than 0.15 time units (indicating poor predictability), yellow if $0.15 \leq \tau_2 < 0.5$, followed by green, turquoise, and blue for τ_2 between 0.5 and 0.8, 0.8 and 1.0, or 1.0 and 2.0. Magenta points reflect doubling times between 2.0 and 4.0, while a few points with doubling times larger than 4 time units are black (indicating enhanced predictability). For the Lyapunov orientation (figure 4(a)), the variation of $\tau_2(\mathbf{x}, \mathbf{l}_1)$ is nicely organized in a banded structure with the highest predictability (black and magenta) at the outer and inner margins of the attractor and the lowest predictability (red) in the middle of the bands and at small z-values. Unlike the Lyapunov orientation, $\mathbf{l}_1$, computing the singular orientation, $\mathbf{v}_1$, requires choosing an optimization time Δt_{opt}; one could choose a constant Δt_{opt} for all $\mathbf{x}$, but the wide variation in $\tau_2(\mathbf{x}, \mathbf{l}_1)$ on the attractor reflected in figure 4(a) (see also figure 6 below), suggests that no global optimization time exists which reflects the relevant segment of each trajectory. Taking local optimization times $\Delta t_{opt} = \tau_2(\mathbf{x}, \mathbf{l}_1)$ allows the optimization time to vary over 3 orders of magnitude in response to variations in uncertainty growth; this at least ensures that the segment of trajectory over which $M(\mathbf{x}, \Delta t_{opt})$ is evaluated reflects time scales over which q-pling could occur*. The $\tau_2(\mathbf{x}, \mathbf{v_1})$ for this choice are shown in figure 4(b). For these finite values of q, note that there is often an orientation which q-ples *faster* than the Lyapunov direction: the Lyapunov orientation is *not* optimal if the goal is to find the "worst case scenario" growth, as is often the case in meteorology. Qualitatively, the picture for the random orientations (figure 4(c)) is similar, suggesting that "doubling" shows larger variation with initial condition than with initial orientation for the Lorenz attractor.

As indicated by the red points in figure 5, the distribution of points at which q-pling occurs [*i.e.* $\mathbf{x}(t + \tau_2(\mathbf{x}, \Theta))$ in figure 5] does not reflect the natural measure. Figure 5 shows this distribution to be similar for all three initial orientations. Each case reveals a "high predictability" region at large z where there are no points at all; this region is considered in detail in the next section. There are differences between the panels within the "low predictability" region (small z): doubling in $\mathbf{v_1}$ will almost always occur "earlier" in time than in Lyapunov direction when the singular direction is defined with $\Delta t_{opt} = \tau_2(\mathbf{x_0}, \mathbf{l}_1)$. In the Lorenz systems, this leads to a contraction in the final point distribution in state space for $\mathbf{v}_1$ relative to the distribution for $\mathbf{l}_1$. There is also a small band consisting only of green points in figure 5(a) just below the "eyes" of the attractor, while in figures 5(b) and 5(c) the same region includes both red and green points. Also note locations where doubling occurs in the random orientation but not the Lyapunov direction. The local orientation of the first global Lyapunov vector is sometimes more stable than the randomly chosen direction, and thus completely fails to warn of the worst case.

In stochastic systems, waiting times like τ_2 typically display (shifted) exponential distributions (see for example Papoulis 1991); the organization of predictability within deterministic chaos is reflected by the non-exponential character (particularly in the peaks) in the distribution of τ_2, shown for each of the three orientations in figure 6. Much weaker peaks are found in the Rössler and Moore-Spiegel systems (not shown); for the Lorenz system these peaks are especially sharp, because there are regions where doubling is not only unlikely but impossible, as proven in the next section. A collection

* The shortest possible doubling time for infinitesimals corresponds to the smallest Δt such that the largest singular value exceeds 2, and we call the corresponding $\mathbf{v}_1$ the nightmare direction (see section 6). Barkmeijer (1996) discusses the construction of rapidly growing finite perturbations.

of initial conditions which enter this region without having doubled *cannot* double until they emerge from it, forming a well-defined subset of initial conditions which will have "delayed" doubling events.

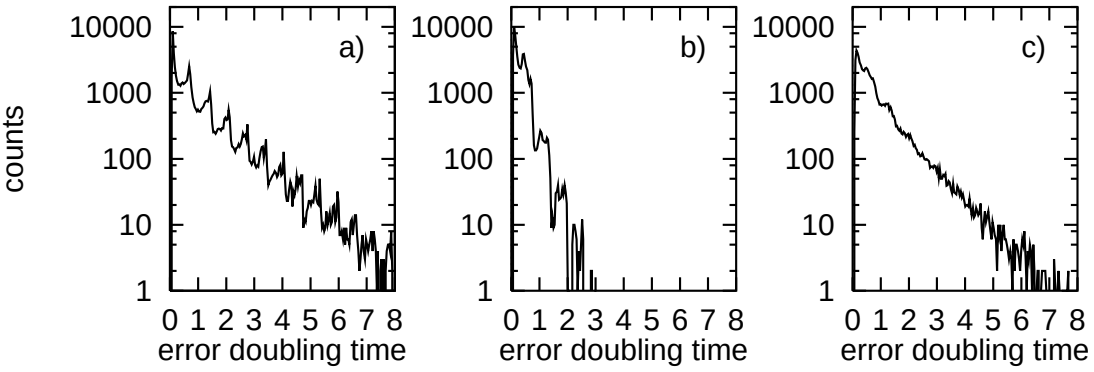

Figure 6. Histogram of uncertainty doubling times shown in figure 4 for Lorenz system in (a) the Lyapunov direction, (b) the maximum growing direction, and (c) the random direction.

Global average predictability measures: The decay of predictability is often presented through an "error growth" curve. "Error" in this case is defined via the distribution of the relative increase of infinitesimal uncertainties with time; specifically, for each initial condition we consider the relative uncertainty as a function of time, $m(t) = |\boldsymbol{\epsilon}(t)|/|\boldsymbol{\epsilon}(0)|$. $m(t)$ is also be referred to as the magnification or amplification of the initial (infinitesimal) uncertainty with specified orientation. The distribution is always considered over initial conditions, but we stress here that the initial orientation is crucial as well. First, consider local orientations to be in $\mathbf{l}_1$. Inasmuch as the true equations are used this is a perfect model, there is no prediction error *per se*: the curves in figure 7 reflect the growth of infinitesimal uncertainties given a perfect model.

Figure 7 shows both the arithmetic average (panel a) and geometric average (panel b) of the relative uncertainty as a function of time in the Lorenz system, where the initial uncertainty has (local) orientation $\mathbf{l}_1$. The relationship (or lack thereof) between the slope of these curves and Λ_1, the largest Lyapunov exponent, has created some confusion in the literature. We stress that Λ_1 determines the slope *only* if four conditions are met, namely: (a) the model is perfect *and* (b) the initial uncertainties are both infinitesimal *and* aligned in the local orientation of the global Lyapunov vector (*i.e.* $\mathbf{l}_1$) *and* (c) the geometric mean is taken *and* (d) the sample of initial conditions is large enough to effectively cover the attractor. The loss of any one of these conditions can break the connection between this slope and Λ_1.

A number of authors have reported "oscillations" in the error growth curves for the Lorenz system. Our estimate of the geometric average, the solid curve in Figure 12b, suggests that no such oscillations exist in this case, and this is the only case where the slope is determined by Λ_1. The oscillating dashed lines in Figure 12b show the effect dividing initial conditions into two groups (defined in the next section) each dashed line oscillates, but the oscillations cancel exactly. This suggests spurious oscillations would be

observed if the initial conditions were not chosen uniformly with respect to the natural measure: if all four of the conditions listed in the previous paragraph are met, then oscillations are forbidden by construction, and one is to expect a straight line with a slope given by Λ_1, as observed. If the arithmetic average is taken, then oscillations are to be expected, but the growth no longer reflects the Lyapunov exponents of the system over the short time scales of interest here. Indeed, if the geometric average is to reflect the value of Λ_1 and the growth rate is not constant (globally), then the arithmetic average *must* exceed the geometric average. In short, exponential growth in the geometric mean all but requires so-called "super-exponential" growth in the arithmetic mean.

Note that the global average of τ_2 *cannot* be determined from the slope of this graph, as the doubling time of this average is not the average doubling time. In inhomogeneous systems, global measures of predictability will always be of limited utility, nevertheless some will be more limited than others. The mean τ_2 is the expected time for the uncertainty in a randomly chosen initial condition to double. This is independent of the time at which the "averaged uncertainty" has doubled (whichever average is taken).

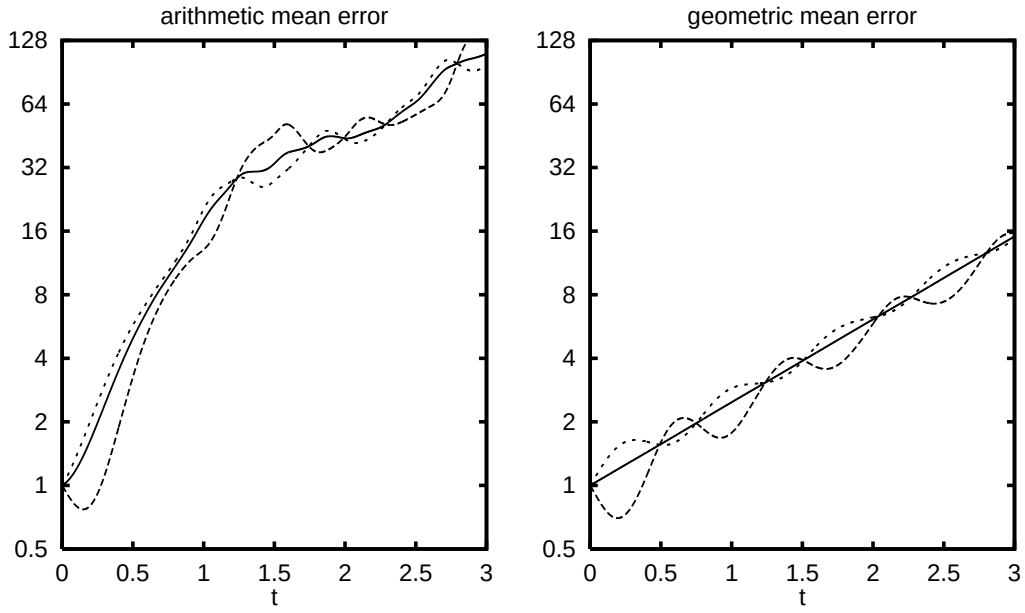

Figure 7. "Error growth" curves for infinitesimal uncertainties oriented in the Lyapunov direction. Panels show the arithmetic mean (a) and geometric mean (b) as a function of time. Each panel shows the global average (solid line) over $N = 2^{15}$ intial conditions, while the short dashed (long dashed) curves represent averages taken over those initial condition which lie in the growing (shrinking)region of the Lorenz attractor. Note that only the slope of the solid curve in the log-linear plot of panel (b) reflects the Lyapunov exponent.

Practically, the statistics of the uncertainty doubling times are of greater use in predicting error growth than the timing of the average uncertainty statistics. But the utility of the τ_q is also limited due to the inhomogeneity of the attractor and the assumption of infinitesimal uncertainty. To relate the distribution of τ_2 to some "Limit of Predictability" we must assume that the linearized dynamics also describe finite uncertainties *and* that the uncertainty growth is only weakly correlated over consecutive time segments

of order τ_2; that is, for example, that $\tau_4 \simeq 2\tau_2$. Figure 8 reveals this is not the case for any of the systems considered. As discussed in Smith (1996), the relation $\tau_{q^2} \simeq 2\tau_q$ may not hold, except for q so large that the assumption the initial uncertainty is still small becomes ludicrous.

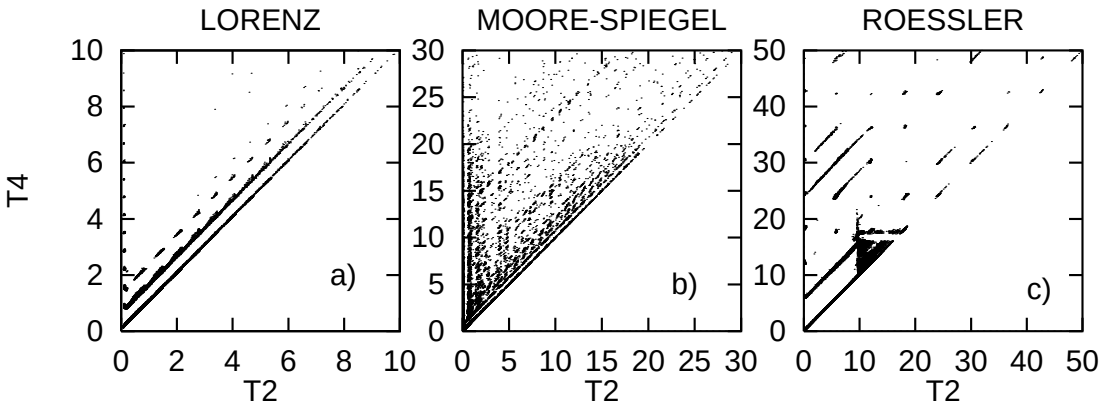

Figure 8. A comparison of the doubling time and quadrupling time of $N = 2^{16}$ initial conditions for the Lorenz, the Moore-Spiegel and the Rössler systems. In each case, the initial uncertainty is in the Lyapunov orientation.

Returning to Figure 8, note that in all three systems many of the points lie just above the diagonal indicating that the uncertainty growth tends to be episodic; here τ_4 only slightly exceeds τ_2. In the case of Lorenz, points also fall roughly along lines parallel to the diagonal, this suggests that those points which do not quadruple immediately after doubling tend to quadruple a (roughly) fixed time after they have doubled. This structure is due to the macroscopic structure of the Lorenz attractor which leads to almost periodic alternations between regions of uncertainty growth and uncertainty decrease, as established in the next section. Similar, if less well defined structure is seen in the Moore-Spiegel and Rössler attractors. Note in Rössler, however, there are also points which double quickly ($\tau_2 < 10$), but do not quadruple for a relatively long time ($\tau_4 > 50$). The lesson to draw from these distributions is that even in these simple systems, growth of infinitesimal uncertainties is very inhomogeneous. The inhomogeneity in operational NWP models is discussed in Smith and Gilmour (1998).

Table 1 summarizes statistics for all three chaotic systems. First compare the arithmetic averages and note that the average quadrupling-time is never twice the average doubling time. Second, note that the medians differ from the arithmetic averages indicating that the q-pling time distributions are far from Gaussian (as shown in figure 6). For the arithmetic average: in Lorenz system $\langle \tau_2 \rangle_{arith.} > \frac{1}{\Lambda_1}$ while for both the other systems $\langle \tau_2 \rangle_{arith.} < \frac{1}{\Lambda_1}$, whereas $\langle \tau_4 \rangle_{arith.} < \frac{2}{\Lambda_1}$ in all three systems. This result reflects the importance of regions of negative uncertainty growth play in the Lorenz system.

Finally, recall that Λ_1 is measured in bits per unit time, yet $\frac{1}{\Lambda_1}$ does not determine $\langle \tau_2 \rangle$. To see that this must be the case, recall that Lyapunov exponents are defined as

an averaged rate in the limit of an infinitely long fiducial trajectory which will cover the attractor uniformly. To compute $\langle\tau_2\rangle$ we take initial conditions uniformly distributed on the attractor - we may consider for example points $\mathbf{x}$ equally spaced in time along the same (infinitely long) fiducial trajectory with separation Δt. The Lyapunov exponent is then the growth rate after time Δt, averaged over each $\mathbf{x}$, with the appropriately oriented initial uncertainty. τ_q involves the same initial conditions and orientations, but the trajectory is sampled only *until* the uncertainty doubles; for some $\mathbf{x}$ this is longer than Δt, for some it is less. Even though the initial points are uniform with respect to the natural measure in both cases, the calculation of the τ_q samples the attractor differently than the calculation of the Λ_i. In short, Lyapunov exponents do not reflect predictability* except in the simplest of systems.

	Lorenz	Moore-Spiegel	Rössler
$\rho(G_1)$ in %	31.2	0.6	4.5
$\rho(G_2)$ in %	13.4	-	-
Λ_1 in [bits / unit time]	1.31	0.24	0.13
$\frac{1}{\Lambda_1}$	0.77	4.17	7.75
$\langle\tau_2\rangle_{arith.}$	1.00	1.22	6.26
$\langle\tau_2\rangle_{geom.}$	0.54	0.25	4.36
$\langle\tau_2\rangle_{median}$	0.63	0.24	5.58
$\langle\tau_4\rangle_{arith.}$	1.32	2.18	8.33
$\langle\tau_4\rangle_{geom.}$	0.87	0.49	5.78
$\langle\tau_4\rangle_{median}$	0.87	0.39	7.02

TABLE 1. Fractions of measure in G_1, G_2, maximum Lyapunov exponent, arithmetic mean, geometric mean and the median of both doubling times and quadrupling times are given for each of the three systems. In all cases the uncertainties are oriented in Lyapunov direction. Trajectories exceeding 10^7 time units were used in estimating Λ_1 to two decimal places.

5. Exact results for instantaneous uncertainty growth

While it is not possible to integrate the augmented equations analytically, exact bounds on uncertainty dynamics may still be obtained if we can place a restriction on every trajectory within a region of state space. For example, locating a region within which all instantaneous growth rates are negative [for all points and all orientations] proves that all uncertainties must decrease for as long as the fiducial trajectory remains within that region. Such regions must display enhanced predictability for finite times. In this section, we establish such regions in the Lorenz system, and related regions for all three systems. Denote by G_1 any region within which the Jacobians, J, at each point have eigenvalues with negative real parts. For non-normal J, this is not sufficient to rule out positive growth rates (see section 3), yet such regions are observed (numerically) to be dominated by decreasing uncertainty with time. If, on the other hand, each eigenvalue of $J + J^T$ is negative within a region, say G_2, then *no* instantaneous growth rate within G_2 is positive: all infinitesimal uncertainties will decrease with time for (at least) as long as the trajectory remains within G_2. In general $G_2 \in G_1 \in R^m$. General conditions for the existence of such regions are given in the next two paragraphs; explicit formulas for the surfaces shown in figures 9 are presented in Appendix B.

Initial conditions with negative real parts of the eigenvalues of J: G_1

* The Baker's Apprentice Maps (see Smith (1994a,1997)) provide a simple but dramatic illustration of a family of maps which includes members with an arbitrarily large average doubling time, while every member has a largest Lyapunov exponent greater than one.

The eigenvalues of a 3×3 matrix are the roots of the characteristic polynomial:

$$a_0\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0 \tag{22}$$

Taking $a_0 = 1$, the *Routh-Hurwitz-Criterion* (see e.g. Jetschke (1989)) states that all λ_i have *negative* real parts, if and only if T_0,T_1,. . .,T_m are all positive, where in this case

$$T_0 = a_0 > 0, \quad T_1 = a_1 > 0, \quad T_2 = \begin{vmatrix} a_1 & 1 \\ a_3 & a_2 \end{vmatrix} > 0$$

$$T_3 = \begin{vmatrix} a_1 & 1 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix} > 0 \longleftrightarrow a_3 > 0, \quad \text{because} \quad a_4, a_5 = 0 \quad \text{for} \quad m = 3$$

with $a_k = 0$ if $k > m$. Then the subset of states which have negative real parts of the eigenvalues of the Jacobian is given by:

$$G_1 = \{\mathbf{x} | T_0, \ldots, T_m > 0\} \tag{23}$$

Initial conditions with negative eigenvalues of $(J + J^T)$: G_2

For symmetric matrices like $(J + J^T)$, the λ_i are negative definite only if all the principal sub-matrices of $[-(J + J^T)]$ have positive determinants V_i, $i = 1, m$. For m=3:

$$\begin{aligned} V_1 &= -2J_{1,1} > 0 \\ V_2 &= 4J_{1,1}J_{2,2} - (J_{1,2} + J_{2,1})^2 > 0 \\ V_3 &= \det[-(J + J^T)] > 0 \end{aligned}$$

Therefore, the subset of states with $\lambda_i < 0$ of $(J + J^T)$ is given as

$$G_2 = \{\mathbf{x} | V_1, \ldots, V_m > 0\} \tag{24}$$

Figures 9 show, for each system, surfaces defining G_1 and G_2 and their relation to that attractor. Of the three, only the Lorenz system has a nonempty region G_2, within which all uncertainties decay. (The Ikeda Map also has a nonempty G_2 region; note that the relevant condition for a map is that all eigenvalues of $J^T J$ are less than one.) We now relate these surfaces of figures 9 to the observed dynamics.

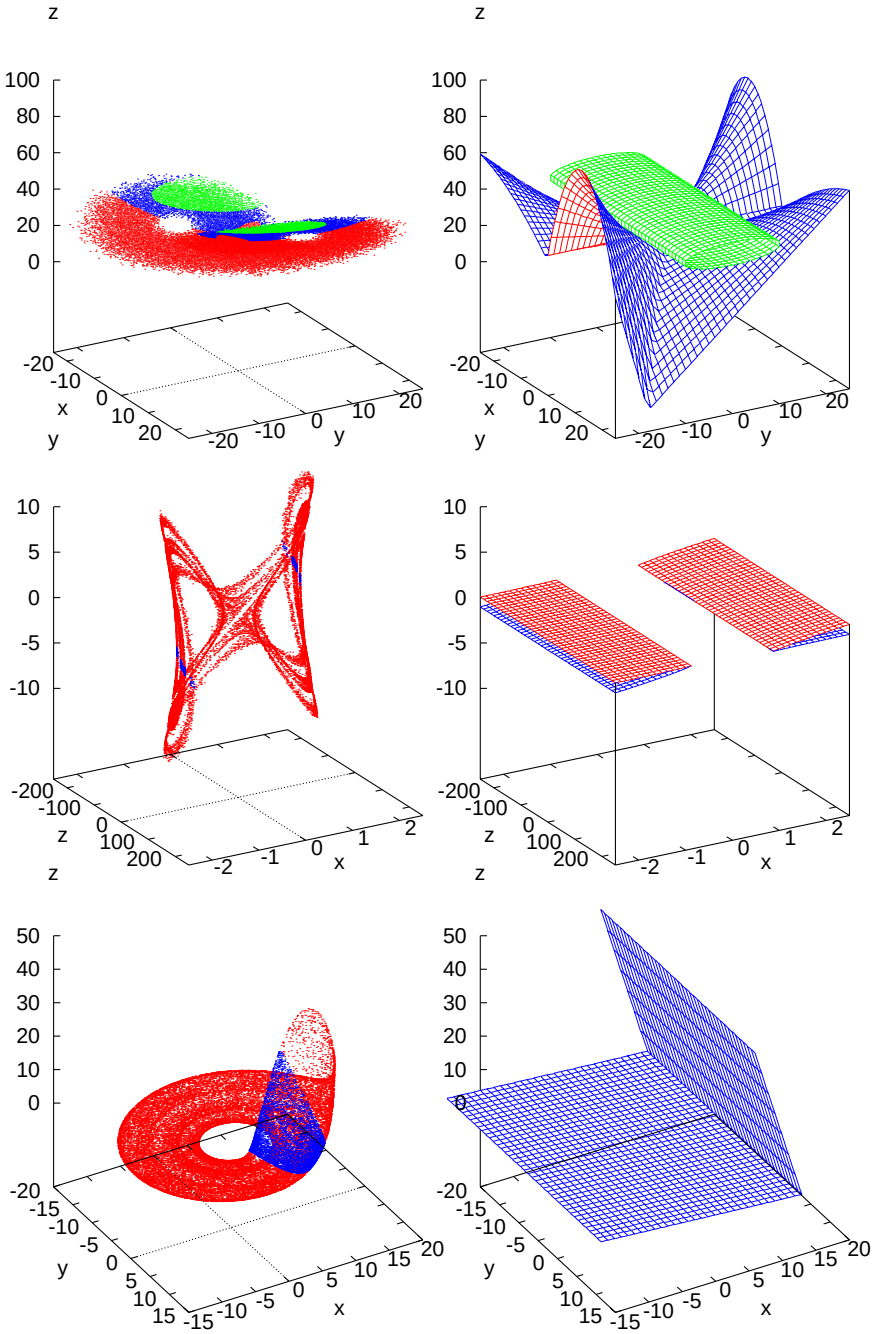

Figure 9. The panels on the right show the surfaces bounding G_1 (the blue side facing the stable, the red side facing the unstable region) and G_2 (green) in (from top to bottom) the Lorenz, the Moore-Spiegel, and the Rössler systems. The location of the attractors in each system is indicated in the corresponding panel on the left, where points are plotted as blue dots within $G_1 \setminus G_2$, as red dots outside G_1 and as green dots within G_2.

For any system, fixed points within a region G_1 give rise to simple attractors (which may, of course, coexist with strange attractors). In the Lorenz system this occurs* for small values of the parameter r; as r increases, h, the height of either of the pair† of fixed points $\mathbf{x}_{f,\pm} = (\pm x_f, \pm y_f, z_f)$ above the surface defining G_1, is

$$h = z_f - z^*(x_f, y_f). \tag{25}$$

Noting that $z_f = r - 1$ (see section 2) and using equation 23 to determine the z coordinate of the separating surface $z^*(x, y)$ given x and y explicitly (see Appendix B), we have

$$h(r, b, \sigma) = \frac{b}{\sigma^2 + \sigma}[(1 + b - \sigma)r + (3 + b + \sigma)\sigma]. \tag{26}$$

Figure 10 shows h as a function of r, for $\sigma = 10$, $b = \frac{8}{3}$. The $\mathbf{x}_{f,\pm}$ become unstable as h goes through zero at $r = \frac{470}{19}$ and the largest Lyapunov exponent becomes positive even for initial conditions near $\mathbf{x}_{f,\pm}$. The fraction of the measure of the attractor within G_1 is also shown. Note that for small negative h, Λ_1 tends to increase with increasing r (decreasing $\rho(G_1)$); this general tendency fails, of course, in any windows of r corresponding to stable periodic attractors found within this chaotic parameter regime.

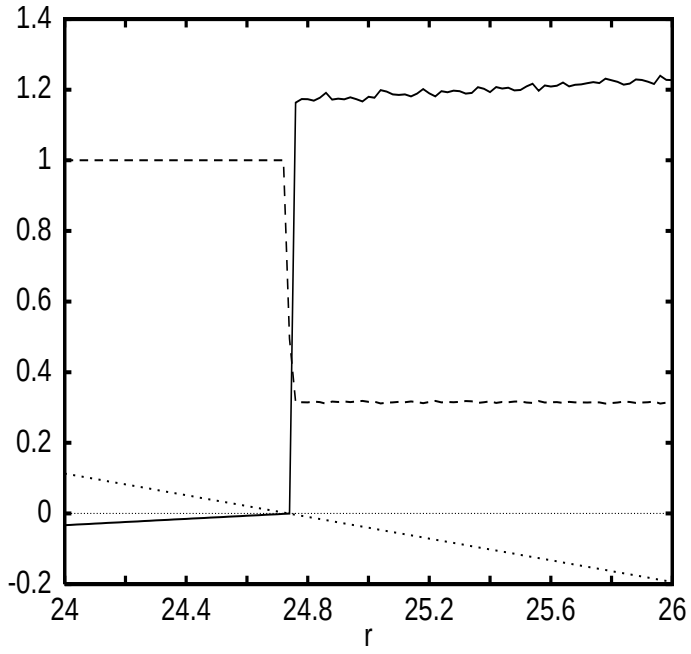

Figure 10. The variation in the Lorenz system of several statistics for initial conditions near $\mathbf{x}_{f,\pm}$ as a function of the parameter r. Namely, the fraction of measure in G_1 (long dashed) and maximum Lyapunov exponent (solid). The height of the fixed points above the surface defining G_1 is also shown (dotted). Note the changes that occur as the fixed points pass through this surface.

The fraction of time a system spends in G_1 or (G_2) is related to the fraction of the attractor in this region shown in Table 1. While this fraction reflects predictability,

* We thank T. Frisius for pointing out the additional example $r = 50, \sigma = 2, b = 0.8$

† By symmetry, these two fixed points cross the surface simultaneously.

few conclusions can be drawn from any globally averaged predictability measures. It is, nevertheless, of interest to know how often short term prediction of a chaotic system will be simple and whether return of skill is to be expected.

These exact results also support the numerically-based conjecture from figure 5 that uncertainties increase only in certain regions: it is only possible to exceed an uncertainty threshold (in this case doubling) in a region where uncertainties can increase; hence all the points in figure 5 must lie outside G_2. For smaller $q > 1$ (*i.e.* 1.05-pling times) the location of q-pling boundaries approach G_2 (figures not shown). For all initial orientations, those points lie outside G_2. For the Lyapunov direction they are also outside G_1! This leads to the conjecture that the uncertainties are closely aligned to maximum eigen-directions (or subspaces spanned by these eigen-directions) (Ziehmann-Schlumbohm *et al.* 1995); this demonstrates once again the importance of determining which orientation is most relevant in the context of ensemble forecasting. While this analytical approach is attractive, its general application is limited to low order systems (as there exists no general solution for polynomials with order larger than three), although results on subspaces may often be obtained.

6. Finite uncertainties and ensemble forecasting

While a variety of statistics for quantifying predictability have been contrasted in the preceding sections, those that invoke linearized dynamics share a common flaw: the assumption that uncertainties remain effectively infinitesimal for the time scales of interest. Clearly, as long as an uncertainty is infinitesimal it can place no limit on predictability, and as soon as it becomes finite, the linearized dynamics are indefensible, although one may verify the extent to which nonlinear terms dominate the dynamics by explicit calculation on a case by case basis. In order to determine the extent to which these infinitesimal τ_q reflect true limits to predictability, we contrast these time scales with those determined using a perfect ensemble of finite radius, evolved under a perfect model. Within the perfect model/perfect ensemble scenario (Smith 1996), forecasts are accountable, meaning that deviations from the forecast probability density function decrease towards zero as the sample size increases. These forecasts provide the standard against which we measure other ensemble forecasts, as well as the utility of the time-scales determined from the linearized dynamics. We first discuss the formation of perfect ensembles, interpret one example, shown in figure 11, then turn to the efficacy of the τ_q as an indicator of the growth of perfect ensembles, and finally contrast the performance of imperfect ensembles against this standard.

The true state of the system, $\mathbf{x}$ will lie on the attractor, while an observation, $\mathbf{x}_{obs}$, need not. Given $\mathbf{x}_{obs}$ and knowledge of the type of observational uncertainty, we generate a perfect ensemble by collecting analogues from a long integration of the model from those times when the trajectory returns near $\mathbf{x}$, where "near" is defined by the statistical form of the observational uncertainty. The ensemble is perfect in that its members are not only near $\mathbf{x}$, but are also on the attractor. For simplicity, we will assume that the only observational uncertainty is due to quantization noise, $\mathbf{x}$ then corresponds to a particular (truncated) value $\mathbf{x}_{obs}$ and the perfect ensembles are drawn from all the other points on the attractor which correspond to the *same* value of $\mathbf{x}_{obs}$. Perfect ensembles are, of course, also well-defined for more complex observational uncertainties, the key point is that given a particular value of $\mathbf{x}_{obs}$ the members of a perfect ensemble are realizable states of the system weighted by their likelihood as candidates for $\mathbf{x}$. In the case of quantization noise, each point in the ensemble has equal weight; the relative weights under other observational noises can easily be determined. Regardless of the

details of the measurement uncertainty, the essential idea here is that if $\mathbf{x}$ lies on a manifold (or attractor) with dimension less than that of the model-state space, then the formation of perfect ensembles requires more knowledge than a complete specification of the measurement uncertainty: given this specification we can compute the probability of an observation given the true state, but *not* the probability of a point being the true state given the observation. This follows from the fact that the true state must be on the attractor, as well as consistent with the observation (Smith 1998).

The evolution of one perfect ensemble from the Lorenz system is shown in figure 11, where the observed distribution of values of the variable x is shown as a function of time. At $t = 0$ each of the 512 members of this ensemble are within the same 6-bit box in state space: each of these points corresponds to exactly the same finite precision observation. The distribution initially spreads out, but shows true return of skill both for large positive and for large negative values of x, corresponding to times when the ensemble is in G_1 or G_2. While the ensemble quickly develops a spread comparable with the diameter of the attractor, it is obvious that the information content in this distribution decays at a much slower rate than is suggested by it standard deviation. At $t = 3.5$, it is clear that system will have either a large positive or large negative value of x, yet the optimal root-mean-square (RMS) forecast at this time is near $|x| = 0$; this is a fundamental shortcoming of the RMS forecast error as a cost function for nonlinear systems. In the limit of large t, the distribution will approach the climatological distribution of x, that is, the projection of the natural measure onto the variable of interest. Indeed an ultimate limit of predictability is the time required for a perfect ensemble to become statistically indistinguishable from the climatology; this limit will be a function of the initial condition, the observational accuracy, and the size of the ensemble (Smith 1996). In the current paper, we are interested in earlier stages of the loss of predictability; for instance, we may compute the time required for the average deviation of the ensemble to increase by a factor of two, and contrast this with the uncertainty doubling time. In fact there are two questions here, first whether to compute the chosen statistic in full state space, or only in a particular variable. And second, whether to use the fiducial trajectory (control) or the ensemble average as the reference trajectory when computing the average deviation. The last question is relevant since in practice we do not know the true state (the control); table 2 shows both.

For each of 4096 initial conditions chosen at random on the attractors, a perfect ensemble was constructed with $N_{ens} = 256$ analogues within a distance of $\epsilon = 2^{-7}$ corresponding to length scale just under 1% of the "diameter" of the attractor*. The ensemble q-plying time was defined as the smallest time at which the average deviation, α, of the ensemble exceeds q times its inital value α_0. In addition, the first time at which any member of the ensemble exceeds a distance of q times its initial distance from the reference trajectory was recorded as the "fastest member" time.

* For each of the attractors considered in this section, the variables were rescaled so that the attractor had extent approximately one in each coordinate direction.

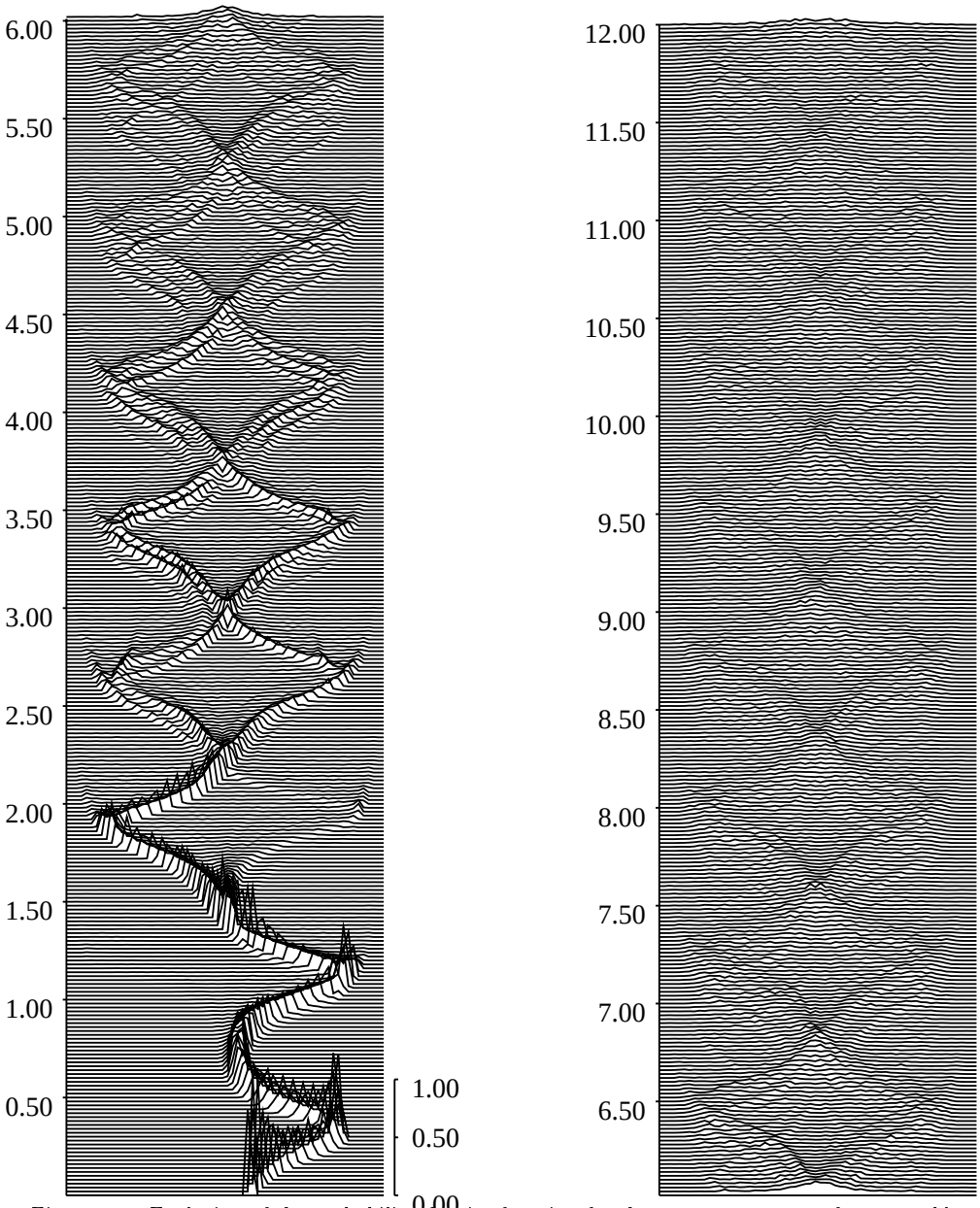

Figure 11. Evolution of the probability density function for the variable x in a perfect ensemble under the Lorenz equations. Time is plotted in the vertical; the right panel follows from the top of the left. As time increases (upwards), the initially sharp distribution at $t = 0$ initially decays, but then shows true return of skill (at $t \approx 0.4$) at the time when the ensemble crosses the G_1 surface. Although the ensemble soon has members at each extreme of the attractor ($t \approx 2.75$), it clearly differs from the asymptotic distribution at the top of the left panel ($t = 6.0$); this difference is still statistically significant at the top of the right panel ($t = 12.0$).

The q-plying times for infinitesimals in several different orientations are contrasted with that of the perfect ensemble in Tables 2 and 3. Defining the q-pling time in the singular vector orientation requires the choice of an optimisation time which should, of course, vary with the value of q. To determine a relevant time scale for each q, we first compute the average τ_q for the Lyapunov orientation, the singular vectors $\mathbf{v}_1$ are then computed using this average τ_q as the optimization time. We also consider the orientation in which an infinitesimal perturbation will first increase by a factor q, and define this as the nightmare (or worst case) orientation, $\mathbf{n}_1$. The shortest possible q-pling time for infinitesimals at $\mathbf{x}_0$ may be found by computing the singular values of $M(\mathbf{x}_0, \Delta t)$ as a function of Δt as in Greene and Kim (1987). The nightmare direction at $\mathbf{x}_0$ is simply the $\mathbf{v}_1$ corresponding to this time. For $q = 8$, the average τ_q for the Lyapunov, first singular, and nightmare orientations are 1.74, 1.10 and 1.05, respectively.

control	Lyapunov Vector	Singular Vector($[\tau_2]_{LV}$)	Nightmare
Deviation	0.765	0.630	0.610
Fastest Member	0.615	0.573	0.950
mean			
Deviation	0.772	0.627	0.610
Fastest Member	0.572	0.534	0.929
control	Lyapunov Vector	Singular Vector($[\tau_8]_{LV}$)	Nightmare
Deviation	0.802	0.887	0.859
Fastest Member	0.852	0.889	0.979
mean			
Deviation	0.804	0.887	0.860
Fastest Member	0.797	0.839	0.926

TABLE 2. Correlations between τ_q times from a perfect ensemble and those from infinitesimal uncertainties oriented in Lyapunov, Singular and Nightmare directions in the Lorenz system. The upper panel is for $q = 2$, the lower $q = 8$. Deviations from the control and from the ensemble mean are considered.

The linear correlations were computed between infinitesimal τ_q and ensemble q-plying times for $q = 2, 4, 8, 16$ and 32. The results in the Lorenz system for both $q = 2$ and $q = 8$ are given in Table 2; repeating the experiment yields correlations within 0.02 of those shown, and usually within 0.01. The τ_q of the infinitesimal singular vectors consistently yield higher correlations with τ_q of the average deviation of the finite uncertainties in all cases *except* $q = 2$; this holds both for deviations from the control and for deviations from the ensemble mean. If, however, we are interested in the worst case, namely an estimate of time required for the fastest member of the ensemble to increase by a factor of q, the nightmare vector best captures this behaviour. When estimating the growth of the average deviation, similar results are obtained whether we calculate the distances to the control or the mean, however for the worst case the difference is crucial. This is not surprising because the tangent model is set up using the trajectory through the control and not the ensemble mean (which need not correspond to a realizable trajectory).

control	Lyapunov Vector	Singular Vector($[\tau_2]_{LV}$)	Nightmare
Deviation	0.413	0.229	0.441
Fastest Member	0.286	0.109	0.618
control	Lyapunov Vector	Singular Vector($[\tau_8]_{LV}$)	Nightmare
Deviation	0.147	0.431	0.478
Fastest Member	0.104	0.719	0.869

TABLE 3. Correlations in the Moore-Spiegel system between τ_q times from a perfect ensemble and those of infinitesimal uncertainties oriented in Lyapunov, Singular and Nightmare directions. Deviations are from the control. The upper panel is for $q = 2$, the lower for $q = 8$.

Given the general arguments against generalizing from the behaviour of infinitesi-

mal uncertainties, the correlations in table 2 might be considered surprisingly large. To investigate further we contrast the results from the Lorenz system with those of the Moore-Spiegel system (see table 3). The correlations in the Moore-Spiegel system are much lower than those in the Lorenz system. The high correlations of the Lorenz system provide a misleading indication of the general quality of infinitesimal τ_q, and probably result from the extremely simple dynamics on this attractor. This leads us to distinguish two distinct limiting factors: the breakdown of the linear regime due to nonlinear dynamics at small (finite) amplitudes and an ultimate limit of uncertainty growth due to the finite size of the attractors.

Given an attractor of diameter one, we attempt to avoid finite diameter effects by considering finite perturbations of magnitude ϵ, such that $q\epsilon < 0.5$. With this upper bound on ϵ, we wish to take ϵ large enough so that the local nonlinearities of the flow play a role in the dynamics. Note that these length scales are in the state space of the dynamical system, not physical space; we are concerned with finite amplitude effects where $0 < |\boldsymbol{\epsilon}| << 1$. For the Lorenz system these two conditions conflict: examining contours of the correlation between infinitesimal q-pling times and finite q-pling time in the Lyapunov orientation as a function of q and ϵ (not shown), we see that the linear approximation remains accurate* all the way out to length scales comparable with the diameter of the attractor. In general this is not the case; the Moore-Spiegel system shows a range of length-scales and q-pling times at which the nonlinearities break the correlation before the size of the attractor is reached. We choose $\epsilon = 2^{-7}$ as representative of these length scales, and stress that the level of observational accuracy required to exploit the infinitesimal time scales will depend on the particular system and cannot be determined *a priori* .

Next we turn to the performance of constrained ensembles with finite amplitude perturbations. Figure 12 reflects how well the perfect ensemble distributions are captured by small operational ensembles, where initial conditions "on the attractor" are not available. For each of 512 initial conditions, several ensembles consisting of the true initial condition (control) and two perturbations of equal and opposite magnitude ϵ are evolved and their q-pling times are recorded. In addition, the shortest time at which any member of the ensemble exceeds $q\alpha_0$ is recorded. The figure shows the correlation between the perfect ensemble and each of the four constrained ensembles; in each of the latter, finite perturbations are constrained to lie with a certain orientation. In order to remove any effects of initial conditions "off the attractor", the correlation between the full 256 member perfect ensemble and a two-member perfect ensemble is also shown.

The most noticeable aspect of figure 12 is how large *all* the correlations remain in the Lorenz case. We first consider the estimated growth of the average deviation. For small q, the two-member perfect ensembles consistently perform well in both systems, indicating the importance of being "on the attractor" in the early stages of uncertainty growth. In the Moore-Spiegel system, nightmare ensembles outperform the two-member perfect ensemble for $q \geq 4$, as do the singular vector ensembles, which are of similar quality for large q. The Lyapunov ensembles show the smallest correlations. In the Lorenz system, the nightmare and singular vector ensembles perform worst for small q, while for $q \geq 4$ they dominate the Lyapunov ensemble, although all correlations exceed 0.7 in this case. In terms of indicating a warning of "worst forecast bust," the nightmare ensemble is generally the best. The singular vector ensembles generally outperform the Lyapunov ensembles, except at small q.

Lastly, we explored the effect of increasing the size of small perfect ensembles, with ensembles of 16 members, we consistently get correlations in excess of 0.9 for the average

* Specifically, the correlation exceeds 0.90 for $q\epsilon = 0.5$

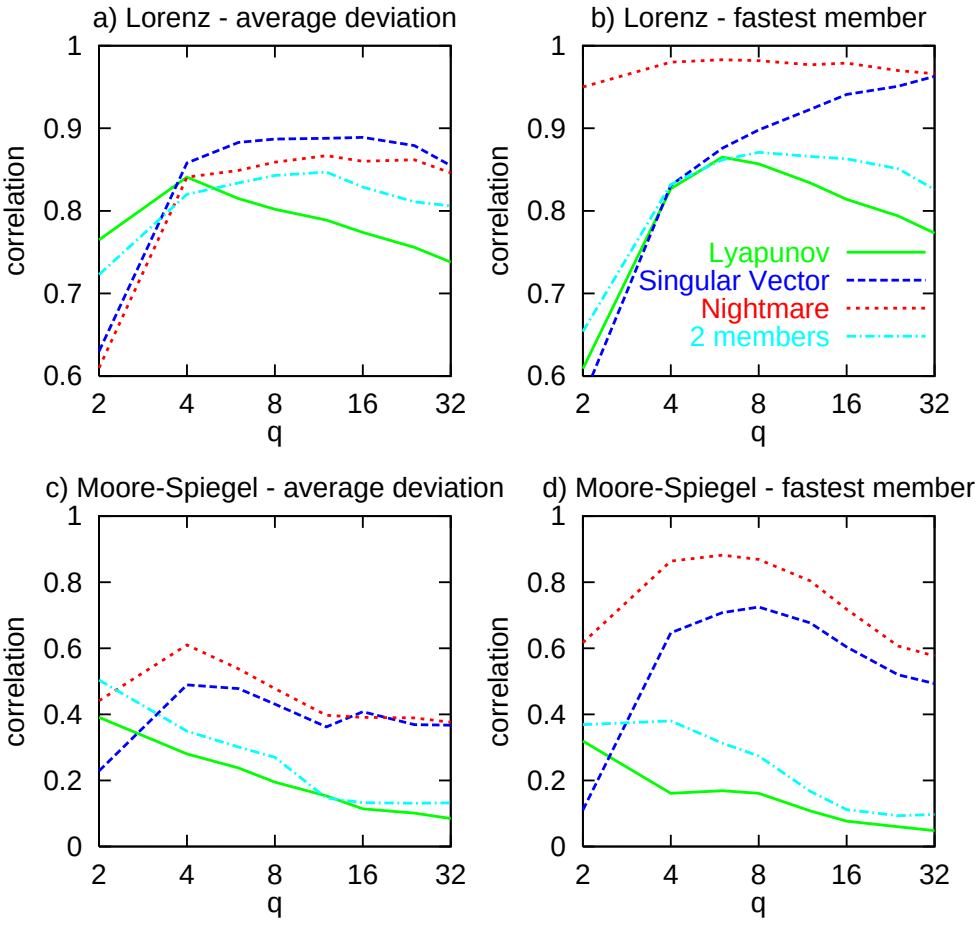

Figure 12. Panels (a) and (c) show the correlations between the time required for the average deviation to increase by a factor of q in a perfect ensemble with that in each of the 2-member ensembles. The ability of an ensemble to reflect the "worse case" scenario is shown in panels (b) and (d), similar graphs which reflect the correlation between the time required for a member of the ensemble to exceed a distance $q\epsilon_0$ in the each two member ensemble with that of the perfect ensemble. In each case, the average deviation and fastest member evolution of finite perturbations in the Lyapunov (solid), singular vector (dash), nightmare (dot) directions, and a 2-member perfect ensemble (dash-dot) are shown.

deviation, and between 0.8 and 0.9 for the fastest member over the range $2 < q < 16$ for the Moore-Spiegel system. This is a significant improvement over the performance of the two-member ensembles shown in Figure 12.

In the context of NWP, it is commonly stated that given the current observational error, the linearized dynamics hold for about two days (see, however, Smith and Gilmour (1998), Gilmour(1998)); we note that for all the low-dimensional systems we have considered this statement is, at best, misleading: the time scale varies widely with initial condition on the attractor, usually in an organized manner. In such cases one cannot usefully employ *one* time scale over which linearity is assumed to hold, even for a fixed observational uncertainty.

Observational uncertainties are always finite; in this section we have contrasted

infinitesimal time scales with those determined from finite precision observations. The correlation between the τ_q derived from the finite uncertainties estimates in the ensemble and that of the infinitesimals can vanish suddenly as observational uncertainty increases. One might respond by including higher order terms in a local analysis, but the macroscopic structures evident in Figure 11 suggest that no local analysis can be expected to hold once the ensemble develops macroscopic structure. Smith (1996) considers the interplay of these effects with long term prediction, the results above indicate that ensembles are equally necessary to quantify short term predictability, in the absence of all but infinitesimal observation error.

7. Discussion of related results and conclusions

Given a perfect model, predictability is limited only by the growth of the uncertainty in the initial condition. If this initial uncertainty is sufficiently small, then its evolution will be reflected in the linearized equations of motion, and may be quantified through a singular value decomposition of the linear propagator, as argued by Lorenz (1965) over 30 years ago. Lorenz quantified the *variability of predictability* with initial condition in a 28-variable atmospheric model using what are now called "finite-time Lyapunov exponents"; other statistics with this aim include the "non-uniformity factor" of J.S. Nicolis *et al.* (1983) and the difference between the most probable error and average error as considered by Benzi and Carnevale (1989). In the present paper, we have shown that this variability severely restricts the use of Lyapunov exponents as a measure of predictability, this restriction is enhanced by the fact that both finite-time and global Lyapunov exponents are defined as average rates, not average times. Explicit time-scales for uncertainty growth times have been estimated and used to visualize the organized variation of predictability in state-space. They provide a useful measure for the growth of either infinitesimal or finite uncertainties. The organization of predictability has previously been quantified in two-dimensional maps through variations in local divergence rates by Haubs and Haken (1985), while Nese (1989) estimated the average of the local divergence rate numerically at several locations on the attractor of the Lorenz system, noting regions of convergence. We believe estimated times provide a better measure of predictability than effective rates, in part because estimating an *effective* rate requires selecting a time scale *a priori* .

In Section 5 bounds are placed on local growth rates analytically, suggesting a method to prove the existence of finite volumes of state-space within which *all* perturbations *decrease* with time for a finite period. Following Farrell (1994) and Trefethen *et al.* (1993), we stress the importance of the non-normality of the Jacobian matrix (or linear propagator) in determining whether uncertainty will grow or decay in a chaotic system; regions of certain uncertainty decrease are proven to exist in the Lorenz system and in the Ikeda map, while numerical evidence indicates that there are regions in many chaotic systems where uncertainties with the most likely orientations decrease with time. Other measures of finite time stability include the Lorenz index advanced by Mukougawa *et al.* (1991); Figure 7 of Mukougawa *et al.* (1991) provides clear numerical evidence in support of the analytical results of section 5. Also of interest are the "effective Lyapunov exponents" and associated graphics computed by Doerner *et al.* (1991) which, like the Lorenz index, are a function of a particular prediction time-scale.

In addition to the references above, this work has been significantly influenced by the predictability studies of Trevisan (1993), Trevisan and Legnani (1995), Palmer (1988, 1993a), Palmer *et al.* (1992), Eckhardt and Yao (1993), C. Nicolis *et al.* (1992, 1995), Yoden and Nomura (1993), Houtekamer and Derome (1994), and eventually the 1995 ECMWF workshop on predictability 1995. Some of these papers (for example Trevisan

(1993) and Houtekamer and Derome (1994)) report *oscillations in "error growth curves"* for low-dimensional systems . We replicate these observations in figure 7 and discuss the extent to which this is simply a question of taking an arithmetic mean or a geometric mean; we then employ the analytical results of section 5 to distinguish what is to be expected from a chaotic system from those behaviours particular to the 1963 Lorenz system, and within this model, behaviours further restricted to the "standard" parameter values. The symmetry of the Lorenz attractor, and the roughly periodic sampling of regions in which uncertainties decrease with time, along with the extremely good approximation of macroscopic uncertainty dynamics provided by the linearization about a fiducial trajectory, combine to make to this Lorenz system a weak straw-man for the evaluation of forecast schemes.

In this article we have demonstrated the variation of uncertainty dynamics with respect to initial condition, orientation and magnitude in state space. These results are of use not only in the first principles models discussed above, but also in the data based models commonly constructed from observations of nonlinear dynamical systems (Broomhead and Lowe 1988, Casdagli 1989, Farmer and Sidorowich 1987, Mees 1993, Smith 1992, Sugihara and May 1990). The need for clarifying the true sources limiting the prediction of nonlinear systems has been argued by May (1996): while the "non-uniformity" of uncertainty growth has been recognized for some time, it is still common to find statements to the effect that "an estimate of the mean e-folding time is provided by the inverse of the sum of the positive Lyapunov exponents." The theoretical arguments of Sections 5 and 6 show that this is not true in general. The Baker's Apprentice Maps in Smith (1994a,1997) provide a family of systems which allow arbitrarily large average uncertainty doubling times, yet every member has $\Lambda_1 \geq 1$, and hence $\frac{1}{\Lambda_1} \leq 1$. The contents of Table 1 compare the inverse of the largest Lyapunov-exponent and average doubling time for several chaotic flows. Tables 2 and 3 and figures 12 show additional results relating local measures based on the finite time linearized dynamics with the behaviour of small but finite initial uncertainties.

Linear intuitions suggest that forecast errors will steadily increase with lead time; the *decrease* in forecast error with increasing forecast time was dubbed the "return of skill" problem by Anderson and van den Dool (1994). While the statistical significance of such results in meteorology has been questioned by Anderson and van den Dool (1994), we stress that return of skill is not unexpected in nonlinear systems (see Tong and Moeanaaddin (1988)). Return of skill is to be expected in any system which explores a region of state space in which all uncertainties decrease with time. Numerical evidence for a true return of skill is observed not only in the Lorenz system*, but also in a variety of other chaotic systems, including the Moore-Spiegel and Rössler systems, again in the regions of state space suggested by the exact results of section 5. The extent to which these observations are due to the simple structure of these systems is unknown; however we note that some evidence for return of skill has been noted in forecasts of the thermally driven rotating fluid annulus (Smith 1992, Smith 1994, Read *et al.* 1991), which has a somewhat more complex state space structure. It is not known whether "return of skill" plays a significant role in forecasting the Earth's atmosphere, in large part because the observational record is short relative to the recurrence time of the system. This may always be the case, as the recurrence time of the system has recently been estimated to be vast compared with the lifetime of the system (indeed, even when compared with the age of the Universe (van den Dool 1994)), another indication of the importance of

* As stressed by Palmer (1993,1993a), the Lorenz system also shows an artificial return of skill (*e.g.* $t \approx 3.0$ in figure 11) due to the symmetry of the attractor, which we agree is "manifestly spurious."

physical analogue laboratory experiments.

The central focus of this paper has been to demonstrate that "the" relevant effective growth rate of a small initial uncertainty depends strongly on both its location and orientation in state space, its magnitude, and even the verification time. Our results hold implications for ensemble selection for Numerical Weather Prediction. Operational ensemble forecast programs often selectively consider uncertainty growth in particular orientations; two competing methods contrast the subspace defined by breeding vectors (see Toth and Kalnay (1993) and references thereof) with that defined by singular vectors (see Palmer (1993) and references thereof). In a perfect model context, the singular vector subspace corresponds to fixed optimization time singular vectors discussed in Section 2, while the breeding vectors are related to the Lyapunov vectors; it is crucial to remember that operational breeding vectors differ from Lyapunov vectors in that operational breeding vectors are neither infinitesimal nor determined in a perfect model environment, hence operational breeding vectors incorporate information on the analysis error implicitly. That being said, the results in section 6 suggest an advantage of singular vector based ensembles over Lyapunov vector based ensembles. We note that allowing the optimization time to decrease in regions of particularly rapid growth can improve the results further. A major source of interest in ensemble prediction lies, of course, in the medium range beyond the time scales for which the linear approximation is relevant. The method of ensemble formation employed should reflect the goals of the forecaster (*e.g.* to obtain a bounding forecast, or an accurate probability density function).

Advances in understanding the dynamics of uncertainty will help to clarify the crucial role the magnitude of the initial perturbations plays in constructing ensembles. We hope these results will serve as motivation for the analysis of a hierarchy of increasingly complex mathematical systems and physical laboratory analogues, in both perfect and imperfect model regimes, to be executed in parallel with current studies on full-blown numerical weather forecasting models.

Appendix A

Details of the Error Growth Experiments

The numerical integrations discussed in this paper were computed under a fourth order Runge-Kutta scheme (Press *et al.* 1987) with a fixed time step (0.005 for the Lorenz and Moore-Spiegel systems, 0.004 for the Rössler system)

For the uncertainty growth, and especially the doubling experiments, initial points were chosen by integrating for a randomly chosen time ($\tau_{decorr} + \gamma\tau$) after a given measurement, or by integration of an independent trajectory. Such an approach is required in order to avoid bias or correlations in the test points. If, for example, the next initial condition for an uncertainty doubling experiment is chosen as the location at which the previous test point doubled, one will tend to get initial conditions with the distribution reflected by the red points in figure 5, which is *not* uniform with respect to the natural measure. There is evidence of this type of bias in the literature.

Appendix B

Explicit formulas for the surfaces

Explicit formulas for the surfaces shown in figure 9 are given in this Appendix; the derivation of these results can be found in Ziehmann-Schlumbohm (1994).

In the Lorenz system (2) the coefficients of the characteristic polynomial are $a_0 = 1$, $a_1 = \sigma + 1 + b$, $a_2 = \sigma(1 + b - r + z) + b + x^2$ and $a_3 = \sigma(1 - rb + x^2 + bz + xy)$. Given the standard parameters $\sigma, b, r > 0$, we have $T_0 = a_0 = 1 > 0$ and $T_1 = a_1 = \sigma + 1 + b > 0$. Therefore we only need to find the conditions for $T_2 = a_1 a_2 - a_3$ and a_3 to be positive.

$$\{\mathbf{x}|T_2 > 0\} = \{\mathbf{x}|z > z_1(x, y) = b_1 + b_2 x^2 + b_3 xy\} \tag{B.1}$$

where

$$\begin{aligned} b_1 &= \frac{\sigma^2 r + \sigma r - \sigma^2 - \sigma^2 b - 2\sigma b - b - \sigma b^2 - b^2 - \sigma}{\sigma^2 + \sigma} \\ b_2 &= -\frac{1 + b}{\sigma^2 + \sigma} \\ b_3 &= \frac{1}{1 + \sigma} \end{aligned}$$

$$\{\mathbf{x}|a_3 > 0\} = \{\mathbf{x}|z > z_2(x, y) = c_1 + c_2 x^2 + c_3 xy\} \tag{B.2}$$

with $c_1 = r - 1$, $c_2 = -\frac{1}{b}$, $c_3 = -\frac{1}{b}$ leading to

$$G_1 = \{\mathbf{x}|T_0, T_1, T_2, a_3 > 0\} = \{\mathbf{x}|z > z^\star = \max(z_1, z_2)\} \tag{B.3}$$

The determinants of the three principal sub-matrices of $[-(J + J^T)]$ are

$$\begin{aligned} V_1 &= 2\sigma \\ V_2 &= 4\sigma - (\sigma + r - z)^2 \\ V_3 & \quad 2b[4\sigma - (\sigma + r - z)^2] - 2y^2. \end{aligned}$$

All initial conditions $\mathbf{x}$ within an elliptic tube build the subset of states with negative eigenvalues of $(J + J^T)$

$$G_2 = \{\mathbf{x}|V_1, V_2, V_3 > 0\} = \{\mathbf{x}|\frac{(z - (r + \sigma))^2}{4\sigma} + \frac{y^2}{4b\sigma} < 1\}. \tag{B.4}$$

The Moore-Spiegel equations

$$\begin{aligned} \frac{dx}{dt} &= y \\ \frac{dy}{dt} &= z \\ \frac{dz}{dt} &= -z - (t - r + rx^2)y - tx \end{aligned} \tag{B.5}$$

describe the motion of a parcel of ionized gas in the atmosphere of a star. With $t = 26$ and $r = 100$ the system is chaotic (Moore and Spiegel (1966)). Its Jacobian is

$$J(\mathbf{x}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -(t + 2rxy) & -(t - r - rx^2) & -1 \end{pmatrix} \tag{B.6}$$

and leads to

$$G_1 = \{\mathbf{x}| -\frac{t}{2rx} < y < \frac{x^2-1}{2x} \text{ if } x > 0 \wedge \frac{x^2-1}{2x} < y < -\frac{t}{2rx} \text{ if } x < 0\} \tag{B.7}$$

and from

$$J(\mathbf{x}) + J^T(\mathbf{x}) = \begin{pmatrix} 0 & 1 & -(t+2rxy) \\ 1 & 0 & 1-(t-r-rx^2) \\ -(t+2rxy) & 1-(t-r-rx^2) & -2 \end{pmatrix} \tag{B.8}$$

follows immediately $G_2 = \emptyset$.

The Rössler system (Rössler 1976)

$$\begin{aligned} \frac{dx}{dt} &= -y - z \\ \frac{dy}{dt} &= x + ay \\ \frac{dz}{dt} &= b + zx - cz \end{aligned} \tag{B.9}$$

is also chaotic with the parameters $a = 0.15$, $b = 0.20$ and $c = 10.0$. Its Jacobian

$$J(\mathbf{x}) = \begin{pmatrix} 0 & -1 & -1 \\ 1 & a & 0 \\ z & o & x-c \end{pmatrix} \tag{B.10}$$

leads to

$$G_1 = \{\mathbf{x}|x < c - a \wedge \frac{ax^2}{c-x} - \frac{(2ac-a^2)x}{c-x} + \frac{ac^2+a-a^2c}{c-x} < z < \frac{c-x}{a} \tag{B.11}$$

so, as in the Moore-Spiegel-system, $G_2 = \emptyset$.

Acknowledgement

LAS is supported by a Senior Research Fellowship from Pembroke College, Oxford and gratefully acknowledges additional support as a Guest Professorship at the University of Potsdam made posible by the INK "Kognitive Komplexität" (DFG). We have had particulary useful disucssions with A. Fowler, I. Gilmour, J. Hansen, J. Kurths, P. McSharry, T. Palmer and Z. Toth, and are also indebted to an anonymous reviewer for numerous constructive criticisms on a number of earlier drafts. Completion of this work was made possible by support from Predictability DRI grant N00014-99-1-0056.

References

Abarbanel, H. D. I.,Brown, R. and Kennel, M. B.	1991	Variation of Lyapunov exponents on a strange attractor. *J. of Nonlinear Sci.*, **1,** 175–199
Anderson, J. L. and van den Dool, H. M.	1994	Skill and return of skill in dynamic extended-range forecasts. *Monthly Weather Review*, **122,** 507–516
Barkmeijer, J.	1996	Constructing fast-growing perturbations for the nonlinear regime. *J. Atmos. Sci.*, **53,** 2838–2851
Benzi, R. and Carnevale, G. F.	1989	A possible measure of local predictability. *J. Atmos. Sci.*, **46,** 3595–3598
Broomhead, D. S. and Lowe, D.	1988	Multivariable functional interpolation and adaptive networks. *J. Complex Systems*, **2,** 321–355
Buizza, R.	1996	Optimal perturbation time evolution and sensitivity of ensemble prediction to perturbation amplitude. *Q.J.R. Meteorol. Soc.*, **121,** 1705–1738
Buizza, R. and Palmer, T. N.	1995	The singular-vector structure of the Atmospheric Global Circulation. *J. Atmos. Sci.*, **52,** 1434–1455
Casdagli, M.	1989	Nonlinear prediction of chaotic time series. *Physica D*, **35,** 335–356
Doerner,R., Hübinger, B., Martiensen, W., Grossmann, S. and Thomae, S.	1991	Predictability portraits for chaotic motions. *Chaos, Solitons and Fractals*, **1,** 553–571
Eckhardt, B. and Yao, D.	1993	Local Lyapunov exponents in chaotic systems. *Physica D*, **65,** 100–108
ECMWF	1995	*Predictability*, Seminar Proceedings, Shinfield Park, Reading RG2 9AX, UK, 4 - 8 September 1995.
Farmer, J. D. and Sidorowich, J.	1987	Predicting chaotic time series. *Phys. Rev. Lett.*, **59,** 845–848
Farrell, B. F.	1994	*Evolution and revolution in cyclogenesis theory.*, In "The Life Cycles of Extratropical Cyclones" Vol. 3, pages 101–110
Farrell, B. F.	1990	Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci.*, **47,** 2409–2416
Gilmour, I.	1998	Nonlinear model evaluation: $\iota-$shadowing, probabilisitic prediction and weather forecasting D. Phil. thesis, University of Oxford (submitted)
Greene, J. and Kim, J.-S.	1987	The calculation of Lyapunov Spectra. *Physica* **D 24,** 213–225
Haubs, G. and Haken, H.	1985	Quantities describing local properties of chaotic attractors. *Zeitschrift für Physik B*, **59,** 459–468
Houtekamer, P. L. and Derome, J.	1994	Prediction experiments with two-member ensembles. *Monthly Weather Review*, **122,** 2179–2191
Ikeda, K.	1979	Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system. *Opt. Commun.*, **30,** 257
Jetschke, G.	1989	*Mathematik der Selbstorganisation* , Friedr. Vieweg & Sohn, Braunschweig, Wiesbaden
Kostelich, E. and Yorke, J.	1990	Noise reduction: Finding the simplest dynamical system consistent with the data. *Physica D*, **41,** 183–196
Kurths, J. and H. Herzel, H.	1987	An attractor in a solar time series. *Physica D*, **25,** 165–172
Lacarra, J.-F. and Talagrand, O.	1988	Short range evolution of small perturbations in a barotropic model. *Tellus*, **40 A,** 81–95
Lorenz, E. N.	1963	Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20,** 130–141
Lorenz, E. N.	1965	A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17,** 321–333
May, R.	1996	Necessity and chance. *Bulletin A. Math. Soc. Journal*, **32(3),** 291–308

Mees, A. I. 1993 Parsimonious dynamical reconstructions. *Int. J Bif. and Chaos*, **3,** 669–675

Molteni, F and Palmer, T.N. 1993 Predictability and finite-time instability of the northern winter circulation. *Q.J.R.Meteorol.Soc.*, **119,** 269–298

Moore, D. W. and Spiegel, E. A. 1966 A thermally excited nonlinear oscillator. *Astrophys. J.* , **143(3),** 871–887

Mukougawa, H., Kimoto, M. and Yoden, S. 1991 A relationship between local error growth and quasi-stationary states. *J. Atmos. Sci.*, **48,** 1231–1237

Nese, J. M. 1989 Quantifying local predictability in phase space. *Physica D*, **35,** 237–250

Nicolis, C. 1992 Probabilistic aspects of error growth in atmospheric dynamics. *Q. J. R. Meteorol. Soc*, **118,** 553–568

Nicolis, C., Vannitsem, S. and Royer, J.-F. 1995 Short-range predictability of the atmosphere: mechanism for superexponential error growth. *Q. J. R. Met. Soc.*, **121,** 705–722

Nicolis, J.S., Meyer-Kress, G. and Haubs, G. 1983 Non- uniform chaotic dynamics with implications to information processing. *Zeitschrift für Naturforschung*, **38 a,** 1157–1169

Orr, W. 1907 Stability of instability of the steady motions of a perfect fluid. *Proc. Roy. Irish Acad.*, **A27,** 9–138

Oseledec, V. I. 1968 A multiplicative ergodic theorem. *Trans. of the Moscow Math. Soc.*, **19,** 197–231

Ott, E. 1994 *Chaos in Dynamical Systems*, CUP, Cambridge

Palmer, T. N. 1988 Medium and extended range predictability and stability of the Pacific/North American mode. *Q. J. R. Met. Soc.*, **114,** 691–713

Palmer, T. N. 1993 Extended-range atmospheric prediction and the Lorenz model. *Bulletin A. Meteorol. Soc.*, **74,** 49–64

Palmer, T. N. 1993a A nonlinear dynamical perspective on climate change. *Weather*, **48,** 314–326

Palmer, T.N., Molteni, F., Mureau, R., Buizza, R., Chapelet, P. and Tribbia, J. 1992 Ensemble prediction. *Research Department Tech. Memo. No. 188, 45 pp.* (Available from Director, ECMWF, Shinfield Park, Reading, Berkshire, England, RG29AX

Papoulis, A. 1991 *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill International Editions, New York.

Parker, T. S. and Chua L. O. 1989 *Practical numerical algorithms for chaotic systems*, Springer-Verlag, Berlin, Heidelberg, New-York

Press, W. H., Flannery B. P., Teukolsky S. A. and Vetterling W. T. 1987 *Numerical Recipes*, Cambridge University Press, Cambridge

Pires, C., Vautard, R. and Talagrand,O. 1996 On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus*, **48 A**, 96–121

Read, P., Bell, M. J., Johnson, D. W. and Small, R. M. 1992 Quasi-periodic and chaotic flow regimes in a thermally driven, rotating fluid annulus. *J. Fluid Mech.*, **238,** 599–632

Rössler, O. E. 1976 An equation for continuous chaos. *Phys. Lett. A*, **57A(5),** 397–398

Smith, L. A. 1992 Identification and prediction of low-dimensional dynamics. *Physica D*, **58,** 50–76

Smith, L. A. 1994 Visualising predictability with chaotic ensembles. In: *Advanced Signal Processing: Algorithms, Architectures and Implementations*,F.T. Luk, editor, volume 2296, pages 293–304, Bellingham, WA, 1994. SPIE.

Smith, L. A. 1994a Local optimal prediction. *Phil. Trans. R. Soc. Lond.*, **A 348,** 371–381

Smith, L. A. 1996 *Accountability and error in ensemble prediction in baroclinic flows.* In *Predictability*, volume 1 of *Seminar Proceeding*, pages 351–368. European Centre for Medium-Range Weather Forecasts, September 1995.

Smith, L. A.	1997	*The Maintenance of Uncertainty*, In Proceedings of the International School of Physics "Enrico Fermi", Course **CXXXIII**, pages 177–246, Italian Physical Society, Bologna.
Smith, L. A. and Gilmour, I.	1998	Accountability and internal consistency in ensemble formation. In: *Proceedings of the ECMWF Workshop on Predictability*, ECMWF, Shinfield Park, Reading RG2 9AX, UK.
Sparrow, C.	1982	*The Lorenz Equations*, Springer-Verlag, New York
Strang, G.	1988	*Linear algebra and its application*, Hartcourt Brace Jovanovich, San Diego
Sugihara, G. and May, R. M.	1990	Nonlinear forecasting as a way of distinguishing chaos from measurement error in a time series. *Nature*, **344,** 734–741
Talagrand, O. and Vautard, R. and Strauss, B.	1997	Evaluation of probabilistic prediction systems In *Predictabilty*, ECMWF Workshop Proceedings, Shinfield Park, Reading, UK, 1997. ECMWF
Tennekes, H.	1991	*Karl Popper and the accountability of numerical forecasting.* In *New Developments in Predictability*, ECMWF Workshop Proceedings, Shinfield Park, Reading, UK, 1991. ECMWF
Thompson	1957	Uncertainty of initial state as a factor in the predictability of large-scale atmospheric flow patterns. *Tellus*, **9,** 275–295
Tong, H. and Moeanaddin, R.	1988	On multi-step non-linear least squares prediction. *The Statistician*, **37,** 101–110
Toth Z.	1991	Estimation of atmospheric predictability by circulation analogs. *Monthly Weather Review*, **119**, 65–72
Toth Z. and Kalnay, E.	1993	Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society*, **74(12),** 2317–2330
Trefethen, L. Trefethen, A. Reddy, S. and Driscoll, T	1993	Hydrodynamic Stability without eigenvalues. *Science*, **261**, 578–584
Trevisan, A.	1993	Impact of transient error growth on global average predictability measures. *J. Atmos. Sci.*, **50,** 1016–1028
Trevisan, A. and Legnani, R.	1995	Transient error growth and local predictability: a study in the Lorenz system. *Tellus*, **47A,** 103–117
van den Dool, H.M.	1994.	Searching for analogues, how long must we wait? *Tellus*, **46 A**, 314–324
van den Dool, H.M. and Rukhovets, L.	1994.	On the weights for an ensemble-averaged 6-10-day forecast. *Weather and Forecasting (NMC notes)*, **9,** 457–465
Wolf, A. Swift, J. Swinney, H. L. and Vasrano, J. A.	1985	Determining Lyapunov exponents from a time series. *Physica*, **16 D,** 285
Yoden, S. Nomura, M.	1993	Finite-time Lyapunov stability analysis and its application to atmospheric dynamics. *J. Atmos. Sci.*, **50,** 1531–1543
Ziehmann, C., Smith, L. A. and Kurths, J	1999	The Bootstrap and Lyapunov exponents in Deterministic Chaos. *Physica*, **D,** (in press)
Ziehmann-Schlumbohm, C.	1994	*Vorhersagestudien in chaotischen Systemen und in der Praxis* . PhD thesis, Freie Universität Berlin. Meteorologische Abhandlungen. Neue Folge Serie A. Band 8 Heft 3
Ziehmann-Schlumbohm, C., Fraedrich, K. and Smith, L. A.	1995	Ein internes Vorhersagbarkeitsexperiment im Lorenzmodell. *Meteorologische Zeitschrift N.F.*, **4,** 16–21