

State dependent predictability: Impact of uncertainty dynamics, uncertainty structure and model inadequacies

V. M. Khade and J. A. Hansen

Dept. of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received: 2 September 2003 – Revised: 12 June 2004 – Accepted: 28 July 2004 – Published: 2 August 2004

Abstract. The implications of state dependent, finite time error growth has been studied using singular values in a chaotic 2-dimensional map. Earlier studies have demonstrated the superiority of the singular values over the Lyapunov number in representing error growth over finite time scales, since they take state dependency into account. In this work, linearized error growth as given by singular values under operational constraints like non-isotropic initial uncertainty and model error is considered. It is demonstrated that the relevant singular values in the case of non-isotropic initial uncertainty are quite different from the isotropic case. The effect of model inadequacy on error growth is delineated.

1 Introduction

It has long been recognized that inaccuracies in weather forecasts result from uncertainties in the initial conditions and model imperfections. The initial uncertainty has a number of sources, including the paucity of observations and noise in observations, making it impossible to accurately determine the state of the atmosphere. Though the meteorological community was aware of the initial condition uncertainty problem (Thompson, 1957), its gravity was fully appreciated after the discovery of “chaos” (Lorenz, 1963). Ever since, there has been a spate of scientific papers in almost all branches of science devoted to understanding, quantifying and analyzing properties of chaotic systems. Techniques to improve the prediction of chaotic systems has formed a sizeable subset of these efforts.

In chaotic systems, an error in the initial condition grows, on average, exponentially fast. Any small but finite error in the initial condition tends to amplify rapidly, thus frustrating attempts to accurately forecast a chaotic system. Many “geophysical systems” appear to be chaotic. The critical importance of determining the initial state accurately has prompted

the application of techniques to pin down the initial state with as much certainty as possible. Prominent among these techniques are 4 dimensional variational analysis (Dimet and Talagrand, 1986), the Extended Kalman Filter (Evensen, 1992) and the Ensemble Kalman Filter (Evensen, 1994).

One of the most popular diagnostics for quantifying the evolution of error in initial conditions for chaotic systems is the Lyapunov exponent (Oseledec, 1968). Several studies have used the Lyapunov exponent formulation to analyze the growth of error in dynamical systems (Smith et al., 1999). The Lyapunov exponent as defined by Oseledec gives the average growth rate of error over the whole state space for infinite time. However, from a practical standpoint this growth rate is not very useful since in real life we are generally interested in short term forecasts of some localized part of state space. The concept of singular vectors and values (Strang, 1988) gives a more relevant measure for error growth than the Lyapunov exponent. E. N. Lorenz (Lorenz, 1965) pioneered the interpretation of singular vectors and values in the context of meteorology. The singular vectors are a function of localized dynamics and finite time intervals rather than global dynamics and infinite time intervals.

Much work on singular values and vectors in operational Numerical Weather Prediction (NWP) has been done. But the sensitivity of singular values and their distribution over state space to the shape of initial uncertainty and to model inadequacy remains by and large unexplored, though some work has been done (Barkmeijer et al., 1998). The typical assumptions of isotropic uncertainty and perfect models fail in the real world, resulting in misleading estimates of singular error growths.

In the current study, the state dependent singular error growth over finite time has been studied in a 2-dimensional chaotic map, the Ikeda system (Ikeda, 1979). The Ikeda system has been chosen to allow visual demonstration of the impact of practical constraints on the distribution of the singular values. To begin with, the advantage of singular values over Lyapunov numbers is demonstrated. Next, the impact of the above mentioned simplifications (isotropic initial uncertainty

and perfect model) on singular error growth has been analyzed.

Section 2 below outlines the dynamical system employed, the Ikeda map. It also introduces the idea of linearization and the linear propagator. Section 3 delineates the theoretical framework of Lyapunov and singular values. It is meant to introduce the mathematics behind these concepts. Section 4 explains the methodology and results of the current study. Finally, Sect. 5 summarizes the results.

2 Ikeda system and linearization

Though the exact nature of initial error growth will depend on the system under consideration, some of its generic features can be understood and possible issues elucidated by studying low-dimensional chaotic systems. Of course, the solutions obtained by using these systems need not scale to high dimensional and more complex models, but the generic results obtained can be illuminating. The Ikeda system has been used by several authors (Smith et al., 1999; Hansen, 2001) to elucidate predictability studies. All work in this study is restricted to the Ikeda system. A description of the Ikeda system and its linearization is given below.

The equations of the Ikeda system are,

$$x_{i+1} = 1 + \mu(x_i \cos \theta - y_i \sin \theta) \quad (1)$$

$$y_{i+1} = \mu(x_i \sin \theta + y_i \cos \theta), \quad (2)$$

where

$$\theta = a - \frac{b}{(x_i^2 + y_i^2 + 1)}$$

$a = 0.4$, $b = 6.0$, $\mu = 0.9$ and “ i ” denotes the step number. Note that this system is a “map” as opposed to a “flow”. A map is discrete in time while a flow (e.g. Lorenz, 1963) is continuous.

Linearization of the Ikeda system is a crucial concept since singular and Lyapunov error growth hinge on its validity. Linearization can be explained as follows. The state space is defined by (x, y) . Consider a point (x_i, y_i) in the phase space. Consider the perturbed point given by $(\tilde{x}_i, \tilde{y}_i) = (x_i + \epsilon_{x(i)}, y_i + \epsilon_{y(i)})$ where $\epsilon_{x(i)} = \tilde{x}_i - x_i$ and $\epsilon_{y(i)} = \tilde{y}_i - y_i$ denote the perturbations (or the errors). The evolution of (x_i, y_i) and $(\tilde{x}_i, \tilde{y}_i)$ is described by the nonlinear equations, Eqs. (1) and (2). Since x_{i+1} and y_{i+1} are functions of x_i and y_i one can introduce the notation,

$$x_{i+1} = F_1(x_i, y_i)$$

$$y_{i+1} = F_2(x_i, y_i).$$

Using the Taylor series, the evolution of linearized error is given by,

$$\begin{bmatrix} \epsilon_{x(i+1)} \\ \epsilon_{y(i+1)} \end{bmatrix} = \begin{bmatrix} F'_{1x}(x_i, y_i) & F'_{1y}(x_i, y_i) \\ F'_{2x}(x_i, y_i) & F'_{2y}(x_i, y_i) \end{bmatrix} \begin{bmatrix} \epsilon_{x(i)} \\ \epsilon_{y(i)} \end{bmatrix}, \quad (3)$$

where

$$F'_{1x} = \frac{\partial F_1}{\partial x_i}$$

and so on.

The matrix on the right hand side (which is the Jacobian of Eqs. 1 and 2) is called the linear propagator,

$$\mathbf{M}_i = \begin{bmatrix} F'_{1x}(x_i, y_i) & F'_{1y}(x_i, y_i) \\ F'_{2x}(x_i, y_i) & F'_{2y}(x_i, y_i) \end{bmatrix}.$$

Suppose the error at the i^{th} step is

$$\epsilon^{(i)} = \begin{bmatrix} \epsilon_{x(i)} \\ \epsilon_{y(i)} \end{bmatrix}.$$

Then the error at the $(i + 1)^{\text{th}}$ step is

$$\epsilon^{(i+1)} = \mathbf{M}_i \epsilon^{(i)} \quad (4)$$

assuming that the magnitude of $\epsilon^{(i)}$, $\|\epsilon^{(i)}\|$, is small enough so that the linearity assumption holds. Though this relation holds approximately for finite perturbations, it holds exactly for infinitesimal perturbations. Equation (4) is the linearization of the system equations about a non-linear trajectory.

The number of steps (or the time, in the case when the system is a flow) over which \mathbf{M} is constructed is called the optimization time τ . \mathbf{M} is called the “tangent linear model”. \mathbf{M} depends on the initial conditions and the optimization time τ i.e. $\mathbf{M} = f(x_i, y_i, \tau)$. The tangent linear model is particularly useful because of this property which states that the linearized error at the n^{th} step can be evaluated by knowing the initial error and the linear propagator over each of the intermediate steps. The validity of this linearization (i.e. the proximity of the linear error evolution to the non-linear error evolution) will depend on the magnitude and direction of the initial error and τ . In general, for a given τ , the smaller the initial error, the better the validity. The number of steps over which linearity holds has to be ascertained before the linear propagator is actually utilized. In the current work the initial magnitude of errors has been chosen small enough so that linearity holds for the number of steps over which \mathbf{M} is constructed.

3 Lyapunov numbers, singular values and norms

A finite error in initial conditions grows nonlinearly. One needs to define “indices” of error growth to quantify this error growth. Lyapunov numbers and singular values are such “indices” of error growth, defined under linearized dynamics. These accurately quantify the growth of errors insofar as linearization holds.

3.1 Lyapunov numbers

Lyapunov vectors and numbers have been discussed in many papers. The following discussion has been adapted from (Ziehmann et al., 1998).

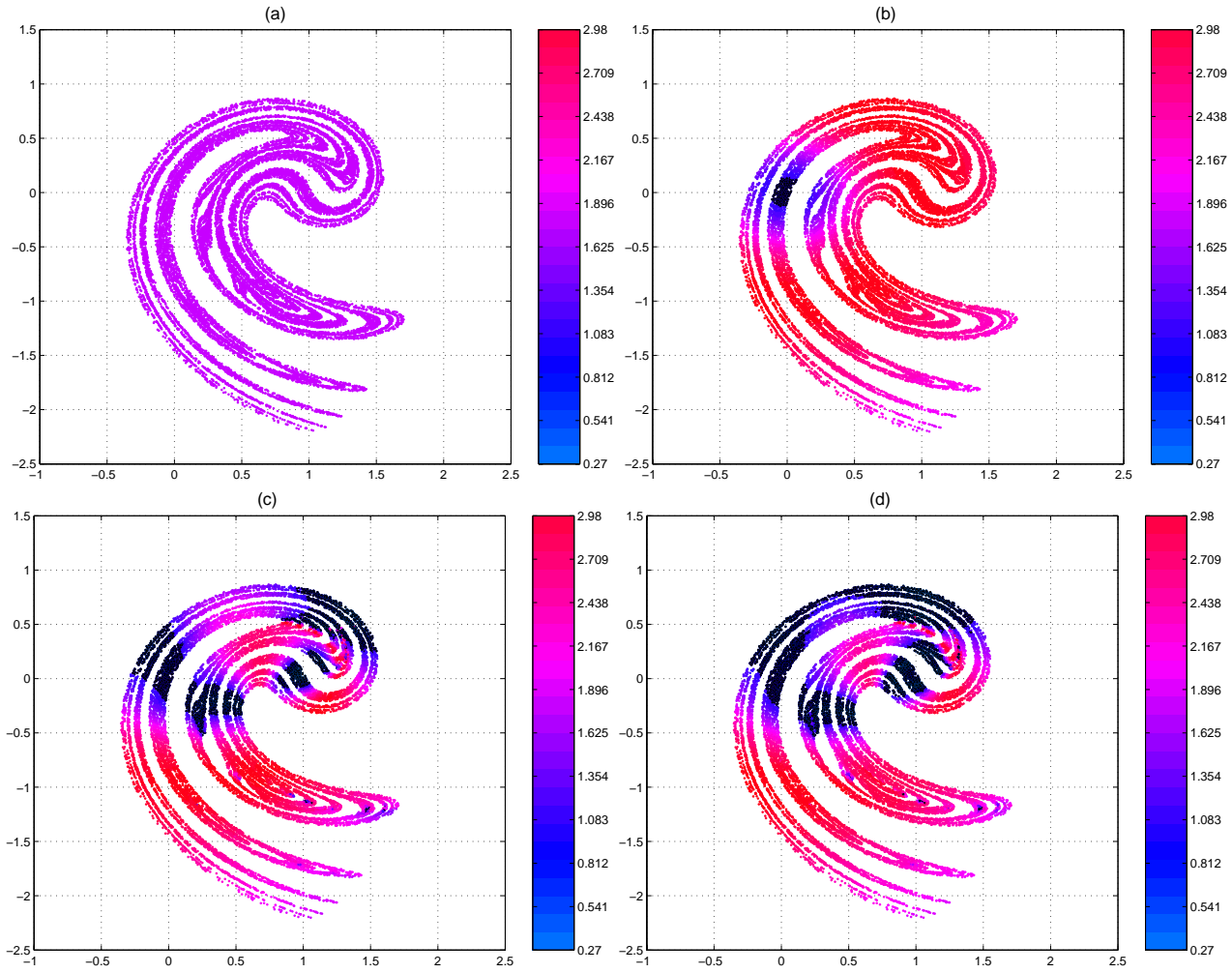


Fig. 1. Distribution of the leading singular value (Lyapunov number in case of (a)) over the Ikeda attractor for $\tau = 1$ for a perfect model (black color indicates $\sigma_1 < 0$). (a) Lyapunov number=1.66, (b) σ_1 , initial isotropic distribution, (c) σ_1^r , initial non-isotropic distribution, superscript r represents “relevant”, (d) σ_1^r initial non-isotropic distribution, \mathbf{P}^a rotated by an angle drawn from $N(0, 10^\circ)$.

Consider an m -dimensional state space. The Lyapunov exponents are defined assuming infinitesimal perturbations at time $t = 0$. As discussed in Sect. 2, the linear propagator \mathbf{M} evolves forward the initial error (or perturbation) linearly along the nonlinear trajectory over time τ . Consider the matrix,

$$\mathbf{O}(\mathbf{x}, \tau) = [\mathbf{M}(\mathbf{x}, \tau)^T \mathbf{M}(\mathbf{x}, \tau)]^{1/2\tau}$$

where the superscript T indicates the transpose of the matrix. Oseledec (Oseledec, 1968) showed that if the limit of $\tau \rightarrow \infty$ exists, then the m eigenvalues of,

$$\mathbf{O}(\mathbf{x}) = \lim_{\tau \rightarrow \infty} \mathbf{O}(\mathbf{x}, \tau)$$

are independent of the state \mathbf{x} . Then the Lyapunov exponents are defined as,

$$\Lambda_i = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \log_e(p_i)$$

for $i = 1, m$; where $p_i, i = 1, m$ are the eigenvalues of \mathbf{O} . By convention, $\Lambda_i > \Lambda_j$ for $i < j$. The leading Lyapunov number is defined as,

$$L_1 = e^{(1/\tau)\Lambda_1} = p_1.$$

L_1 gives the average “factor” of growth of the fastest growing direction over time τ . i.e. if Λ_1 has been calculated for a finite time $\tau = k\Delta t$, say where k is the number of steps and Δt is the time step, then at each step the error grows by a factor of L_1 , so that at the end of k steps it has grown by a factor of $L_1^k \Delta t$.

Figure 1a shows the one step growth factor as given by the Lyapunov number. The points shown are on the attractor and each one is colored by the single leading Lyapunov number. According to this picture, an error at any point in the state space would grow by a factor of 1.66 over $\tau = 1$. In general it would grow by a factor of 1.66^τ over τ steps. Note that by definition the error growth given by the Lyapunov number is uniform over the whole state space.

3.2 Singular vectors and values

The subject of singular vectors has been treated by numerous authors in linear algebra texts and the scientific literature (Strang, 1988). The singular vectors and values are defined for the linearized system of equations. Having calculated \mathbf{M} over optimization time τ , the singular vectors and values are defined as follows.

$$\mathbf{M} = \mathbf{U}\Sigma\mathbf{V}^T, \quad (5)$$

where

$$\Sigma = \sqrt{\Lambda}.$$

\mathbf{U} and \mathbf{V} are called the left (or final time) and right (or initial time) singular vectors, respectively and the elements of Λ , which is a diagonal matrix, are the square of the singular values of \mathbf{M} . These can be obtained by either the singular value decomposition i.e. $SVD(\mathbf{M})$ or through the eigenvectors and eigenvalues of $\mathbf{M}^T\mathbf{M}$ and $\mathbf{M}\mathbf{M}^T$. Note that \mathbf{U} and \mathbf{V} define orthonormal bases.

In the context of the Ikeda system, given a linear propagator constructed over a particular τ , \mathbf{U} is 2×2 , \mathbf{V} is 2×2 and there are 2 singular values. From Eq. (5),

$$\mathbf{M}\mathbf{v}_i = \sigma_i\mathbf{u}_i,$$

where $i = 1, 2$. The singular values can be written as,

$$\sigma_i = \frac{\|\mathbf{M}\mathbf{v}_i\|}{\|\mathbf{v}_i\|},$$

where $\sigma_1 > \sigma_2$.

These give the ratio of magnitudes of vectors aligned along \mathbf{u}_i at final time to those aligned along \mathbf{v}_i at initial time. The interpretation of these vectors and values is of immense importance from the standpoint of error growth under linearized dynamics. Given an isotropic (i.e. having the same magnitude in all directions) distribution of uncertainty at initial time, the errors aligned along \mathbf{v}_1 grow largest and at optimization time align along \mathbf{u}_1 (note that this assumes a norm, which is discussed in more detail in the next subsection). The corresponding growth factor is given by σ_1 . Those aligned along \mathbf{v}_2 grow the least with the corresponding growth factor given by σ_2 . All other directions grow by factors that are between σ_1 and σ_2 . The leading singular value places an upper limit on the prediction error of the system under linearized dynamics and isotropic uncertainty; it provides an estimate of the maximum error one can expect in the forecast.

In the case of n dimensions, the n singular values give the n growth factors associated with the n singular directions which form an orthonormal basis. An isotropic n -dimensional sphere evolves into an ellipsoid at final time whose n axes are given by the \mathbf{U} vectors.

3.3 Norms

There are several ways in which ‘‘distance’’ can be measured between two points in a phase space. This distance is defined by a function known as the norm (Lipschutz, 1991). A

given uncertainty distribution may be isotropic with respect to one norm, while being non-isotropic with respect to another. Thus it is very important to specify the norm when discussing uncertainty structure.

Let \mathbf{X} be a real or complex vector space. Suppose there is a function $\|\mathbf{x}\|$ which assigns to each $\mathbf{x} \in \mathbf{X}$ a real number. This function is called a ‘‘norm’’ on \mathbf{X} if it satisfies the following axioms:

$$\|\mathbf{x}\| \geq 0$$

$$\|\mathbf{x}\| = 0 \text{ iff } \mathbf{x} = 0$$

$$\|k\mathbf{x}\| = |k|\|\mathbf{x}\| \text{ where } k \text{ is a scalar}$$

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$$

The vector space \mathbf{X} with a norm is called a ‘‘normed vector space’’. There can be several functions which define different norms as long as they satisfy the above axioms. For instance, if x_1, x_2, \dots, x_n are components of a vector \mathbf{x} , then 3 possible norms are,

$$\|\mathbf{x}\| = \max(|x_i|), i = 1, n$$

$$\|\mathbf{x}\| = |x_1| + |x_2| + \dots + |x_n|$$

$$\|\mathbf{x}\| = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}.$$

These are called the L_∞ , L_1 and L_2 (or Euclidean) norms, respectively. The distance between two vectors $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ denoted by the function $d(\mathbf{x}, \mathbf{y})$ is dictated by the norm employed. If the L_2 norm is used then,

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + \dots + |x_n - y_n|^2}.$$

It appears that the terms ‘‘norm’’ and ‘‘metric’’ have been used synonymously in meteorology literature (Palmer et al., 1998). Predictability studies using singular vector analysis aim to find the directions in phase space at initial (i.e. analysis) time, that grow into the directions of largest error after some specified interval of time (called the optimization time). There are several choices available for the ‘‘physical dimension’’ of these perturbations such as total energy, enstrophy and streamfunction variance. The state vectors can be operated on by appropriate operators to obtain the perturbations in a particular ‘‘physical dimension’’. (The operator chosen defines a norm on the space e.g. L_2 norm, energy norm, enstrophy norm, etc., Palmer et al., 1998). In general, irrespective of the norm chosen, the analysis error structure may be non-isotropic. $SVD(\mathbf{M})$ where \mathbf{M} is the linear propagator, does not give the singular directions and values which are relevant to predictability, since the non-isotropic nature of the analysis error structure is not taken into consideration. This drawback can be overcome by using the covariance, \mathbf{P}^a , of the analysis errors in the chosen norm’s space. If $\sqrt{\mathbf{P}^{a-1}}$ is used to transform the analysis errors then the resulting transformed analysis errors are isotropic. Thus

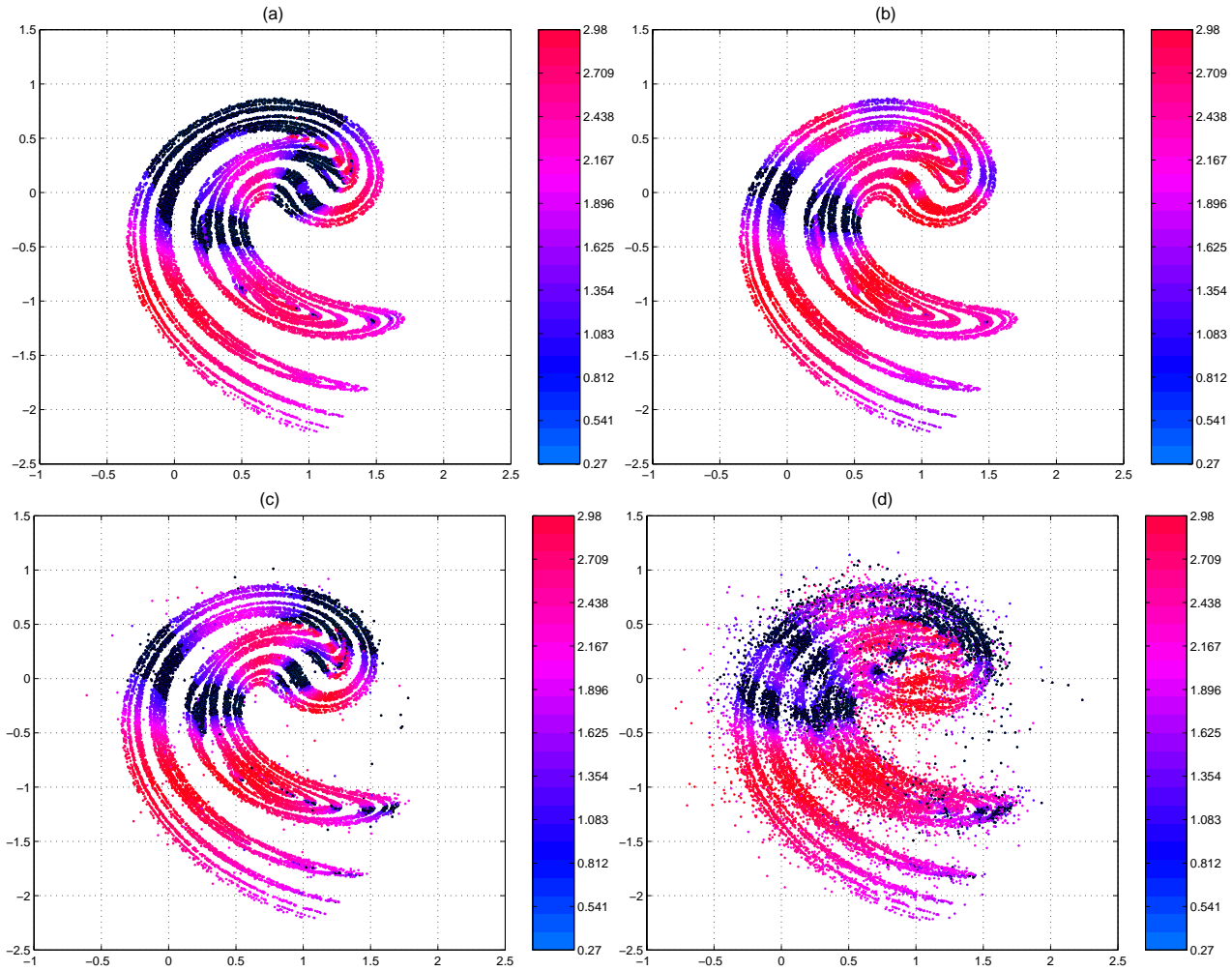


Fig. 2. Distribution of σ_1^r (where superscript r represents “relevant”) over the Ikeda attractor for $\tau = 1$ and initial non-isotropic distribution (black color indicates $\sigma_1 < 0$). (a) perfect model, \mathbf{P}^a rotated by an angle drawn from $N(0, 30^\circ)$, (b) perfect model, off diagonal terms of \mathbf{P}^a neglected i.e. only variance is considered, (c) parametrically imperfect model, $a = 0.404$, (d) structurally imperfect model, first 6 terms of sine and cosine series included.

$SVD(\mathbf{M}\sqrt{\mathbf{P}^{a-1}}^{-1})$ gives the singular directions and values relevant to predictability. It accounts for situations in which large initial uncertainty (i.e. error) combined with small error growth lead to larger final time errors than small initial uncertainty combined with large error growth. Using $\sqrt{\mathbf{P}^{a-1}}$ as a transform is equivalent to using \mathbf{P}^{a-1} as a norm. \mathbf{P}^{a-1} is called the analysis error covariance norm (Barkmeijer et al., 1998). It is also called the Mahalanobis metric (Palmer et al., 1998). Usually in realistic atmospheric models it is numerically very expensive to calculate \mathbf{P}^{a-1} . Barkmeijer et al. (1998) explores some methods of obtaining the estimates of \mathbf{P}^{a-1} . It so happens that if perturbations are measured in energy then the analysis error distribution appears to be isotropic (Palmer, 1994), which is advantageous since then \mathbf{P}^{a-1} need not be calculated. In the work that follows, the L_2 norm is used and $\sqrt{\mathbf{P}^{a-1}}$ is employed to transform the non-isotropic analysis error distribution to an isotropic one.

4 Singular vectors and values in the Ikeda system

The concept of singular vectors and values presented in Sect. 3.2 has been applied to the Ikeda system. Basically, the question addressed is, do the (leading) singular values change over the state space? This represents the state dependence of error growth under numerous operational constraints.

The concept of singular error growth in the Ikeda system can be illustrated with the following example. Figure 3 illustrates singular error growth at the point $(0, -1)$ in the Ikeda phase space. A Gaussian uncertainty distribution centered at $(0, -1)$, which is the truth, is generated at the initial time. The perturbations to be evolved by \mathbf{M} (not shown in the figure) are sampled from this distribution. The green circle in Fig. 3a shows the 2ξ bound of this distribution where ξ stands for standard deviation. This distribution is propagated using the tangent linear model. At each step the truth and 2ξ bound of the ensemble are plotted. The tangent linear

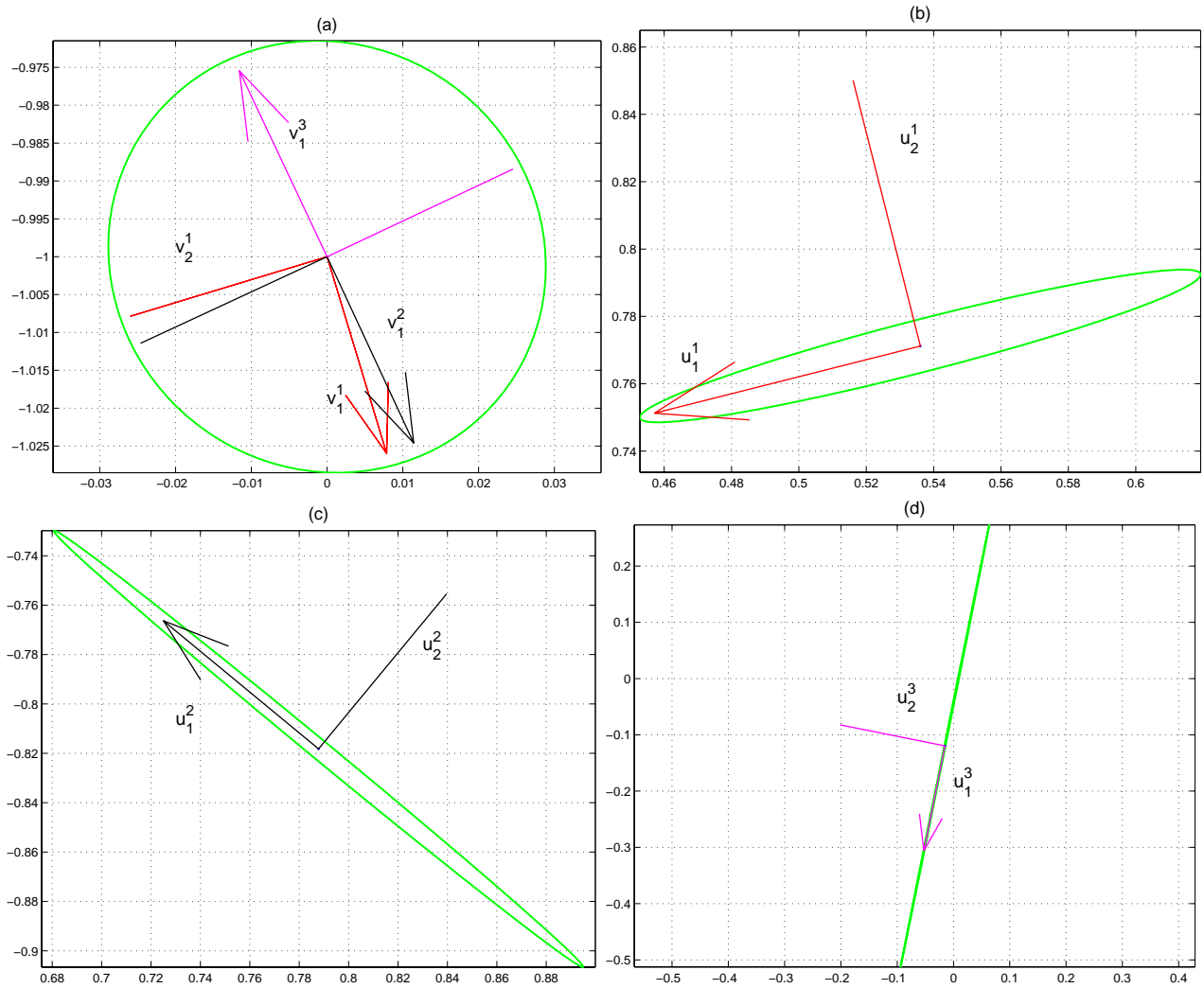


Fig. 3. Evolution of isotropic distribution over 3 time steps. \mathbf{V} (shown in (a)) are the initial time vectors. \mathbf{U} are the final time vectors which are shown in (b), (c) and (d) for $\tau = 1, 2, 3$, respectively. Superscript indicates the time step. The 2ξ bounds of distribution are shown in each figure. Note that \mathbf{U} vectors are not scaled by σ . **(a)** Initial time singular vectors are for different τ , **(b)** $\tau = 1, \sigma_1 = 2.07, \sigma_2 = 0.27$, **(c)** $\tau = 2, \sigma_1 = 4.77, \sigma_2 = 0.13$, **(d)** $\tau = 3, \sigma_1 = 13.77, \sigma_2 = 0.03$.

model stretches (or shrinks) and rotates each of the perturbation vectors so that at each step an ellipse is generated. Figures 3b, 3c and 3d show the ellipses at $\tau = 1, \tau = 2$ and $\tau = 3$, respectively. Firstly, one is interested in the direction in the future which has the maximum variance at the forecast time (or optimization time τ) and the variance in this direction relative to the initial variance can be regarded as a quantitative measure of predictability. Secondly one is interested in knowing the direction at initial time that grows into direction of maximum variance at τ . This knowledge can be used to help one to reduce the error in the relevant initial direction so that forecast errors at final time can be reduced.

The \mathbf{U} , \mathbf{V} and Σ are calculated by the SVD of \mathbf{M} constructed over each τ , using the L_2 norm on the space and assuming an isotropic distribution of uncertainty at initial time. The major and minor axis of these ellipses are given by \mathbf{u}_1 and \mathbf{u}_2 , respectively. The directions at initial time that

evolve into these directions at final time are given by \mathbf{v}_1 and \mathbf{v}_2 . Note that each time step has a unique set of \mathbf{U} and \mathbf{V} vectors, with the corresponding growth factors given by the σ 's. Hence, for instance, over one step any error aligned along \mathbf{v}_1^1 direction (shown in red, with arrowhead) will increase in magnitude by 2.97 (which is the maximum) while that aligned along \mathbf{v}_2^1 direction (shown in red) will shrink by a factor of 0.27 (which is the minimum). Note that the superscripts indicate the time step while the subscripts give the index of the singular vector (subscript 1 indicates leading singular vector). At final time, these directions will align themselves along \mathbf{u}_1^1 and \mathbf{u}_2^1 , respectively as shown in Fig. 3b. For $\tau = 2$ errors aligned along the direction shown in black (with arrowhead) in Fig. 3a, will grow the most (by a factor of 4.77) and align itself along the direction shown in Fig. 3c. Similarly the directions shown in magenta color in Fig. 3a, evolve to those in Fig. 3d over $\tau = 3$, with the maximum

growth factor of 13.77 and a minimum of 0.03. As $\tau \rightarrow \infty$ the \mathbf{U} and \mathbf{V} vectors converge to the same vectors.

Thus the singular vector formulation is very remarkable in that it gives (1) the directions that would grow the most (2) the directions to which they would grow and (3) the factors by which they grow. The leading singular value represents the worst case scenario for forecast error under the linearization assumption.

4.1 Isotropic initial uncertainty

The above procedure of calculating the leading singular value has been applied to a large number of points in the state space so that its variation can be inspected. This motivates the superiority of singular values over the Lyapunov number insofar as finite time forecast error studies are concerned. Consider Fig. 1a. The ‘‘average over space via an average over time’’ implicit in the definition of the Lyapunov number, gives this uniform distribution, which belies the true nature of finite time error growth, as is evident in Fig. 1b.

Figure 1b, shows the leading singular value for $\tau = 1$ at many points, as calculated assuming isotropic initial uncertainty. This work has been done earlier (Smith et al., 1999); but has been presented as a background to the results to follow. The L_2 norm, as defined in Sect. 3 has been used to measure distances in the phase space. The error growth factor as given by the singular value in Fig. 1b shows significant structure and variation. The errors in some parts of the state space could grow by as much 2.97 while in other parts they could actually shrink (these areas having $\sigma_1 < 1$ are colored in black). A comparison of Fig. 1a and Fig. 1b shows that the Lyapunov number is in fact quite a dubious measure of finite time error growth. Note that the Lyapunov number has the same value of 1.66 everywhere in the state space. In areas like those around the origin, the forecast launched for $\tau = 1$ would be very good, as is evident from Fig. 1b. The Lyapunov number gives a false impression that the forecast will be poor in this region. On the other hand in regions like those around (1, 0.5) the Lyapunov number artificially enhances the forecast quality, while in reality the forecast will be quite poor as is evident from Fig. 1b. The Lyapunov number is all the more misleading for longer optimization times (results not shown). For instance over 2 steps the Lyapunov factor would be 1.66^2 while it is seen that singular value varies quite a lot over the state space. This comparison gives cogent reasons to decisively discard Lyapunov number as a measure of finite time error growth and adopt the singular value as a far better substitute. The error growth as given by the singular value is state and time dependent. The information given by the singular values not only diagnose the forecast quality over state space but also could be useful for data assimilation or targeting. Thus if the initial distribution of errors is isotropic (in the L_2 norm on the space) and the model is perfect then Fig. 1b indeed gives a genuine picture of forecast quality.

Operationally, data assimilation is used to generate initial uncertainty distribution which renders it non-isotropic with

respect to the L_2 norm. The next sub-section gives a brief overview of data assimilation using the Ensemble Kalman filter (EnKF).

4.2 Data assimilation

In this work the EnKF data assimilation (Evensen, 1994; Houtekamer et al., 1998) will be used to produce an analysis and associated uncertainty. Let n be the dimension of the state vector \mathbf{x} . Let the number of ensemble members be N . The EnKF equations can be described as follows. For $i = 1, N$,

$$\mathbf{x}_i^f(t) = \mathbf{F}[\mathbf{x}_i^f(t-1)] \quad (6)$$

$$\mathbf{P}^f(t) = \frac{1}{N-1}[\mathbf{A}^f(t) - \mathbf{X}^f(t)][\mathbf{A}^f(t) - \mathbf{X}^f(t)]^T \quad (7)$$

$$\mathbf{K}(t) = \mathbf{P}^f(t)\mathbf{H}(t)^T[\mathbf{H}(t)\mathbf{P}^f(t)\mathbf{H}(t)^T + \mathbf{R}(t)]^{-1} \quad (8)$$

$$\mathbf{x}_i^a(t) = \mathbf{x}_i^f(t) + \mathbf{K}(t)[\mathbf{y}_i^o(t) - \mathbf{H}(t)\mathbf{x}_i^f(t)] \quad (9)$$

$$\mathbf{P}^a(t) = \frac{1}{N-1}[\mathbf{A}^a(t) - \mathbf{X}^a(t)][\mathbf{A}^a(t) - \mathbf{X}^a(t)]^T.$$

$\mathbf{x}_i^f(t)$ is the first guess generated by a short term forecast. $\mathbf{X}^f(t)$ and $\mathbf{X}^a(t)$ are $n \times N$ matrices which give the ensembles before and after assimilation, respectively. Thus each column of $\mathbf{X}^f(t)$ and $\mathbf{X}^a(t)$ is an ensemble member i.e. a different realization of the state and each row is a state component. $\mathbf{A}^f(t)$ and $\mathbf{A}^a(t)$ are $n \times N$ matrices having the corresponding ensemble mean of $\mathbf{X}^f(t)$ and $\mathbf{X}^a(t)$, respectively, in each of its column. The mean is calculated by averaging over each row of $\mathbf{X}^f(t)$ and $\mathbf{X}^a(t)$. Thus a particular element in each column of $\mathbf{A}^a(t)$ (which is the same as each of other N columns) is an average over the corresponding row in $\mathbf{X}^a(t)$. $\mathbf{P}^f(t)$ (dimension $n \times N$) gives the uncertainty in the first guess in the form of its covariance matrix. $\mathbf{R}(t)$ (dimension $m \times N$, where m is the number of state variables observed; for this work $m = n$) gives the uncertainty in the observations. The Kalman gain term $\mathbf{K}(t)$ (dimension $n \times m$) gives a measure of confidence one should place in the first guess and the observations depending on their respective uncertainties. $\mathbf{H}(t)$ is a map from the model space to observation space. $\mathbf{y}_i^o(t)$ are the perturbed observations. The perturbed observations are obtained by renoising the observation mean $\mathbf{y}_o(t)$ i.e. $\mathbf{y}_i^o(t) = \mathbf{y}_o(t) + \sqrt{\mathbf{R}}N(0, 1)$, where $N(0, 1)$ is the standard Normal distribution (see Burgers et al. (1998) for details). $\mathbf{x}_i^a(t)$ is the so called analysis which combines the first guess uncertainty and the observation uncertainty, thus providing the best estimate of initial condition. Finally $\mathbf{P}^a(t)$ is the covariance matrix that provides an estimate of uncertainty in the analysis. The Eqs. (6) through (9) assume a perfect model.

All NWP centers employ data assimilation to produce the best possible initial conditions. It is difficult to imagine that the initial condition produced operationally by data assimilation is isotropic. This is because the first guess is produced by

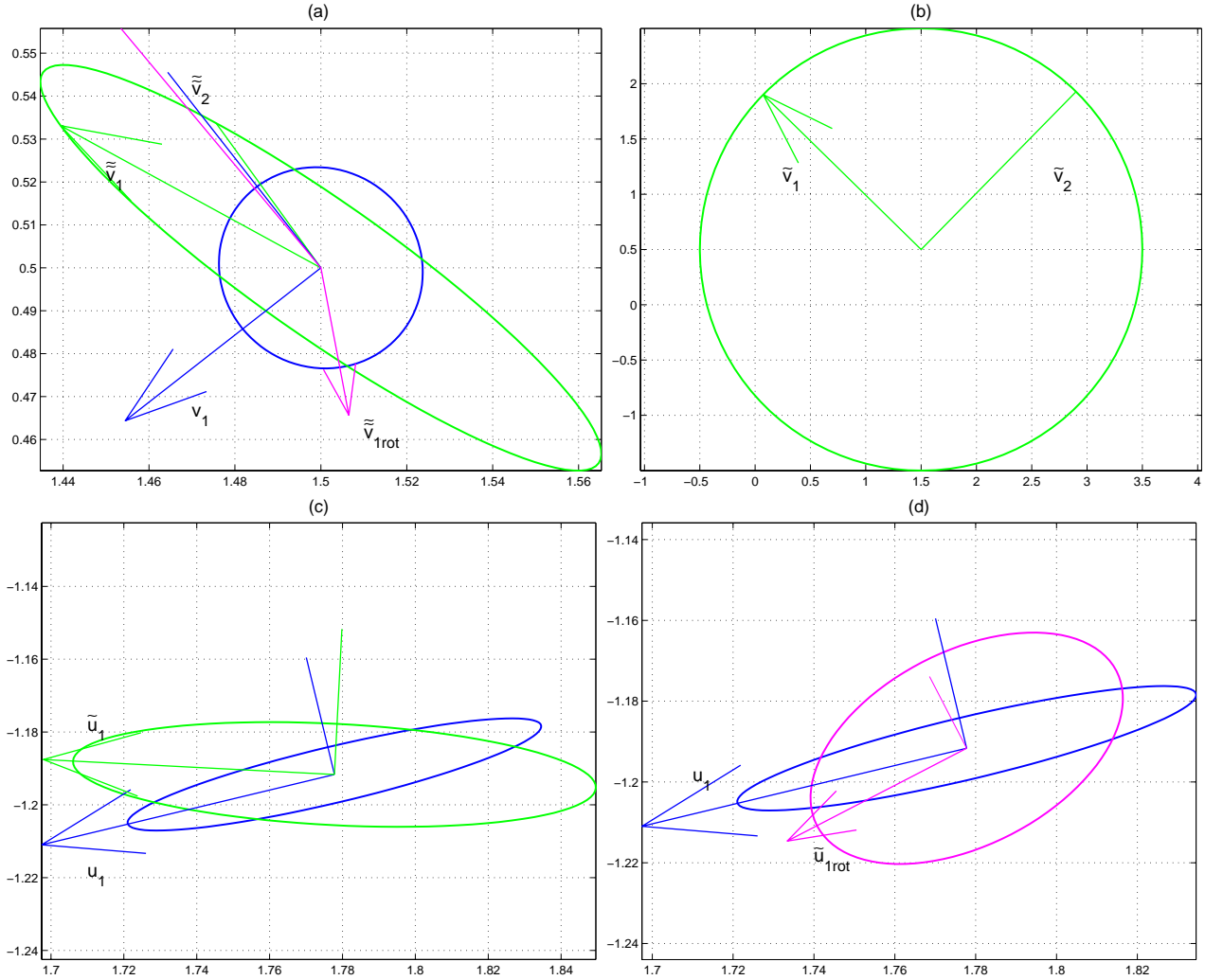


Fig. 4. Singular vectors for non-isotropic distribution over $\tau = 1$. The 2ξ bounds of distribution are shown. **(a)** isotropic (blue) and non-isotropic initial (green) distributions and corresponding initial vectors. $SVD(\mathbf{M})$ gives the \mathbf{V} (shown in blue), \mathbf{U} vectors and $\sigma_1 = 2.52$, $\sigma_2 = 0.32$ are the singular values. These are relevant only for the isotropic distribution. The relevant singular directions for the non-isotropic distribution are \tilde{v}_1 and \tilde{v}_2 , shown in green. **(b)** each vector in the non isotropic distribution in (a) is operated on by $\sqrt{\mathbf{P}^{a-1}}$. The distribution (in green) which appears isotropic in this transformed space is shown. The relevant initial singular directions $\tilde{\mathbf{V}}$ are given by $SVD(\mathbf{M}\sqrt{\mathbf{P}^{a-1}}^{-1})$ which are shown in green. The orthogonal $\tilde{\mathbf{V}}$ in (b) are transformed back ($\tilde{\mathbf{V}} = (\sqrt{\mathbf{P}^{a-1}}^{-1})\tilde{\mathbf{V}}$) to the original space and plotted as $\tilde{\mathbf{V}}$ in (a). Note that they are no longer orthogonal. **(c)** the evolved isotropic and non-isotropic distributions after $\tau = 1$ and final time singular vectors. \mathbf{U} and $\tilde{\mathbf{U}}$ are relevant to the isotropic and non-isotropic distributions, respectively. The relevant leading singular value for non isotropic distribution is $\sigma_1^r = 1.04$, where the subscript r represents “relevant”. $\sigma_1 = 2.52$ for the isotropic distribution. **(d)** the evolved distribution after $\tau = 1$ from rotated ($\theta = 20^\circ$) initial distribution (not shown in (a)) is shown in pink. The singular values remain almost the same but the singular directions (green in (c) and pink in (d)) are very different. The pink arrows in (a) show the initial singular vectors for the distribution with rotated \mathbf{P}^a .

propagating the model state forward and the evolved uncertainty distribution will be elliptical. It would take a very special observational uncertainty distribution to render the analysis uncertainty distribution isotropic. Hence if one wants to carry out the singular analysis in a realistic scenario then a non-isotropic error distribution has to be assumed rather an isotropic one (with respect to the L_2 norm in this case).

4.3 Impact of non-isotropic distribution

As stated above, when data is assimilated the initial uncertainty distribution does not remain isotropic. The question then is whether the directions given by the $SVD(\mathbf{M})$ still give the “relevant” singular values and initial and final time singular vectors. The methodology to find these “relevant” values is discussed below and illustrated in Fig. 4.

Figure 4a shows an isotropic (blue circle) and a non-isotropic (green ellipse) distribution of uncertainty at initial time (the circle and ellipse shown are the 2ξ bounds for the ensembles). The initial time singular vectors as given by the $SVD(\mathbf{M})$ satisfy,

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

so that,

$$\mathbf{M}\mathbf{v}_i = \sigma_i \mathbf{u}_i,$$

where $i = 1, 2$.

The uncertainty distribution is propagated over one step. The resulting distribution (2ξ bound) is shown in Fig. 4c. The blue arrows, \mathbf{v}_1 and \mathbf{v}_2 in Fig. 4a are the initial time singular vectors for an isotropic distribution (i.e. by using $SVD(\mathbf{M})$) (the one with the arrow head is the leading). The corresponding singular values are given by $\sigma_1 = 2.52$ and $\sigma_2 = 0.32$. Any errors (whether in the isotropic or non-isotropic distribution) aligned along \mathbf{v}_1 in Fig. 4a will grow by a factor of 2.52 and align themselves along \mathbf{u}_1 in Fig. 4c. Accordingly, the blue ellipse in Fig. 4c corresponding to the initial isotropic distribution in Fig. 4a has its major and minor axis along the \mathbf{U} vectors. The initial time non-isotropic (green) ellipse in Fig. 4a evolves into the final time (green) ellipse in Fig. 4c. But the (green) major and minor axis of this evolved ellipse in Fig. 4c “are not” along the \mathbf{U} vectors. In fact the major axis of the evolved green ellipse in Fig. 4c is aligned along a different direction labeled $\tilde{\mathbf{u}}_1$. The variance which is maximum along this direction characterizes the operationally relevant predictability, rather than that along \mathbf{u}_1 . The next question is what is the direction at initial time in Fig. 4a that evolved into the $\tilde{\mathbf{u}}_1$ direction at final time? The procedure to find these directions is delineated below.

Firstly, the non-isotropic distribution in Fig. 4a is operated on by $\sqrt{\mathbf{P}^{a-1}}$, where \mathbf{P}^a is the covariance matrix of the non-isotropic distribution. This operation converts the non-isotropic distribution to an isotropic distribution (Fig. 4b, green circle is 2ξ bound).

The singular vectors and values are then given by the $SVD(\mathbf{M}\sqrt{\mathbf{P}^{a-1}})$. Thus,

$$\mathbf{M}\tilde{\mathbf{v}}_i = \tilde{\sigma}_i \tilde{\mathbf{u}}_i,$$

where $i = 1, 2$.

Hence the initial direction (in the transformed space) that grows most over $\tau = 1$ is given by $\tilde{\mathbf{v}}_1$. But this direction is valid only in the transformed space. The corresponding direction in the original space is given by,

$$\tilde{\tilde{\mathbf{v}}}_1 = (\sqrt{\mathbf{P}^{a-1}})^{-1} \tilde{\mathbf{v}}_1.$$

It follows that,

$$\sigma_1^r = \frac{\tilde{\sigma}_1}{\|\tilde{\tilde{\mathbf{v}}}_1\|}.$$

Similarly,

$$\sigma_2^r = \frac{\tilde{\sigma}_2}{\|\tilde{\tilde{\mathbf{v}}}_2\|}.$$

Here the superscript r stands for “relevant”.

The $\tilde{\mathbf{v}}_1$ and $\tilde{\mathbf{v}}_2$ are shown as green arrows in Fig. 4a. Note that they are not orthogonal. These directions can be interpreted as follows. Any perturbation along $\tilde{\mathbf{v}}_1$ grows by a factor of $\sigma_1^r = 1.04$ which is smaller than the factor by which an error aligned along \mathbf{v}_1 will grow ($\sigma_1 = 2.52$). But, the error along $\tilde{\mathbf{v}}_1$ which ends up aligning itself along $\tilde{\mathbf{u}}_1$ at final time happens to be larger than that along \mathbf{u}_1 . This is because though \mathbf{v}_1 has a much larger growth factor than the $\tilde{\mathbf{v}}_1$ direction, the error in that direction is smaller at initial time. So the direction at initial time (in Fig. 4a) that ends up as direction of maximum variance at final time (in Fig. 4c), is dictated both by the model dynamics and the structure of the initial uncertainty. In fact there are many directions in Fig. 4a which have higher growth factors than $\tilde{\mathbf{v}}_1$ and there are also directions which have larger initial uncertainty (e.g. the major axis of the green ellipse in Fig. 4a), but $\tilde{\tilde{\mathbf{v}}}_1$ is the direction that has the “right” combination of these two properties and thus it aligns itself in the direction of maximum variance at the final time. Thus the relevant singular directions and values when the “initial uncertainty distribution is non-isotropic” is given by the $SVD(\mathbf{M}\sqrt{\mathbf{P}^{a-1}})$ rather than $SVD(\mathbf{M})$. The $\tilde{\tilde{\mathbf{V}}}$, $\tilde{\tilde{\mathbf{U}}}$ and $\tilde{\tilde{\Sigma}}_r$ take into account both the uncertainty structure and the dynamics of error growth. Note that the singular vector and values given by $SVD(\mathbf{M})$ are correct, in the sense that perturbations in the non-isotropic distribution aligned along \mathbf{v}_1 grow by a factor of σ_1 ; it is just that they are not relevant from the stand point of finite time forecasting for this case.

The fact that the singular vectors and values relevant for the non-isotropic case are different from those for isotropic case has important implications for forecasting. Assuming an isotropic distribution, when it is actually not, misinforms one about the direction at initial time that will end up having maximum variance at final time and the singular values could actually lead to inordinate estimates of forecast error. $SVD(\mathbf{M})$ gives the largest growing initial time direction while $SVD(\mathbf{M}\sqrt{\mathbf{P}^{a-1}})$ (Ehrendorfer and Tribbia, 1995) gives the initial time direction which results in the largest error at forecast time. $SVD(\mathbf{M})$ considers the dynamics but neglects the uncertainty structure i.e. assumes it to be isotropic. $SVD(\mathbf{M}\sqrt{\mathbf{P}^{a-1}})$ gives the relevant information from the point of view of predictability since it tells what current errors are going to impact the forecast the most. $SVD(\mathbf{M}\sqrt{\mathbf{P}^{a-1}})$ takes into account both the dynamics and non-isotropic structure of the initial uncertainty distribution.

The operational reality of non-isotropic initial uncertainty has been accounted for in Fig. 1c. The EnKF procedure outlined in Sect. 4.2 is implemented at each step with $N = 100$. To begin with an initial isotropic ensemble is constructed around a point (i.e. truth) with a prescribed standard deviation ($\epsilon^0 = 1\%$ of attractor size, which is given by the average of the standard deviations of x and y over the attractor) in the x and y components, by drawing from a standard normal distribution. Then the step given by Eq. (6) is implemented

by propagating each of the ensemble members forward using Eqs. (1) and (2). The truth is propagated forward as well. Data is generated artificially by first displacing the truth in a random direction by ϵ^0 , and then an ensemble of observations is generated around this displaced value of truth, with a prescribed standard deviation (ϵ^0), drawing from a standard normal distribution. The matrix \mathbf{R} has the variances of x and y observational errors along its diagonal. The gain term \mathbf{K} is evaluated by using \mathbf{R} and the covariance of the ensemble, \mathbf{P}^f (\mathbf{H} is set equal to the identity matrix). Finally data is assimilated using Eq. (9), to produce the “best” ensemble i.e. assimilated ensemble. The data is assimilated at each step. At each step a forecast is launched from the mean of the assimilated ensemble and the resulting singular vectors and values are calculated according to the method outlined above (i.e. $SVD(\mathbf{M}\sqrt{\mathbf{P}^{a-1}}^{-1})$ is used to obtain the leading singular value). The ensemble mean analysis is then colored by the leading singular value.

The picture in Fig. 1c is profoundly different from the isotropic case in Fig. 1b. There are far more regions where the growth factor associated with the initial direction that grows into the major axis of the final time ellipse is less than 1. (This information can be used, for example, for deciding the direction in which the errors need to be reduced in order to get better forecasts). The differences are more striking for higher values of τ (results not shown). This is the relevant picture of the variation of 1 step forecast error over state space for a perfect model and imperfect observations, not the isotropic case of Fig. 1b.

By considering the non-isotropic uncertainty distribution one has taken an important step toward simulating operational reality, but there are other issues that remain to be addressed. For instance it would be overtly cavalier to assume that one has access to the correct covariance matrix in operational forecasting. In all operational forecasting schemes the covariance matrix used is erroneous. Because the ways in which the operational covariance matrices are wrong are unknown, we explore the implications of an incorrect covariance matrix by introducing a very simple-minded error: we randomly rotate the eigenvectors of the covariance matrix. This issue has been treated in Fig. 4d. The direction (in pink, with arrowhead) in Fig. 4a denoted by $\tilde{\mathbf{v}}_{1rot}$ corresponds to the leading initial time singular vector as calculated by rotating \mathbf{P}^a through an angle of $\theta = 20^\circ$ (the 2ξ bound of this rotated distribution is not shown in Fig. 4a). This rotated distribution is propagated forward by using the tangent linear model and their 2ξ bound is shown in Fig. 4d as the pink ellipse. The corresponding final time singular vector is denoted by $\tilde{\mathbf{u}}_{1rot}$. It is observed that though the singular directions change dramatically, the singular values remain almost the same ($\sigma_1^r \simeq \sigma_{1rot}^r$ and $\sigma_2^r \simeq \sigma_{2rot}^r$). Figures 1d and 2a show the distribution of σ_1 when \mathbf{P}^a is rotated through an angle drawn from $N(0, 10^\circ)$ and $N(0, 30^\circ)$, respectively. This means at each point from which the forecast is launched, \mathbf{P}^a is rotated through a different angle chosen at random from a Gaussian distribution with standard deviation 10° and 30° .

Note that for smaller rotations the distribution remains the same (Figs. 1c and 1d), but it changes dramatically for larger rotations (Figs. 1c and 2a). This demonstrates that for this system small angle errors can be tolerated well, but we again emphasize that this is only one simple form of error in the covariance matrix.

A variant of this operational issue is when the error variances are available but the error covariance are not. In this case the distribution of singular values will look like those in Fig. 2b. Note that it looks intermediate to the isotropic and non-isotropic case. For this model using variances only is better than ignoring the uncertainty structure.

The above discussion took into consideration some important aspects of finite time operational forecasting, which mainly concerned the initial condition uncertainty. Model inadequacy remains one of the most important stumbling blocks in the path towards accurate forecast. In fact it is speculated that model inadequacy could be vitiating the forecast more seriously than the initial condition uncertainty (Orrell et al., 2001). The next section deals with the impact of model inadequacy on forecast errors.

4.4 Impact of model inadequacy

An idea about the impact of model inadequacy on predictability can be obtained by looking at its effect on the singular value distribution. The model error can be broadly classified into parametric and structural categories. The parametric model error pertains to having the “wrong” value of constants in a given model but the correct structure. In the Ikeda system, the constants are given by a , b and μ . The sensitivity of singular values to change in these parameters has been studied. The Ikeda system given by Eqs. (1) and (2) is treated as perfect model (or the system). Figure 2c shows the singular values calculated from the Ikeda system in which the value of parameter a is higher by 1% as compared to that in the perfect model. Data assimilation is performed as in Sect. 4.3, so that Fig. 2c shows the impact of having both a initial non-isotropic distribution and an imperfect model.

Comparing Fig. 1c with Fig. 2c one notes that the singular values do not change much with small changes in a . Similarly, it is noted that for small changes in values of μ and b the singular values do not change appreciably (result not shown). Hence, one can conclude that the parametric error does not affect the singular values to a large extent, for this range of parametric error for this model. The impact is of course dependent on τ . Larger impact is seen for longer τ which means that models with parametric error are “good” models for this system for short τ .

Structural model error pertains to a situation in which the functional dependence of one or more terms in the model is different from the “correct” dependence. In the Ikeda model such an error has been incorporated by replacing the cosine and sine terms in the model (Eqs. 1 and 2) by the first 6 terms in their series expansion. Thus the model with structural

error is given by,

$$x_{i+1} = 1 + \mu \left(x_i \sum_{n=0}^6 \frac{(-1)^n \theta^{2n}}{2n!} - y_i \sum_{n=0}^6 \frac{(-1)^n \theta^{2n+1}}{(2n+1)!} \right) \quad (10)$$

$$y_{i+1} = \mu \left(x_i \sum_{n=0}^6 \frac{(-1)^n \theta^{2n+1}}{(2n+1)!} + y_i \sum_{n=0}^6 \frac{(-1)^n \theta^{2n}}{2n!} \right), \quad (11)$$

where

$$\theta = a - \frac{b}{(x_i^2 + y_i^2 + 1)}$$

and $a = 0.4$, $b = 6.0$, $\mu = 0.9$.

Data is assimilated as in the case of parametric error discussed above. The observations for assimilation are drawn from the system (i.e. truth), while the ensemble members are propagated using Eqs. (10) and (11).

Results are shown in Fig. 2d. Comparing Fig. 1c to Fig. 2d one sees that there is large difference between the distribution of singular values. Particularly in the region centered around (1, 0.5), the regions of shrinking errors have decreased considerably. One sees similar behavior around (0, 0). Also note that the structural error changes the attractor structure and its effect is seen clearly around the region (1, 0). For longer optimization times the impact of structural model error is more spectacular. Figure 2d shows the impact of having a structurally imperfect model and non-isotropic initial condition. This is the scenario which is closest to the operational reality. Note the large change in the distribution of the leading singular value in going from theoretical simplification (Fig. 1a to the operational reality (Fig. 2d).

5 Conclusions

A cascade of predictability scenarios was presented in the context of a 2-dimensional chaotic map, ranging from theoretical idealization to something akin to operational reality, thus eliciting some of the issues impending accurate forecasts.

The singular value does a far better job in representing finite time forecast errors than the Lyapunov number. The leading singular value calculated assuming isotropic initial uncertainty distribution changes over the phase space. Thus forecast errors under “linearized dynamics” is “state dependent”, over “finite time scales”. In some regions of the state space all initial errors actually “decay”, showing local enhanced forecast quality. The regions of decaying initial errors persist, (though they change location) for longer optimization times. The singular error growth in most parts of the attractor is rapid, in that errors can grow almost by a factor of 3 over one step.

Operationally, due to noisy observations and assimilation of data, the initial uncertainty is always non-isotropic. The leading singular direction gives the direction of maximum growth of initial errors. It is shown that these relevant directions are quite different from those calculated assuming

an isotropic distribution. These directions depend both on the growth factor given by the local dynamics and the error magnitude given by the uncertainty structure. The method to obtain these relevant directions and growth factors is the $SVD(\mathbf{M}\sqrt{\mathbf{P}^{a-1}}^{-1})$ rather than $SVD(\mathbf{M})$. The leading singular value distribution is quite different in the non-isotropic case. Given the fact that calculation of covariance matrix might be very expensive, using only the variances is a reasonable approximation to using the full covariance matrix. It is found that the singular values are robust to small rotations of the covariance matrix while the singular vectors are fragile to such rotations for this system.

Model error changes the distribution of singular values as compared to the perfect model case. Qualitatively, it appears that for the level of parametric and structural errors specified in Sect. 4.4 the structural error induces more change in the distribution of singular values as compared to the perfect model than the parametric error for this model. It is, of course, impossible to generalize this result to other models. For longer forecast times, the perfect model assumption gives a highly misleading picture of error growth.

Though a method to find the relevant singular values and vectors for initial non-isotropic uncertainty is available, it seems difficult to get around the model inadequacy problem. When only an imperfect model is at hand one is doomed to get the wrong singular values. The comparison of scenarios presented here clearly shows that idealizations (isotropic initial uncertainty and perfect model) can seriously hamper accurate forecasts and motivates research into ways to reduce model inadequacy.

Acknowledgements. The work was supported by NSF/ITR Grant ACI-0121182 and ONR TIP N00014-01-0473. The authors thank L. Smith.

Edited by: A. Osborne

Reviewed by: two referees

References

- Barkmeijer, J., VanGijzen, M., and Bouttier, F.: Singular Vectors and Estimates of the Analysis-error Covariance Metric, Quarterly Journal of the Royal Meteorological Society, 124, 1695–1713, 1998.
- Burgers, G., van Leeuwen, P. J., and Evensen, G.: Analysis Scheme in the Ensemble Kalman Filter, Monthly Weather Review, 126, 1719–1724, 1998.
- Dimet, F. X. L. and Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, Tellus, 38A, 97–110, 1986.
- Ehrendorfer, M. and Tribbia, J.: Optimal prediction of forecast error covariances through singular vectors, Journal of the Atmospheric Sciences, 54, 286–313, 1995.
- Evensen, G.: Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model, J. Geophys. Res., 97, 17 905–17 924, 1992.

- Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99(C5), 10 143–10 162, 1994.
- Hansen, J. A.: Probabilistic noise reduction, *Tellus*, 53A, 587–598, 2001.
- Houtekamer, P. L., Mitchell, P., and Herschel, L.: Data Assimilation Using an Ensemble Kalman Filter Technique, *Monthly Weather Review*, 126, 796–811, 1998.
- Ikeda, K.: Multiple-valued stationary state and the instability of the transmitted light by a ring cavity, *Opt. Commun.*, 30, 257–261, 1979.
- Lipschutz, S.: *Linear algebra*, McGraw-Hill, 1991.
- Lorenz, E. N.: Deterministic nonperiodic flow, *J. Atmos. Sci.*, 20, 130–141, 1963.
- Lorenz, E. N.: A study of the predictability of a 28-variable atmospheric model, *Tellus*, 17, 321–333, 1965.
- Orrell, D., Smith, L., Barkmeijer, J., and Palmer, T.: Model error in weather forecasting, *Nonlinear Processes in Geophysics*, 8, 357–371, 2001.
- Oseledec, T. N.: A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems, *Trans. Moscow Math. Soc.*, 19, 197–231, 1968.
- Palmer, T. N.: Singular Vectors and the Predictability of Weather and Climate, *Philosophical Transactions of The Royal Society of London A*, 348, 459–475, 1994.
- Palmer, T. N., Gelaro, R., Barkmeijer, J., and Buizza, R.: Singular vectors, metrics and adaptive observations, *Journal of Atmospheric Sciences*, 55, 633–653, 1998.
- Smith, L. A., Ziehmann, C., and Fraedich, K.: Uncertainty dynamics and predictability in Chaotic systems, *QJRM*, 125, 2855–2886, 1999.
- Strang, G.: *Linear algebra and its Applications*, Penguin, 1988.
- Thompson, P. D.: Uncertainty of initial state as a factor in the predictability of large scale atmospheric flow patterns, *Tellus*, 9, 275–295, 1957.
- Ziehmann, C., Smith, L. A. and Kurths, J.: The bootstrap and Lyapunov exponents in deterministic chaos, *Physica. D*, 126, 49–59, 1998.