

The role of large-scale spatial patterns in the chaotic amplification of perturbations in a Lorenz'96 model

By S. HERRERA^{1*}, D. PAZÓ¹, J. FERNÁNDEZ² and M. A. RODRÍGUEZ¹, ¹*Instituto de Física de Cantabria (IFCA), CSIC-UC, 39005 Santander, Spain;* ²*Department of Applied Mathematics and Computer Science, Universidad de Cantabria, 39005 Santander, Spain*

(Manuscript received 18 January 2011; in final form 27 June 2011)

ABSTRACT

The preparation of perturbed initial conditions to initialize an ensemble of numerical weather forecasts is a crucial task in current ensemble prediction systems (EPSs). Perturbations are added in the places where they are expected to grow faster, in order to provide an envelope of uncertainty along with the deterministic forecast. This work analyses the influence of large-scale spatial patterns on the growth of small perturbations. Therefore, we compare Lyapunov vector (LV) definitions, used in the initialization of state-of-the-art EPSs, with the so-called characteristic LVs. We test the dynamical behaviour of these LVs in the two-scale Lorenz'96 system. We find that the commonly used definitions of LVs include non-intrinsic and spurious effects due to their mutual orthogonality. We also find that the spatial locations where the small-scale perturbations are growing are 'quantized' by the large-scale pattern. This 'quantization' enhances the artificial disposition of the LVs, which is only avoided using the characteristic LVs, an unambiguous basis which may also be of great use in realistic models for assessing or initializing EPSs.

1. Introduction

The general circulation of the atmosphere leads to the development of a patterned (non-random) spatial structure in meteorological fields at large scales. Smaller-scale features develop within the large-scale flow and their ability to grow—or even exist—depends on the large-scale environment. Moreover, small-scale features may amplify and contribute to the development of the large-scale patterns observed at a given time. As an example, upper-level geopotential height in mid-latitudes shows a large-scale wavy pattern associated with the meandering of the jet stream. Smaller-scale pressure systems can only develop where the large-scale troughs and ridges allow. In this sense, the weather forecast error growth at the small scales is affected by the large-scale spatial patterns. This work analyses the effect of the large-scale spatial patterns on the spatio-temporal growth of perturbations at small scales and the consequences for state-of-the-art probabilistic forecasting techniques.

In recent decades, modern forecasting techniques included some measure of uncertainty allowing for a probabilistic forecast that improves the deterministic results. In particular, the ensemble prediction system (EPS; Epstein, 1969; Toth and Kalnay,

1993; Palmer et al., 1993) is the most used method to achieve this probabilistic forecast. EPSs involve the evolution of an ensemble of suitable initial conditions with one or more numerical weather prediction models. The spread of the different members of the ensemble provides a measure of the uncertainty associated with the initial conditions or the model used. An appealing view of the EPS is also that the ensemble mean usually performs better than any of the ensemble members (Leith, 1974), since many of the smaller-scale low-predictable features of the flow are smoothed away by averaging. Differences between the members of the ensemble evolved with the same model and considered at the same time are, in general, finite fluctuations reflecting the amplification of small initial errors in chaotic models. The study of these fluctuations is important, not only to analyse the chaotic behaviour of the system, which can be of academic interest, but also to prepare optimal initial conditions for the ensemble, which has practical applications.

A basic study of fluctuations comprises the computation of the Lyapunov exponents (LEs). A further step is the characterization of the associated Lyapunov vectors (LVs) including the analysis of their dynamical behaviour. Previous studies of this type are scarce, especially with realistic models. The main problem in using realistic models for this kind of analysis is the difficulty in dealing with the associated linearized version of the model, which is necessary to compute LEs and vectors. A realistic model of the atmosphere includes small-scale

*Corresponding author.

e-mail: herreras@unican.es

DOI: 10.1111/j.1600-0870.2011.00545.x

dynamical and microphysical phenomena in parametric form, often including conditional terms which are non-differentiable and non-invertible. Simple dynamical models, referred to as 'toy models', are customarily used in place of realistic meteorological models when one requires the linearized model, the adjoint model, or a large number of ensemble members (Lorenz and Emanuel, 1998; Annan, 2004; Orrell, 2005; Baek et al., 2006; Fertig et al., 2007). Toy models retain certain aspects of the behaviour of the atmosphere (Lorenz, 1996) while being much more tractable than operational models. Basic aspects of the amplification of small errors in toy models resemble those of the real atmosphere (Orrell, 2002; Gutiérrez et al., 2008). However, even with toy models, only recently Lyapunov-like vectors have been systematically compared (Pazó et al., 2010).

Toy models with several time scales are good candidates to fill the gap between the most simplistic toy models and realistic models. In this work, we consider a toy model put forward by Lorenz (1996) that exhibits a patterned large-scale variable and a small-scale variable interacting with the large scale. The two-scale Lorenz'96 model has been extensively used in the literature to study the influence of multiple spatio-temporal scales on the predictability of atmospheric flows. For instance, Wilks (2005) studied stochastic parametrization, Lorenz and Emanuel (1998) analysed targeted observations, and there are also studies testing model error growth at different scales (Orrell, 2003) and applying newly developed methodologies to reduce it (Danforth and Kalnay, 2008). The two-scale Lorenz'96 model possesses two sets of coupled variables (see Section 2) which have been used to represent a generic large-scale variable along a latitude circle in mid-latitudes and a small-scale variable coupled to it (Wilks, 2005). The model mimics some crucial aspects of the atmosphere as a conservative advection, internal dissipation and constant forcing. However, it does not arise from any simplification of the governing equations and, thus, the variables represent unspecified meteorological variables. In the literature, these variables have been said to represent large-scale static stability and small-scale convection (Lorenz, 2006), 500 hPa geopotential height and surface high-resolution variables such as temperature or wind (Orrell, 2002), and several others. For our purposes, the specific meteorological quantity is irrelevant since all of them show some spatially organized structure as does the large-scale variable of the Lorenz'96 system.

LVs identify the locations (directions in phase space) where a perturbation experiences a maximal growth. While the direction of maximal growth can be easily identified, the determination of other important directions leads to the definition of different kinds of LVs: backward, forward and characteristic (Legras and Vautard, 1996). Some of these LVs are related to modern techniques to build initial conditions for EPS: Backward LVs are related to 'bred vectors' (Toth and Kalnay, 1993) and 'singular vectors' are a finite-time version of forward LVs. In a recent work, Pazó et al. (2010) characterized the spatio-temporal dynamics of backward, forward and characteristic LVs, along with

bred and singular vectors in a one-scale version of the Lorenz'96 model. They concluded that characteristic LVs provide the most adequate properties for EPS initialization, since (1) they are dynamically adapted to the system (i.e. they lay on the attractor) and capture the spatial locations where the perturbation is growing; (2) they contain information about the future trajectory, so they capture the locations where the perturbation will grow; (3) their spatial structure (degree of localization) remains stable in time, thus keeping diversity, and (4) they do not have an imposed orthogonality constraint, which may lead to artificial diversity and localization along sub-optimal directions. It is worth remarking that all the dynamic vectors used (including bred vectors) are dependent on the norm type, and only the characteristic LVs, computed by combination of the backward and forward dynamics, are independent of the scalar product definition used in their construction. Hence, the calculation of characteristic LVs is essential to know the effect of the scalar product chosen on the computation of the other dynamic vectors.

A recent work (Herrera et al., 2010) analyses initial fluctuations in a two-scale Lorenz'96 model from a spatio-temporal point of view using the mean variance of the logarithms (MVL; Gutiérrez et al., 2008) diagram. Taking the spatial localization as a simple quantifier of chaos, Herrera et al. (2010) report on a complicated behaviour of fluctuations when the effect of the two scales is well resolved. The authors also show that the effect is dependent on the strength of the coupling between the large and small scales. Another important effect, not studied by Herrera et al. (2010), concerns the interplay between the spatial patterns exhibited by the model and the spatio-temporal amplification of perturbations. The study of this effect is one of the main goals of this paper, which can be summarized as follows:

- (1) Study the interplay between the large-scale spatial patterns and the amplification of perturbations in a two-scale Lorenz'96 system, focusing on the spatial component, thus, resorting to the computation of different LVs.
- (2) Analyse the dependence of the previous interplay on the strength of the coupling between the two layers with different time and spatial scales.
- (3) Check whether current operational approaches in ensemble initialization suffer from shortcomings stemming from the mutual orthogonality constraint.

To achieve our goals, we perform a complete analysis of the two-scale Lorenz'96 model. After defining the model in Section 2, we compute the Lyapunov spectrum for different degrees of coupling in Section 3, and define the LVs to be used. Section 4 is devoted to the computation of characteristic LVs, comparing their spatial structure with that of the other dynamic vectors. In Section 5, we show the spurious effects which appear in backward and forward LVs due to its orthogonal arrangement. Finally, in Section 6, we make a detailed analysis of the degree of transversality of characteristic LVs by means of the

ensemble dimension. We summarize the main conclusions reached in Section 7.

2. The Lorenz'96 model

2.1. One-scale model

The L96 model (Lorenz, 1996) is the simplest model to simulate the chaotic behaviour of the atmosphere. It was originally introduced to mimic mid-latitude atmospheric dynamics. In its one-scale version, it is a set of ordinary differential equations coupled in a ring geometry:

$$\frac{dz_i}{dt} = z_{i-1}(z_{i+1} - z_{i-2}) - z_i + F \quad i = 1, \dots, m, \quad (1)$$

where z_i represents a scalar meteorological quantity at m equidistant grid points on a latitude circle. The terms on the right-hand side of this nonlinear equation are, respectively, modeling nonlinear advection, linear damping and constant forcing. A linear analysis of eq. (1) around the homogeneous solution $z_i = F$ indicates instability with respect to perturbations amplifying like waves with a given length scale and velocity for $F > 8/9$ (Lorenz and Emanuel, 1998; Majda and Wang, 2006). Moreover, for $m > 10$ the dynamics is generally chaotic above $F \approx 6$ (Lorenz, 2006). Lorenz and Emanuel (1998) adjusted the parameters of the model to optimize the similitude with mid-latitude atmospheric dynamics. With $F = 8$ and $m = 40$, we obtain waves with a length of 6 or 7 grid units, a velocity of 1.4 with arbitrary time units and a maximum (i.e. leading) LE of about 1.5. Finally, to reach the expected similitude Lorenz associated one time unit with 5 d.

2.2. Two-scale model

Following with the same scheme and in order to study the influence of multiple spatio-temporal scales on the predictability of atmospheric flows, Lorenz (1996) introduced a two-scale model. This model consists of m slow variables x_i coupled with $m \times n$ fast variables $y_{j,i}$ (Fig. 1) whose evolution is governed by the following nonlinear equations:

$$\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F - \frac{hc}{b} \sum_{j=1}^n y_{j,i} \quad (2a)$$

$$\frac{dy_{j,i}}{dt} = cb y_{j+1,i}(y_{j-1,i} - y_{j+2,i}) - c y_{j,i} + \frac{hc}{b} x_i, \quad (2b)$$

where $i = 1, \dots, m, j = 1, \dots, n$; h is the coupling constant, b controls the (inverse) amplitude of the fast variables, c is a time scaling constant and F is a constant forcing. Here, both the x_i 's and the $y_{j,i}$'s are assumed to be periodic, that is, $x_{m+1} = x_1$; $x_0 = x_m$ and $y_{n+1,i} = y_{1,i+1}$; $y_{0,i} = y_{n,i-1}$.

There are a few important differences between eqs. (2a) and (2b). The first one is the presence of a constant c in (2b) that controls the time scale of variable y ; in our simulations $c = 10$.

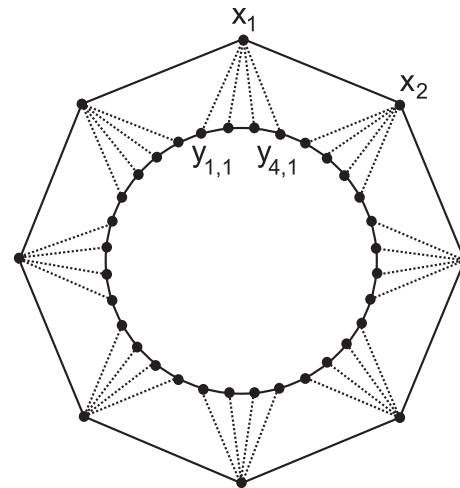


Fig. 1. Schematic illustration of the L96 system with $m = 8$ and $n = 4$; $m \times n$ grid points lay on the inner circle.

This choice makes y the fast-varying small-scale variable. The second important difference is the absence of a constant forcing term in (2b). This makes variable x play the role of 'effective' F in eq. (2b). Finally, note that the advection term in the (fast) variables y is reversed with respect to the (slow) variables x .

In this paper, we considered the parameter values $F = 8$ and $m = 32$, and $n = 8, b = 10, c = 10$, leading to a two-scale model where the fast variables fluctuate 10 times faster than the slow ones, but with a smaller amplitude.

The parameter h controls the coupling between the two layers, and the study of its influence on the results is one of the goals of this paper. Notice that by setting $h = 0$, the large-scale variable of the two-scale system, eq. (2a) falls back to the one-scale system, eq. (1), and the small-scale variable vanishes after an initial transient state due to the internal dissipation term and the absence of external forcing.

2.3. Numerical simulations of the L96 flow

Figure 2(a) shows in grey scale a typical spatio-temporal pattern obtained by integrating eq. (1) from a random Gaussian initial condition $z_i(t=0)$. Figures 2(b) and (d) show the spatio-temporal dynamics of fields x and y , respectively, for the two-scale model (2) with weak coupling ($h = 0.1$). The simulation starts with random initial conditions of amplitude 10 times smaller for the y 's than for the x 's. The latter variable has been initialized as the z variable in panel (a): $x_i(0) = z_i(0)$. It can be seen that the evolution of x is very similar to that of z due to the weakness of the coupling in this example. In the case where the coupling is strong, Fig. 2(c), the large-scale dynamics rapidly deviates from the one-scale model. The westward propagation of the large-scale wavy pattern is clearly appreciated in Figs 2(a), (b) and (c) (Rossby-like wave). Taking the natural time unit, which is defined by the time relaxation of dissipative effects, and the

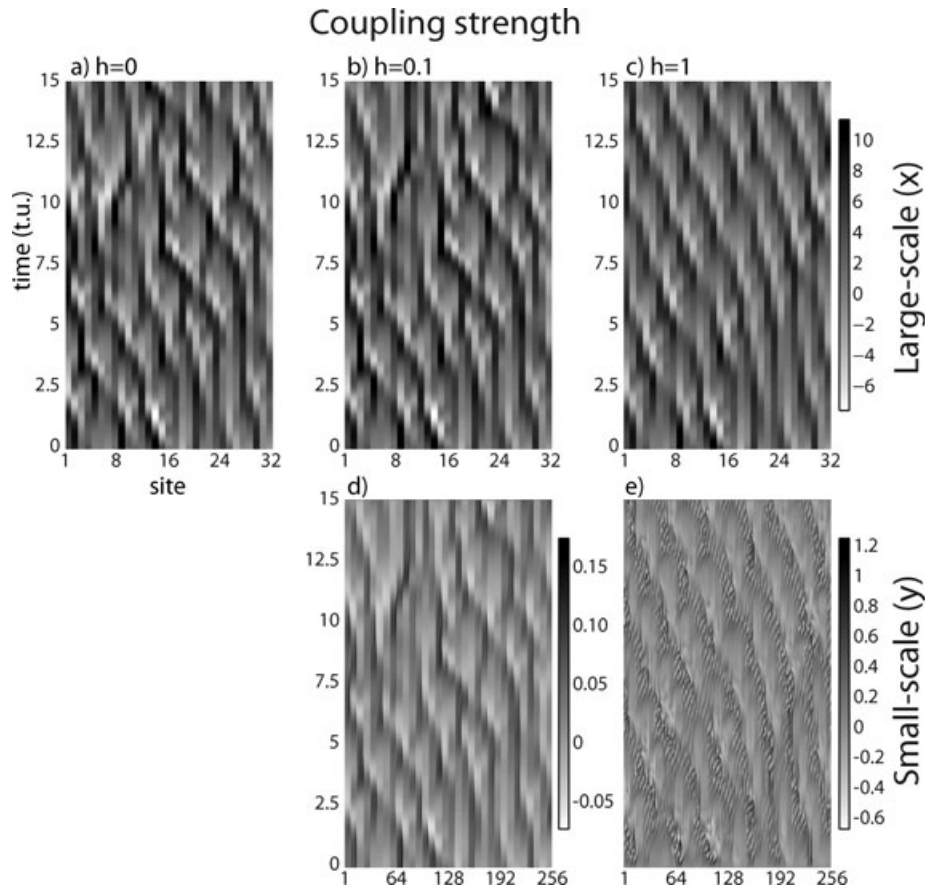


Fig. 2. Spatio-temporal dynamics of (a) the one-scale L96 model (1); (b),(d) slow-fast variables of the L96 model (2) with weak coupling, $h = 0.1$; (c),(e) slow-fast variables of the L96 model (2) with strong coupling, $h = 1$. Both systems were integrated numerically using a fourth-order Runge–Kutta method with time-step $dt = 0.005$. The natural time unit is equivalent to five days in an equivalent atmospheric model. Note the different greyscale level in each panel.

space unit of the fast variable, the velocity of this Rossby-like wave, as observed in Figs 2(d) and (e), is approximately 12. The wave propagation of the fast variable is more difficult to quantify.

Complementing the space–time plots, we show in Fig. 3(a) a snapshot of the fields x and y for $h = 1$. In order to improve the visualization, we use the same representation as in Lorenz and Emanuel (1998), where each grid point of the slow variable is ‘inflated’ by a factor of $n = 8$ and represented as a bar of eight spatial units in the space coordinates of the fast variables (i.e. $\text{Space} = (i - 1) \times 8 + k$, with $k = 1, \dots, 8$). The slow variable has a patterned structure consisting of six local maxima (implying a pseudo-wavelength of $nm/6$ in the fast-variable space).

The spectral analysis shown in Figs 3(c) and (d) provides more detailed information of the propagating patterns. To eliminate spurious effects at high wavenumbers in the spectral density stemming from inflating the large-scale variable x , we show only the spectra of this variable for wavenumbers k in the range $2\pi/(nm) \leq k < k_n^* = \pi/8$. This analysis is performed by averaging 100 single spectral densities, taken at the times when

the system is statistically stationary, that is when the memory of the initial condition has disappeared (after the final time of Fig. 2). As expected, the slow variable power spectrum in the weak coupling ($h = 0.1$) case, see Fig. 3(c), is very similar to the spectrum of the single scale model in Fig. 3(b). Both spectra show a peak at $k \approx 6 \frac{2\pi}{nm}$, which is the mode of the Rossby-like wave. These Rossby-like waves are better defined in the strong coupling case [see the sharp peak in Fig. 3(d)].

For the fast variable, a peak at high wavenumbers corresponding to small-scale eastward-travelling waves can be recognized only in the strong coupling case ($h = 1$).

3. Lyapunov spectrum and Lyapunov vectors

Probabilistic predictions are usually carried out by means of ensembles whose members are suited perturbations in order to attain a degree of diversity that captures errors in the initial conditions. For the sake of simplicity, one of the members, called control, is usually taken as reference, and then, differences between members and control are finite fluctuations. In

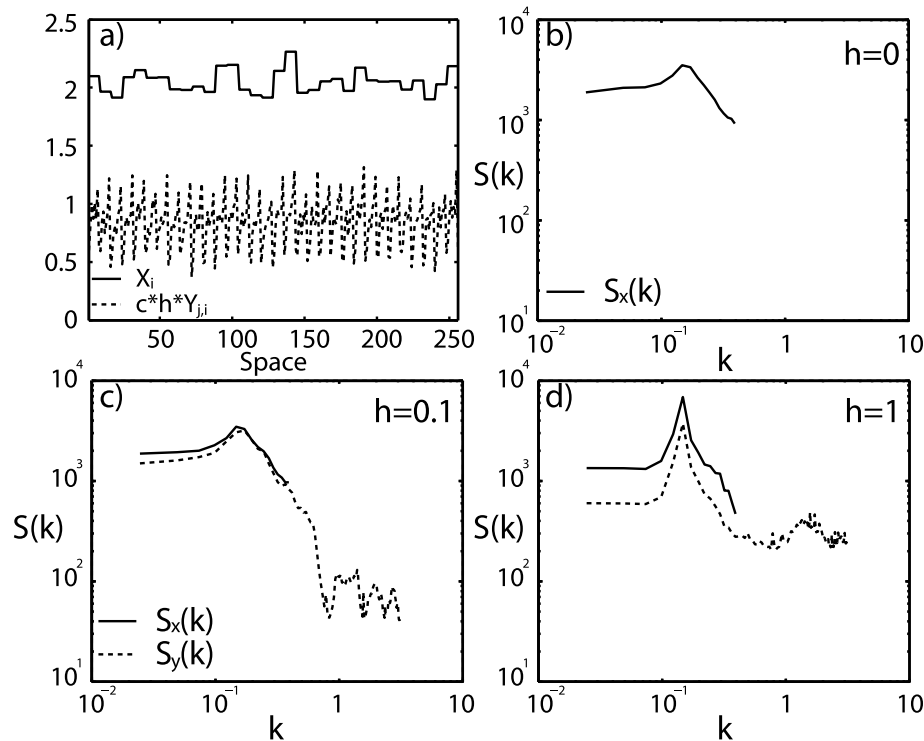


Fig. 3. (a) Snapshot of the slow (x) and fast (y) variables for $h = 1$. Spectral density of a stationary state in (b) the one-scale model, (c) two-scale with weak coupling and (d) two-scale with strong coupling.

general, these fluctuations are not infinitesimal, but as they are dynamically assimilated they should be well projected on the most-expanding subspace of the tangent linear space of the control trajectory. The degree of projection and the diversity of the ensemble define the quality of the ensemble for a given task. Hence, the analysis of this tangent space is necessary either to characterize or to generate ensembles with some physical soundness. Lyapunov methods are the standard tools to perform this analysis.

Small (ideally infinitesimal) perturbations δx_i , $\delta y_{j,i}$ around a control trajectory are governed by the tangent linear model, obtained by linearizing the system (eq. 2)

$$\begin{aligned} \frac{d\delta x_i}{dt} &= (x_{i+1} - x_{i-2})\delta x_{i-1} + x_{i-1}(\delta x_{i+1} - \delta x_{i-2}) \\ &\quad - \delta x_i - \frac{hc}{b} \sum_{j=1}^n \delta y_{j,i} \end{aligned} \quad (3a)$$

$$\begin{aligned} \frac{d\delta y_{j,i}}{dt} &= cb\delta y_{j+1,i}(y_{j-1,i} - y_{j+2,i}) \\ &\quad + cb y_{j+1,i}(\delta y_{j-1,i} - \delta y_{j+2,i}) - c\delta y_{j,i} + \frac{hc}{b}\delta x_i. \end{aligned} \quad (3b)$$

These equations combine the original phenomena of diffusion and dissipation with a kind of forced advection guided by the control trajectory. Hence, infinitesimal perturbations in this chaotic system can be seen as a stochastically driven field in

which the (correlated) noise is generated by the control system. When the system exhibits a spatial pattern (as is the case here and in the atmosphere), fluctuations may exhibit the same pattern depending on the strength of the coupling. Figures 3(c) and (d) show the spatial spectral density of the fields x and y for weak ($h = 0.1$) and strong ($h = 1$) coupling, respectively. In the weak coupling case, as in the one-scale model, the broad peak indicates that the spatial patterns are only slightly distinguishable from the noisy chaotic background, so it is not expected that the pattern will translate into the dynamics of the linearized model. On the contrary, in the strong coupling case, the signal to noise ratio of the power spectrum is very high, suggesting that the spatial pattern should be an important element of the chaotic fluctuations as well. Note that this is an unexplored situation in current studies appearing in more mathematical literature, where fluctuations originated by chaos are roughly homogeneous (in the sense of their spatio-temporal spectral properties), see for example, Pikovsky and Politi (1998) and Szendro et al. (2007). Then, it is interesting to know the effects induced by spatial patterns on both the Lyapunov spectrum and the LVs.

3.1. Lyapunov spectrum

The Lyapunov spectrum was computed integrating the tangent linear model, eq. (3), using the standard method proposed by Benettin et al. (1980). This method comprises a simultaneous

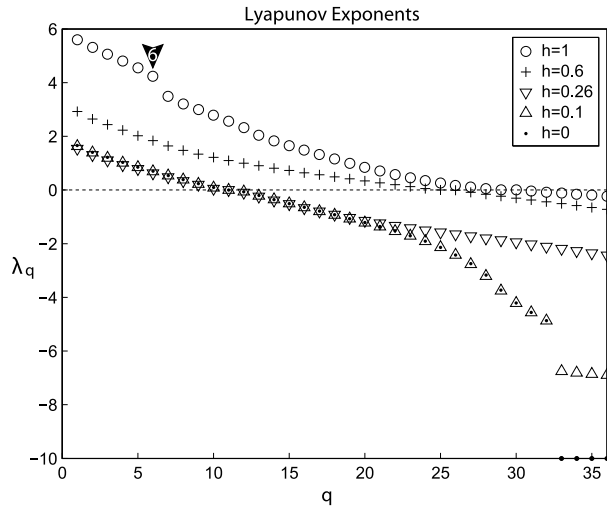


Fig. 4. The 36 largest LEs of the two-scale L96 model for different values of the coupling strength h .

integration of a set of perturbation fields that are periodically orthonormalized to avoid the collapse of all elements along the most unstable direction [pointing along the leading Lyapunov vector (LLV)]. The Lyapunov spectrum for several coupling coefficients, including the one-scale case ($h = 0$), is shown in Fig. 4, where we plot only the 36 largest LEs λ_q ($q = 1, \dots, 36$) ordered in decreasing value. Separation of slow and fast dynamics is complete for the case $h = 0$, where from eq. (3b) one immediately obtains $\lambda_{q>32} = -c = -10$. This separation is very abrupt for weak couplings and becomes almost continuous from moderated values on ($h > 0.26$). As the coupling coefficient h increases, there are two effects that decrease predictability: (1) the increase of the largest LE—leading to a faster error growth rate—and (2) the increase in the number and value of positive exponents—leading to increased Kolmogorov entropy (Schuster, 1988). In the case of strong coupling ($h = 1$), a subtle feature can be observed for the largest LEs, which can be defined as a kind of packaging: At every sixth LE there is a step in the spectrum. This is an effect of the wave-like spatial pattern with wavenumber 6 shown before (Fig. 2). The steps indicate a tendency of the system to self-split into six chaotic cells.

As already shown by Pazó et al. (2010) with the one-scale L96 model (1), and shown in the following subsection, a LV of order q (associated with the q th largest LE) has a characteristic length scale $\sim nm/q$. Then, the descending step observed in the Lyapunov spectrum occurs when the characteristic length of the LV is a new divisor of the pattern size, as a resonance-like phenomenon between the intrinsic Rossby-like wavelength and this characteristic length.

3.2. Lyapunov vectors

LEs inform us about the average exponential growth rate of infinitesimal perturbations, but they are scalar quantities that

lack any information concerning the directions in tangent phase space where perturbations are amplified or reduced. This information is very relevant in spatially extended systems such as the atmosphere. In this section, we summarize some of the basic definitions and properties of the LVs. We also point the interested reader to a detailed review by Legras and Vautard (1996).

The term leading Lyapunov vector (LLV) is usually applied to the direction $\mathbf{g}_1(t) = (\delta x_1, \delta x_2, \dots, \delta x_m; \delta y_{1,1}, \delta y_{2,1}, \dots, \delta y_{n,m})$, which is asymptotically approached evolving, via eq. (3), an arbitrary initial perturbation imposed far enough in the past. Given a certain realization of the flow, eq. (2), the LLV depends only on t . Note that the LLV provides only a direction, since the size is irrelevant due to the linearity of the tangent linear model, eq. (3). This fact allows us to periodically rescale the perturbation to avoid numerical overflows in numerical simulations.

The LLV size grows in time as $\sim \exp(\lambda_1 t)$ and, for long time averages, the LE λ_1 is independent of the definition of the norm. LEs smaller than the largest one also have associated LVs but one must be more explicit as several definitions for the LVs exist. In our case, we followed the definitions by Legras and Vautard (1996):

(1) ‘Backward’ LVs, $\{\mathbf{b}_q(t)\}$, are the orthogonal set of LVs that is a byproduct of computing the LEs with the usual method, which comprises periodic Gram–Schmidt orthonormalization of the set of perturbations. An infinitesimal perturbation along the q th backward LV will shrink as $\sim \exp(-\lambda_q t)$ if integrated backwards in time (in practice this will only occur for a time interval due to numerical roundoff errors). The same experiment but now integrating forward will yield the same exponential amplification rate λ_1 for all backward LVs. A feature of backward LVs usually overlooked is that they depend on the scalar product used in the orthogonalization. We adopt the usual scalar product with the identity matrix as metric, and its associated Euclidean norm. Note that $\mathbf{b}_1 = \mathbf{g}_1$ is the only vector of the set that is not orthogonalized, and at the same time is the only backward LV whose forward and backward amplification rates coincide (with opposite signs).

(2) ‘Forward’ LVs, $\{\mathbf{f}_q(t)\}$, are the time-reversed counterpart of backward LVs. They form another orthogonal basis and a perturbation along one of its members will amplify as $\sim \exp(\lambda_q t)$. However, backward dynamics will expand with an exponential growth rate given by minus the smallest LE: $\sim \exp(-\lambda_{(n+1)m} t)$. To compute forward LVs, one has to evolve the nonlinear system, eq. (2), backwards in time while remaining inside the attractor. To prevent the flow from leaving the attractor (now repeller because of the backward integration), one has to save the state of the system periodically in the computer (we did it every 0.3). This allows a reliable traceback of the trajectory. For the perturbations, we have not used the Jacobian matrix but its transpose. This allows us to obtain the forward LVs from $q = 1$ up to $q_{\max} = 36$ [instead of $q = (n + 1)m$ up to $(n + 1)m - 35$ if the

Jacobian matrix were not transposed], see Legras and Vautard (1996) for details. It is worth noting that forward LVs are the limiting case of (initial) singular vectors as the optimization time goes to infinity.

(3) ‘Characteristic’ LVs, $\{\mathbf{g}_q(t)\}$, are a set of LVs that do not form an orthogonal basis, but instead have remarkable dynamical properties: A perturbation along the q th characteristic LV will grow with exponent λ_q under forward integration. It will shrink with exponent $-\lambda_q$ under backward integration and, therefore, they are said to be covariant with the dynamics. The characteristic LVs are unambiguously defined (up to an irrelevant nonzero constant factor), so no conventions like the scalar product are needed to define them. Computation of characteristic LVs in high-dimensional systems has been possible only recently, thanks to a method proposed by Wolfe and Samelson (2007). The method involves computing the first q backward LVs and the first $q - 1$ forward LVs, and solving a linear system of equations. Backward and forward LVs are needed because the method searches for a set of coefficients $\{\alpha_i^{(q)}(t)\}$ measuring the projection of the q th characteristic LV onto the first q backward LVs $\mathbf{g}_q = \sum_{i=1}^q \alpha_i^{(q)} \mathbf{b}_i$; with the constraint of a transversality condition $\mathbf{g}_q \perp \mathbf{f}_l, \forall l = 1, \dots, q - 1$.

4. Spatial structure of Lyapunov vectors

4.1. The leading Lyapunov vector

In the one-scale L96 model, the LLV is highly localized (Pazó et al., 2008). By localization we mean that the vector achieves the largest magnitude around a few sites of the system. This is a general phenomenon observed also in other dynamically generated vectors such as bred vectors (Hallerberg et al., 2010) or singular vectors (Pazó et al., 2009). Localization arises as a consequence of the multiplicative character of error growth and, therefore, is a generic feature of the LLV in extended systems (Pikovsky and Politi, 1998).

In more quantitative terms, the LLV obeys a log-normal statistics, implying that only after taking the logarithm does the statistics become Gaussian. In particular, spatio-temporal correlations follow scaling laws only if we consider a ‘surface’ vector of the form $(\ln |\delta z_1|, \ln |\delta z_2|, \dots, \ln |\delta z_m|)$ (Pikovsky and Politi, 1998; López et al., 2004). Moreover, there exists universality in the spatial spectral density $S(k)$ of the LLV after this transformation; $S(k)$ follows a power law $\sim 1/k^2$ in one-dimensional systems (such as the one-scale L96 model).

In the two-scale L96 model, we can expect that a part of the features observed in the one-scale L96 model remains valid. For the sake of illustration, we plot in the top panels of Fig. 5 a snapshot corresponding to the LLV, including both the slow and fast variables and for weak and strong coupling cases. In the two-scale L96 model, it is reasonable to work with the two ‘layers’ of variables x and y separately, and thus, we split the LVs into two parts: $\mathbf{g}_q(t) = (\mathbf{g}_q^{(x)}; \mathbf{g}_q^{(y)})$, likewise for backward and forward

LVs. The phenomenon of localization can be very clearly seen in Fig. 5, better in the strong coupling case. Logarithmic values of the LLV are plotted in the second row of Fig. 5. As expected from the theory (Pikovsky and Politi, 1998), the pattern resembles a random walk (or a rough surface).

Finally, we plot the spatial spectral density (averaged over 2000 samples) in the bottom panels of Fig. 5. The expected $1/k^2$ dependence of the spectrum is only observed in the strong coupling case since the chaos in the weak case is not strong enough (the largest LE is small). We can see in Fig. 5 that besides the expected behaviour inherited from the one-scale model there is a superposed peak due to the inclusion of the Rossby-like waves in the structure of fluctuations. Again, the intensity of this effect increases with the coupling. In the case of strong coupling, we can see how the localization is affected by the existence of spatial patterns.

4.2. Sub-leading Lyapunov vectors

An analysis of the spatial structure of sub-LLVs (i.e. $q > 1$ in the one-scale L96 model) can be found in recent works by Pazó et al. (2008, 2010). After defining associated ‘surfaces’ via a logarithmic transformation of the sub-LLVs, they observed that these lower-order LVs are piecewise copies of the LLV ‘surface’. The typical length of the replication intervals defines an intrinsic length l_q for the q th LV. l_q decreases as q grows, and eventually, beyond a certain value of q , the replication property can be no longer observed.

Figure 6 shows the spectral density of the fast component of the LLV and some lower-order backward LVs, $\{\mathbf{b}_q^{(y)}\}_{q=1,2,4,8,16}$, in the cases of weak and strong coupling. Given a value of q , $S(k)$ separates from the reference curve, $S(k)$ for $q = 1$, at a certain crossover wavenumber k_q that defines the intrinsic length $l_q = \frac{2\pi}{k_q}$ mentioned above.

For $h = 0.1$, the k^{-2} scaling of the LLV power spectrum is not apparent, as expected for poorly developed chaos. In contrast, for strong coupling ($h = 1$) the LLV power spectrum exhibits a clear $\sim k^{-2}$ envelope at small wavenumbers (i.e. large spatial scales). In addition, the Rossby-like wave manifests as a distinctive peak (plus the harmonics) superposed on the background spectrum. The background spectra show the aforementioned crossover between the $1/k^2$ and $1/k$ scalings at the characteristic length l_q (cf. Pazó et al., 2008).

5. Localization sites of LVs on large-scale patterned fields

The instantaneous position of the LVs’ localization site in the system is completely different for each set (backward, forward or characteristic). Szendro et al. (2007) compared the relative position of localization sites for different LVs in a coupled map lattice, finding that at a given time each backward LV is localized at a different place in the system. This is not surprising due

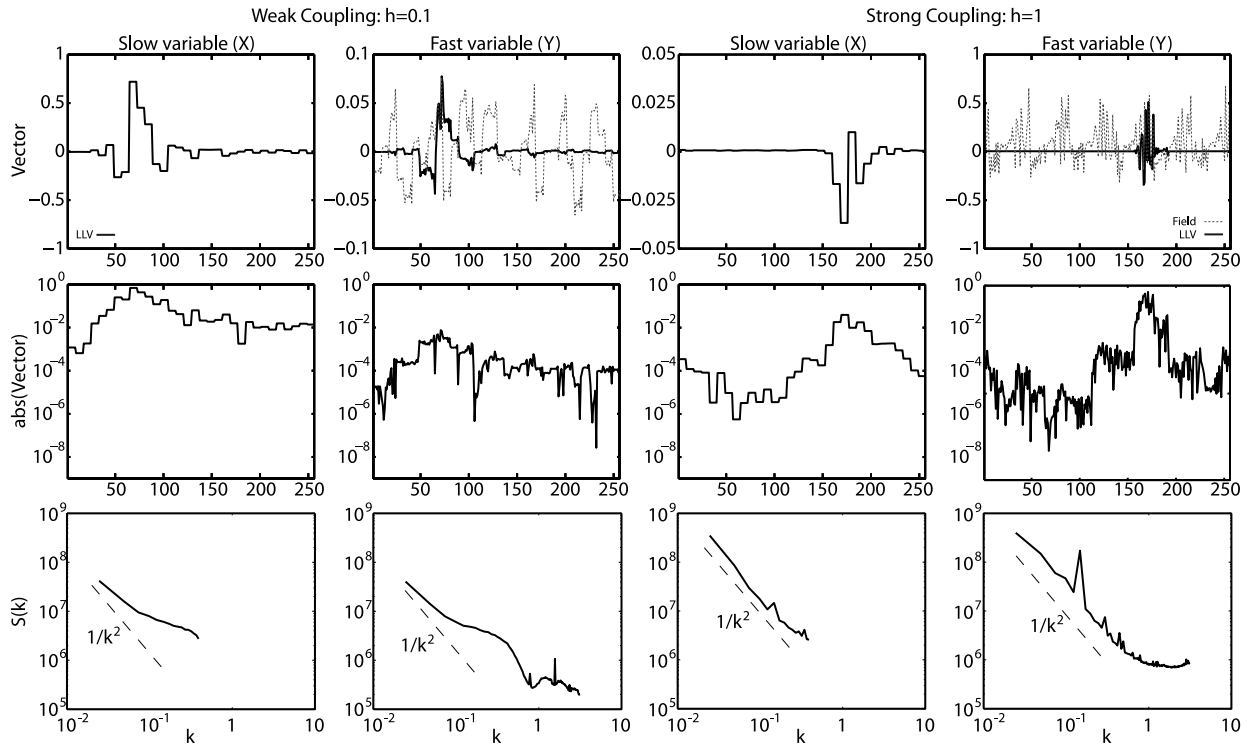


Fig. 5. Representation of the leading LLV g_1 for weak ($h = 0.1$) and strong ($h = 1$) coupling. Each component is split into two panels representing the slow (x) and fast (y) variables. The different rows represent the LLV (top row), its absolute value in the logarithmic scale (see the text for details, middle row) and its spatial spectral density (bottom row). A snapshot of the field is depicted for reference on the top row (grey). The power law $\sim 1/k^2$ is also depicted for reference in the spectra.

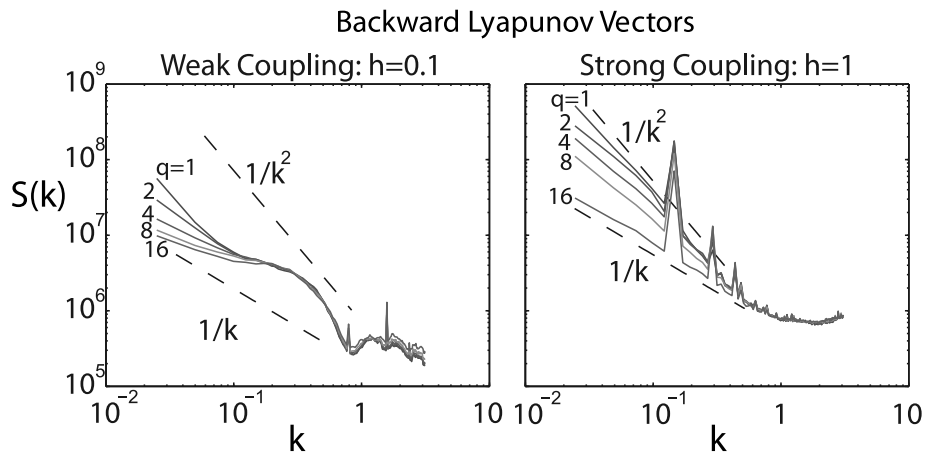


Fig. 6. Power spectra of $q = 1, 2, 4, 8, 16$ (log-transformed) backward LVs for the weak (left) and strong (right) coupling systems. The power laws $1/k$ and $1/k^2$ are depicted for reference.

to their common orthogonality. This has been claimed to be a consequence of the extensivity of spatio-temporal chaos (see, e.g. Egolf et al., 2000). However, Szendro et al. (2007) observed that characteristic LVs have a tendency to cluster their localization sites. This is also observed in chaotic Hamiltonian lattices (Romero-Bastida et al., 2010), and therefore, it is expected to be a generic situation in spatio-temporal chaotic systems. We

shall see now how these features observed in one-scale systems translate into the two-scale L96 model where, at least for $h = 1$, in which case a strong modulation of the fields exist [recall the spectral density in Fig. 3(d)].

Figure 7 shows a snapshot of the small-scale (fast) component of the ten LLVs for different LV types and strong coupling ($h = 1$). Both backward and forward LVs with $q = 1$ to 6 have their

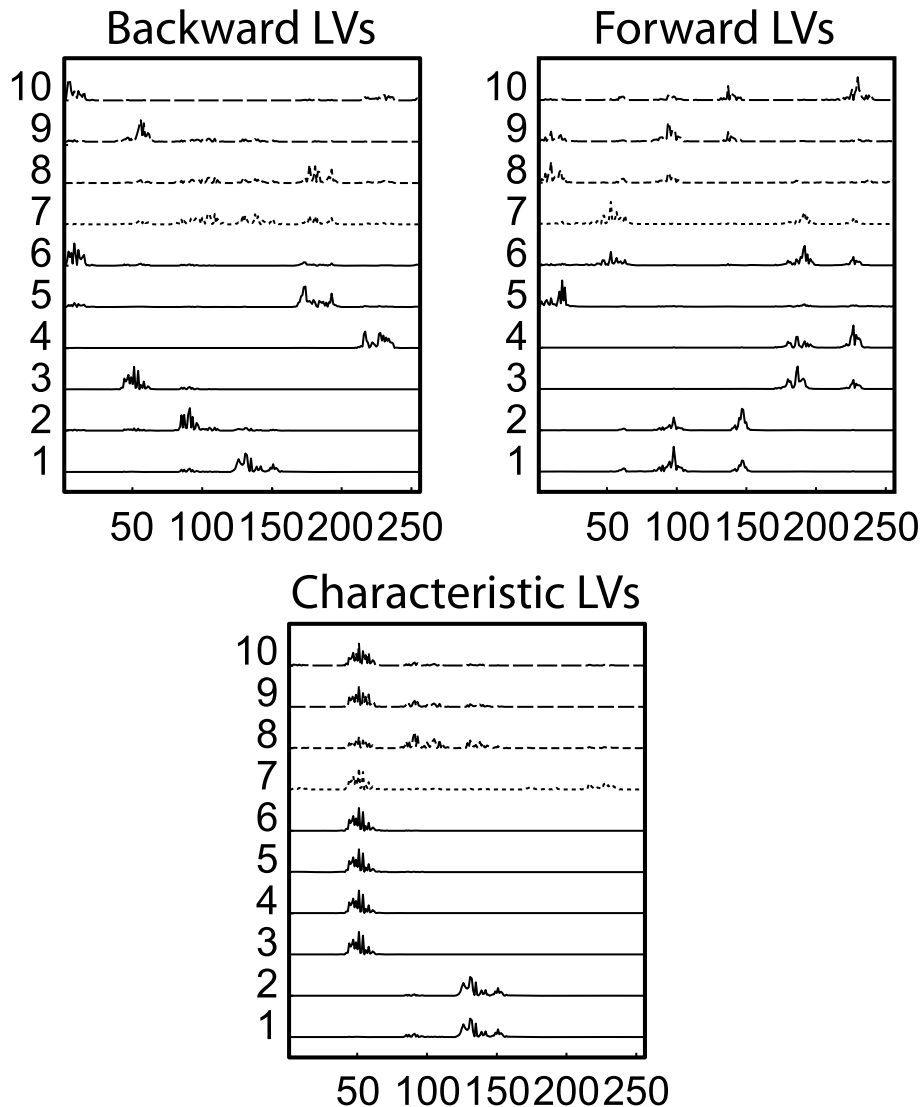


Fig. 7. Absolute value of the y components of the LLVs from $q = 1$ (bottom) to $q = 10$ (top): (a) backward LVs, $b_q^{(y)}$, (b) forward LVs, $f_q^{(y)}$ and (c) characteristic LVs, $g_q^{(y)}$. The three panels correspond to the same time (marked by a rectangle in Fig. 8), and the curves have been shifted upwards $+q$ to improve their visibility.

localization sites roughly uniformly distributed. This reflects the fact that their localization sites of each LV is at a different Rossby-like crest. Note also that backward and forward LVs have little in common in the sense that both sets are apparently uncorrelated (except indirectly by the Rossby-like waves). This can be explained by the fact that backward LVs receive information only from the past evolution of the system, while forward LVs contain only information from the future. Under chaotic dynamics the memory of the system extends of the order of λ_1^{-1} (almost insensitive to the system size), while LVs convergence occurs in a much larger temporal scale, of order $\sim(\text{size})^\zeta$ with $\zeta = 3/2$ in one-dimensional systems with extensive chaos (Pikovsky and Politi, 1998). The slow variables show a similar

behaviour, localizing around the same sites as the fast ones in Fig. 7 (not shown).

The most striking feature of Fig. 7 is the completely different relative localization of characteristic LVs in comparison with backward and forward LVs. Since characteristic LVs are not constrained to be orthogonal they can localize about the same sites, what is not possible for backward and forward LVs. Obviously this picture is dynamic, as shown in Fig. 8, which depicts the temporal evolution of localization sites of the six LLVs on top of the flow. In this figure, we plot the location of the absolute maximum of each LV on the corresponding control trajectory. In the strong coupling case, each maximum follows the motion of the Rossby-like wave for a given time until a new maximum appears

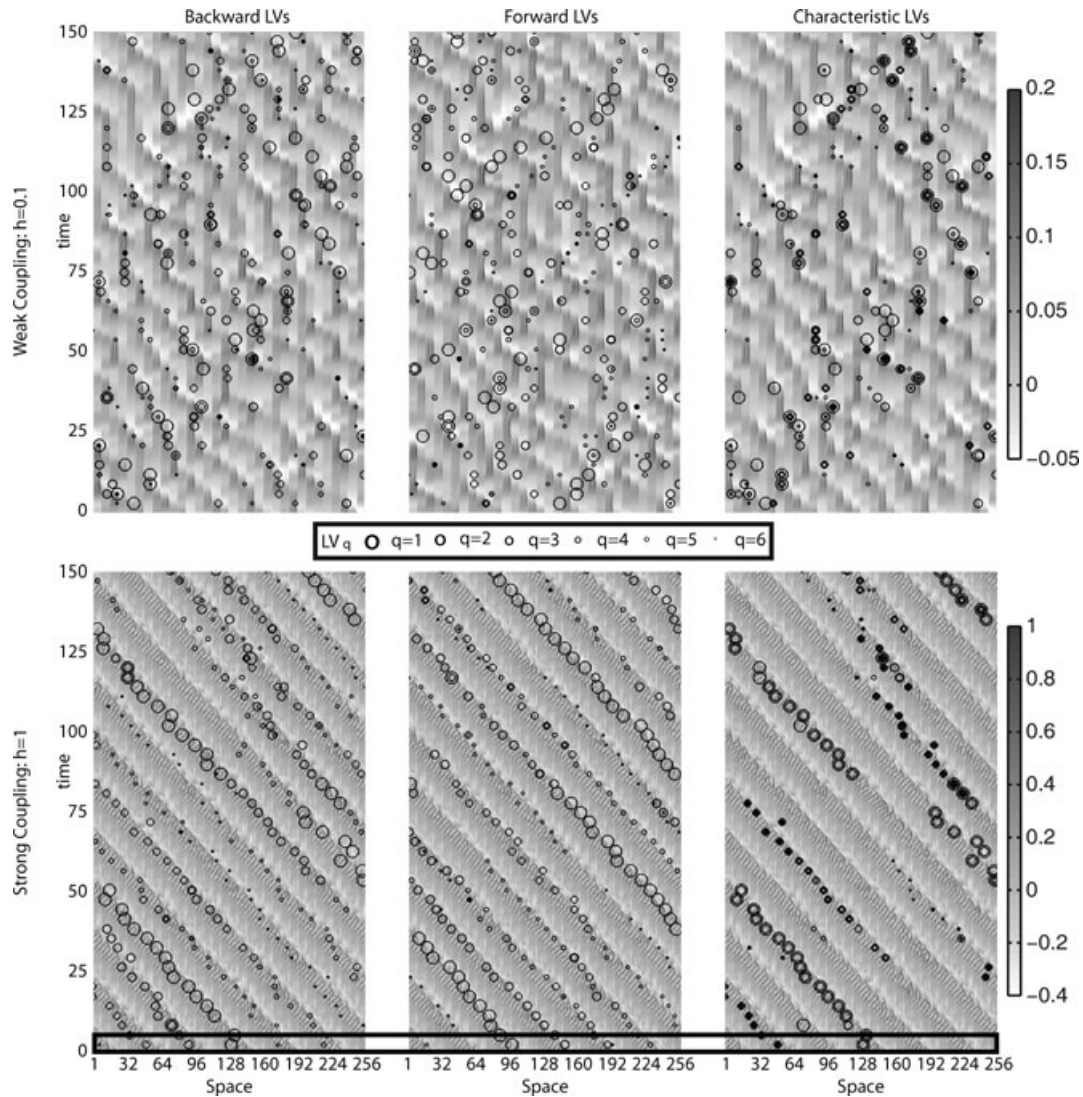


Fig. 8. Space-time evolution of the maximum (in absolute value) of the y component of backward, forward and characteristic LVs ($q = 1, \dots, 6$). The background represents in gray scale the field y . Top (bottom) panels correspond to $h = 0.1$ ($h = 1$). The rectangle on top of the bottom panels indicates the time corresponding to Fig. 7.

at another node. Note that the LLV follows a path different from the $q = 1$ forward LV. For LVs of lower order, each backward (or forward) LV exhibits its peak in a node different from the other vectors, as shown in the snapshots of Fig. 7. Contrastingly, characteristic LVs localize on a few sites with considerable overlapping. The situation depicted for $h = 1$ in Fig. 7 corresponds to the time where a rectangle is superimposed in Fig. 8, and it is qualitatively stationary in time. In the next section, we analyse the statistics of the overlapping in quantitative terms.

In the weak-coupling case (Fig. 8, top row), the flow is still patterned but the wavelength is less clear than in the strong-coupling case [hence, the broader spectral density peak in Fig. 3(c)]. The paths of the maximum value of each LV cannot be easily identified since they jump drastically in space as time evolves.

However, the localization sites are still near the ridges of the pattern at each time, and remarkably the characteristic LVs again show a much lower diversity than backward or forward LVs do.

6. Ensemble dimensions

Figure 7 shows that the characteristic LVs have a remarkable tendency to overlap (i.e. to become almost aligned). For instance, we can mention that the second characteristic LV is localized 60% of the time in the same crest as the LLV, while 40% of the time it is localized at the place of the second backward LV. Moreover, the first forward LV and the second characteristic LV seldom localize over the same ridge ($\sim 2\%$ of the time). These

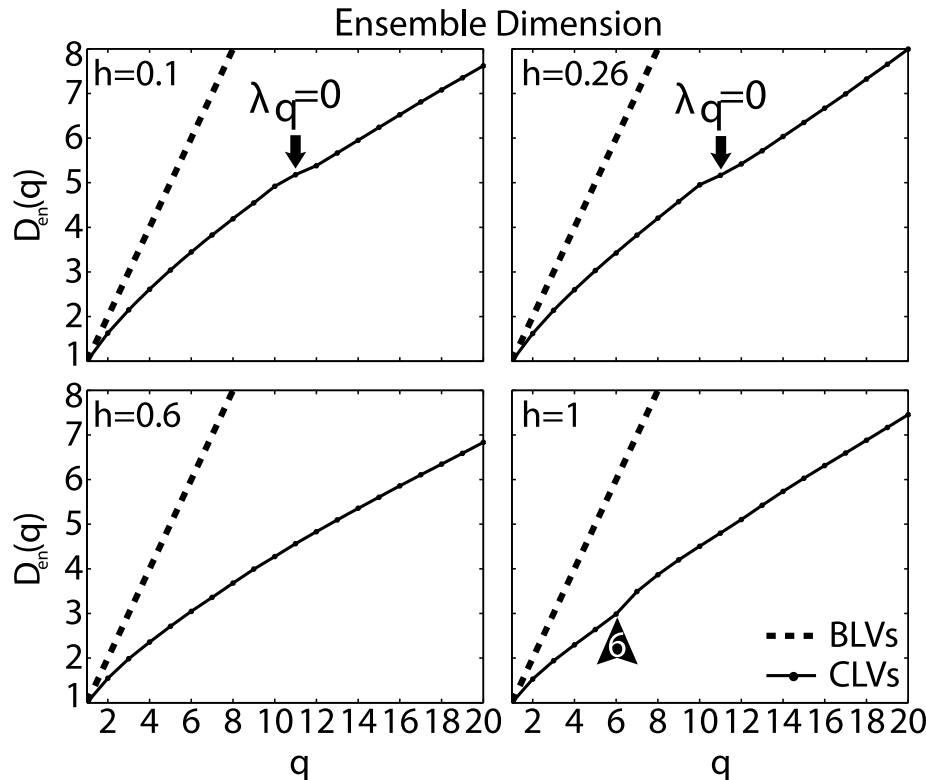


Fig. 9. Average ensemble dimension as a function of q for four different values of h .

facts are interesting but give an incomplete view of what is occurring. The aim of this section is to quantify the degree of independence of the LVs and, at the same time, to assess their dependence on h .

We resort here to a statistic known as ensemble dimension (Oczkowski et al., 2005) or bred vector dimension (Patil et al., 2001), that measures the degree of ‘transversality’ of an ensemble, here composed of the q leading characteristic LVs. The ensemble dimension measures the number of effective degrees of freedom that explains most of the total ensemble variance. The ensemble dimension at a given time is obtained from the set of eigenvalues $\{\mu_i(t)\}_{i=1,\dots,q}$ of the $q \times q$ covariance matrix \mathbf{C} with elements

$$C_{ij}(t) = \mathbf{g}_i \cdot \mathbf{g}_j \quad (4)$$

where the standard scalar product is used. The average ensemble dimension is

$$D_{\text{en}}(q) = \left\langle \frac{\left(\sum_{i=1}^q \sqrt{\mu_i} \right)^2}{\sum_{i=1}^q \mu_i} \right\rangle \quad (5)$$

where the denominator ($\sum_i \mu_i$) equals q since we take normalized LVs ($\|\mathbf{g}_i\| = 1$). The statistic (5) typically returns a number between two limiting values: $D_{\text{en}}(q) = 1$ (if all q vectors are aligned) and $D_{\text{en}} = q$ (if all q vectors are orthogonal, as in the

case of backward and forward LVs). Notice that another similar measure of effective dimensions has been used in the literature. Bretherton et al. (1999) use the eigenvalues μ_i instead of their square root in eq. (5). This measure assigns different intermediate values, but the limiting values and interpretation are the same as in eq. (5).

Figure 9 shows the dependence of $D_{\text{en}}(q)$ on q for four different values of h . It is remarkable that the data deviate significantly from the bisectrix (indicated by a dashed line), which means that the q th characteristic LV is on average far from being orthogonal to the set of $q - 1$ leading characteristic LVs. This is in agreement with the results in the bottom panel of Fig. 7, and it apparently holds for all values of h .

The curves of $D_{\text{en}}(q)$ for $h = 0.1$ and $h = 0.26$ are quite smooth, except at $q = 11$, which is precisely the value of q corresponding to the null LE, which exists because of the invariance of the flow under a time shift (because the system is autonomous), see Fig. 4. The vanishing LE does not fit in the rest of the spectrum because \mathbf{g}_{11} is a peculiar vector, parallel to the velocity vector of the flow. Remarkably, there is a low transversality between this vector and the LLVs as already noticed in other geophysical models with low-dimensional chaos (Samelson and Wolfe, 2003). For $h = 0.6$ and $h = 1$, the vanishing LE is located at a value of q , out of the range of Fig. 9. The curve for $h = 1$ exhibits a jump precisely at the same point as in the Lyapunov spectrum, see Fig. 4. This is a remarkable result because it shows that the characteristic

LVs indeed convey a large amount of information of the tangent dynamics that cannot be recovered from the other vector types: Forward and backward LVs result in trivial ensemble dimensions ($D_{\text{en}}(q) = q$, dashed line shown in Fig. 9).

Finally, we note that the jumps observed in $D_{\text{en}}(q)$ are located at the same values of q if the ensemble dimension is computed only from either the x components of the characteristic LVs, $g_q^{(x)}$, or the y components $g_q^{(y)}$.

7. Summary and conclusions

We performed an analysis of the spatial dynamics of the amplification of perturbations in the Lorenz'96 model. We used this toy model as a surrogate of the atmosphere in the sense that this model presents a large-scale variable with a patterned structure and a small-scale variable coupled to it. The simplicity of this toy model allowed us to compute different types of LVs, including the characteristic LVs, which are independent of the definition of the scalar product and not subject to orthogonality constraints. These remarkable properties come at the cost of a larger computational effort.

The configuration ($m = 32$, $n = 8$) and parameter values ($F = 8$) used in the model causes the large-scale variable to develop a structure consisting of a wave pattern with 6 ridges.¹ According to the spatial spectrum, this wave pattern is reinforced by the small-scale variable as the coupling increases. The Lyapunov spectrum also develops, as the coupling strengthens, a packeting related to wavenumber 6 of the large-scale pattern.

Different centres around the world use EPSs operationally in order to obtain an upper limit to the spread between members by perturbing the initial conditions in the fastest-growing directions given by LV-like perturbations. We have shown, using the L96 model, how different LV types have different properties regarding their spatial localization and diversity. This localization occurs where the large-scale pattern allows it, along the six ridges of the large-scale flow. Backward and forward LVs show a spurious localization induced by the orthogonalization. This spurious effect will propagate to an ensemble made from them since, without the help of characteristic LVs, spurious and genuine localization peaks are indistinguishable. Hence, in a real forecast, there can be many sources of uncertainty that, if not avoidable, should at least be identifiable. Preparation of ensembles with vectors other than the characteristic LVs may lead to the inclusion of spurious effects. In a system with a large-scale spatial pattern, the available growth sites are 'quantized' by the patterned structure. A positive outcome of this spatial 'quantization' is that some information from the future, which is difficult to obtain, is more easily used than in the case of homogeneous

chaos. This is exemplified by the existing anticorrelation between the first forward LV and second characteristic LV.

The characteristic LVs localize over places where the large-scale variable allows but, at the same time, they span only a few of the available large-scale ridges. This is a result of the lack of any orthogonality constraint, implying that the rest of the locations where backward and forward LVs localize are in many cases only mere byproducts of the orthogonality (with a chosen metric). This also implies that the diversity of LVs other than the characteristic ones is overestimated and leads to perturbation in directions which do not grow maximally. The unphysical character of patterns obtained from a procedure imposing orthogonality is not new and has been recognized in other vectors such as the empirical orthogonal functions (Richman, 1986), which are obtained from an eigenvalue problem from the covariance matrix of empirical data. Ensembles initialized with characteristic LVs do not show spurious effects induced by an imposed orthogonality. Two issues prevented the use of these LVs in the initialization of EPS: Their computation has only been possible recently, and the computational effort is larger than that required for the backward and forward LVs. With the method developed by Wolfe and Samelson (2007) and the increasing computer power available, it would be feasible to test in a realistic model of the atmosphere the ability of the characteristic LVs to capture the spatial locations of interest for perturbing and initializing an EPS and quantify their improvement over the current state-of-the-art techniques. These techniques are limited to the use of methods that take information either from the past (bred vectors, backward LVs, etc.) or from the future (singular vectors) but not from both simultaneously. Also, current techniques increase the diversity of the ensemble members by using, implicit or explicitly, some method of orthogonalization.

8. Acknowledgments

The authors thank two anonymous referees for their constructive remarks. We also thank Tony Owen for a revision of the manuscript. The authors are grateful to the Spanish Ministry of Science and Innovation (FIS2009-12964-C05-05, CGL2007-64387/CLI, CGL2010-21869 and CGL2010-22158-C02-01 grants) for partial support of this work. D.P. acknowledges support by CSIC under the Junta de Ampliación de Estudios Programme (JAE-Doc).

References

- Annan, J. 2004. On the orthogonality of bred vectors. *Mon. Wea. Rev.* **132**, 843–849.
- Baek, S. J., Hunt, B. R., Kalnay, E., Ott, E. and Szunyogh, I. 2006. Local ensemble Kalman filtering in the presence of model bias. *Tellus* **58A**, 293–306.
- Benettin, G., Galgani, L., Giorgilli, A. and Strelcyn, J.-M. 1980. Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems. *Meccanica* **15**, 9.

¹ The number of ridges is proportional to the system size m , and grows monotonically with the forcing term F . Our simulations also indicate that the number of ridges is not very sensitive to parameter h .

- Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M. and Bladé, I. 1999. The effective number of spatial degrees of freedom of a time-varying field. *J. Climate* **12**, 1990–2009.
- Danforth, C. M. and Kalnay, E. 2008. Using singular value decomposition to parameterize state-dependent model errors. *J. Atmos. Sci.* **65**, 1467–1478.
- Egolf, D. A., Melnikov, I. V., Pesch, W. and Ecke, R. E. 2000. Mechanisms of extensive chaos in Rayleigh-Bénard convection. *Nature* **404**, 733–736.
- Epstein, E. 1969. Stochastic dynamic prediction. *Tellus* **21A**, 739–759.
- Fertig, E. J., Harlim, J. and Hunt, B. R. 2007. A comparative study of 4D-VAR and a 4D ensemble Kalman filter: perfect model simulations with Lorenz-96. *Tellus* **59A**, 96–100.
- Gutiérrez, J. M., Primo, C., Rodríguez, M. A. and Fernández, J. 2008. Spatiotemporal characterization of ensemble prediction systems - the mean-variance of the logarithms (MVL) diagram. *Nonlin. Processes Geophys.* **15**(1), 109–114. URL: <http://www.nonlin-processes-geophys.net/15/109/2008>.
- Hallerberg, S., Pazó, D., López, J. M. and Rodríguez, M. A. 2010. Logarithmic bred vectors in spatiotemporal chaos: structure and growth. *Phys. Rev. E* **81**, 066204.
- Herrera, S., Fernández, J., Rodríguez, M. A. and Gutiérrez, J. 2010. Spatio-temporal error growth in the multi-scale Lorenz96 model. *Nonlin. Processes Geophys.* **17**, 329–337.
- Legras, B. and Vautard, R. 1996. A guide to liapunov vectors. In: *Proc. Seminar on Predictability Vol. 1* (ed. T. Palmer. ECWF Seminar, ECMWF, Reading, UK, pp. 135–146.
- Leith, C. E. 1974. Theoretical skills of Monte Carlo forecasts. *Mon. Wea. Rev.* **102**, 409–418.
- López, J. M., Primo, C., Rodríguez, M. A. and Szendro, I. 2004. Scaling properties of growing noninfinitesimal perturbations in space-time chaos. *Phys. Rev. E* **70**, 056224.
- Lorenz, E. N. 1996. Predictability, a problem partly solved. In *Proceedings of ECMWF seminar on Predictability*, ECMWF, Reading, UK, pp. 1–19.
- Lorenz, E. N. 2006. Regimes in simple systems. *J. Atmos. Sci.* **63**, 2056–2073.
- Lorenz, E. N. and Emanuel, K. 1998. Optimal sites for supplementary weather observations: simulation with a small model. *J. Atmos. Sci.* **55**, 399–414.
- Majda, A. J. and Wang, X. 2006. *Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows* (eds. Majda, A. J. and Wang, X.). Cambridge University Press, Cambridge, UK.
- Oczkowski, M., Szunyogh, I. and Patil, D. J. 2005. Mechanisms for the development of locally low-dimensional atmospheric dynamics. *J. Atmos. Sci.* **62**, 1135–1156.
- Orrell, D. 2002. Role of the metric in forecast error growth: how chaotic is the weather?. *Tellus* **54A**, 350–362.
- Orrell, D. 2003. Model error and predictability over different timescales in the Lorenz'96 systems. *J. Atmos. Sci.* **60**, 2219.
- Orrell, D. 2005. Ensemble forecasting in a system with model error. *J. Atmos. Sci.* **62**, 1652–1659.
- Palmer, T. N., Molteni, F., Mureau, R., Buizza, R., Chapelet, P. and co-authors. 1993. Ensemble prediction. In: *Proc. Seminar on Validation of Models Over Europe: vol. 1* (ed. T. N. Palmer), ECMWF Seminar, ECMWF, Reading, UK.
- Patil, D. J., Hunt, B. R., Kalnay, E., Yorke, J. A. and Ott, E. 2001. Local low dimensionality of atmospheric dynamics. *Phys. Rev. Lett.* **86**, 5878–5881.
- Pazó, D., López, J. M. and Rodríguez, M. A. 2009. Exponential localization of singular vectors in spatiotemporal chaos. *Phys. Rev. E* **79**(3), 036202. doi:10.1103/PhysRevE.79.036202. URL: <http://link.aps.org/abstract/PRE/v79/e036202>.
- Pazó, D., Rodríguez, M. A. and López, J. M. 2010. Spatio-temporal evolution of perturbations in ensembles initialized by bred, Lyapunov and singular vectors. *Tellus* **62A**, 10–23.
- Pazó, D., Szendro, I. G., López, J. M. and Rodríguez, M. A. 2008. Structure of characteristic Lyapunov vectors in spatiotemporal chaos. *Phys. Rev. E* **78**(1), 016209. doi:10.1103/PhysRevE.78.016209. URL: <http://link.aps.org/abstract/PRE/v78/e016209>.
- Pikovsky, A. and Politi, A. 1998. Dynamic localization of Lyapunov vectors in spacetime chaos. *Nonlinearity* **11**(4), 1049–1062. doi:10.1088/0951-7715/11/4/016.
- Richman, M. B. 1986. Rotation of principal components. *J. Climatol.* **6**, 293–335.
- Romero-Bastida, M., Pazó, D., López, J. M. and Rodríguez, M. A. 2010. Structure of characteristic Lyapunov vectors in anharmonic Hamiltonian lattices. *Phys. Rev. E* **82**(3), 036205. doi:10.1103/PhysRevE.82.036205.
- Samelson, R. and Wolfe, C. 2003. A nonlinear baroclinic wave-mean oscillation with multiple normal mode instabilities. *J. Atmos. Sci.* **60**(9), 1186–1199.
- Schuster, H. G. 1988. *Deterministic Chaos: An Introduction* Schuster, H. G. VCH, Germany.
- Szendro, I. G., Pazó, D., Rodríguez, M. A. and López, J. M. 2007. Spatiotemporal structure of Lyapunov vectors in chaotic coupled-map lattices. *Phys. Rev. E* **76**(2), 025202. doi:10.1103/PhysRevE.76.025202. URL: <http://link.aps.org/abstract/PRE/v76/e025202>.
- Toth, Z. and Kalnay, E. 1993. Ensemble forecasting at NMC: the generation of perturbations. *Bull. Amer. Meteor. Soc.* **74**, 2317–2330.
- Wilks, D. S. 2005. Effects of stochastic parameterization in the Lorenz96 system. *Quart. J. Roy. Meteor. Soc.* **131**, 389–407.
- Wolfe, C. L. and Samelson, R. M. 2007. An efficient method for recovering Lyapunov vectors from singular vectors. *Tellus* **59A**(3), 355–366. doi:10.1111/j.1600-0870.2007.00234.x.