Nonlin. Processes Geophys. Discuss., 1, 1509–1537, 2014
www.nonlin-processes-geophys-discuss.net/1/1509/2014/
doi:10.5194/npgd-1-1509-2014


Nonlinear Processes in Geophysics
Discussions
Open Access



# Self-breeding: a new method to estimate local Lyapunov structures

**J. D. Keller[1,2] and A. Hense[3]**

[1]Hans-Ertel-Centre for Weather Research, Bonn, Germany
[2]Deutscher Wetterdienst, Offenbach, Germany
[3]Meteorological Institute, University of Bonn, Bonn, Germany

Received: 2 July 2014 – Accepted: 19 August 2014 – Published: 9 September 2014

Correspondence to: J. D. Keller (jkeller@uni-bonn.de)

Published by Copernicus Publications on behalf of the European Geosciences Union & the American Geophysical Union.





Self-breeding – a new method to estimate local Lyapunov structures

J. D. Keller and A. Hense

## Abstract

We present a new approach to estimate local Lyapunov vectors. The so called self-breeding method is based on the breeding of growing modes technique from medium range weather forecasting and consists of a continuous forecasting and rescaling cycle. Using the Lorenz96 model we test and characterize the behavior of the algorithm regarding error growth, spatial perturbation structure estimates and orthogonalization. The results indicate that the method can be used to generate error growing modes optimized for a certain rescaling interval, thus enabling the user to target specific scales of error growth. When an additional orthogonalization procedure is applied, the method is able to produce structures representing the error growth subspace spanned by the largest Lyapunov vectors.

## 1 Introduction

Forecasting future states of chaotic systems such as the atmosphere is prone to errors due to the implied exponential growth of small (or infinitesimal) errors in the initial state (Lorenz, 1963). These initial state errors, i.e. the difference between the estimation of the state and the *true* state of the system in model space, therefore inherently generate uncertainty in the forecast. Generally the spatio-temporal development of errors is not homogeneous but depends on the state of the system itself, i.e. so called laminar phases with less rapid error growth rates can interchange with chaotic phases exhibiting rapid and large error growth. In weather forecasting this is called flow-dependent predictability (e.g. Berner et al., 2009).

However, the chaotic nature of a system does not imply that the forecast will leave the attractor of the system but that its trajectory will be part of the same attractor while deviating from the *true* trajectory. For short time periods, the shape of these deviations is determined through a linear combination of the error modes of the system, i.e. the directions of the major error growth. These directions are called Lyapunov vectors





NPGD
1, 1509–1537, 2014



J. D. Keller and A. Hense

(LVs) and the associated exponential error growth rate in each direction is determined by the corresponding Lyapunov exponent (LE) $\lambda_i$ (after Lyapunov, 1992). The largest Lyapunov exponent $\lambda_i$ is called the leading LE and the corresponding spatial structure is the leading LV. The existence of the LVs and LEs has been proven by Oseledec (1968).

When investigating a chaotic system, an overall set of LVs can be estimated which determines the average error growth regarding the system's attractor. We will denote these as the global LVs. However, the current state of the system, i.e. its actual dynamical properties at a certain point in phase space characterizing a specific flow or weather configuration in case of atmospheric forecasting, determine the current growing error modes. Therefore, the actual sensitivity of the system to small errors in the state vector is not based on one set of global LVs but on a different set of LVs at each point of its attractor. These structures are called the local LVs (e.g. Fujisaka, 1983).

To illustrate the idea, one could think of two typical weather regimes over the North-Atlantic European regions like a westerly flow and a blocking high. The variability on the one to five day scale is very large in the former case with intensive baroclinic developments and a high degree of sensitivity to the initial state while in the latter, the variability on the larger scales is low without notable developments. In the first case, one could imagine to obtain larger local Lyapunov exponents than in the second case corresponding to different spatial characteristics. An even more complicated case could arise when there is a change between these to regimes. Classical Lyapunov exponents and vectors would represent an average behavior between all these cases whereas their local counterparts could be more appropriate for forecasting applications. The example also shows that the specification of the time scale or time period respectively, might be crucial when deriving local Lyapunov characteristics. Complex non-linear models may have not one but multiple regimes of error growth that can be attributed to different time scales. For instance, atmospheric large scale flow regimes on time scales of one to five days may embed convective instability regimes on sub-daily time scales with corresponding error growth. From a longer time scale perspective those smaller scale regimes appear "saturated" due to the non-linear effects keeping the system on its attractor.

Global Lyapunov structures are useful when describing the non-linear characteristics of a system. Further, the identification of regions with a high average sensitivity to deviations of the *true* state is especially important when planning new observing systems or incorporating new observations into a data assimilation scheme. To maximize the impact of such a permanent addition of observational information, the assessment of the spatio-temporal error growth structures associated with LVs can be used to identify locations for an optimal gain in analysis quality (e.g. Lorenz and Emanuel, 1998; Bishop and Toth, 1999).

For a single forecast, however, local LVs are much more important as they determine the actual error growth of a system depending on its current position in the attractor (e.g. Szunyogh et al., 1997). Therefore, it would be optimal to know the structure of local LVs when predicting future system states or more precisely, the probability density function (PDF) of the future system state.

The general problem when estimating LVs for a system is the number of degrees of freedom of the system. For large complex models such as weather forecasting models there is no feasible way to calculate LVs. Therefore, one has to turn to approximate methods. The breeding of growing modes or bred vector method described in Sect. 2 is one of these methods which uses the full nonlinear model to efficiently approximate uncertainty structures in initial states. A different approach to determine the local LV structures of a system is the Singular Vector (SV) method (Buizza and Palmer, 1995) which makes use of a tangent linear as well as an adjoint model to estimate the phase space directions of largest error growth for a predefined optimization time.

The main difference between the BV and SV methods is that BVs describe the uncertainty structures corresponding to the leading local LV in relation to the past evolution of the system using the full non-linear model while the SV method approximates the LVs for the future with respect to the optimization time using a tangent-linear and adjoint model approach. The BV method therefore project onto the so called *backward* LVs





NPGD
1, 1509–1537, 2014

while the SVs are estimates for the *forward* LVs (Legras and Vautard, 1995). For applications such as numerical weather prediction (NWP) where the PDF of a future system state is to be predicted, the *forward* LVs are more desireable as they are designed to optimize the error growth with respect to the future evolution of the system.

In this paper we present an adaptation of the bred vector method called self-breeding to estimate *forward* local LVs using the full non-linear model. The approach chosen for self-breeding also enables us to apply the breeding method to systems with non-cyclical model space, i.e. boundary conditions are provided at the borders of the model space. Section 2 contains a detailed description of the bred vector technique as well as our self-breeding approach. Section 3 gives an overview of the model used and its respective experiment setup. Sections 5 and 6 show the results of our experiments and the conclusions drawn.

## 2 The Self-breeding method

In this section we present the adaptation of a method to estimate uncertainty structures in non-linear models called breeding of growing modes. The structures determined by this technique are called bred vectors (BVs).

### 2.1 Breeding of growing modes technique

The BV technique was first described by Toth and Kalnay (1993, 1997) and is related to the method presented by Wolf et al. (1985). The technique estimates uncertainty structures in dynamical systems corresponding to the so called *errors of the day* and approximate the leading local LVs. The uncertainty structures are *bred* by the system by an iterative procedure called the *breeding cycle*:

1. Perturb the state of the system using random perturbations.
2. Advance the perturbed and unperturbed state in time using the non-linear model of the system.





NPGD
1, 1509–1537, 2014

3. Rescale the perturbations such that their amplitude equals that of the initial perturbations.

4. Add the perturbations to the unperturbed initial state.

5. Repeat steps 2 to 5.

In the following, we describe this procedure in detail.

Let $\boldsymbol{x}_{\mathrm{a}}(t_0)$ be an analysis state vector, i.e. the best approximation of the *true* state vector (e.g. the output from a data assimilation scheme), of a given model at time $t_0$. Random perturbations $z(t_0)$ are then added to $\boldsymbol{x}_{\mathrm{a}}(t_0)$ leading to a perturbed state $\boldsymbol{x}'_{\mathrm{a}}(t_0)$. The full, non-linear model $\mathcal{M}$ is then used to advance the analysis state as well as the perturbed state along their respective trajectories in the phase space to the next analysis time $t_1 = t_0 + \Delta t$:

$$\boldsymbol{x}_{\mathrm{f}}(t_1) = \mathcal{M}\boldsymbol{x}_{\mathrm{a}}(t_0) \tag{1}$$

$$\boldsymbol{x}'_{\mathrm{f}}(t_1) = \mathcal{M}\boldsymbol{x}'_{\mathrm{a}}(t_0). \tag{2}$$

To determine the error growth in these forecasts, a norm $\|\cdot\|$ is defined to measure the amplitude of the model state perturbations $z_{\mathrm{f}} = \boldsymbol{x}_{\mathrm{f}} - \boldsymbol{x}'_{\mathrm{f}}$. With that, we can calculate the error growth of the perturbations from $t_0$ to $t_1$

$$\mathrm{eg}(t_1) = \frac{\left\|\boldsymbol{x}'_{\mathrm{f}}(t_1) - \boldsymbol{x}_{\mathrm{f}}(t_1)\right\|}{\|\boldsymbol{x}'_{\mathrm{a}}(t_0) - \boldsymbol{x}_{\mathrm{a}}(t_0)\|} = \frac{\|z_{\mathrm{f}}(t_1)\|}{\|z_{\mathrm{a}}(t_0)\|}. \tag{3}$$

Using the error growth value $\mathrm{eg}(t_1)$, the perturbations $z_{\mathrm{f}}(t_1)$ are rescaled to match the initial perturbation amplitude $\|z_{\mathrm{a}}(t_0)\|$

$$z_{\mathrm{a}}(t_1) = \frac{z_{\mathrm{f}}(t_1)}{eg(t_1)}. \tag{4}$$

The perturbation vector $z_{\mathrm{a}}(t_1)$ is then added to the unperturbed state or a current analysis state $\boldsymbol{x}_{\mathrm{a}}(t_1)$, e.g. retrieved from a data assimilation scheme, forming the new





NPGD
1, 1509–1537, 2014

Self-breeding – a new method to estimate local Lyapunov structures

J. D. Keller and A. Hense

perturbed state $\boldsymbol{x}'_a(t_1)$. Then the process is repeated by advancing both – the unperturbed state $\boldsymbol{x}_a(t_1)$ as well as perturbed state $\boldsymbol{x}'_a(t_1)$ – by $\Delta t$ to the next analysis time step, rescaling the perturbations and again, adding them to the current analysis state and so on.

With the iterative rescaling of the growing error structures, the trajectory of the perturbed state is relocated in the neighbourhood of the phase space location of the unperturbed or analysis state. Thus the errors are again small in relation to the new initial state of the model. Therefore it is expected that they will grow exponentially and that the resulting error structures – the BVs – will quickly converge towards the leading LV (Kalnay et al., 2002).

One of the advantages of the method is that it makes use of the full non-linear model to estimate the uncertainty structures of the state. Therefore no linear assumptions have to be made and no tangent linear version of the model has to be determined. The method is easy to implement and has been used in multiple systems (e.g. Carassi et al., 2007; Magnusson et al., 2008; Keller et al., 2008).

In addition, the BV method can be tuned to the specific needs of the user. For example, the use of different norms (e.g. Pazó et al., 2013) can lead to different approximations of uncertainty structures, i.e. the leading local Lyapunov vector with respect to the properties of the breeding cycle.

### 2.2 Ensemble approach

However, the user is often interested not only in the strongest growing error modes, i.e. the leading local LV, but in forecasting the PDF of the future state of the system given the current state. To estimate the full expansion of the uncertainty in phase space, one would need to explore every part of uncertainty contributing to the PDF of the current state. However, for higher dimensional models, such an undertaking is futile due to the large number of degrees of freedom.

A common approach is the so called ensemble method. A sample is taken from the PDF of the current state and advanced in the future using $\mathcal{M}$. However, depending on the size of the sample, the effects of the strongest growing modes upon the ensemble uncertainty may likely not be included in the sample and thus the ensemble forecast will not comprise the errors giving the largest contribution to the future state PDF. If we assume an ensemble size $N$, we would want our ensemble to include the local LVs corresponding to the $N$th largest local Lyapunov exponents $\lambda_1, \ldots, \lambda_N$. Such initial condition perturbations will provide us with the most probable forecast error structures given our sample from the initial state PDF. Such an approach will at best make use of *forward* LVs as these project the initial state of the forecast onto the most likely directions of growing errors in the future.

The breeding cycle as described in Sect. 2.1 is commonly extended to not only use one initial perturbation but $N$ different perturbations. Due to the fact, that the BVs converge towards the leading local LV, we are provided with $N$ only slightly different states. Corazza et al. (2003) show that additional realizations of that breeding cycle will not result in an effective increase of the dimension of the perturbation subspace. However, the small differences among the BV realizations define a subspace of error growing modes which can be used to its full extend by orthogonalizing the BVs (see Sect. 2.4 and Keller et al., 2010, for details).

Such orthogonalized BVs are still only based on the past evolution of error growth and do not neccessarily project onto the strongest growing error modes in the system's immediate future states. In the next section, we present an adaptation of breeding – the self-breeding method – which allows for the estimation of BVs associated with *forward* LV structures. In addition, this technique also provides the means to generate state uncertainty estimates for local area models (LAMs).

### 2.3 Self-breeding

We propose an alternate approach to implement the breeding cycle which we call self-breeding. In our method the breeding cycle is applied to the same time period over and over again until characteristic perturbations for that specific time have evolved. Different norms and rescaling periods can be used to target specific scales or phenomena. The perturbations estimated with the self-breeding method represent uncertainty structures associated with local Lyapunov vectors. Note that these estimates correspond to the specific time and to the targeted scale.

The method can be applied several times with different initial perturbations leading to a set of different estimates of uncertainty structures, i.e. the BVs. These structures can serve as state-dependent initial conditions for ensemble forecasts which allow us to predict the PDF of the near future model state – again, for a specific time and scale. The BVs estimated with the self-breeding method are comparable with the perturbations estimated using the SV method in a sense that they are the estimates of the future error growth, i.e. they are associated with the *forward* LVs. In comparison to the SV method, self-breeding generates these perturbations by applying the full non-linear model.

Our method is implemented as follows. We choose an arbitrary rescaling cycle time interval $\delta t$ for which we want to generate a set of localized BVs. Our breeding cycle will then start at the initial time $t_{\text{initial}}$. We denote the unperturbed or control state at the initial time step as $\boldsymbol{x}(t_{\text{initial}})$.

The initialization of the self-breeding cycle can be done in analogy to the standard breeding cycle with random perturbations. However, complex non-linear models usually have several different parameters which are linked to each other as defined by the underlying system. Therefore, an alternative approach would be to generate reasonable initial conditions for the cycle which are balanced between the model parameters. One possibility to accomplish this is to conduct an initial run prior to $t_{\text{initial}}$ which is also initiated using random perturbations but allows the model to adapt to these perturbations and generate balanced perturbations as initial conditions for the self-breeding cycle.

The perturbed state is denoted as $\widetilde{\boldsymbol{x}}_0(t_{\text{initial}})$ with the subscript 0 indicating the initial cycle of the self-breeding process. In a next step the full non-linear model $\mathcal{M}$ is then used to advance the states $\boldsymbol{x}(t_{\text{initial}})$ as well as $\widetilde{\boldsymbol{x}}_0(t_{\text{initial}})$ to time $t_{\text{end}}$. Then, a norm is used to measure the perturbation growth





NPGD
1, 1509–1537, 2014

$$eg = \frac{\|z_0(t_{\text{end}})\|}{\|z_0(t_{\text{initial}})\|} \tag{5}$$

over the elapsed time with $z_0(t_{\text{initial}}) = \tilde{\boldsymbol{x}}_0(t_{\text{initial}}) - \boldsymbol{x}_0(t_{\text{initial}})$ and $z_0(t_{\text{end}}) = \tilde{\boldsymbol{x}}_0(t_{\text{end}}) - \boldsymbol{x}_0(t_{\text{end}})$ respectively. $z_0(t_{\text{end}})$ is then rescaled to the initial perturbation amplitude and added to the unperturbed control state at $t_{\text{initial}}$ thus leading to the initial state for the next breeding cycle $\tilde{\boldsymbol{x}}_0(t_{\text{initial}})$. The process is then repeated to "breed" the perturbations. The cycling can be conducted for a fixed number of cycles or with some terminating condition, e.g. until an approximation of the error growth to a saturation value can be observed.

## 2.4 Orthogonalization

As mentioned before, BVs tend to converge towards the leading local LV hence exhibiting only small variations among their state vector structure. In order to maximize the subspace spanned by the BVs an orthogonalization is applied as an ensemble transform (ET, e.g. Bishop and Toth, 1999). The orthogonalization is implemented as a singular value decomposition (SVD) of the similarity matrix of the single BVs (see Keller et al., 2010, for details). In this implementation we calculate the matrix over all timesteps in a breeding cycle such that

$$\mathbf{C}_{ij} = \sum_{t=1}^{N_t} \sum_{i=1}^{N_{\text{BV}}} \sum_{j=1}^{N_{\text{BV}}} \left( z_i^{(c,t)} \cdot z_j^{(c,t)} \right) \tag{6}$$

is one element of the similarity matrix with $N_t$ the number of timesteps in a breeding cycle, $N_{\text{BV}}$ the number of BVs generated, $z_i^{(c,t)}$ the pertubation of BV $i$ of cycle $c$ and time step $t$.

In this way the orthogonalization is only performed in the subspace of the BVs which is $N_{\text{BV}}$-dimensional. For common applications in geophysics such as weather





NPGD
1, 1509–1537, 2014

Self-breeding – a new method to estimate local Lyapunov structures

J. D. Keller and A. Hense

forecasting $N_{BV}$ is much smaller compared to the dimension of the model phase space itself hence making the orthogonalization feasible at very low computational costs.

## 3 Model and experiments

We conduct a series of experiments in order to test the proposed self-breeding algorithm using a simple model nevertheless exhibiting chaotic behavior.

### 3.1 Model

For our experiments on the self-breeding algorithm, we use the Lorenz96 model (Lorenz, 1995) which is a toy model developed in order to investigate non-linear system behavior in a simple environment. The Lorenz96 model can be seen as a representation of a latitude circle of the earth's atmosphere reproducing continuous time variations in its parameter while exhibiting chaotic behavior. For each of the $N_x$ parameters in the model, the dynamics are described by

$$\frac{dx_k}{dt} = (x_{k+1} - x_{k+2}) \cdot x_{k-1} - x_k + F. \quad (7)$$

The dynamics comprise quadratic non-linear advection, dissipation as well as a forcing term ($F$). The model is integrated using a 4th-order Runge–Kutta scheme with the time step as well as the forcing being set to the values originally proposed by Lorenz ($\Delta t = 0.005$, $F = 8$) but for a system dimension of $N = 50$.

### 3.2 Experiment setup

To thoroughly examine the characteristics of the perturbation structures determined with the self-breeding algorithm, we employ several experiments with different setups. These are designed to illustrate the characteristics and sensitivities of the method regarding initial perturbation, rescaling interval and diversity. The basis for the





NPGD
1, 1509–1537, 2014

experiments is an unperturbed control run generated as reference with a spin up of 1000 time steps.

Then 50 sets of random perturbations are generated as deviations from the control representing our initial conditions for the self-breeding process. This yields 50 realizations for the same initial time step. Further, the procedure is conducted for 10 different initial time steps.

To investigate the impact of different rescaling interval lengths, the aforementioned experiments are repeated for different settings for the self-breeding cycle rescaling interval $\delta t$ from 10 to 100 time steps with an increment of 10. In each case, the end of the cycling is set to the same time step of the control run in order to compare the characteristics from the generated breeding modes.

## 4 Local Lyapunov estimates

In order to investigate the representation of error growth characteristics defined by the phase space of the Lyapunov vectors, the set of local Lyapunov vectors has to be determined, i.e. the spatial structures representing the directions of error growth for the section of the model's attractor corresponding to the spatio-temporal state of the self-breeding interval. To analyze the subspace of the attractor, 50 random perturbations are generated for each initial time step of each self-breeding experiment with the maximum perturbation amplitude varying from 0.005 to 0.1 and simulations are conducted using these perturbations on the initial conditions.

Then the linearized propagator for the 4th-order Runge–Kutta-scheme

$$\mathbf{M}_t = \mathbf{I} + \frac{1}{6}\left[\left(\mathbf{I} + \left(\mathbf{I} + \left(\mathbf{I} + \frac{1}{2}\Delta t \mathbf{J}_3\right)\frac{1}{2}\Delta t \mathbf{J}_2\right)\Delta t \mathbf{J}_1\right)\Delta t \mathbf{J}_0 \right. \\ \left. + \left(2\mathbf{I} + \left(\mathbf{I} + \frac{1}{2}\Delta t \mathbf{J}_3\right)\Delta t \mathbf{J}_2\right)\Delta t \mathbf{J}_1 + (2\mathbf{I} + \Delta t \mathbf{J}_3)\Delta t \mathbf{J}_2 + \Delta t \mathbf{J}_3\right] \quad (8)$$





NPGD
1, 1509–1537, 2014

is calculated at each time step of each simulation with $\mathbf{J}_i$ the Jacobians of the system at the (intermediate) Runge–Kutta time steps, $\mathbf{I}$ the identity matrix and $\Delta t$ the model time step. These propagators are then sequentially applied (i.e. the matrices are multiplied) to generate an overall propagator for the simulations period

$$\mathbf{M} = \mathbf{M}_{N_t}\mathbf{M}_{N_t-1}\ldots\mathbf{M}_1. \qquad (9)$$

In order to determine the Lyapunov exponents and vectors, a matrix

$$\mathbf{M}_2 = \mathbf{M}\mathbf{M}^{\mathrm{T}} \qquad (10)$$

is caluculated. Then the average of these matrices $\overline{\mathbf{M}_2}$ is determined over all 51 simulations (1 control plus 50 random perturbed runs). Finally, a singular value decomposition is used obtain a set of eigenvalues $e_i$ and eigenvectors $\boldsymbol{e}_i$. Then, $\lambda_i = \ln e_i / N_t$ and $\boldsymbol{e}_i$ represent the local Lyapunov exponents and vectors corresponding to the evolution of the system over the corresponding self-breeding cycling interval. The procedure is then repeated for all self-breeding cycling periods and all target time steps. In theory, the global Lyapunov exponents and vectors could be obtained by repeating this procedure for an inifinite number of target time steps and averaging over the resulting structures.

Results for the local Lyapunov estimation can be found in Figs. 1 and 2. The local Lyapunov spectra for all experiments are shown in Fig. 1 with the exponent decreasing from left to right. The small black vertical line in each row represents the numerical shift from positive to negative exponents.

The differences between the breeding cycling intervals (represented by the different rows in the diagrams ascending from top to bottom) exhibit a flattening of the spectra from smaller to larger rescaling intervals. Similarly, the number of positive and negative exponents is shifted from more negative to more epositive exponents with increasing rescaling intervals for most target time steps. However, the change of the spectra with the rescaling intervals is not the same for all target time steps, thus indicating the dependence of the local Lyapunov exponents on the system's state.





NPGD

1, 1509–1537, 2014

Figure 2 shows the first four Lyapunov vector estimates for the target time step 2000 with each column representing the vector in model space determined for a specific trajectory length (rescaling interval). While the structures exhibit certain similarities among the rescaling intervals especially for the vectors corresponding to the first two local Lyapunov exponents (right panels), the structures still differ even though they are estimated for the same target time step. This variability is more pronounced for vectors 3 and 4 (left panels).

## 5 Results

### 5.1 Self-breeding implementation

First, the evolution of the BVs developing when using the self-breeding technique is investigated. Figure 3 shows a Hovmüller diagram of the self-breeding cycles for the first five realizations for 25 cycles with a 20 time steps rescaling interval.

After approximately ten cycles, all realizations (also those not shown) have attained a steady four-dimensional state, i.e. the model perturbations converge to the same perturbed state at every iteration of the self-breeding cycle. While the model state evolution in the first cycles is different depending on the initial random perturbations, their spatial structure is the same among all realizations after the convergence except for the sign and minimal variations among the realizations. This structure would be expected to be in correspondence with the leading local Lyapunov vector. However, a comparison with the leftmost columns in the panels of Fig. 2 suggests that the BVs are in accordance with the second local Lyapunov vector. In the initial cycles, realizations 1, 2, 4 and 5 seem to converge towards the structure associated with the leading local Lyapunov exponent but are then drawn to the second.

The average error growths (for each cycle) corresponding to the BV structures (lower panel in Fig. 3) also demonstrate the convergence of the uncertainty estimates, with all realizations exhibiting the same value after the 10th iteration. During the initial phase,





NPGD
1, 1509–1537, 2014

the error growth exhibits divergent behavior for the five realizations shown. All realizations exhibit a reduced error growth compared to the saturation error growth sooner or later during the initial phase (up to cycle 11).

While realizations 1 and 4 quickly converge to the saturation error growth after a few self-breeding cycles, realization 5 shows a pronounced error growth in the second cycle which is larger than that during the saturation phase. Here, the spatial structures strongly correspond to the leading Lyapunov vector, thus explaining the higher error growth rate. During the following transition phase of this realization towards the steady state, the error growth strongly decreses (with a minimum in cycle 8) before rising again and finally reaching the saturation value in cycle 11. This behavior indicates that the self-breeding technique may be able to reproduce other modes of error growth, although not by applying the simple implementation. This is further discussed in Sect. 5.3.

To compare the perturbation structures for different rescaling intervals, further experiments are conducted with a time shift of the initial state for each rescaling interval such that all self-breeding cycles end at the same time step relative to the control run. As discussed in Sect. 4, the local Lyapunov vectors differ for the different rescaling lenghts. Therefore, it is expected that the self-breeding algorithm also produces different BV structures.

As can be seen from Fig. 4, the converged error mode structures from all experiments exhibit similar spatial structures. However, the structures become more complex and their spatial extent in model space is growing from shorter to longer rescaling intervals (top to bottom). This indicates that the spatial self-breeding BV structures are also not independent from the length of the rescaling interval, similar to the corresponding local Lyapunov vector estimate in Fig. 2 (second panel from the left).

The estimated BVs are specifically dependent on the processes corresponding to the temporal scales of the chosen rescaling interval and can therefore identify uncertainty structures, i.e. error growing modes, in accordance with the local Lyapunov estimates corresponding to the respective rescaling interval. Therefore, the length of the rescaling





NPGD
1, 1509–1537, 2014

interval can be seen as an optimization time to target specific processes and their uncertainty or error growth characteristics with respect to the phase space state, i.e. to the error growth characteristics of the attractor sub-space in consideration.

In the experiments for target time step 2000, the maximum average error growth is obtained for a 40 time step rescaling interval with the growth rates being smaller for longer as well as shorter cycling periods (cf. Fig. 4). This also supports the hypothesis, that specific scales of error growth can be targeted by using the length of the rescaling interval as the different perturbation growth amplitudes can be attributed to processes at different temporal scales when considering the same region in the system phase space.

### 5.2 Estimation of local Lyapunov vector structures

To further investigate the representation of the phase space of model error growth by the BVs, one has to compare two vector spaces spanned by the respective two sets of bred and LVs. The LVs are calculated according to the method described in Sect. 4 and the BVs are taken as the mature state at the end of each self-breeding cycle. One can not expect a one-to-one correspondence between each vector tuple. However one can look for an orthogonal rotation matrix **U** applied to the BV and another orthogonal rotation **V** applied to the LVs such that the corresponding expansions coefficients form a diagonal matrix. This is a variant of the classical canonical correlation analysis (e.g. von Storch and Zwiers, 1999). The entries of the diagonal matrix are called canonical correlations and indicate that an orthogonal rotation of one specific BV can be found which projects perfectly onto a rotated LV in case that the canonical correlation is one. Canonical correlation less than one indicate that such a pair of rotations can only be found up to an angle $\varphi$ between a pair of rotated BVs and LVs such that the $\cos\varphi$ is equal to the canonical correlation.

The results of this procedure can be found in Fig. 5. The left panel shows the canonical correlation coefficients of the 50 components of the BVs ($y$ axis) to the estimated local Lyapunov vectors for the different self-breeding cycling intervals ($x$ axis) with a





NPGD
1, 1509–1537, 2014

separate diagram for each of the 10 different target time steps. While using the full set of BVs, the correlation coefficients are either 0 or 1. The right panel shows the same diagram but for the average over the ten different target time steps.

It can be seen that the representation of the 50-dimensional Lyapunov space by the BVs varies considerably among the different system states. The simple self-breeding BVs cover at least seven dimensions of the Lyapunov space (deepest dips in the histograms in Fig. 5 for target time step 2500 and breeding cycle length of 30 time steps). In other combinations of target time step and cycling interval length the full Lyapunov space is spanned by the BVs (highest levels of histograms). On average (Fig. 5b) either a short cycling length interval (20 time steps) or a longer cycling interval of 80 to 100 timesteps leads to BVs spanning the desired Lyapunov space while intermediate values provide BVs with less information about the LVs of the Lorenz96 model. This indicates that higher error growth rates which have been found to be largest at these intermediate cycling interval lengths lead to a more uniform development of the perturbations thereby further reducing variation and enhancing similarity among the BV structures.

With the number of canonical correlation coefficients being equal to the dimension of the model space, the set of BVs allows for a full representation of the phase space of error growth. However, in real world scenarios, the possible number of realizations (or ensemble members) is limited to a small fraction of the dimensions of the system. Therefore, canonical correlations between BVs and Lyapuov vectors are calculated with an increasing number of BVs. An example of the results of this procedure for a rescaling interval of 20 time steps and target time step 2000 can be found in the left panel of Fig. 7.

Using the first 11 BVs to estimate the Lyapunov space (left column) leads to a canonical correlation of 0.5 or higher for only 3 dimensions indicating a large projection of BVs onto the calculated LVs. This number increases with the number of BVs used but – for this example – reaches a saturation value of 16 dimensions. The large white area denotes the dimensions of Lyapunov phase space which are not represented through the simple self-breeding BVs. Thus, such a set of BVs would not be able to reproduce the full spectrum of error growth dimensions.

### 5.3 Ensemble transform

As can be seen in Fig. 3, the perturbations exhibit the strong similarities common for the breeding technique. This is due to the characteristic of the method that all members tend to converge towards one (local) Lyapunov vector. In order to account for this behavior, an ensemble transform (ET) approach is used to maximize the diversity among the realizations and to optimize the usage of the subspace of the perturbation phase space spanned by the bred vectors (cf. Sect. 2.4). The structures determined by this approach are hereafter called ensemble transform bred vectors (ETBVs).

Figure 6 shows the results of the ET implementation for the first five realizations with a rescaling interval of 20 time steps. In comparison to the simple breeding implementation, the orthogonalized perturbation structures exhibit a strong diversity. Here, each of the single members represents a different uncertainty structure corresponding to a different Lyapunov exponent. A comparison with the left column of the first four local Lyapunov vectors shown in Fig. 2 shows that the spatial structures of these Lyapunov estimates can be obtained from the first five ETBVs. The swapping of the ETBV structures between different realizations is a result of the orthogonalization process.

The lower panel of Fig. 6 is an error growth diagram for the ET self-breeding process. The results show that the maximum error growth is larger (approx. 5.5 % per time step) than that achieved with the simple self-breeding implementation (approx. 4 % per time step) which is to be expected as the BVs in the simple approach converge towards the local Lyapunov vector corresponding two the second largest Lyapunov exponent while the ETBV of realization 1 matches the spatial structures associated with the largest local Lyapunov exponent. Further, the error growth for the other realizations in the ET version varies considerably, representing the different directions of error growth in the subspace spanned by the local Lyapunov vectors corresponding to the system





NPGD
1, 1509–1537, 2014

state. Thus, the ETBVs allow for assessing a much larger spectrum of error growth characteristics than the simple BVs.

This is further supported by the canonical correlation analysis in Fig. 7. The right panel shows the canonical correlations for subsets of the ETBVs and the estimated local Lyapunov vectors. Compared to the simple self-breeding BVs in the left panel, the ETBVs are able to fully represent the respective subspace of error growth, i.e. the subspace of the local Lyapunov phase space. The missing dimension in each column results from the canonical correlation procedure itself as the calculations are based on patterns which have to sum to zero.

## 6 Conclusions

The self-breeding method is a new approach to estimate local Lyapunov structures in non-linear models. Thereby, the technique makes no simplifications or assumptions of linearity but uses the full non-linear model to estimate modes of error growth.

The experiments using the Lorenz96 model show that the simple self-breeding algorithm without orthogonalization is able to estimate growing modes which are independent from the initial perturbation. While the time necessary to identify the fastest growing error modes does not seem to be dependent on the number of breeding cycles, the spatial structures itself vary considerably among different rescaling intervals – even for the same target time step.

Further, the different error growth rates between experiments with different optimization times suggest that the BVs generated with self-breeding are not only dependent on the system state but also on the temporal scale of the self-breeding cycling. Hence, the self-breeding technique allows for targeting the specific uncertainties originating from the underlying non-linear process dominant on the temporal scale in consideration.

The BVs are therefore estimates for the perturbation structures specific to the rescaling interval which in turn can be seen as the error growth optimization time. When initializing an ensemble forecast at the initial time of the self-breeding cycle, the ensemble





NPGD
1, 1509–1537, 2014

will then exhibit an optimal error growth for this temporal scale. In that sense, the approach is similar to the singular vector method but without the assumption of linearity, i.e. no tangent linear or adjoint model is necessary to generate the growing error modes.

The study also shows that the self-breeding approach produces BVs that can at least partially represent the phase space of error growth. The extent of representation strongly depends on the system state (target time step) as well as the optimization time (rescaling interval length) such that for some combinations, the full local Lyapunov phase space can be reproduced.

However, a major shortcoming of simple BVs is the uniformity among the estimated error mode structures which could well be the result of the very regular Lorenz96 model. The inclusion of an ensemble transform implementation into the self-breeding cycle does not only abate the homogeneity among the BVs, but constitutes a complete estimate of the first $N_{BV}$ dimensions of the Lyapunov vector phase space. A comparison to the estimated first Lyapunov vectors shows that these structures can even be reproduced by the ETBVs.

Thus, self-breeding provides a simple way to estimate *forward* error growing modes which can be used for covariance estimation as well as ensemble initialization. The results presented are strictly valid only for the Lorenz96 model which exhibits a higher number of degrees of freedom compared to models usually employed in such experiments. However, the results are not yet representative for complex real world models, e.g. weather forecasting models.

*Acknowledgements.* The authors want to thank the German Federal Ministry for Transportation and Digital Infrastructure for funding this work in the framework of the Hans-Ertel-Centre for Weather Research.





NPGD
1, 1509–1537, 2014



Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

## References


Berner, J., Shutts, G. J., Leutbecher, M., and Palmer, T. N.: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system, J. Atmos. Sci., 66, 603–626, 2009. 1510

Bishop, C. H. and Toth, Z.: Ensemble transformation and adaptive observations, J. Atmos. Sci., 56, 1748–1765, 1999. 1512, 1518

Buizza, R. and Palmer, T.: The singular-vector structure of the atmospheric global circulation, J. Atmos. Sci., 52, 1434–1456, 1995. 1512

Carassi, A., Trevisan, A., and Francesco Uboldi, F.: Adaptive observations and assimilation in the unstable subspace by breeding on the data-assimilation system, Tellus A, 59, 101–113, doi:10.1002/qj.2115, 2007. 1515

Corazza, M., Kalnay, E., Patil, D. J., Yang, S.-C., Morss, R., Cai, M., Szunyogh, I., Hunt, B. R., and Yorke, J. A.: Use of the breeding technique to estimate the structure of the analysis "errors of the day", Nonlin. Processes Geophys., 10, 233–243, doi:10.5194/npg-10-233-2003, 2003. 1516

Fujisaka, H.: Statistical dynamics generated by fluctuations of local Lyapunov exponents, Prog. Theor. Phys., 70, 1264–1275, 1983. 1511

Kalnay, E., Corazza, M., and Cai, M.: Are bred vectors the same as Lyapunov vectors?, in: AMS Symposium on Observations, Data Assimilation and Probabilistic Prediction, Orlando, FL, USA, 173–177, 2002. 1515

Keller, J. D., Kornblueh, L., Hense, A., and Rhodin, A.: Towards a GME ensemble forecasting system: Ensemble initialization using the breeding technique, Meteorol. Z., 17, 707–718, 2008. 1515

Keller, J. D., Hense, A., Kornblueh, L., and Rhodin, A.: On the Orthogonalization of Bred Vectors, Weather Forecast., 25, 1219–1234, 2010. 1516, 1518

Legras, B. and Vautard, R.: A guide to Liapunov vectors, in: Proceedings of the Seminar on Predictability, vol. 1, ECMWF, Reading, UK, 141–156, 1995. 1513

Lorenz, E. N.: A deterministic non periodic flow, J. Atmos. Sci., 20, 130–141, 1963. 1510

Lorenz, E. N.: Predictability: a problem partly solved, in: Proceedings of the Seminar on Predictability, vol. 1, 1–18, ECMWF, Reading, UK, 1995. 1519

Lorenz, E. N. and Emanuel, K. A.: Optimal sites for supplementary weather observations: simulation with a small model, J. Atmos. Sci., 55, 399–414, 1998. 1512






NPGD
1, 1509–1537, 2014

Self-breeding – a new method to estimate local Lyapunov structures

J. D. Keller and A. Hense

Lyapunov, A. M.: The general problem of the stability of motion, Int. J. Control, 55, 531–773, 1992. 1511

Magnusson, L., Leutbecher, M., and Källén, E.: Comparison between Singular Vectors and Breeding Vectors as Initial Perturbations for the ECMWF Ensemble Prediction System, Mon. Weather Rev., 4092–4104, doi:10.1175/2008MWR2498.1, 2008. 1515

Oseledec, V. I.: Multiplicative ergodic theorem: characteristic Lyapunov exponents of dynamical systems, Trans. Moscow Math. Soc., 19, 197–231, 1968. 1511

Pazó, D., López, J. M., and Rodríguez, M. A.: The geometric norm improves ensemble forecasting with the breeding method, Q. J. Roy. Meteorol. Soc., 139, 2021–2032, doi:10.1002/qj.2115, 2013. 1515

Szunyogh, I., Kalnay, E., and Toth, Z.: A comparison of Lyapunov and optimal vectors in a low-resolution GCM, Tellus A, 49, 200–227, doi:10.1034/j.1600-0870.1997.00004.x, 1997. 1512

Toth, Z. and Kalnay, E.: Ensemble forecasting at NMC: the generation of perturbations, B. Am. Meteorol. Soc., 74, 2317–2330, 1993. 1513

Toth, Z. and Kalnay, E.: Ensemble forecasting at NCEP and the breeding method, Mon. Weather Rev., 125, 3297–3319, 1997. 1513

von Storch, H. and Zwiers, F. W.: Statistical Analysis in Climate Research, Cambridge University Press, 1999. 1524

Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A.: Determing Lyapunov exponents from a time series, Physica D, 16, 285–317, 1985. 1513





NPGD

1, 1509–1537, 2014

Self-breeding – a new method to estimate local Lyapunov structures

J. D. Keller and A. Hense

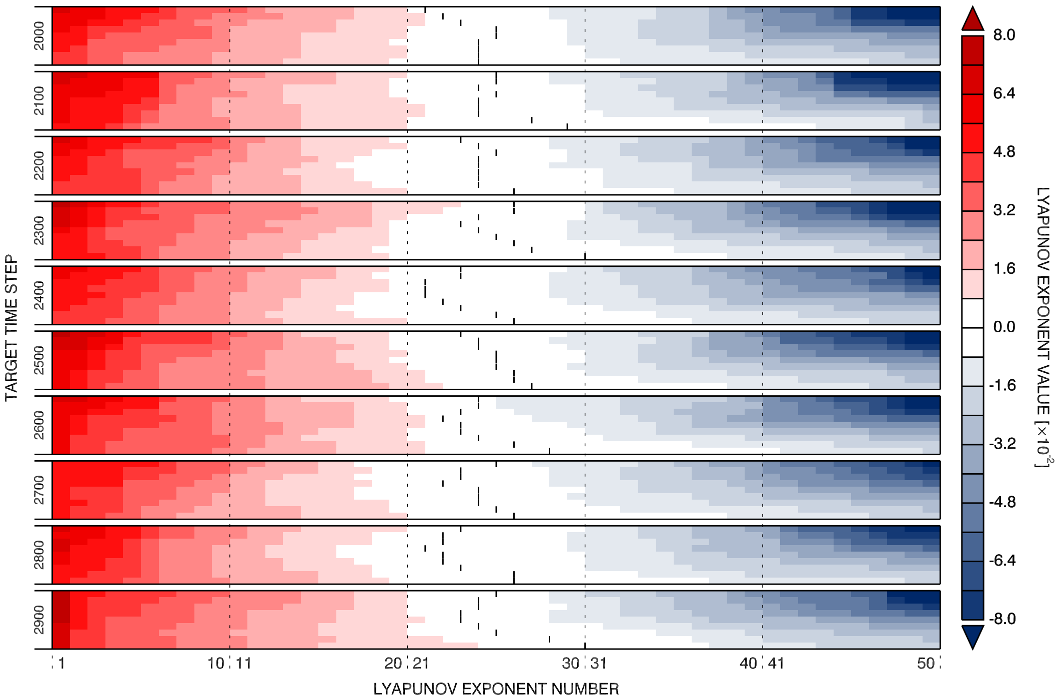


**Figure 1.** Estimated local Lyapunov spectra for all experiments. The 10 panels depict the Lyapunov exponents for the target time steps (as denoted on the left) and the different values for the 9 self-breeding rescaling intervals (20 to 100 from top to bottom) within each diagram. The vertical black bar in each row denotes the change of sign from positive to negative Lyapunov exponent.





NPGD

1, 1509–1537, 2014

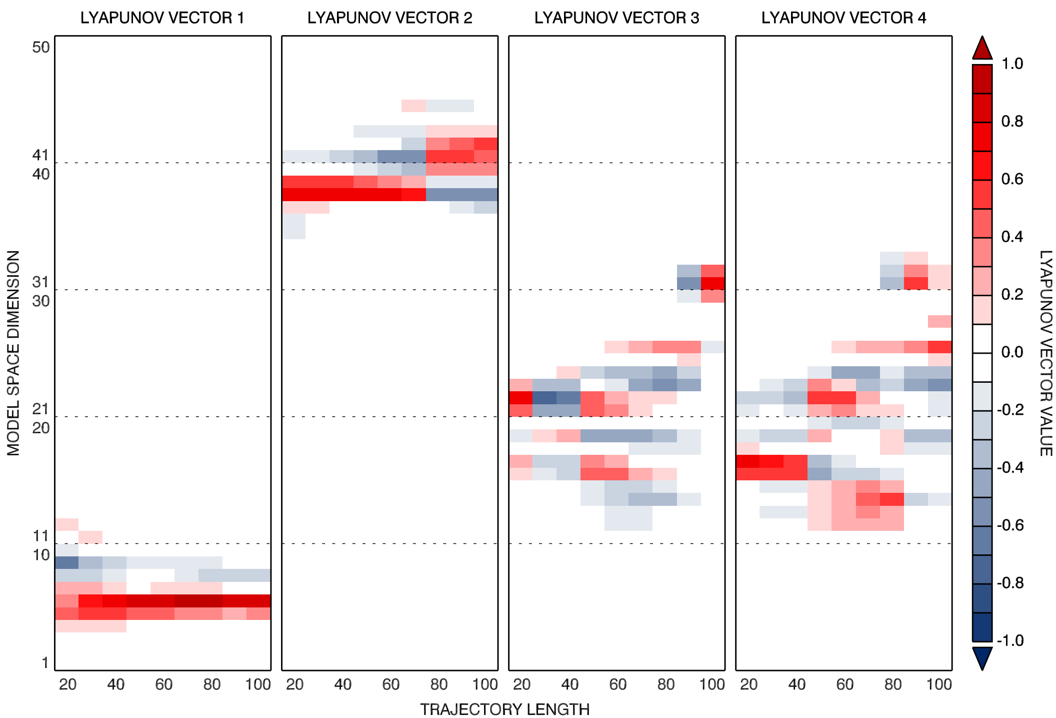


**Figure 2.** The estimated local Lyapunov vectors corresponding to the largest 4 Lyapunov exponents for the first target time step (2000). The model space dimensions are denoted on the $y$ axis. The columns in each diagram represent the estimated Lyapunov vectors for a self-breeding cycling interval (left 20, right 100).





NPGD
1, 1509–1537, 2014



J. D. Keller and A. Hense

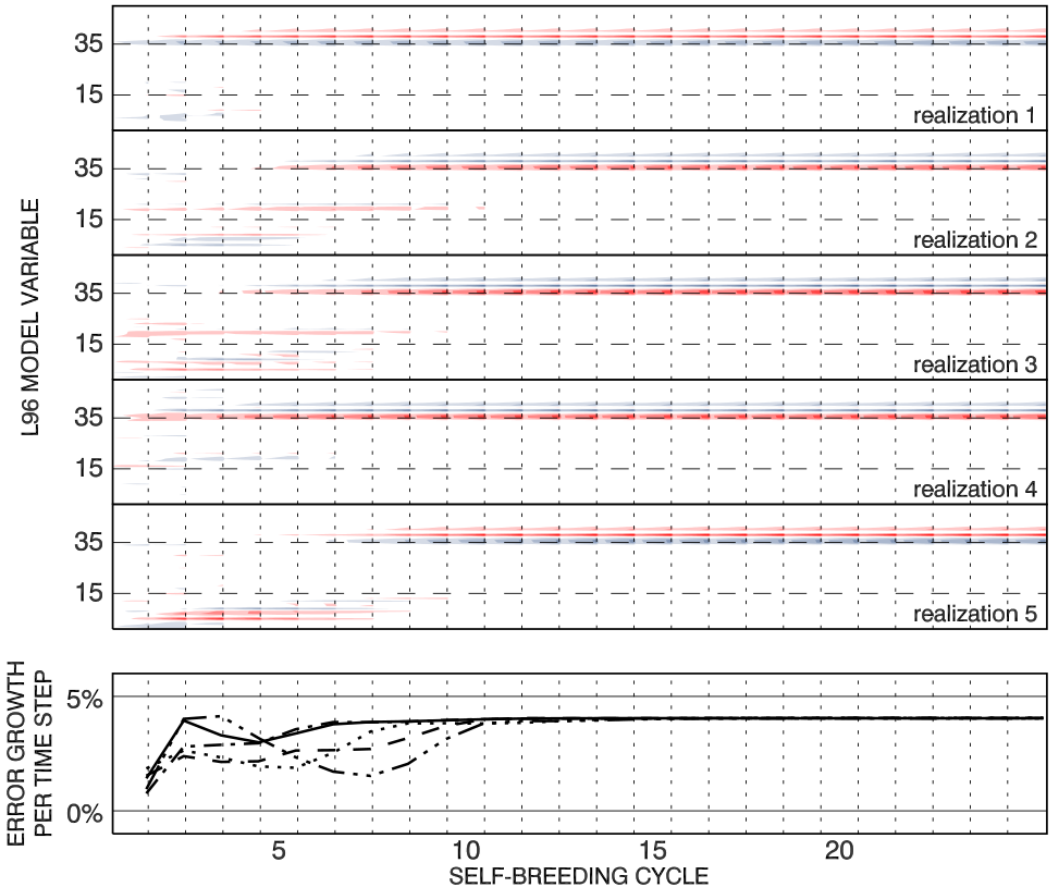


**Figure 3.** Hovmüller diagrams (upper panel) of the perturbations for the realizations 1 to 5 with rescaling intervals of 20 time steps and the corresponding mean error growth per time step (lower panel) for each rescaling interval (solid realization 1, dotted 2, dashed 3, dash-dotted 4 and dash-dot-dot-dotted 5).

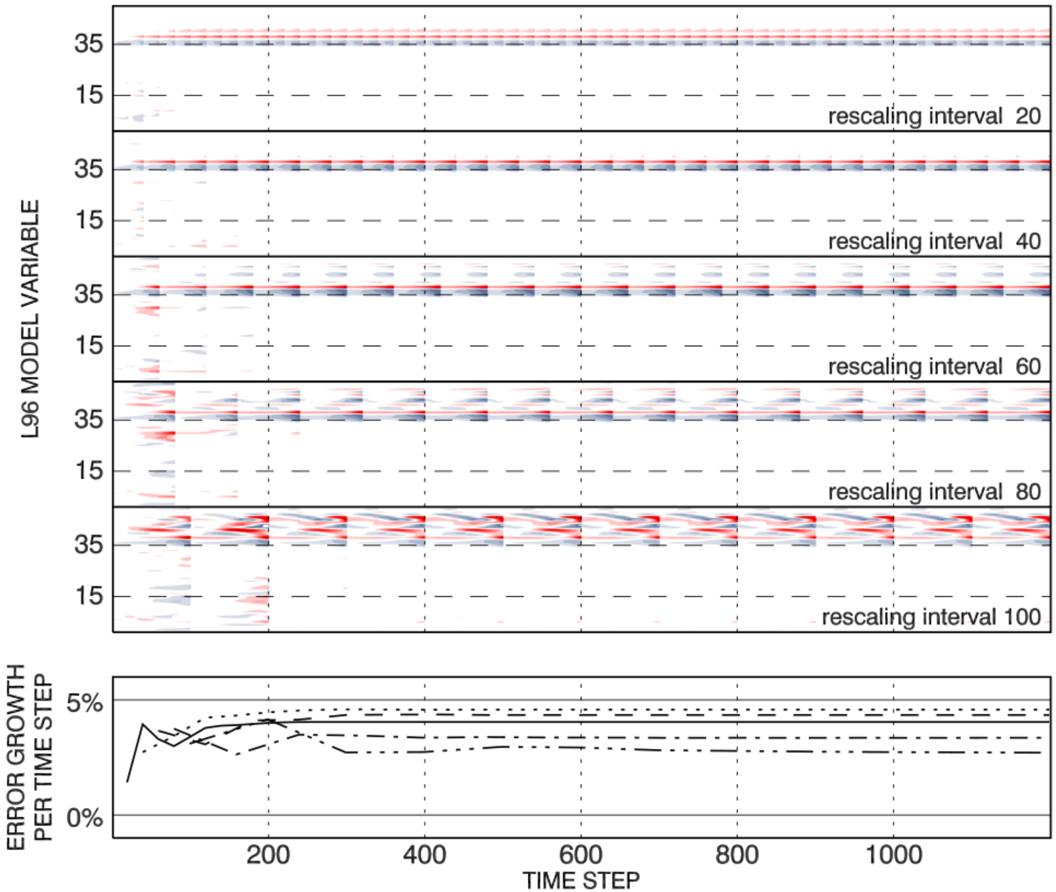


**Figure 4.** Hovmüller diagrams (upper panel) of the perturbations for the first realization from the set of experiments with different rescaling intervals and the corresponding mean error growth per time step (lower panel) for each rescaling interval (solid 20, dotted 40, dashed 60, dash-dotted 80 and dash-dot-dot-dotted 100 time steps).





NPGD
1, 1509–1537, 2014

Self-breeding – a new method to estimate local Lyapunov structures

J. D. Keller and A. Hense

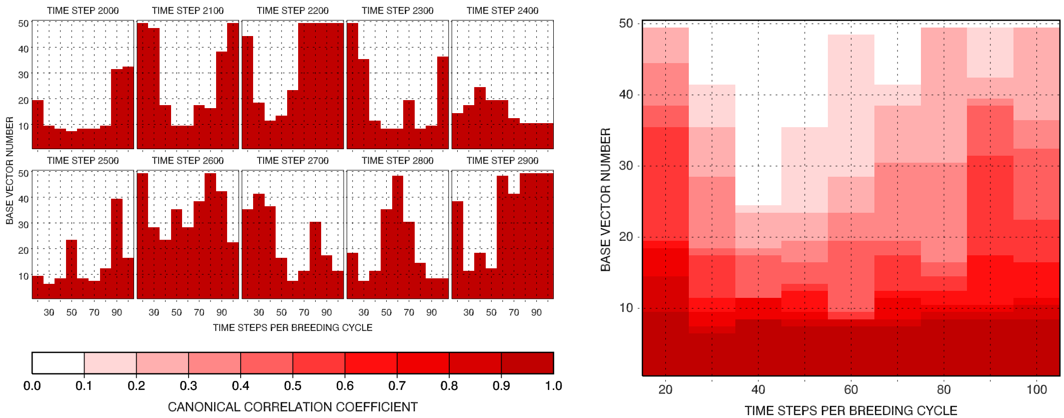


**Figure 5.** Canonical correlations between the set of simple self-breeding BVs and the full set of estimated Lyapunov vectors for the 10 different target time steps (left panels) and the average over these 10 cycling intervals (right panel).

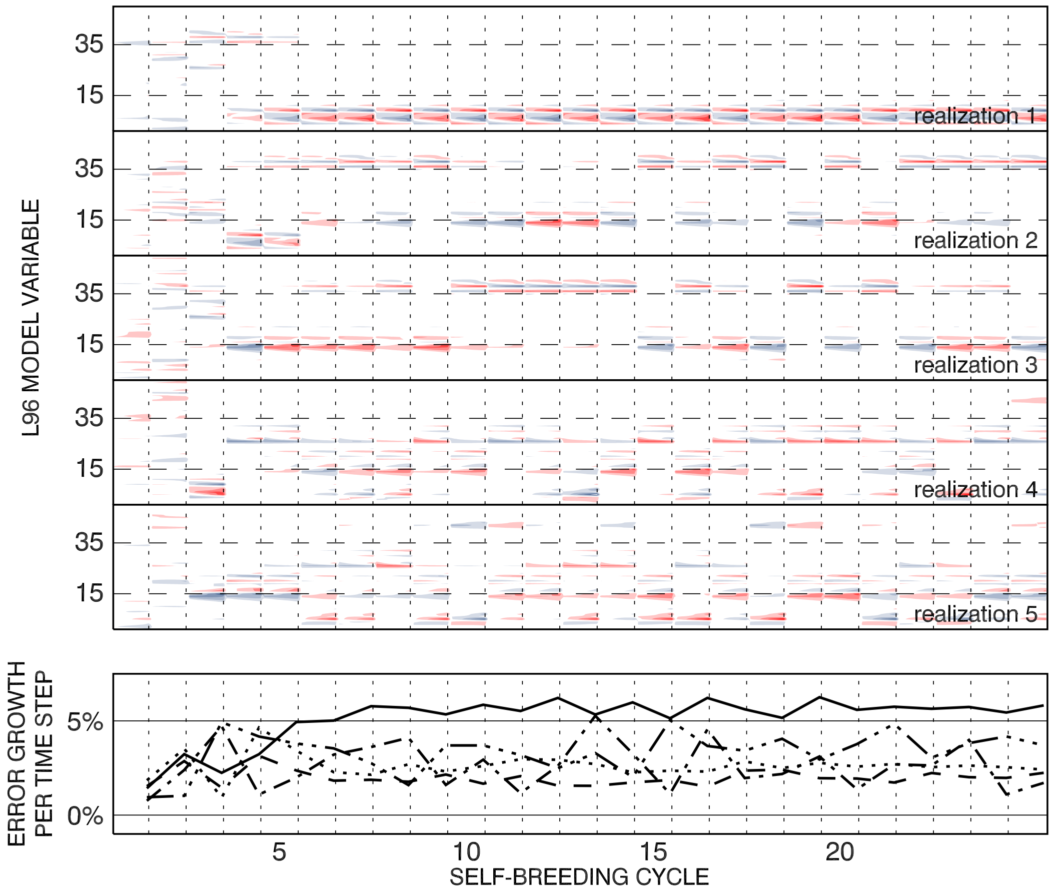


**Figure 6.** Same as Fig. 3 but for the implementation using the ensemble transform to maximize the usage of the uncertainty phase space.

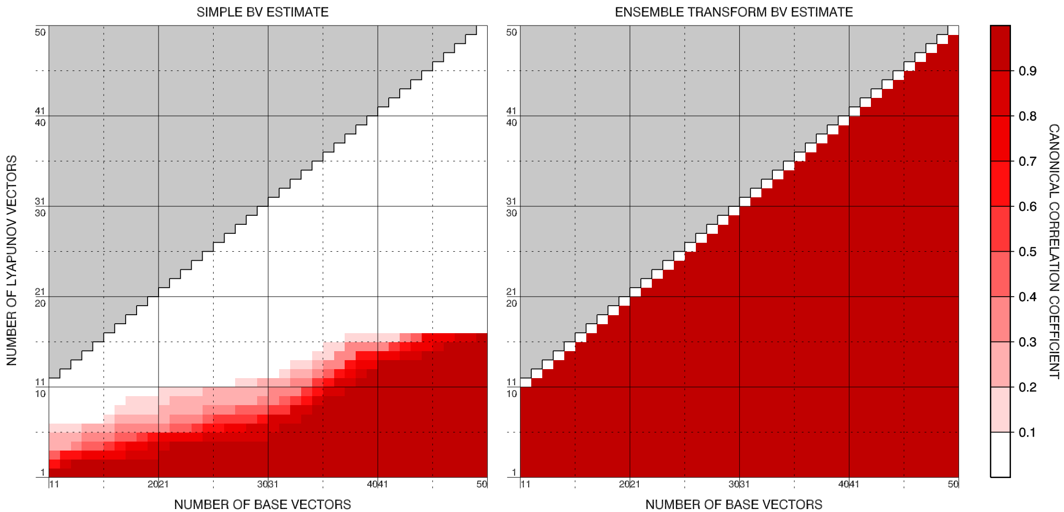


**Figure 7.** Canonical correlations between estimated Lyapunov vectors and simple BVs (left panel) and BVs from the orthogonalized self-breeding procedure (right panel).





NPGD
1, 1509–1537, 2014



J. D. Keller and A. Hense