ELSEVIER

Physica D 126 (1999) 49–59

PHYSICA D

The bootstrap and Lyapunov exponents in deterministic chaos

Christine Ziehmann [a,*], Leonard A. Smith [b,a], Jürgen Kurths [a]

[a] *Universität Potsdam, Institut für Physik, Nichtlineare Dynamik, Am Neuen Palais 10, 14415 Potsdam, Germany*

[b] *Mathematical Institute, University of Oxford, Oxford, UK*

Received 17 September 1997; received in revised form 3 March 1998; accepted 12 September 1998

Communicated by A.M. Albano

Abstract

Inasmuch as Lyapunov exponents provide a necessary condition for chaos in a dynamical system, confidence bounds on estimated Lyapunov exponents are of great interest. Estimates derived either from observations or from numerical integrations are limited to trajectories of finite length, and it is the uncertainties in (the distribution of) these finite time Lyapunov exponents which are of interest. Within this context a bootstrap algorithm for quantifying sampling uncertainties is shown to be inappropriate for multiplicative-ergodic statistics of deterministic chaos. This result remains unchanged in the presence of observational noise. As originally proposed, the algorithm is also inappropriate for general nonlinear stochastic processes, a modified version is presented which may prove of value in the case of stochastic dynamics. A new approach towards quantifying the minimum duration of observations required to estimate global Lyapunov exponents is suggested and is explored in a companion paper. ©1999 Elsevier Science B.V. All rights reserved.

Keywords: Bootstrap; Chaos; Lyapunov exponent; Uncertainty; Observational noise; Dynamical noise; Nonlinear systems

1. Introduction

The fundamental role played by Lyapunov exponents in defining chaotic dynamics [1] has stimulated the search for a statistical framework within which the accuracy of estimated exponents might be quantified. In this paper, we demonstrate that the framework proposed in [2], if interpreted as a bootstrap approach, fails to provide reliable bounds for estimates of either finite time Lyapunov exponents [3] or global Lyapunov exponents; both of these exponents are defined in Section 2. Section 3 introduces several dynamical systems and the bootstrap, a general approach to estimate the remaining uncertainty in an estimated statistic by employing resampling methods. Following Künsch [4], we note that the application of the bootstrap is restricted to data from a proper underlying process (stationarity, weak dependence or short term dependence) and a well-behaved functional. Both the conditions on the process and those on the statistic are violated in the case of Lyapunov exponents in deterministic chaotic systems. First, successive observations of these chaotic processes are neither independent nor weakly dependent; certainly, whenever Takens' Theorem [5] can be applied, the use of a bootstrap must be carefully justified. Second, Lyapunov exponent estimates rely on the product of matrices. In general, matrix multiplication does not commute (except in one dimension). Thus the bootstrap is inappropriate for multiplicative ergodic functionals of deterministic dynamical systems.

* Corresponding author; Tel.: +49-331-997-1302; fax: +49-331-977-1142; e-mail: chriss@agnld.uni-potsdam.de

0167-2789/99/$ – see front matter
PII: S0167-2789(98)00256-5

The successful application of the bootstrap hangs both on the nature of the underlying process and on the properties of any noise. If the noise is simply independent and identically distributed (IID) measurement uncertainty (observational noise), then the bootstrap approach remains inappropriate, as shown in Section 3.1. In Section 4, it is argued that it is also inappropriate for the general class of stochastic systems discussed by Bailey et al [6], although a modified bootstrap approach is suggested which may prove useful in these stochastic systems with positive Lyapunov exponents (these systems are called aleatoric [7] in order to distinguish them from deterministic chaos).

An alternative approach for quantifying whether finite time Lyapunov exponents might have converged towards their global counterparts is suggested in the last section. This approach is investigated for a variety of dynamical systems in a companion paper [8] hereafter referred to as *Paper II*. Both papers are concerned with "small" time scales given information along a trajectory which is itself of only finite duration. The main results of the current paper are that the framework of Gençay [2] read as a bootstrap is inappropriate for deterministic chaos, and that this is also the case for a wide class of stochastic systems. In addition, two conflicting definitions of "Local Lyapunov Exponents" are distinguished in Section 3.1.1, the importance of distinguishing between observational noise and dynamic noise is reiterated, and a new approach towards identifying whether the given observations might be of sufficient duration to obtain a good estimate of a global Lyapunov estimate is suggested.

2. Finite-time Lyapunov exponents

Lyapunov exponents are defined through the infinite time behaviour of infinitesimal uncertainties ϵ about the state $\mathbf{x}$ of a dynamical system in an m-dimensional state space [5]. The linear propagator $\mathcal{M}(\mathbf{x}, \Delta t)$, also called the tangent map, evolves any infinitesimal initial uncertainty $\epsilon_0 \in \mathbb{R}^m$ about $\mathbf{x}(0)$ forward for a time Δt along the nonlinear trajectory $\mathbf{x}(t)$,

$$\epsilon_{\Delta t} = \mathcal{M}(\mathbf{x}, \Delta t)\epsilon_0. \tag{1}$$

In continuous time systems (flows) where $\dot{\mathbf{x}} = F(\mathbf{x})$, $\mathcal{M}(\mathbf{x}, \Delta t)$ is defined by simultaneous integration of the nonlinear and the first variational equations; in discrete time systems (maps) where $\mathbf{x}_{i+1} = F(\mathbf{x}_i)$, the one step tangent map is simply the Jacobian: $\mathcal{M}(\mathbf{x}, \Delta t = 1) = \mathcal{J}(\mathbf{x})$. If we interpret Δt as the fundamental propagation time of the problem at hand (be it a sampling time, iteration time, or integration step size) and k as an integer number of propagation steps, then for each k and each x we have the $m \times m$ matrices:

$$\mathcal{O}(\mathbf{x}, k\Delta t) = [\mathcal{M}(\mathbf{x}, k\Delta t)^{\mathrm{T}} \mathcal{M}(\mathbf{x}, k\Delta t)]^{1/(2k\Delta t)} \tag{2}$$

and

$$\mathcal{O}(\mathbf{x}) = \lim_{k\to\infty} \mathcal{O}(\mathbf{x}, k\Delta t), \tag{3}$$

where $\mathcal{A}^{\mathrm{T}}$ indicates the transpose of $\mathcal{A}$. Oseledec [9] proved that if the limit of $k \to \infty$ exists, then under a wide range of conditions the eigenvalues of $\mathcal{O}(\mathbf{x})$ are independent of $\mathbf{x}$ for almost all $\mathbf{x}$ in the same basin of attraction. An m dimensional dynamical system thus has m Lyapunov exponents, $\Lambda_i, i = 1, 2, \ldots, m$ defined [5] via e_i, the eigenvalues of $\mathcal{O}(\mathbf{x})$, as $\Lambda_i = \log_2(e_i), i = 1, 2, \ldots, m$; by convention $\Lambda_i \geq \Lambda_j$ for $i < j$. The eigenvalues of the matrix $\mathcal{M}^{\mathrm{T}}\mathcal{M}$ are simply the squares of the singular values of $\mathcal{M}$. Thus in the limit $k \to \infty$, the logarithm of the first singular value of $\mathcal{M}(\mathbf{x}, k\Delta t), \sigma_1$, approaches the first global Lyapunov exponent, Λ_1:

$$\Lambda_1 = \lambda_1^{(\infty)} \equiv \lim_{k\to\infty} \frac{1}{k\Delta t} \log_2(\sigma_1). \tag{4}$$

Note that all logarithms are taken as base two, thus the exponents have units of bits per unit time (Δt).

Following Lorenz [3] and Abarbanel et al. [10], we use the $\sigma_i^{(k)}$ (the singular values of the tangent map $\mathcal{M}(\mathbf{x}, k\Delta t)$) to define [1] *finite time Lyapunov exponents* as

$$\lambda_i^{(k)}(\mathbf{x}) = \frac{1}{k\Delta t} \log_2(\sigma_i^{(k)}(\mathbf{x})). \tag{5}$$

For clarity, we shall often drop the explicit dependence on $\mathbf{x}$ for $\lambda_i^{(k)}(\mathbf{x})$ and $\sigma_i^{(k)}(\mathbf{x})$. In the 1960s, Lorenz [3] used the $\lambda_i^{(k)}$ to quantify variations in the predictability of a 28-variable atmospheric model, and the associated singular vectors are currently used in several of the competing methods for the construction of ensembles employed in operational numerical weather forecasting [11–14].

[1] This terminology is justified in Section 3.2.

There is a vast literature on the estimation of the Lyapunov spectrum (see, for example [5,15–25]). The so-called "Jacobian method" proposed by Eckmann and Ruelle [5] and further developed in [19,20], requires three steps: (i) selecting a model state-space, (ii) estimating the one step tangent maps $\mathcal{M}(\mathbf{x}, \Delta t)$, and (iii) computing Lyapunov exponents from the series of tangent maps (computational issues are discussed in Appendix A). The picture is most easily drawn for time discrete systems where the tangent map for k steps from an initial condition $\mathbf{x}_1$ is simply the product of the Jacobians along the trajectory, that is,

$$\mathcal{M}(\mathbf{x}_1, k) = J(\mathbf{x}_k) \cdots J(\mathbf{x}_2) J(\mathbf{x}_1). \qquad (6)$$

Step (iii) provides the point of departure for the current paper: we will assume that a series of N_{obs} tangent propagators $\mathcal{M}(\mathbf{x}, \Delta t)$ is in hand.

3. Deterministic chaos and the bootstrap

In this section we employ three common chaotic systems to demonstrate that even a perfect knowledge of N_{obs} tangent maps along a trajectory does not allow the estimation of confidence bounds for either Λ_1 or $\langle\lambda_1^{(N_{\mathrm{obs}})}\rangle$ in the case of deterministic dynamics. The systems are:

(i) The Henon map [26] allows a direct comparison with previous results on the bootstrap [2,27]; the equations are

$$x_{i+1} = 1 - a x_i^2 + y_i, \qquad y_{i+1} = b x_i, \qquad (7)$$

with the parameters $a = 1.4$ and $b = 0.3$. The Jacobian of the Henon map [2] has a particularly simple structure which is independent of y:

$$\begin{pmatrix} -2ax & 1 \\ b & 0 \end{pmatrix}$$

(ii) The Ikeda map [28]:

$$x_{i+1} = 1 + \mu[x_i \cos(t) - y_i \sin(t)],$$
$$y_{i+1} = \mu[x_i \sin(t) + y_i \cos(t)], \qquad (8)$$

with $t = 0.4 - 6/(x^2 + y^2 + 1)$ and $\mu = 0.9$ provides a slightly more complex Jacobian. Lyapunov exponents of both systems are well documented (see [5,10,19,22,29] and references therein) and our results below are supported by those presented in these references.

(iii) The Sinai map [30]:

$$x_{i+1} = x_i + y_i + \delta \cos(2\pi y_i) \mod 1,$$
$$y_{i+1} = x_i + 2y_i \mod 1, \qquad (9)$$

with $\delta = 0.10$. In this system the Jacobian depends only on y; further the distribution of y appears to be uniform on the interval $(0, 1)$, which may allow the development of analytic arguments to support the conclusions of this section.

In each case the Jacobian is known analytically. Given a (numerical) trajectory, the linear propagators over k steps can be determined; thus each propagation time k implies a true distribution of $\lambda_i^{(k)}$, with one value for each initial condition $\mathbf{x}_1$ and $\mathbf{x}_1$ is uniformly distributed with respect to the natural measure on the attractor, $\mu(\mathbf{x})$.

3.1. *Demonstrating the inappropriateness of the bootstrap*

When analysing observational data, the number of observations, N_{obs}, is usually fixed; there is one realization of the data, and no more. For a given statistic, only one estimate over the complete data set exists and there is no access to the distribution of $\langle\circ\rangle_{N_{\mathrm{obs}}}$. We adopt the notation $\langle\circ\rangle$ to denote the mean of the true distribution, and $\langle\circ\rangle_N$ to denote a mean estimated from N consecutive realizations along a single trajectory; thus in a deterministic system, $\langle\circ\rangle_N$ is defined by the value of $\mathbf{x}_1$. The bootstrap [31] provides a powerful statistical methodology aimed at quantifying the uncertainty resulting from a finite sample. This is accomplished by using resampling methods to construct empirical distributions of the statistic (or functional) in question [4,32–34]. Drawing randomly, with replacement, from the original N_{obs} elements, the functional of interest is calculated yielding a so-called *bootstrap replicate*. This procedure is then repeated, leading to an empirically based estimate of the distribution of the statistic via that of the bootstrap replicates. This method is inappropriate, however, for dependent data, for which moving blocks bootstraps have been proposed (see, for example, [4]). In that case, one draws

[2] Note that two misprints in [2] yield the wrong value both for the Jacobian (Eq. (9) of [2]) and for the Λ_1 (Table 1 of [2]) of the Henon map.

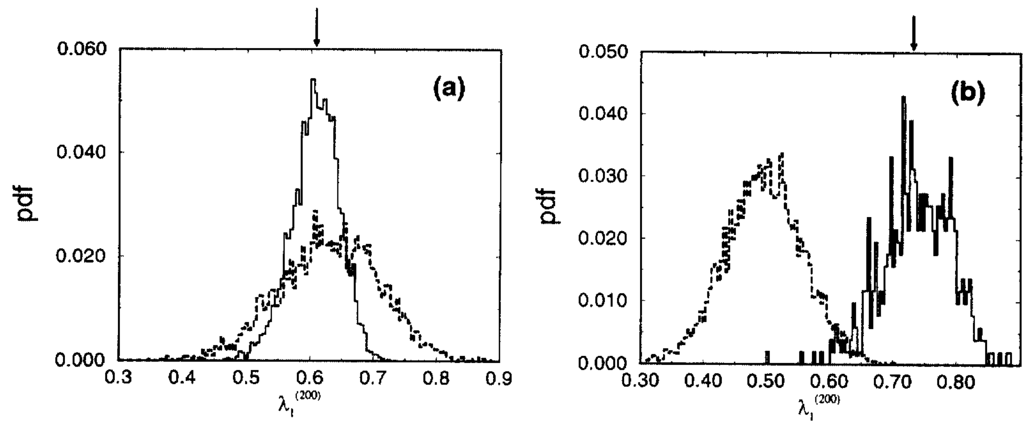

Fig. 1. Distributions of $\lambda_1^{(200)}$ for deterministically (solid) and randomly (dotted) sampled Jacobians drawn from the invariant measure (a) of the Henon map and (b) of the Ikeda map. The distributions are different. An estimate of Λ_1 is indicated by the arrow at the top axis in each panel.

randomly with replacement blocks of length l from the original data in an attempt to take the dependence in the data into account.

The authors of this paper (and its referees) read Gençay [2] as promoting the moving block approach to preserve the distribution of delay reconstruction points, and then computing the bootstrap replicates $\tilde{\lambda}_1^{(N_{\text{obs}})}$ from the products of randomly ordered Jacobians drawn from this distribution in the hope that "the moments of the largest Lyapunov exponent estimator obtained from the bootstrap procedure matches the distribution of the true largest Lyapunov exponent". When matrix manipulation does not commute, this procedure leads to a distribution of replicates which differs significantly from the true distribution even in the limit of large samples. This is illustrated in Fig. 1, which contrasts the true distribution $\lambda_1^{(k)}$, and that of the bootstrap replicates $\tilde{\lambda}_1^{(k)}$, for the case $k = 200$ for both the Henon map and the Ikeda map. This figure establishes the main result of this paper: the distribution of these bootstrap replicates is significantly different from that of the true distribution even in the noise-free case with huge data sets. This is illustrated in Table 1 for $k = 2^{16}$; clear difference remains in both the Henon system and the Ikeda system: the moments of the distribution obtained by the bootstrap procedure do not match those of the true finite time Lyapunov exponent in these deterministic chaotic systems. An alternative reading of Gençay [2] interprets the algorithm as merely sub-sampling different series of Jacobians in their observed order, it is this second reading that was intended by Gençay (Gençay, 1998, private communication). Under this second reading, the algorithm does no resampling at all, and hence not a bootstrap approach; while the issue of commutation of matrix multiplication discussed in this section is avoided, errors in estimating Λ_i due to deterministic dynamics discussed in Section 3.2 are not moderated

Table 1
Finite time Lyapunov exponents $\lambda_i^{N_{\text{obs}}}$ and corresponding bootstrap replicates for three initial conditions in each system; $N_{\text{obs}} = 2^{16}$.

	$\lambda_1^{(N^{\text{obs}})}$	$\lambda_2^{(N^{\text{obs}})}$	$\sum_{i=1}^{2} \lambda_i^{(N^{\text{obs}})}$
Henon (deterministic)	0.60389	−2.34086	−1.73697
	0.60484	−2.34180	−1.73697
	0.60705	−2.34401	−1.73697
Henon (bootstrap)	0.63450	−2.37146	−1.73697
	0.62742	−2.36438	−1.73697
	0.62682	−2.36378	−1.73697
Ikeda (deterministic)	0.73219	−1.03620	−0.30401
	0.72343	−1.02744	−0.30401
	0.73404	−1.03805	−0.30401
Ikeda (bootstrap)	0.49590	−0.79991	−0.30401
	0.49585	−0.79985	−0.30401
	0.49371	−0.79772	−0.30401
Sinai (deterministic)	1.38441	−1.55638	−0.17197
	1.38423	−1.55388	−0.16964
	1.38438	−1.55410	−0.16971
Sinai (bootstrap)	1.38432	−1.55491	−0.17059
	1.38391	−1.55187	−0.16795
	1.38434	−1.55407	−0.16973

For both the Henon and Ikeda systems, the variation within either the true values or the replicates is small compared to the differences between them; variations of order 1% remain in $\lambda_1^{(2^{16})}$. The Sinai case is discussed in the text.

in any way. It is the bootstrap interpretation, not the sub-sampling interpretation, that is deployed in [27].

It is clear from Fig. 1 that the distributions of $\lambda_1^{(200)}$ for the deterministically ordered and randomly sampled Jacobians are different. In the Ikeda system both distributions hardly overlap and the true Lyapunov exponent Λ_1 lies beyond the 99.9 percentile of this distribution. While in the Henon map and Sinai map, a "shuffling" of matrices does not affect the finite time Lyapunov exponents dramatically. For $k = 200$ in Henon, the mean in the random case is only slightly larger than in the deterministic case and the true Lyapunov exponent lies well within both distributions; nevertheless the distribution of the replicates differs substantially from the true distribution. This difference increases with increasing k, as is reflected in Fig. 2 which shows the evolution of distributions of the $\lambda_1^{(k)}$ (left-hand panels, deterministic cases) and the $\tilde{\lambda}_1^{(k)}$ (right-hand panel, bootstrap replicates) as a function of k for all three systems. Fig. 2 illustrates that the bootstrap methodology fails even for relatively large N_{obs}, this is most obvious for the Ikeda map.

Why would the distribution of bootstrap replicates $\tilde{\lambda}_1^{(k)}$ bear any resemblance to that of the $\lambda_1^{(k)}$? This will occur whenever there tends to be a systematic alignment between the left singular vectors of $\mathcal{M}(\mathbf{x}(t), \Delta t)$ and the right singular vectors of $\mathcal{M}(\mathbf{x}(t + \Delta t), \Delta t)$, independent of $\mathbf{x}$ and Δt. In both the Baker's map and the Baker's apprentice maps [7,35], corresponding right and left singular vectors are parallel; in these very special cases, the product of Jacobians does commute. There is a similar, though less dramatic, tendency towards alignment in the Sinai map, illustrated by the "similarity" of Fig. 2 (e) and (f). Clearly, such global organisation cannot be assumed a priori in an unknown nonlinear dynamical system, and we expect $\langle\lambda_1^{(k)}\rangle > \langle\lambda_1^{(2k)}\rangle > \Lambda_1$ for all k. Also note that for both the Henon and Ikeda systems, the determinant of the Jacobian is a constant, independent of $\mathbf{x}$. Thus the third column of Table 1 is constant (to within questions of numerical precision), independent of how the Jacobians are sampled. As argued elsewhere [3,13,35], the same mathematical simplicity that makes systems like the Baker's map and the Henon map attractive for analytical analysis, also makes them weak straw men for the evaluation of data analysis techniques.

The discussion above considered deterministic maps, yet the shortcomings in the distribution of the bootstrap replicates is expected to be more obvious in a flow, where continuity will introduce more obvious structure into $P(\mathbf{x}(t + \Delta t) \mid \mathbf{x}(t))$. Consider a trajectory from $\mathbf{x}(0)$ for a fixed duration T. The linear propagator over time T is estimated from the series of Jacobians at $\mathbf{x}(i\Delta t)$, $i = 0, 1, 2, \ldots, k - 1$, that is every integration time step Δt, where $T = k\Delta t$. For a fixed time interval T, $\lambda_i^{(T)}(\mathbf{x}(0))$ converges as $\Delta t \to 0$ while the bootstrap replicates generated by randomly resampling from these k Jacobians become less and less relevant in the same limit.

Ideally, we are interested in the distribution of $\lambda_1^{(k)}(\mathbf{x})$, where $\mathbf{x}$ is uniformly distributed with respect to the natural measure on the attractor. Two difficulties arise in time series analysis: first, if the entire sample is taken along a single trajectory then the average of N values of $\lambda_1^{(k)}$, $\langle\lambda_1^{(k)}\rangle_N$, provides a biased estimate of $\langle\lambda_1^{(k)}\rangle$. Second, if the observations are inexact then each of the individual tangent maps may be inexact. We deal with the second issue next, while the first will be discussed in Section 3.2 below.

3.1.1. Observational noise

Successful application of the bootstrap depends both on the nature of the underlying process and on the properties of any observational noise. Gençay [2] and Bask and Gençay [27] considered deterministic chaotic processes where the data includes independent and identically distributed (IID) observational noise added to a system variable. The noise is solely a measurement uncertainty, and does not effect the dynamics in any way. Thus the addition of observational noise does not save the bootstrap. Uncertainty in the data results in imperfect estimates of each of the Jacobians, thus in place of Eq. (6) we have

$$\tilde{\mathcal{M}}(\mathbf{x}_1, k) = \tilde{J}(\mathbf{x}_k) \cdots \tilde{J}(\mathbf{x}_2)\tilde{J}(\mathbf{x}_1). \qquad (10)$$

Even though each Jacobian is inexact, each $\mathbf{x}_i$ is completely determined by $\mathbf{x}_1$, and this time ordering of the Jacobians will distinguish $\lambda_i^{(k)}(\mathbf{x}_1)$ from the bootstrap replicates of randomly ordered k-products of the same Jacobians. A valid approach to confidence limits akin to the bootstrap methodology would consist of altering the original data with small perturbations within the observational uncertainty [3] and then to re-fit the

[3] Alternatively, one might either employ different model structures on the identical data set or multiple realizations as in [6].

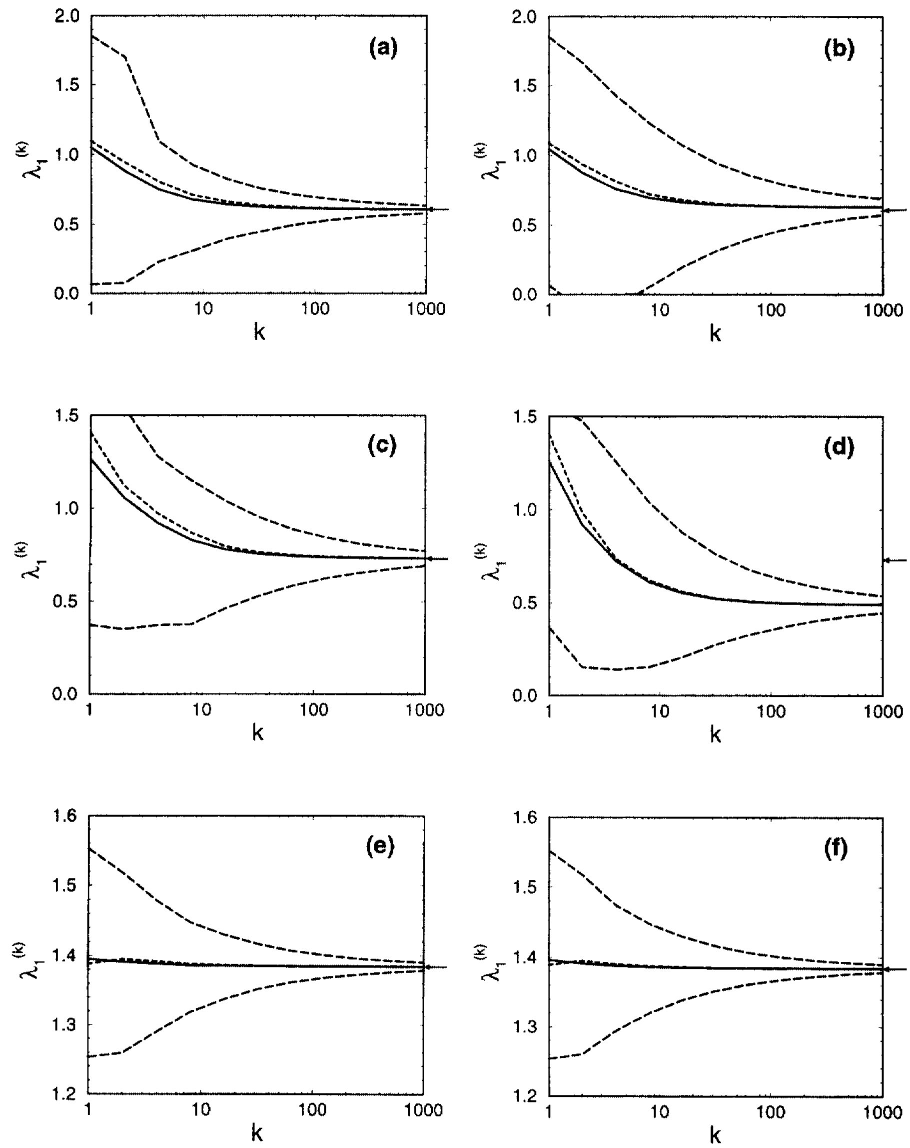

Fig. 2. Distribution of the finite time Lyapunov exponents $\lambda_1^{(k)}$ as a function of k for the deterministic (left) and the random (right) sampling of Henon (a) and (b), Ikeda (c) and (d), and Sinai (e) and (f) systems. The mean (solid), median (short-dashed), and the 5% and 95% quantiles (long-dashed) are shown. For each k the distributions are based on sample sizes of 2^{15} Jacobians. The arrow to the right of each panel indicates an estimate of Λ_1 for the corresponding deterministic system.

model, compute new Jacobians, and obtain a distinct estimate of $\lambda_i^{(k)}(\mathbf{x}_1)$. While this would allow a distribution of replicates for estimating the uncertainty in $\lambda_i^{(k)}(\mathbf{x}_1)$ for a particular (if unknown) value of $\mathbf{x}_1$, it sheds no light on the distribution of $\lambda_i^{(k)}$ over the attractor, nor on the relation of $\langle\lambda_i^{(k)}\rangle$ to Λ_i.

3.2. Implications of deterministic dynamics

Given $N_{\rm obs}$ tangent maps along a trajectory, there is only one estimate of $\lambda_1^{(N_{\rm obs})}$; similarly, there are $N = N_{\rm obs}/k$ values from the distribution of $\lambda_1^{(k)}$. This partitioning yields a series of disjoint finite time Lyapunov exponents which are "independent" only in the sense that no particular Jacobian is used twice in the calculation. They are, in general, not independent in a statistical sense since they lie along a deterministic trajectory. For each and every one of these $N_{\rm obs}/k$ values, the corresponding $\lambda_1^{(k)}$ is completely defined by the initial state of the system, $\mathbf{x}_1$. Thus $\langle\lambda_1^{(k)}\rangle_N$ is also defined by $\mathbf{x}_1$. Observational noise does not destroy this dependence.

If the goal is to estimate Λ_1 from the particular observed value of $\langle\lambda_1^{(k)}\rangle_N$, one must cope with three distinct difficulties: (1) the systematic error due to a single finite trajectory, (2) the sampling error in the estimated mean due to finite N, and (3) the simple fact that for finite k, $\langle\lambda_1^{(k)}\rangle \neq \Lambda_1$ except in special cases (as noted above). As stressed by Ellner et al. [29], each of these effects persist, even when the particular $\lambda_1^{(k)}$ (or the Jacobians) are known exactly; Ellner et al. refer to the difference between Λ_1 and $\langle\lambda_1^{(k)}\rangle$ as "block error", but we note that there is no "error" here per se, these are simply different quantities. Obtaining $N_{\rm rep}$ samples with the aim of quantifying the variation in $\langle\lambda_1^{(k)}\rangle_N$ will require $N_{\rm rep}$ disjoint blocks each of length Nk; a bias will still remain if the new $N_{\rm obs} = N_{\rm rep}Nk$ maps lie along the same trajectory. The central point here is that $\langle\lambda_i^{(k)}\rangle_N$ provides a biased estimate of $\langle\lambda_i^{(k)}\rangle$ when the N samples are taken along a single length of trajectory, this significantly limits applicability of the bootstrap.

The asymptotic convergence of $\langle\lambda_1^{(k)}\rangle$ to Λ_1 as $k \to \infty$ is widely discussed (see, for example, [10,22,25,29] and references thereof); in contrast, here we are concerned with small k. Indeed, we are interested in the evolution of $\langle\lambda_1^{(k)}\rangle$ for precisely those k prior to the onset of the asymptotic scaling range documented by those authors. It is for these small values of k that confidence bounds, if available, would be of most value. Of course, we cannot say a priori what constitutes small k; it will be system dependent. Although $k = 200$ (the case considered by Gençay) is often considered to be "large", Fig. 1 shows that the width of the (true) distribution of $\lambda_1^{(k)}$ still exceeds 10% of its mean value for both the Henon map and the Ikeda map. The companion paper *paper II* investigates the distributions of $\lambda_i^{(k)}$ for small k, the estimation of their means, and the evolution of the mean values $\langle\lambda_1^{(k)}\rangle$ and $\langle\lambda_m^{(k)}\rangle$ towards Λ_1 and Λ_m, respectively. Note that the convergence may depend not only on k, but also upon the initial condition; for example, there exist long chaotic transients to a periodic orbit or fixed points [36,37].

In deterministic systems there is a fundamental ambiguity in estimating positive Λ_i from a finite trajectory, namely that the observations may represent only transient behaviour, or a subset of the full dynamics. Thus it is nontrivial to approximate the limit of infinite time and assess the global Λ_i, even when the dynamical equations are known, and a plethora of definitions (and associated names) for "local" Lyapunov exponents exist [3,10,38–45]. While most of these converge to the Lyapunov exponents Λ_i as $k \to \infty$, they can differ considerably for the finite times which are of interest here. Among these definitions, it is important to distinguish clearly between two distinct classes of exponent, both of which are commonly referred to as "local Lyapunov exponents". Both classes consider the effective growth rate of infinitesimal uncertainties over a fixed (finite) time interval; they differ in the initial orientation of the uncertainty. Eq. (5) defined *finite time Lyapunov exponents* with respect to the singular vector basis arising from the singular value decomposition (SVD) of $\mathcal{M}(\mathbf{x}, \tau)$ for finite τ; these exponents were dubbed local Lyapunov exponents by Abarbanel et al. [10] and this usage is now common. Alternatively, one may sample the growth of infinitesimal perturbations along some other well-defined basis; for example, the local orientations of a basis which defines the global Lyapunov exponents. Effective growth rates sampled in this very different basis are also referred to as "local Lyapunov exponents" [38,46–50]. To avoid increasing the existing confusion, we would refer to the second as *finite sample Lyapunov exponents* and avoid the moniker local Lyapunov exponent all together. The distributions of finite time exponents are contrasted with those of finite sample exponents of the same system in *Paper II*; the results here concern only finite time Lyapunov exponents. We stress that neither the finite sample nor finite time exponents imply exponential growth: any nonzero singular value defines a corresponding "exponent" regardless of whether the growth is linear or exponential or intermittent (or otherwise). Any increase in separation will yield a positive finite time exponent: exponential growth is not required. These facts severely limit the utility of Lyapunov exponents in quantifying predictability [7,12], as does the requirement that the duration of interest ($k\Delta t$) must be specified a priori. Alternatives to Lyapunov exponents, such as local uncertainty doubling times and state dependent optimisation times for defining singular vectors, are discussed in [7,12,13].

4. Bootstraps in nonlinear stochastic systems

The applicability of the bootstrap to nonlinear stochastic systems is also of interest. One class of stochastic systems consists of some deterministic map, $\mathbf{F}(\mathbf{x})$, with dynamic noise, $\epsilon_{\rm dyn}$; thus $\mathbf{x}_{i+1} = \mathbf{F}(\mathbf{x}_i) + \epsilon_{\rm dyn}$. Here the noise affects the current state (and thus the future evolution) of the system, whereas observational noise affects the measurements but leaves the underlying dynamics unscathed. Recently, Bailey et al. [6] and Yao and Tong [51] have made significant progress in understanding the Lyapunov exponents of members in this class of systems where $\mathbf{F}(\mathbf{x})$ is chaotic. We will refer to these nonlinear stochastic dynamical systems with positive Lyapunov exponents as aleatoric [7]; they differ considerably from the case of observational noise on the same deterministic chaotic system. The difference proves crucial to the applicability of the bootstrap. Bailey [52] has proven a generalisation of the central limit theorem (CLT) under fairly general assumptions for these stochastic systems. In contrast, it can be proven that the central limit theorem does not apply to convergence of Lyapunov exponent estimates in some deterministic systems (see [53]), and there is widespread numerical evidence [10,22,29,41] of "anomalous convergence" in deterministic systems for which analytic results are not available. The crucial distinction lies in whether or not "the noise" influences the dynamics: in the aleatoric case (Bailey) it does, justifying a Markov chain approach which leads to the CLT. In the chaotic case it does not; Taken's theorem applies to the underlying trajectory which remains deterministic even if the observations are inexact, leading to Eq. (10).

Interpreting the algorithm of Gençay [2] as a bootstrap also produces inappropriate replicates for the aleatoric systems; here $\mathbf{x}_1$ no longer uniquely defines $\mathbf{x}_k$, although the conditional probability distributions $P(\mathbf{x}_{i+1} \mid \mathbf{x}_i)$ may vary greatly from the invariant measure, $\mu(\mathbf{x})$. The crucial point is that even though the trajectory becomes only short-term dependent, matrix multiplication remains noncommutative. Strictly speaking, the bootstrap interpretation of Gençay's algorithm only applies when $P(\mathbf{x}_{i+1} \mid \mathbf{x}_i)$ reflects $\mu(\mathbf{x})$; that is, when the linear propagator is formed from a series of IID Jacobians.

Unlike the case of deterministic chaos, however, there is an alternate bootstrap approach for aleatoric systems above. Dynamically conditioned moving blocks of the Jacobians themselves can provide useful bootstrap replicates (hereafter called block-J replicates) as long as careful match criteria are followed. In particular, one must ensure that if $J(\mathbf{x}_i)$ multiplies $J(\mathbf{x}_j)$ from the left, then $\delta \equiv \mathbf{x}_i - F(\mathbf{x}_j)$ is of the order of the dynamical noise, $\epsilon_{\rm dyn}$; that is one must consider $P(F(\mathbf{x}_j) + \epsilon_{\rm dyn} = \mathbf{x}_i \mid \mathbf{x}_j)$ when selecting an $\mathbf{x}_i$. In the limit $\mid \epsilon_{\rm dyn} \mid \to 0$, this probability distribution approaches a delta function on $\mid \mathbf{x}_i - \mathbf{F}(\mathbf{x}_j) \mid$ and we recover[4] the deterministic case. In fact, demonstrating the (in)effectiveness of these dynamically conditioned bootstrap replicates may provide a test for distinguishing deterministic chaos from aleatoric dynamics.

5. Discussion and conclusion

The traditional bootstrap cannot be applied when estimating multiplicative ergodic statistics; a specific proposal for bootstrap confidence bounds [2] is untenable, mainly due to the fact that matrix multiplication does not commute. Previously reported results for the Henon map were serendipitous; dramatically different behaviour arises in other simple two-dimensional deterministic maps, such as the Ikeda map, and even in the Henon system for longer time series. In this respect the bootstrap is inappropriate for deterministic chaos. Nor can the algorithm be applied to aleatoric systems of the kind discussed by Bailey [52]. A dynamically conditioned moving block bootstrap approach might show promise in this case; adjacent blocks could be chosen probabilistically if conditioned on the current value of $\mathbf{x}(t)$, but not with uniform probability with respect to all observed $\mathbf{x}$. The bootstrap interpretation of the algorithm proposed in [2] is valid only for estimating the distribution of Lyapunov exponents of prod-

[4] As noted by a referee, we are considering the dynamics in the true state space of the system; generalizations to model-state spaces may prove nontrivial.

ucts of randomly ordered (IID) matrices [1]. It fails both in the case of deterministic chaos and in aleatoric stochastic dynamics.

For a particular system, we wish to know if our observations (or integrations) might be sufficient to allow, to a good approximation, $\Lambda_1 \approx \langle\lambda_1^{(k)}\rangle$, and before concluding, we suggest an alternative avenue towards identifying nonconvergence. Our goal is a necessary but not sufficient condition which, if satisfied, would merely indicate that our observations might be sufficient to allow, to a good approximation, $\Lambda_1 \approx \langle\lambda_1^{(k)}\rangle$ for experimentally accessible k. For sufficiently large k, we expect the power law convergence [10,22,29,41], yet as stressed by Ellner et al. [29], knowledge of the asymptotics is of limited utility in any particular case when only finite k is accessible. How small is k? As made clear by an anonymous referee, it is commonly held that $k = 200$ is a case of "large k" and thus the width of the distribution of $\lambda_1^{(200)}$ is often thought to be negligible. Figs. 1 and 2 show this is not the case for either the Henon map or the Ikeda map.

We do not know a priori how large a value of k is required; we can however invert the process and compare the distributions of $\lambda_1^{(k)}$ and $\lambda_1^{(k')}$, with $k' > k$, in order to determine whether k might cover a long enough duration so that the $\lambda_1^{(k)}$ approaches Λ_1. One straightforward approach is to construct a surrogate distribution for $\lambda_1^{(k')}$ based on the observed distribution of $\lambda_1^{(k)}$ under the null hypothesis that consecutive $\lambda^{(k)}$ were effectively independent. If contrasting the relevant moments of the surrogate distribution with those of the sample from the true distribution allows us to reject the null hypothesis, then we have an evidence that k is not large enough. Note that in forming the surrogate distribution, one must manipulate the $\lambda^{(k)}$ themselves; taking $k' = 2k$ for simplicity, the normalised singular values of $\mathcal{M}(\mathbf{x}_i, k\Delta t)\mathcal{M}(\mathbf{x}_j, k\Delta t)$ with randomly chosen $\mathbf{x}_i$ and $\mathbf{x}_j$ need reflect neither $\lambda_1^{(2k)}$ nor Λ_1. Even for large k, matrix multiplication does not commute.

The question of distinguishing "large k" from "small k" for both maps and flows is addressed in *Paper II*, where the algorithm suggested above is deployed. We stress that our test is only a necessary condition that for a given finite k, $\langle\lambda_1^{(k)}\rangle \approx \Lambda_1$ (and $\langle\lambda_m^{(k)}\rangle \approx \Lambda_m$), and hence a lower bound on what might constitute "large" k. The search for a sufficient condition continues; it will not succeed for the general deterministic case.

Acknowledgements

We are happy to acknowledge helpful discussion and comments from I. Gilmour, W. Jansen, I. Jolliffe, H. Künsch, T. Schreiber, and J. Theiler. We are particularly indebted to an anonymous referee for numerous constructive criticisms of an earlier draft, and to M. Bask and R. Gençay for a very informative exchange of views. CZ acknowledges financial support from the Free University of Berlin. LAS is supported through a Senior Research Fellowship at Pembroke College, Oxford and a Guest Professorship at the University of Potsdam made possible by the INK "Kognitive Komplexität" (DFG).

Appendix A

A major computational obstacle to estimating exponents lies in obtaining numerically stable estimates of the singular values of the large matrix product of Eq. (6), inasmuch as the propagator $\mathcal{M}(\mathbf{x}, k)$ may become numerically singular for relatively small k, if the product in (6) is estimated naively by matrix multiplication. The original method to avoid this singularity [5,20] transforms the product of Jacobians in (6) into a product of upper triangular matrices via a QR decomposition. Unfortunately, this approach cannot be applied in the case of the finite time Lyapunov exponents; we employ the iterative method of Abarbanel et al. [10], except for "small" values of k. Abarbanel et al. [10] obtained the singular values of $\mathcal{M}$ via recursive QR decomposition. Initially, take $Q_0 = I$ and define $A^{(1)}$ via

$$\mathcal{M}^{\mathrm{T}}\mathcal{M} = J_1^{\mathrm{T}} \cdots J_k^{\mathrm{T}} J_k \cdots J_1 \equiv A^{(1)}, \tag{A.1}$$

which, by QR decomposition yields

$$A^{(1)} = Q_{2k}^{(1)} R_{2k}^{(1)} R_{2k-1}^{(1)} \cdots R_1^{(1)}. \tag{A.2}$$

The superscript in brackets denotes the number of the iteration steps. Next, we define a matrix $A^{(2)}$ which has the same R factors but the $Q_{2k}^{(1)}$ on the right side:

$$A^{(2)} = R_{2k}^{(1)} R_{2k-1}^{(1)} \cdots R_1^{(1)} Q_{2k}^{(1)}, \tag{A.3}$$

and perform QR again:

$$A^{(2)} = Q_{2k}^{(2)} R_{2k}^{(2)} R_{2k-1}^{(2)} \dots R_1^{(2)}. \tag{A.4}$$

Since $A^{(2)} = Q_{2k}^{(1)\mathrm{T}} A^{(1)} Q_{2k}^{(1)}$, $A^{(1)}$ and $A^{(2)}$ are similar and have the same eigenvalues. One continues to create $A^{(3)}, A^{(4)}, \dots, A^{(n)}$ and, as n increases, $Q_{2k}^{(n)}$ converges to the identity matrix [54]. $A^{(n)}$ is then upper triangular.

In practice, the iteration was continued until each matrix element of $Q_{2k}^{(n)}$ differed by less than 10^{-6} from identity. For small k this may require rather large number of iterations, while for large k typically 2 or 3 iterations are sufficient. This suggests that for small k one might consider the SVD of the direct matrix product, either for the singular vector themselves or for an improved starting estimate of Q_0.

References

[1] D. Ruelle, Chaotic Evolution and Strange Attractors, Accademia Nazionale Dei Lincei, Cambridge University Press, Cambridge, 1989.

[2] R. Gençay, A statistical framework for testing chaotic dynamics via Lyapunov exponents, Physica D 89 (1996) 261–266.

[3] E.N. Lorenz, A study of the predictability of a 28-variable atmospheric model, Tellus 17 (1965) 321–333.

[4] H.R. Künsch, The jackknife and the bootstrap for general stationary observations, Ann. of Stat. 17(3) (1989) 1217–1241.

[5] J.-P. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors, Rev. Mod. Phys. 57 (1985) 617–656.

[6] B.A. Bailey, S. Ellner, D.W. Nychka, Chaos with confidence, Fields Institute Commun. 11 (1997) 115–133.

[7] L.A. Smith, Local optimal prediction: exploiting strangeness and the variation of sensitivity to initial condition, Phil. Trans. R. Soc. London A 348(1688) (1994) 371–381.

[8] L.A. Smith, C. Ziehmann, J. Kurths, Towards confindence estimates for Lyapunov exponents, Physica D, in preparation.

[9] V.I. Oseledec, A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems, Trans. Moscow Math. Soc. 19 (1968) 197–231.

[10] H.D.I. Abarbanel, R. Brown, M.B. Kennel, Lyapunov exponents in chaotic systems: their importance and their evaluation using observed data, Int. J. Mod. Phys. B 5 (1991) 1347–1375.

[11] F. Molteni, R. Buizza, T.N. Palmer, T. Petroliagis, The ECMWF ensemble prediction system: methodology and validation, Q.J.R. Meteorol. Soc. 122 (1996) 73–119.

[12] L.A. Smith, Accountability and error in ensemble prediction, in: Predictability, vol. 1, Seminar Proceeding, European Centre for Medium-Range Weather Forecasts, September 1995, pp. 351–368.

[13] L.A. Smith, C. Ziehmann, K. Fraedrich, Uncertainty dynamics and predictability in chaotic systems, Q.J.R. Meterol. Soc. (1998), in review.

[14] R. Buizza, T.N. Palmer, The singular-vector structure of the atmospheric global circulation, J. Atmos. Sci. 52(9) (1995) 1434–1455.

[15] G. Benettin, L. Galgani, J.-M. Strelcyn, Kolmogorov entropy and numerical experiments, Phys. Rev. A 14 (1976) 2338–2345.

[16] I. Shimada, T. Nagashima, A numerical approach to ergodic problem in dissipative dynamical systems, Prog. Theoret. Phys. 61 (1979) 1605–1616.

[17] J.M. Greene, J.-S. Kim, The calculation of Lyapunov spectra, Physica D24 (1987) 213–225.

[18] A. Wolf, J.B. Swift, H.L. Swinney, J.A. Vastano, Determining Lyapunov exponents from a time series, Physica D16 (1985) 285–317.

[19] M. Sano, Y. Sawada, Measurement of the Lyapunov spectrum from a chaotic time series, Phys. Rev. Lett. 55 (1985) 1082–1085.

[20] J.-P. Eckmann, S. Oliffson Kamphorst, D. Ruelle, S. Ciliberto, Lyapunov exponents from time series, Phys. Rev. A 34(6) (1986) 4971–4979.

[21] J. Kurths, H. Herzel, An attractor in a solar time series, Physica D 25 (1987) 165–172.

[22] R. Brown, P. Bryant, H.D.I. Abarbanel, Computing the Lyapunov spectrum of dynamical system from an observed time series, Phys. Rev. A 43(6) (1991) 2787–2806.

[23] U. Parlitz, Identification of true and spurious Lyapunov exponents, Int. J. Bifur. Chaos 2 (1992) 155.

[24] H.D.I. Abarbanel, R. Brown, J.J. Sidorowich, Lev Sh. Tsimring, The analysis of observed data in physical systems, Rev. Mod. Phys. 65 (1993) 1331–1392.

[25] A.G. Darbyshire, D.S. Broomhead, Robust estimation of tangent maps and Lyapunov spectra, Physica D 89 (1996) 287–305.

[26] M. Henon, A two-dimensional map with a strange attractor, Commun. Math. Phys. 50 (1976) 69.

[27] M. Bask, R. Gençay, Testing chaotic dynamics via Lyapunov exponents, Physica D 114 (1998) 1–2.

[28] K. Ikeda, Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system., Opt. Commun. 30 (1979) 257.

[29] S. Ellner, R. Gallant, D. McGaffrey, D. Nychka, Convergence rates and data requirements for the Jacobian-based estinmates of Lyapunov exponents, Phys. Lett. A 153 (1991) 357–363.

[30] Ya.G. Sinai, Gibbs measures in ergodic theory, Russian Math. Surveys 27(4)(4) (1972) 768.

[31] B. Efron, R.J. Tibshirani, An Introduction to the Bootstrap, Chapman & Hall, New York, 1993.

[32] B. Efron, R. Lepage, Exploring the Limits of Bootstrap, Chapter: Introduction to Bootstrap, Wiley, New York, 1992.

[33] R.Y. Liu, K. Singh, Exploring the Limits of Bootstrap, Chapter: Moving Blocks Jackknife and Bootstrap Capture Weak dependence, Wiley, New York, 1992.

[34] B. Efron, R. Tibshirani, Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, Statist. Sci. 1(1) (1986) 57–77.

[35] L.A. Smith, The Maintanence of uncertainty. In Proceedings of the International Summer School of Physics "Enrico Fermi" Course CXXXIII, Niovo Cimento, Bologna, Italy, Reidel, Dordrecht, 1997.

[36] C. Grebogi, E. Ott, J.A. Yorke, Crisis, sudden changes in chaotic attractors, and transient chaos, Physica D 7 (1983) 181–200.

[37] C.J. Marzec, E.A. Spiegel, Ordinary differential equations with strange attractors, SIAM J. Appl. Math. 38 (1980) 387 (transient choas to limit cycle).

[38] J.M. Nese, Quantifying local predictability in phase, Physica D 35 (1989) 237–250.

[39] J.S. Nicolis, G. Meyer-Kress, G. Haubs, Non-uniform chaotic dynamics with implications to information processing, Zeitschrift für Naturforschung 38a (1983) 1157-1169.

[40] J. Kurths, A. Brandenburg, Lyapunov exponents for hydromagnetic convection, Phys. Rev. A 44 (1991) R3427-R3429.

[41] H.D.I. Abarbanel, R. Brown, M.B. Kennel, Variation of Lyapunov exponents on a strange attractor, J. Nonlinear Sci. 1 (1991) 175–199.

[42] R. Doerner, B. Hübinger, W. Martiensen, S. Grossmann, S. Thomae, Predictability portraits for chaotic motions, Chaos, Solitons and Fractals 1 (1991) 553–571.

[43] H. Mukougawa, M. Kimoto, S. Yoden, A relationship between local error growth and quasi-stationary states: Case study in the Lorenz system, J. Atmos. Sci. 48 (1991) 1231–1237.

[44] S. Yoden, M. Nomura, Finite-time Lyapunov stability analysis and its application to atmospheric dynamics, J. Atmos. Sci. 50 (1993) 1531–1543.

[45] U. Feudel, J. Kurths, A. Pikovsky, Strange non-chaotic attractor in a quasiperiodically forced circle map, Physica D 88 (1995) 176–186.

[46] C. Nicolis, Probabilistic aspects of error growth in atmospheric dynamics, Q.J.R. Meteorol. Soc. 118 (1992) 553–568 .

[47] B. Eckhardt, D. Yao, Local Lyapunov exponents in chaotic systems, Physica D 65 (1993) 100–108.

[48] C. Nicolis, S. Vannitsem, J.-F. Royer, Short-range predictability of the atmosphere: mechanism for superexponential error growth, Q.J.R. Meteorol. Soc. 121 (1995) 705–722.

[49] A. Trevisan, R. Legnani, Transient error growth and local predictability: a study in the Lorenz system, Tellus A47 (1995) 103–117.

[50] Z. Toth, E. Kalnay, Ensemble Forecasting at NMC: The generation of perturbations, Bull. Amer. Meteorol. Soc. 74(12) (1993) 2317–2330.

[51] Q. Yao, H. Tong, Quantifying the influence of initial values Nonlinear prediction, J. R. Statist. Soc. B 56(4) (1994) 701–725.

[52] B. Bailey, Asymptotics and applications of local Lyapunov exponents, Ph.D. Thesis, North Caralina State University, Raleigh, NC, USA 1996.

[53] J. Theiler, L.A. Smith, Anomalous convergence of Lyapunov exponent estimates, Phys. Rev. E 51 (1995) 3738–3741.

[54] J. Stoer, R. Burlisch, Introduction to Numerical Analysis (translated by R. Bartels, W. Gautschi, C. Witzgall), Springer, New York, 1980.