

The role of model dynamics in ensemble Kalman filter performance for chaotic systems

By GENE-HUA CRYSTAL NG^{*†}, DENNIS MCLAUGHLIN, DARA ENTEKHABI and ADEL AHANIN, *Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA*

(Manuscript received 7 June 2010; in final form 9 July 2011)

ABSTRACT

The ensemble Kalman filter (EnKF) is susceptible to losing track of observations, or ‘diverging’, when applied to large chaotic systems such as atmospheric and ocean models. Past studies have demonstrated the adverse impact of sampling error during the filter’s update step. We examine how system dynamics affect EnKF performance, and whether the absence of certain dynamic features in the ensemble may lead to divergence. The EnKF is applied to a simple chaotic model, and ensembles are checked against singular vectors of the tangent linear model, corresponding to short-term growth and Lyapunov vectors, corresponding to long-term growth. Results show that the ensemble strongly aligns itself with the subspace spanned by unstable Lyapunov vectors. Furthermore, the filter avoids divergence only if the full linearized long-term unstable subspace is spanned. However, short-term dynamics also become important as non-linearity in the system increases. Non-linear movement prevents errors in the long-term stable subspace from decaying indefinitely. If these errors then undergo linear intermittent growth, a small ensemble may fail to properly represent all important modes, causing filter divergence. A combination of long and short-term growth dynamics are thus critical to EnKF performance. These findings can help in developing practical robust filters based on model dynamics.

1. Introduction

The ensemble Kalman filter (EnKF) is a flexible data assimilation tool for various geophysical applications. It offers a way to assimilate observations that are often noisy and incomplete with models that describe the non-linear physics of systems such as the atmosphere or ocean. The EnKF approximates the state error covariance with an ensemble of replicates obtained from model simulations. The approximate ensemble covariance is then used to update model state estimates when observations are available. The filter is applied sequentially to produce a series of forecast and analysis estimates. However, a number of challenges plague ensemble data assimilation methods for atmospheric and ocean models. Due to the very large dimension sizes typical of these types of models, only small ensemble sizes are feasible and covariance sampling errors can be severe. The chaotic dynamics of these strongly non-linear systems can then easily cause errors in the estimate to grow rapidly.

Difficulties in applying the EnKF to large chaotic models commonly result in filter divergence (e.g. Houtekamer and Mitchell,

1998; Hamill et al., 2001; Whitaker and Hamill, 2002; Sacher and Bartello, 2009). ‘Divergence’ describes the situation when the filter becomes so overconfident around an incorrect state that subsequent observations are ignored, and the estimate cannot be moved back toward the true state. This occurs if the spread in the ensemble underestimates the error in the estimate. Hamill et al. (2001) suggests that this happens when the variance is too small or cross covariance terms are too large in the ensemble covariance. Past studies have attributed the tendency for divergence to sampling noise in the small ensemble approximation of the error covariance. Focusing on the update step of the EnKF, van Leeuwen (1999), Furrer and Bengtsson (2007) and Sacher and Bartello (2008) have presented theoretical work demonstrating that error in the forecast error covariance results in a downward bias in the analysis error variance. Thus, inevitable sampling error in the forecast ensemble leads to insufficient spread in the analysis ensemble, which can ultimately cause the filter to diverge.

A number of fixes have been developed to enable the application of EnKF in light of these problems. Variance inflation (Anderson and Anderson, 1999) is added to directly counteract underestimated spread in the ensemble. Localization (Houtekamer and Mitchell, 2001) helps remove spurious long-distance covariances that can lead to large analysis errors. The use of multiple ensembles (Houtekamer and Mitchell, 1998) combats

*Corresponding author

e-mail: gng@usgs.gov

†Now at: U.S. Geological Survey, Menlo Park, CA, USA.

DOI: 10.1111/j.1600-0870.2011.00539.x

‘inbreeding’, or the interaction of sampling error with other terms in the EnKF update. These fixes often prove crucial for generating robust (non-divergent) estimates with the EnKF in various geophysical applications. However, they can adversely affect the optimality of the estimates and yield poorer RMS error performance (Whitaker and Hamill, 2002; Sakov and Oke, 2008). Although some efforts have been made for deriving theoretical bases for these approaches (e.g. Furrer and Bengtsson, 2007; Sacher and Bartello, 2008), inflation and localization are usually applied in a heuristic fashion. Inflation and localization parameter values that achieve robust results while maintaining minimal RMSE are problem-specific, and they are usually chosen through trial-and-error. Thus, although effective fixes may exist from an operational point of view, improving understanding of EnKF performance with large chaotic systems can help better direct their application, as well as provide insight into other robust estimation strategies.

This work examines an aspect influencing EnKF performance that has been typically overlooked in most filter divergence studies: the dynamic properties of the forecast model. As mentioned above, past works have generally focused solely on the update step in looking for sources of filter divergence. In one exception, Sacher and Bartello (2009) diagnose signs of divergence over multiple assimilation cycles using numerical experiments, but they point out that the non-linearity of the forecast model hampers more rigorous investigations. When isolating the update step from the forecast, sampling error is generically regarded as a random, unstructured element in most divergence studies. However, the chaotic behaviour of non-linear forecast models also plays an important role in causing errors to grow rapidly and lead to divergence. In this work, we are interested in how EnKF errors, including those that arise from using small ensemble sizes, relate to dynamic features of the model, and how that correspondence impacts filter performance.

Dynamic properties of a system have been considered in geophysical estimation contexts other than EnKF performance. Meteorological ensemble prediction strategies have included the use of singular vectors, bred vectors and Lyapunov vectors (Molteni et al., 1996; Toth and Kalnay, 1997; Pazó et al., 2010). Singular vectors and bred vectors have also provided a basis for adaptive observation networks (Lorenz and Emanuel, 1998; Hansen and Smith, 2000; Trevisan and Uboldi, 2004). Dynamic properties have also served as the motivation for certain reduced rank filter formulations (Cohn and Todling, 1996; Pham et al., 1998; Carme et al., 2001; Farrell and Ioannou, 2001; Heemink et al., 2001; Trevisan and Uboldi, 2004; and Trevisan et al., 2010).

A common aspect of these past studies is that dynamic analyses are typically used to justify a particular ensemble prediction, adaptive observation, or reduced rank estimation procedure. In this paper, we adopt a more diagnostic objective: we assess the impact of dynamic properties on EnKF performance for chaotic systems. It is not readily apparent how the ensemble relates to different types of growth modes in the system, and how filter

divergence depends on the ensemble filter’s ability to track these modes. Because the EnKF makes corrections only in the subspace spanned by the ensemble, must the ensemble include all error growth directions in order to avoid divergence? Should these directions correspond to short-term error growth, or to long-term error growth? If such a critical subspace exists, will the filter find it automatically, and with how many replicates? How does model non-linearity affect these issues? These are the types of questions that we address in this paper.

As noted by Sacher and Bartello (2009), the non-linearity of chaotic models hampers analysis of the EnKF over multiple assimilation cycles. Accordingly, our assessment of the role of model dynamics in EnKF performance will be carried out using numerical experiments. In order to evaluate a range of ensemble configurations, including the case of small sampling error (using a very large ensemble), our tests will be made with the small chaotic model presented in Lorenz and Emanuel (1998). Elucidating how model dynamics affect a data assimilation system is crucial to understanding its limitations and to engineering effective methods. This work is thus an important step in further developing robust data assimilation frameworks.

This paper is organized as follows. Sections 2 and 3 provide background on EnKF and chaos dynamics, respectively. Section 4 further describes past studies that link dynamic analysis with data assimilation. The methods used for our tests will be presented in Section 5, and the results and discussion in Section 6. General conclusions are summarized in Section 7.

2. The ensemble Kalman filter

The EnKF (Evensen, 1994) is a recursive state estimation method that provides sequential state estimates for non-linear systems, by combining model forecasts and observations. Evensen (2003) provides a review of the method. The EnKF alternates between a forecast step with the model and an update (also called analysis) step when observations are available. The update step is designed to produce the linear least-squares estimate of the model states, and thus requires the state mean and covariance in its calculation. The non-linear forecast step consists of Monte Carlo simulations with an ensemble of states that represent independent realizations of the process. Approximations of the covariance from the forecast ensemble are then used in the update calculation.

If the non-linear model is represented by $F(\cdot)$, the forecast step is

$$x^{f,i}(t) = F(x^{a,i}(t-1), \varepsilon^i(t), t), \quad (1)$$

where $x^{f,i}(t)$, $i = [1, \dots, N]$ are the forecast state realizations at time t , and $x^{a,i}(t-1)$, $i = [1, \dots, N]$ are the analysis state realizations at time $t-1$, and $\varepsilon^i(t)$ are random realizations that represent model error, and N is the ensemble size. The forecast mean and error covariance are calculated from the ensemble: $\bar{x}^f(t) = \frac{1}{N} \sum_{i=1}^N x^{f,i}(t)$ and $P^f(t) = \frac{1}{N-1} \sum_{i=1}^N (x^{f,i}(t) - \bar{x}^f(t))(x^{f,i}(t) - \bar{x}^f(t))^T$. The update step

to find the analysis state can be expressed in terms of the Kalman gain $K(t)$ and is

$$x^{a,i}(t) = x^{f,i}(t) + K(t)(y(t) + \delta^i(t) - Hx^{f,i}(t)) \quad (2)$$

$$K(t) = P^f(t)H^T(HP^f(t)H^T + R)^{-1}, \quad (3)$$

where $y(t)$ is the observation with a specified observation error covariance R , H is the observation operator that maps the state to the measurement space, and $\delta^i(t)$ are random realizations from $\mathcal{N}(0, R)$. Note that H is linear here, but it need not be in general. The analysis mean and error covariance are computed from the analysis ensemble, analogous to the forecast case. For linear systems, in the limit as $N \rightarrow \infty$, the ensemble mean of the EnKF matches the classical Kalman filter estimate, and inclusion of the random observation error term in (2) ensures that the ensemble analysis covariance converges to that of the classical Kalman filter.

The classical EnKF formulation in eqs (1)–(3) is a stochastic filter due to the random measurement error term in (3). Ensemble square root filters (EnSRF) avoid the random measurement term by transforming the ensemble forecast to match exactly the analysis error covariance required by the Kalman filter (see review by Tippett et al., 2003). Although divergence problems still affect EnSRFs due to sampling errors in the ensemble representation of the forecast error covariance, these filters produce improved estimates over the classical version (Whitaker and Hamill, 2002). This work employs an EnSRF presented in Sakov and Oke (2008). It is an ensemble transform Kalman filter (ETKF) with a mean-preserving formulation that includes random orthonormal matrix rotations to prevent the buildup of ensemble outliers. It should be noted that the random rotations do not alter the analysis covariance matrix, but only help ensure that variance is distributed among the different replicates. Only the ETKF update step differs from the classical EnKF. For $A = [A^1 | \dots | A^N]$, where $A^i = x^i - \bar{x}$, the update step becomes

$$A^a = A^f T, \quad (4)$$

where T is the ensemble transform matrix

$$T = C\Gamma^{-1/2}C^T U^p. \quad (5)$$

U^p is an arbitrary orthonormal mean-preserving matrix for the rotation, and C and Γ are found from the eigenvalue decomposition

$$C\Gamma C^T = I + \frac{1}{N-1}(HA^f)^T R^{-1} HA^f, \quad (6)$$

with I as the $N \times N$ identity matrix. The analysis mean \bar{x}^a is calculated from the forecast mean by taking the sample mean of eq. (2).

The ETKF is intended produce the exact Kalman analysis error covariance, but this occurs only if the ensemble forecast covariance P^f is equivalent to the true forecast covariance Λ^f . In

practice, the error in the ensemble forecast covariance approximation $\Delta = \Lambda^f - P^f$ can be significant when sample sizes are small, as is the case for large, computationally expensive models. Divergence studies that rigorously analyze the effect forecast covariance sampling error on the update step usually express Δ as a generic random error term (Whitaker and Hamill, 2002; Furrer and Bengtsson, 2007; and Sacher and Bartello, 2008). However, after multiple assimilation cycles, Δ is not purely random and unstructured sampling noise, but instead shows the systematic influence of the non-linear dynamic model through the replicates used to compute P^f with eq. (1). This work examines how errors shaped by the model dynamics may affect filter performance.

3. Chaos dynamics review

A number of geophysical systems, such as the atmosphere and ocean, behave as chaotic attractors. This means that although the system is bounded, perturbations to a state (or errors) can grow very rapidly until they span the entire attractor. As a result, filter divergence in such applications is not marked by unbounded error growth, but instead by the inability to identify where the true state lies within the attractor. To evaluate the EnKF relative to the dynamic properties of a forecast model, we compare ensembles of varying sample sizes by considering vectors that are commonly used to characterize chaotic behaviour, including singular and Lyapunov vectors. In this section, we review how these vectors are defined, and we describe the particular dynamic properties they represent.

3.1. Singular vectors

For small errors $\varepsilon(t) = x(t) - x^{\text{true}}(t)$ and short time τ , error growth in the forecast model can be approximated by the tangent linear model $\varepsilon(t + \tau) \approx M(t, t + \tau)\varepsilon(t)$, where the tangent linear model is calculated from a chain of Jacobian matrices derived from the non-linear model $F(\cdot)$

$$\begin{aligned} M(t, t + \tau) &= M(t, t + kdt) = M(t + (k-1)dt, t + \tau) \\ &\quad \times M(t + (k-2)dt, t + (k-1)dt) \\ &\quad \times \dots \times M(t, t + dt) \end{aligned} \quad (7)$$

$$M(t, t + dt) = \left. \frac{\partial F(x(t), t + dt)}{\partial x} \right|_{x(t)}. \quad (8)$$

$M(t, t + \tau)$ is an $n \times n$ matrix, where n is the state dimension of the model. $M(t, t + \tau)$ has the singular value decomposition

$$M(t, t + \tau) = USV^T, \quad (9)$$

where S is an $n \times n$ diagonal matrix containing the ordered singular values of $M(t, t + \tau)$ (in decreasing order) and U and V are $n \times n$ orthonormal matrices containing left and right singular vectors in their columns, respectively. The singular values found

with (9) optimize the growth of errors (measured in the Euclidean norm), over the period $[t, t + \tau]$. In particular, error along the direction of the i th right singular vector V_i grows according to the i th singular value s_i over $[t, t + \tau]$ and then points in the direction of the i th left singular vector U_i . Thus, the right singular vectors that correspond to singular values greater than 1 span a subspace of the state space in which errors at time t will grow under $M(t, t + \tau)$. The greatest rate of growth is along the direction of V_1 .

It should be noted that the relevance of singular vectors for describing chaotic dynamics depends on a couple of factors. Firstly, singular vectors and values correspond only to a specific time interval, which means for small τ values, they represent short-term growth. Because in the EnKF the state estimate is discontinuous at the update step, we calculate singular vectors of the tangent linear model over a single forecast period. Singular vectors corresponding to singular values greater than 1 indicate the portion of the state space that will grow over that forecast period, irrespective of earlier or later dynamics. Also, singular vectors characterize a system's error growth only when the Jacobian matrices, from which $M(t, t + \tau)$ is derived, are valid or good approximations. Accordingly, Hansen and Smith (2000) showed that singular vectors may only be useful for data assimilation if analysis intervals are sufficiently short and errors sufficiently small so that the tangent linear model adequately describes error dynamics.

3.2. Lyapunov vectors

Like singular vectors and singular values, Lyapunov vectors and Lyapunov exponents describe growth of infinitesimal errors under linearized dynamics. However, they differ in that Lyapunov vectors and exponents characterize growth over an asymptotically long rather than an infinitesimally short time period. The n Lyapunov exponents are defined as follows:

$$\lambda_k = \lim_{\tau \rightarrow \infty} \left(\frac{1}{\tau} \ln(r_k(\tau)) \right), \tag{10}$$

where $r_k(\tau)$ is the square root of the k th ordered eigenvalue of $M(t, t + \tau)M(t, t + \tau)^T$ (in decreasing order). Conceptually, Lyapunov exponents represent the log of the long-term average growth rates of infinitesimal perturbations in the system $F(\cdot)$. The system is considered chaotic if it is not asymptotically periodic and the leading Lyapunov exponent $\lambda_1 > 0$. Under very general conditions, Lyapunov exponents exist and describe long-term error growth independent of both norm and initial condition t (Oseledec, 1968). This makes them robust and global characterizations of the entire system. Numerical calculations of Lyapunov exponents can be complicated by the fast error growth rates of chaotic systems. We follow the procedure described in Legras and Vautard (1996) involving frequent reorthonormalizations to avoid this problem.

Lyapunov vectors L_k (of unit length) are time (and space) dependent directions associated with the Lyapunov exponents:

$$\lambda_k = \pm \lim_{|\tau| \rightarrow \infty} \left(\frac{1}{\tau} \ln \|M(t, t + \tau)L_k(t)\| \right). \tag{11}$$

Lyapunov vectors as defined above are not unique, and the term 'Lyapunov vector' has been used to refer to a number of different types of vectors. Here, we follow Legras and Vautard (1996) and define three types of Lyapunov vectors: backward, forward and characteristic. The key difference among these is the direction in time over which they are defined. Backward Lyapunov vectors at time t are the left singular vectors of the tangent linear model at time t initiated in the asymptotically distant past, $\lim_{\tau \rightarrow -\infty} M(t + \tau, t)$. These vectors follow negative Lyapunov exponent growth rates when integrating the tangent linear model backward (i.e. backward L_k grows on average with rate $-\lambda_k$ in the limit $\tau \rightarrow -\infty$ in (11)). Forward Lyapunov vectors at time t are the right singular vectors of the tangent linear model extending from t to the asymptotically distant future, $\lim_{\tau \rightarrow \infty} M(t, t + \tau)$. These vectors follow Lyapunov exponent growth rates when integrating the tangent linear model forward (i.e. forward L_k grows on average with rate λ_k in the limit $\tau \rightarrow \infty$ in (11)). Backward and forward Lyapunov vectors are calculated using the reorthonormalization method described in Legras and Vautard (1996). Note that though forward Lyapunov vectors calculations typically work backwards from $\tau \rightarrow \infty$ to t (using the adjoint or inverse tangent linear model), we assume that forward Lyapunov vectors are numbered to correspond with Lyapunov exponents $\lambda_1 > \lambda_2 > \dots > \lambda_n$ via (11) in the forward direction.

Backward Lyapunov vectors have the appeal of being adapted to the flow (i.e. positive vectors lie in the tangent space of the attractor) because they have been spun-up from the past, yet all backward Lyapunov vectors will tend toward the first vector at rate λ_1 when propagated forward in time. While the k th forward Lyapunov vectors often point away from the attractor (directions where the attractor is not continuous), they have the advantage of growing or decaying at the long-term average rate corresponding to λ_k , when propagated forward in time. It is useful to define a type of Lyapunov vector that satisfies (11) in both the forward and backward directions, and is adapted to the flow. Legras and Vautard (1996) labeled these as 'characteristic Lyapunov vectors'. By construction as asymptotic left singular vectors, the first J backward Lyapunov vectors are a basis for the subspace in which infinitesimal errors grow, in the long-term, with a rate of at least λ_J . Analogously, the J through n th forward Lyapunov vectors are a basis for the subspace in which infinitesimal errors grow, in the long-term, with a rate of no more than λ_J . Then, as suggested by Legras and Vautard (1996) and detailed by Trevisan and Pancotti (1998), characteristic Lyapunov vectors cL can be calculated by intersecting backward Lyapunov vectors bL and forward Lyapunov vectors fL

$${}^cL_k = {}^bL_{1:k} \wedge {}^fL_{k:n}. \tag{12}$$

The intersecting vector grows (shrinks) over the long-term exactly at the rate $+\lambda_k$ ($-\lambda_k$) under linearized forward (backward) flow. Wolfe and Samelson (2007) provide a more computationally efficient algorithm when fewer than n characteristic Lyapunov vectors are needed. Like backward Lyapunov vectors, characteristic Lyapunov vectors with positive Lyapunov exponents are aligned with the attractor, which corresponds to flow that has been ‘spun-up’ over a long time. Characteristic Lyapunov vectors are also invariant under linearized flow

$${}^c L_k(t + \tau) = M(t, t + \tau) {}^c L_k(t). \quad (13)$$

Characteristic Lyapunov vectors are generally not mutually orthogonal. It can be seen from (12) that backward (forward) Lyapunov vectors correspond to an orthonormalization of characteristic Lyapunov vectors, starting from the first (last) characteristic Lyapunov vector. If there are k_p number of positive Lyapunov exponents, the set of the first k_p backward Lyapunov vectors and the set of the first k_p characteristic Lyapunov vectors at time t both span the subspace that exhibits long-term linearized growth at time t . Analogously, if there are k_q number of negative Lyapunov exponents, the set of the last k_q forward Lyapunov vectors and the set of the last k_q characteristic Lyapunov vectors at time t both span the subspace that exhibits long-term linearized decay at time t . Generally, the long-term unstable and stable subspaces are not orthogonal to each other.

Although Lyapunov vectors are locally defined in that they are a function of time and location, their associated Lyapunov exponents are global descriptors of the system because they correspond to asymptotically long-term average growth rates. Local Lyapunov exponents and vectors are defined as functions of time or phase space position on the system’s attractor. Global Lyapunov exponents are the time (or phase space) average of local Lyapunov exponents and are global descriptors because they correspond to asymptotically long-term average growth rates. Local exponents vary along the orbit and can be larger or smaller than the global exponents. Over short times, certain error directions can grow even faster than the largest global Lyapunov exponent. Trevisan and Pancotti (1998) have attributed this ‘super-Lyapunov’ growth to the non-orthogonality of characteristic Lyapunov vectors in non-periodic orbits. Singular values optimized over finite times can be larger than growth rates dictated by the Lyapunov exponents and singular vectors do not generally correspond to Lyapunov vector directions. Local Lyapunov exponents are used to describe short-term growth of errors along Lyapunov vectors (e.g. Trevisan and Pancotti, 1998). They are defined by the right-hand side of (11), except with τ equal to one time step. The asymptotic long-term average of the local Lyapunov exponents is equal to the global Lyapunov exponents. In the remainder of this paper, the term ‘Lyapunov exponent’ refers to global Lyapunov exponents, following (11). Local Lyapunov exponents will be explicitly named as such when used.

It should be noted that like singular vectors and singular values, Lyapunov vector and exponent theory is defined using

linearized dynamics and infinitesimal perturbations. Thus, even though Lyapunov vectors and exponents describe global and long-term attributes of a chaotic system, unchecked finite perturbations will not maintain these properties over long integration times. Furthermore, Toth and Kalnay (1997) argue that even controlled finite perturbations in the atmosphere may not behave according to Lyapunov properties. Certain error growth dynamics, such as convective instabilities, may be associated with very small perturbations that quickly saturate at magnitudes smaller than those of practical concern with finite perturbations.

To address these problems, Toth and Kalnay (1993, 1997) introduced bred vectors as a finite-perturbation representation of Lyapunov vectors. Bred vectors are derived by propagating and resizing finite perturbations through the full non-linear model over a sufficiently long ‘breeding’ period. Although some breeding procedures involving orthonormalizations have been proposed (Annan, 2004), bred vectors mostly collapse to the single leading backward Lyapunov vector. Because the ensemble in the EnKF is a multidimensional representation of the state space, it is important to be able to resolve the full subspace of instabilities in the system in our work. Thus, the original Lyapunov vector types are more informative for our analysis of EnKF performance than bred vectors and will be used in this work.

4. Dynamical analysis and data assimilation

Although most investigations on filter divergence focus on the role of the update rather than on model dynamics, dynamic properties have been widely considered in derivations of reduced rank filters. Due to large model dimensions, the computational costs for forecasts and update calculations can be prohibitive for geophysical models. This has led to various reduced rank filters that rely on rank-deficient representations of estimation error. In practice, the EnKF is a reduced rank formulation because the number of ensemble replicates is typically less than the state dimension. In this case, the nullspace of the rank-deficient error covariance is affected by random sampling error. Some reduced rank filters decrease dimensionality by reducing the resolution of the model (Fukumori and Malanotte-Rizzoli, 1995). Others construct their error covariance matrices such that their column space coincides with a particular subspace of the model state space. Specifically, many reduced rank filters aim to stabilize errors by focusing on an ‘unstable subspace’ of the state space. There are many ways to define such a subspace for chaotic systems. This flexibility has led to many different types of reduced rank filters. Most of the alternatives share a common dependence on linearized analyses that strictly apply only when errors are small and deviations from the true trajectory can be adequately described by a tangent linear model. Experience with these reduced rank filter implementations provides helpful insight into the system features that affect EnKF performance.

In an early dynamics-based study of reduced rank filtering, Cohn and Todling (1996) tested different frameworks for the

extended Kalman filter, including one using leading singular vectors of the linearized model and one using leading eigenmodes of the error forecast matrix. The singular vector Kalman filter performed more poorly because information gained during previous updates is removed if that information is orthogonal to leading singular vectors of the tangent linear model at later times. Bowler (2006) and Pazó et al. (2010) also found singular vectors to perform poorly in initializing ensemble forecasts because they typically point off the attractor. Cohn and Todling (1996) showed that both filters tested were considerably enhanced if the nullspace of the reduced forecast error matrix is filled using adaptive tuning. These results suggest that short-term growth directions may not be adequate for capturing the continuous dynamics important for Kalman filtering.

To facilitate computations, some reduced rank filters approximate a relevant ‘unstable subspace’ with procedures based on autonomous (time-invariant) systems. For their filter, Farrell and Ioannou (2001) reduced the dimension of an atmospheric model with a balanced truncation of the time-averaged linearized operator. Pham et al. (1998) presented the reduced rank SEEK filter using a derivation that assumes an autonomous, linearized model. SEEK is an extended Kalman filter in which the error covariance is constrained to project on the leading eigenmodes of the tangent linear model. When tested with a time-varying quasi-geostrophic model, SEEK required a ‘forgetting factor’ to emphasize data at the current update time in order to generate good results. The forgetting factor plays a role similar to variance inflation, and its necessity in SEEK indicates that leading model eigenmodes alone may not be a sufficient representation of dynamics to ensure good filter performance.

Carne et al. (2001) modified SEEK for systems with rapidly changing tangent linear models. They developed an algorithm for calculating the time evolution of ‘unstable modes’ using the time derivative of the tangent linear model. Two versions were formulated based on different definitions of the ‘unstable modes’ of the tangent linear model over a forecast period: the leading right singular vectors and the leading eigenmodes. The study is inconclusive as to which version is better, yet the forgetting factor of the original SEEK filter is needed for both. Using the principal components of the error covariance matrix, Lermusiaux and Robinson (1999) also developed a filter that explicitly follows the evolving error subspace.

While many reduced rank filters build upon the linearized extended Kalman filter, some have been used within an EnKF formulation. Verlaan and Heemink (1995) developed a version of the EnKF that drew replicates in directions of leading eigenvectors of the covariance matrix, which are influenced over time by model dynamics. This reduced the randomness found in the classical EnKF. Heemink et al. (2001) further adapted that reduced rank EnKF by adding additional random replicates projected on the nullspace of the reduced rank covariance. The inflation-like fix of the extra replicates was crucial for robust performance in the reduced rank EnKF.

In a recent paper, Trevisan et al. (2010) presented a 4D-Var-AUS (assimilation in the unstable subspace) algorithm that confines the estimate increment to the ‘unstable subspace’, and they showed with a Lorenz (1996) model that their method generated optimal RMSE results. They defined the unstable subspace as the space spanned by backward Lyapunov vectors corresponding to positive and 0-value Lyapunov exponents. Along similar lines, Pazó et al. (2009) compared various ensemble forecasting strategies and proposed characteristic Lyapunov vectors as optimal. Also, Pham et al. (1998) theoretically shows the stability of the SEEK filter for the autonomous case. In this case the growing tangent linear model eigenmodes (on which SEEK is based) span the same subspace as the growing backward and characteristic Lyapunov vectors. Pham’s analysis, together with the results of Trevisan et al. (2010), suggests that a reduced rank Kalman filter based on growing backward or characteristic Lyapunov vectors may be very effective for robust filtering in time-varying, chaotic systems. The work of Trevisan et al. (2010) stands out in that unlike many reduced rank filter formulations, their method did not require additional ad hoc fixes. It was brought to our attention during the review process of this manuscript that Trevisan and Palatella (2011) also recently demonstrated that for small errors, the extended Kalman filter converges to a form constrained to the unstable subspace (EKF-AUS), again without additional fixes.

Understanding the role of dynamic properties in filtering performance is critical for robust filter work. Results obtained with reduced rank filters suggest that it is critical to account for long-term dynamics and model time-dependence when characterizing the subspace of growing errors. However, it is striking that many reduced rank filter approaches require additional fixes such as adaptive tuning, forgetting factors, and random nullspace filling to achieve good performance, regardless of the approach used to define the growing error subspace. This suggests that certain critical properties are missing from those types of approaches.

Toth and Kalnay (1997) argued that analysis errors line up with ‘growing directions’ of the underlying model dynamics in filtering because projections on the stable subspace naturally decay away. If a filter’s error covariance does reduce to some identifiable unstable subspace, the opportunity indeed exists for robust and efficient filter versions that track only a relevant subset of the state space, as sought in reduced rank filter studies. For the 4D-Var method with infinite past observations, Pires et al. (1996) in fact showed theoretically and experimentally (for the perfect model case) that the column space of the error covariance automatically aligns with the subspace spanned by unstable backward Lyapunov vectors. This provided motivation for Trevisan et al.’s (2010) 4D-Var-AUS filter and Trevisan and Palatella’s (2011) EKF-AUS filter. To our knowledge, however, the correspondence of errors with a particular dynamically based subspace has never actually been demonstrated for other filter types, including the EnKF. The conclusions of Pires et al. (1996) and Trevisan and Palatella (2011) cannot directly apply to the

EnKF because of their reliance on the tangent linear model instead of the full non-linear model. For a different ensemble filter (the maximum likelihood ensemble filter), Carrassi et al. (2009) noted an apparent relationship between the necessary ensemble size and the non-linear model's Kaplan-Yorke dimension, and they inferred from that observation that the filter's performance depends on the chaotic dynamics of the system. In this paper, we seek to demonstrate more directly how the performance of the EnKF depends on the dynamic features of the chaotic system and the structure of the ensemble. We note that dynamical behavior of the full closed-loop system (including assimilation of observations) may also be relevant for understanding filter performance. We choose here to focus on the role of open loop model dynamics, building upon previous literature that has demonstrated its utility in ensemble forecasting and assimilation.

5. Experiment set-up

Our investigation of EnKF performance for chaotic systems relies primarily on the strongly non-linear model presented in Lorenz and Emanuel (1998). While much simpler than a full geophysical model, it can be parametrized to exhibit chaotic behaviour similar to that found in atmospheric models. The Lorenz-98 model state $x_j(t)$, $j = 1, \dots, n$, follows the equation:

$$\frac{dx_j}{dt} = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F, \quad (14)$$

with cyclic boundary conditions, and $j = 1, \dots, n$, representing locations along a latitudinal band around the globe, with model dimension n . F is a forcing term and can be tuned to produce chaotic behaviour for a given model dimension. This model is sometimes referred to as the Lorenz (1996) model. In this study, we use $F = 8$. Numerical integrations of (14) are carried out with a fourth order Runge–Kutta scheme with time step $dt = 0.01$, equivalent to 1.2 h in the atmosphere.

A simple Lorenz-98 model is chosen for this study because some tests would be difficult if not impossible with a higher-dimensional and therefore more realistic chaotic model. We are interested in assessing EnKF performance without the interference of ad hoc fixes such as localization and inflation, which potentially alter the properties of the ensemble and mask its correspondence with dynamic features. As a result, we require a model that can be run with a sufficiently large ensemble so that localization and inflation are not needed. Computational costs of calculating Lyapunov and singular vectors could also be challenging for larger models. Tests are carried out for 40- and 6-dimensional versions of the Lorenz-98. Both produce chaotic behaviour with $F = 8$. The 40-D model corresponds to the original configuration of Lorenz and Emanuel (1998), and it has been used in numerous EnKF studies (e.g. Anderson, 2001; Whitaker and Hamill, 2002; Lawson and Hansen, 2004; Ott et al., 2004; and Sakov and Oke, 2008). The 6-D model is a small dimensional

model that still exhibits a diverse range of dynamical properties. This coarse resolution is employed to simplify the multidimensional assessment of the ensemble during filtering. Implications of the Lorenz-98 results for larger and more realistic models will be discussed.

Although model error is often a critical component in estimation problems, we choose to adopt a 'perfect model' assumption to help isolate the effects of dynamics on filter performance. When included, model error can often help inflate underestimated spread in the ensemble during the forecast step. Consequently, divergence is particularly problematic in perfect model formulations without this built-in variance inflation mechanism. How much model error counters underestimated spread depends on various factors such as the magnitude of the model error, and how it is incorporated. Furthermore, imposing model errors can to a certain degree mask the dynamic behaviour of the underlying model. Thus, examining EnKF performance in the most stringent perfect model case is important for establishing a baseline for understanding how model dynamics affect estimates. In the perfect model case, the general forecast step simplifies to

$$x^{f,i}(t) = F(x^{a,i}(t-1), t), \quad (15)$$

where $F(\cdot)$ is the numerical integration of the Lorenz-98 model. Without model error, state errors arise solely from initial condition errors.

The particular experimental configuration used in an EnKF test affects how various dynamic properties impact filter performance. Such dependency was noted by Hansen and Smith (2000) when investigating the performance of different adaptive observation strategies. To detect these differences, we vary ensemble size, number of observations, and observation error magnitude. The observation strategy is particularly important because estimation errors increase and linearization approximations become less accurate when measurement errors increase and/or measurement coverage is reduced.

To meet our objective of better understanding the role of chaos dynamics on EnKF performance, we need quantitative criteria for assessing EnKF runs. A model run is considered to fail if the filter diverges (as defined earlier). An EnKF divergence test is carried out by first generating a truth run with the model. The initial true state is produced by spinning up a randomly selected state for 300 time steps. Observations for the EnKF test are then generated every 10 time steps (equivalent to 12 h in the atmosphere) by adding to the true states random observation noise drawn from $\mathfrak{N}(0, R)$. The ensemble replicate i is initialized by $x_0^i = x_0^{\text{true}} + \eta + \varepsilon^i$, where $x_0^{\text{true}} + \eta$ represents an observation at t_0 , and $\eta, \varepsilon^i \sim \mathfrak{N}(0, R)$. In results not shown here, we found the filter extremely prone to divergence during early assimilation times if the initial ensemble has significant spread, such as that corresponding to climatology. Our initialization procedure is reasonable because in operational contexts, past observations and analyses are available to provide reasonably good initial estimates for ensemble forecasts (compared to climatology).

In order to detect rare divergence events in very long filtering runs, the EnKF is run over 5×10^4 assimilation cycles, using five independent truth runs. A filter run is considered divergent if the error of the ensemble mean becomes high while the trace of the ensemble error covariance matrix remains low (Sacher and Bartello, 2008). The error of the ensemble mean at time t can be quantified by its norm

$$NormErr(t) = \left(\sum_{j=1}^n (\bar{x}_j(t) - x_j^{true}(t))^2 \right)^{1/2} \quad (16)$$

A certain experiment configuration (specific ensemble size and observation configuration) is considered divergent if at least one of the long multiple truth run diverges.

Once a particular EnKF experiment design is found divergent or non-divergent, we compare the ensemble with the spaces spanned by various subsets of the backward, forward, and characteristic Lyapunov vectors and singular vectors of the tangent linear model. The Lyapunov and singular vectors are calculated for Lorenz-98 as described in Section 3, with all perturbations

defined around the true trajectory. The tangent linear model is derived analytically from the numerical discretization equations for solving (14). An analysis of the dynamic features of non-divergent ensembles can reveal whether filtering does in fact cause errors to line up with an ‘unstable subspace’, and it can identify what exactly that ‘unstable subspace’ is. Here we analyze ensemble dynamics by examining projections of error replicates, the error covariance, and the null space of the error covariance on the Lyapunov and singular vectors and by considering the temporal evolution and growth of individual error replicates. Detailed discussions of these tests are provided in the next section.

6. Results and discussion

6.1. Results for 40-D Lorenz-98 model

Figure 1 introduces our discussion of EnKF performance by comparing divergent and non-divergent time series derived from the 40-D Lorenz-98 model. The high open loop RMSE shown in this figure is for a randomly initiated trajectory on the attractor

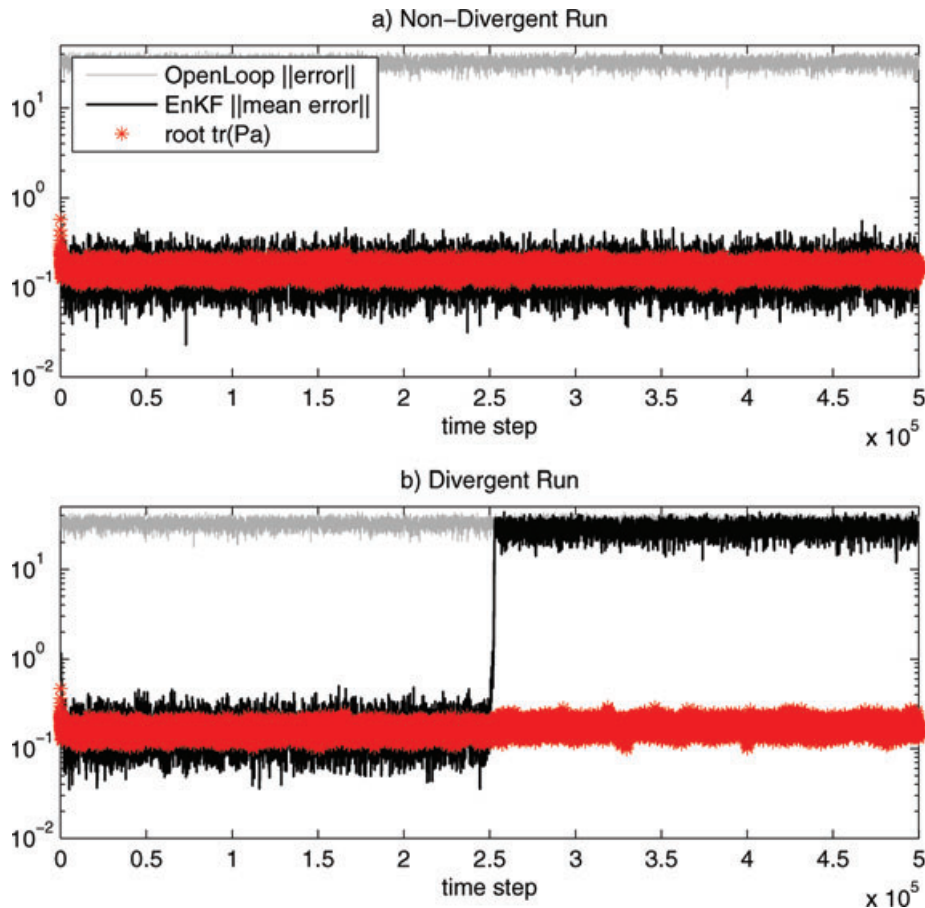


Fig. 1. The norm error of the ensemble mean and the square root analysis covariance trace results for the 40-D Lorenz-98 model, $\sqrt{R} = 0.1$, for (a) a sample non-divergent run ($N = 31$) and (b) a sample divergent run ($N = 21$). The norm error of the ensemble mean and square root of the analysis covariance trace are on average equal for the non-divergent case. In the divergent case, the ensemble spread remains low, as indicated by the trace of the covariance. However, the norm error of the ensemble mean rises to the open loop error level.

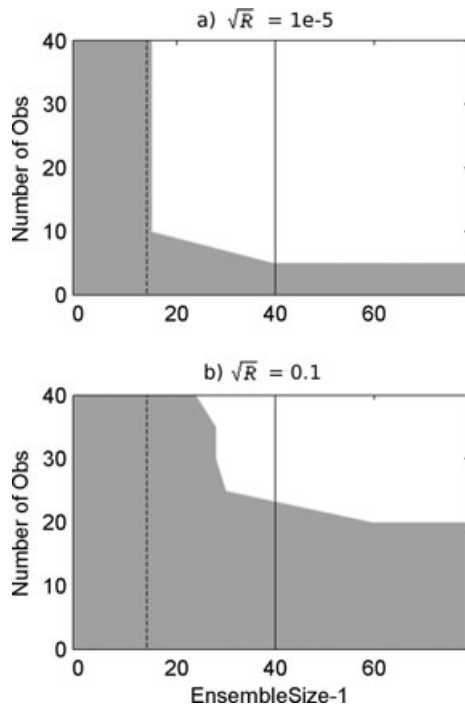


Fig. 2. Divergence test results for 40-D Lorenz-98 model for (a) $\sqrt{R} = 1 \times 10^{-5}$ and (b) $\sqrt{R} = 0.1$. Grey shows experiment configurations (specific ensemble size, number of observations, and observation error magnitude) that diverged; white shows configurations that did not diverge. Experiment configurations were considered divergent if the filter diverged within 5×10^4 assimilation cycles for at least one of five multiple truth tests. The solid vertical reference line corresponds to the model dimension. The dotted vertical reference line corresponds to the dimension of the linearized long-term unstable subspace (i.e. the number of non-negative Lyapunov exponents).

with no data assimilated. The magnitude of the open loop error roughly represents the size of the attractor. Recall that if the filter is working properly, the norm error of the ensemble mean (in black) and the trace of the ensemble covariance (grey asterisks) should on average be equal. The non-divergent example in Fig. 1a aptly shows good correspondence of these two values. In the divergent example of Fig. 1b, the spread of the ensemble remains small, yet the filter error begins to increase. Because the filter is overconfident in the estimate, it cannot be corrected with observations, and the filter diverges so that its estimate is no better than the result obtained from an open loop run with no measurements.

Divergence results for two different observation error variances ($\sqrt{R} = 1e-5$ and 0.1) are summarized in Fig. 2. The filter was also tested using $\sqrt{R} = 1$, but divergence occurred within 5×10^4 assimilation cycles in all tests, up to the largest tested ensemble size of 4001. In Fig. 2, $N - 1$ is plotted on the x -axis. This is the rank of the ensemble covariance (and thus the dimension of the filter update subspace). In order to investigate the effects of sampling error on filter performance, tests

ranged from large ensemble sizes (non-reduced rank) to small ensemble sizes less than the state dimension. The dependence of EnKF performance on observation error size, ensemble size and number of observations is apparent from the divergence results. We find that for very small observation errors, an ensemble with $N - 1 = 15$ is sufficient to prevent filter divergence. However, $N - 1$ must increase to at least 26 if \sqrt{R} is increased to 0.1. For both levels of observation error, the filter will eventually diverge if insufficient states are observed.

A noteworthy result from Fig. 2 is that the filter is capable of producing non-divergent results with $N - 1$ less than the state dimension n (the vertical reference line shows $n = 40$ for comparison). This means that a reduced rank ensemble filter can in fact have robust performance without variance inflation or localization, at least if there are enough high quality measurements available. These results cannot be explained by divergence studies that focus only on the update step (van Leeuwen, 1999; Furrer and Bengtsson, 2007; Sacher and Bartello, 2008). Such studies typically indicate that any sampling error causes underdispersion during the update. In our test, sampling error can be expected to be significant since $N - 1 < n$, and yet the filter still converges. Our results also differ from much of the previous reduced rank filter work (described in Section 4b), which required methods such as adaptive tuning or forgetting factors to avoid divergence.

In order to obtain further insight about the effect of dynamics on filter divergence, it is instructive to determine whether the filter errors automatically align with a particular subspace. To do this we project individual replicates from non-divergent runs onto the sets of Lyapunov and singular vectors described earlier. Lyapunov exponents and singular values of the tangent linear model are used to determine which dynamic vectors within each set are unstable. To facilitate comparisons with Lyapunov exponents, we define 'singular value exponents' as the natural log of the singular values divided by the number of model time units over a forecast period. The spectra of these the Lyapunov and singular value exponents are plotted in Fig. 3. Lyapunov and singular value exponent values greater (less) than 0 represent growing (decaying) modes, but over different time scales, as described in Section 3. The state-dependent singular values are depicted for 100 different attractor points. There are 13 positive Lyapunov exponents and the 14th Lyapunov exponent is calculated to be about 0. Although the extreme ends of the spectrum are of much higher magnitude for singular values, the number of unstable modes corresponding to Lyapunov exponents and singular values are comparable (see the dashed horizontal reference line, which intersects both spectra near $i = 14$). Notably, this common number of unstable modes serves as a minimum threshold for the $N - 1$ value needed to avoid filter divergence in Fig. 2. This alone suggests a strong correspondence between model dynamics and EnKF performance.

Figure 4 shows how the individual error replicates from a non-divergent ensemble with $\sqrt{R} = 0.1$ and $N = 41$ project on

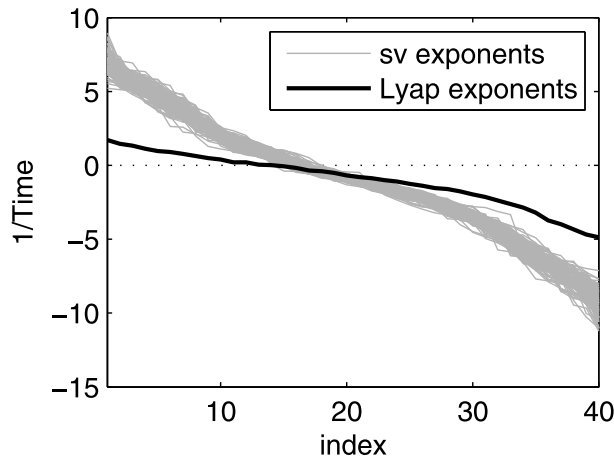


Fig. 3. The black line shows the Lyapunov exponents of the 40-D Lorenz-98 model. There are 13 positive Lyapunov exponents, and the 14th Lyapunov exponent equals 0. The grey lines show singular values of the tangent linear model of the 40-D Lorenz-98 model over forecast periods. To facilitate comparison with Lyapunov exponents, the singular values are expressed as exponents, or the natural log of the singular values, divided by the units of time. Different sets of singular value exponents are shown for different points on the model attractor. Generally, there are 14–16 positive singular value exponents. Positive (negative) exponents signify error growth (decay).

each type of dynamic vector (singular or Lyapunov). Note that in Fig. 4e, non-orthogonal (or oblique) projections on characteristic vectors are shown, because characteristic Lyapunov vectors make up a non-orthogonal basis. Indices above the black dotted lines correspond to unstable modes. The $\sqrt{R} = 0.1$ case is chosen for examination because the ensemble size divergence threshold is less obvious (somewhere between the growing subspace dimension and full rank), so comparisons provide greater insight into the necessary ensemble features for non-divergence. The ensemble size $N = 41$ allows for full-rank ensemble covariance matrices. The analysis ensemble was produced after 600 assimilations with an exhaustive observation network (i.e. all states measured with observation noise) to avoid filter transients.

It is readily apparent from Fig. 4 that short-term growth directions (computed over one forecast period) do not play the dominant role in the ensemble, compared to long-term growth features. There is only slight alignment of the ensemble with growing left singular vectors of $M_{t-dt,t}$ and no obvious match with right singular vectors $M_{t,t+dt}$. As also highlighted by others (e.g. Bowler, 2006; Pazó et al., 2010), this poor correspondence is due to the overlocalization of singular vectors, which cause them to point in directions irrelevant to the flow direction. Similarly, forward Lyapunov vectors do not appear useful for characterizing the ensemble. As discussed in Section 3, forward Lyapunov vectors provide an orthonormalization of characteristic Lyapunov vectors starting from the most decaying one. The lack of any specific pattern in Fig. 4c shows that, unsurprisingly, errors do not preferentially fill the stable subspace.

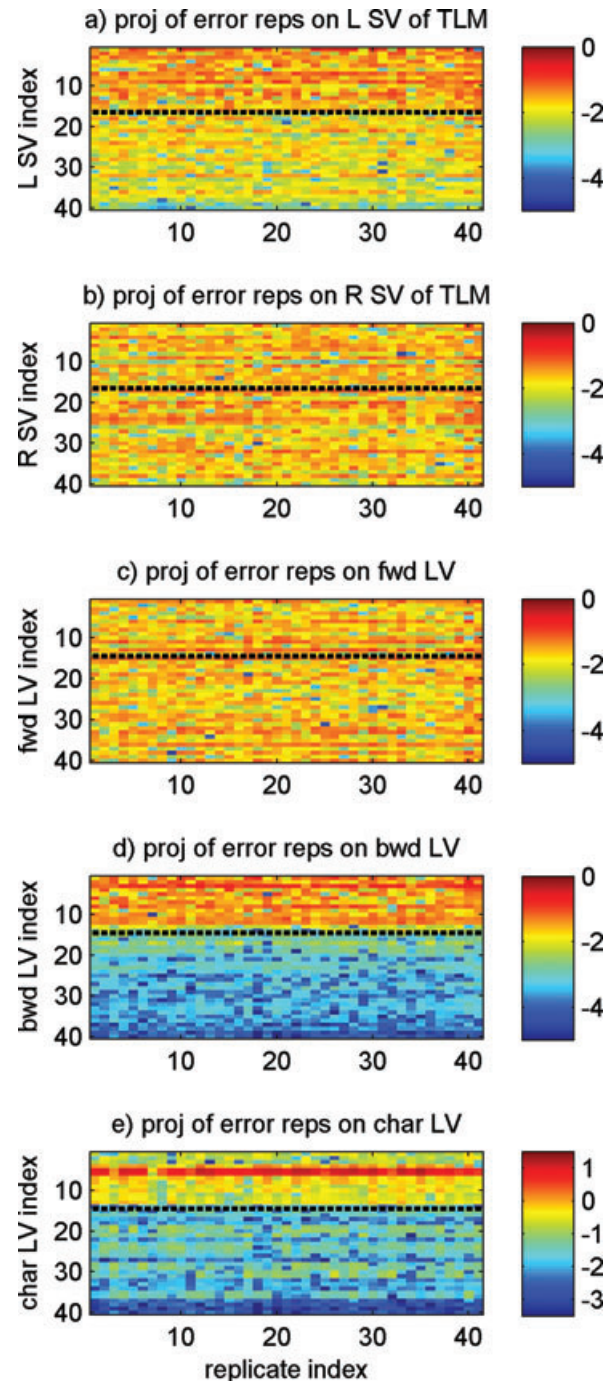


Fig. 4. Error replicates of a non-divergent ensemble are projected on (a) left singular vectors of the tangent linear model (over the subsequent forecast period), (b) right singular vectors of the tangent linear model (over the preceding forecast period), (c) forward Lyapunov vectors, (d) backward Lyapunov vectors and (e) characteristic Lyapunov vectors. Absolute value projections are shown in log-scale. The ensemble is produced with the 40-D Lorenz-98 mode, $N - 1 = 40$, $\sqrt{R} = 0.1$, exhaustive observations, and after 600 assimilation cycles. Above the black dotted line indicates growing modes. The strongest alignment is with growing backward and characteristic Lyapunov vectors.

In contrast, individual error replicates align very strongly with the leading 14 (unstable and neutral) characteristic and backward Lyapunov vectors, with negligible projection on the remaining (stable) vectors. Recall that the first p characteristic and backward Lyapunov vectors span the subspace in which errors exhibit a long-term linearized growth rate of at least λ_p . Like forward Lyapunov vectors, backward vectors are also an orthonormalization of characteristic vectors, but starting with the most unstable mode. Thus, our results show that after a number of assimilation cycles, a well-performing filter ensemble will tend toward the linearized long-term unstable subspace, which can be defined by the space spanned by characteristic or backward Lyapunov vectors corresponding to non-negative Lyapunov exponents. These results agree with findings for 4D-Var by Pires et al. (1996) and Trevisan et al. (2010) and the EKF results of Trevisan and Palatella (2011).

Pazó et al. (2010) proposed that characteristic Lyapunov vectors could be used as optimal directions for ensemble forecasting, because they are adapted to the flow and maintain individual Lyapunov growth rates, rather than collapse toward the single leading Lyapunov vector. In the EnKF, error replicates likely avoid collapse to the single leading Lyapunov vector due to the Kalman update and the random orthonormal matrix rotation, which alter individual replicates within the subspace spanned by the ensemble. For the update step of the EnKF, the ensemble serves to represent the subspace in which the update is performed, and as a result, assessing the update ensemble according to characteristic modes is no more informative than using orthogonalized backward vectors. However, resolving the ensemble according to individual characteristic modes is useful when examining how it evolves during the forecast step. Due to this advantage, we employ characteristic Lyapunov vectors for our analysis when assessing ensemble results over time. The benefit of using characteristic Lyapunov vectors for our analysis will become clearer later in our discussion.

Alignment of the EnKF ensemble with the linearized long-term unstable subspace occurs across a range of \sqrt{R} values and ensemble sizes (not shown here). However, it does not necessarily follow that an ensemble must span this unstable subspace in order to give non-divergent results. To probe this further, we compare ‘low rank’ ensembles (with $N - 1 \leq n$) for non-divergent and divergent cases. We take the filter ensemble after 200 assimilation cycles, with an exhaustive observation network and $\sqrt{R} = 0.1$. An ensemble with $N = 32$ is used to represent the non-divergent case, and an ensemble with $N = 16$ is used to represent the divergent case (see Fig. 2). Note that the smaller ensemble eventually diverges, later than the 200 assimilation cycles. If there is projection of the ensemble nullspace on any of the first 14 backward (unstable) Lyapunov vectors, then the ensemble contains no information about error growth along these vectors. In other words, all replicates are orthogonal to one or more of the system’s long-term unstable directions. Figure 5 shows that in the non-divergent case, there are no un-

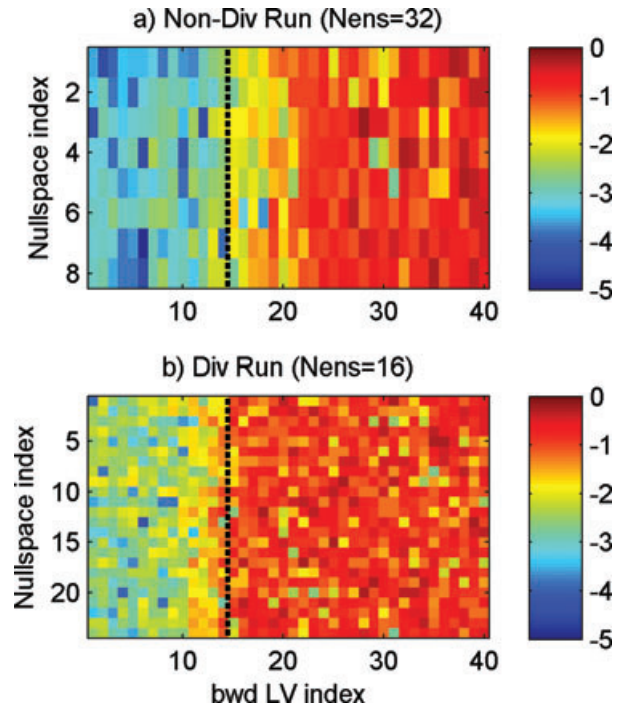


Fig. 5. A basis of the nullspace of the error replicates of a (a) 32-member non-divergent and (b) 16-member divergent ensemble is projected on backward Lyapunov vectors. Absolute values of projections are shown in log-scale. The ensembles are produced with the 40-D Lorenz-98 model, $\sqrt{R} = 0.1$, exhaustive observations, and after 200 assimilation cycles. Left of the dotted black line indicates linearized long-term growing modes. The 32-member ensemble has negligible nullspace projection on growing backward Lyapunov modes and thus spans the subspace spanned by those vectors. In contrast, notable nullspace projection on the 10–15th backward Lyapunov modes in (b) signifies that the 16-member ensemble does not fully span the growing directions.

stable modes outside the ensemble space. In contrast, in the case that eventually diverges, the ensemble does not project on some portions of the unstable subspace (i.e. those spanned by the 10–15th backward Lyapunov vectors). This occurs even though the 16-member ensemble has sufficient degrees of freedom to be able to span the entire unstable subspace.

The results from Fig. 5 strongly suggest that the ensemble must project on all the long-term unstable modes for the filter to avoid divergence. Furthermore, Fig. 4 shows that the replicates of a non-divergent filter with an adequate ensemble size will automatically move to span the long-term unstable subspace. This result can be understood using arguments provided in the literature and summarized in Section 3. For robust estimation, errors that grow must be stabilized during the update step. This happens in the EnKF only if growth directions are represented within the ensemble. It is logical that ‘errors that grow’ must at least include long-term growth directions because the EnKF is a sequential estimation method applied over time. Because

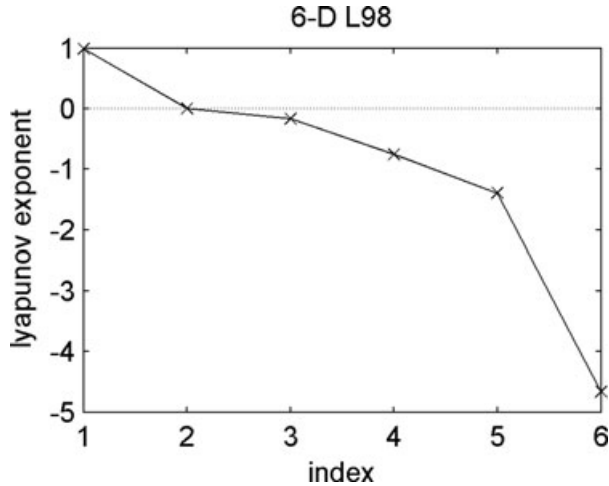


Fig. 6. Lyapunov exponents of the 6-D Lorenz-98 model. There is one positive Lyapunov exponent, and the 2nd Lyapunov exponent equals 0.

these long-term growth directions are ones that will eventually grow most, an ensemble will align with the long-term unstable subspace after some time, regardless of its initial orientation.

Although the importance of the linearized long-term unstable subspace is apparent, the need for ensemble sizes greater than that subspace dimension (unless \sqrt{R} is very small) suggests that other factors are relevant. Why do we need, in the example considered in Fig. 5, significantly more replicates than the number of unstable Lyapunov modes? We will address this question by examining temporal changes in the ensemble with the smaller and more tractable 6-D Lorenz-98 model.

6.2. Results for 6-D Lorenz-98 model

The dynamic behaviour of the 6-D Lorenz-98 system can be anticipated from the Lyapunov exponent spectrum shown in Fig. 6. This simple system has one positive Lyapunov exponent that is about equal to 1, and the second exponent is equal to zero, indicating that the dimension of the linearized long-term non-decaying subspace is 2. We consider the behaviour of this 6-D system for $\sqrt{R} = 1 \times 10^{-5}$, 0.05, and 1. A lower mid-level \sqrt{R} is used compared to the 40-D tests because of the stronger chaotic nature of the model (the sum of Lyapunov exponents is much larger). Our tests show that, as in the 40-D case considered in Fig. 1, an ensemble with $N - 1$ equal to the dimension of the unstable subspace gives non-divergent results when observation errors are very small (1×10^{-5}). For the mid-level \sqrt{R} case, a larger ensemble size is required to avoid divergence, but a reduced rank filter with $N - 1 < n$ still may be adequate if the observation network is dense. However, divergence occurred in all of our tests for $\sqrt{R} = 1$, up to the largest tested ensemble size of 601, which is well beyond full rank. As with the 40-D case, the non-divergent ensemble aligns mostly with linearized long-term unstable subspace.

The need in certain tests for replicates that span portions of the stable subspace as well as the unstable subspace suggest that intermittent errors that grow over short time spans may contribute to long-term divergence. To investigate this hypothesis further, we track the ensemble errors over time. In order to exploit the advantages of the characteristic Lyapunov vectors for both update and forecast steps of the EnKF, we transform the i th error replicate ($\varepsilon^i = x^i - x^{\text{truth}}$) to the Lyapunov coordinate system using non-orthogonal projections as follows:

$$\varepsilon^i = [{}^cL_1 | {}^cL_2 \dots | {}^cL_n] \tilde{\varepsilon}^i \quad (17)$$

$$\tilde{\varepsilon}_j^i = (A^T {}^cL_j)^{-1} A^T \varepsilon^i,$$

$$A = [{}^cL_1 \dots | {}^cL_{j-1} | {}^cL_{j+1} | \dots | {}^cL_n]^{\perp}, \|A\| = 1, \quad (18)$$

where $\tilde{\varepsilon}_j^i$ is the j th element of the i th error replicate in the transformed space ($\tilde{\varepsilon}^i$), cL_j is the j th characteristic Lyapunov vector, and all characteristic vectors are normalized to magnitude 1. Recall that the characteristic Lyapunov vectors are ordered, with index 1 assigned to the vector with the most positive Lyapunov exponent and larger indices assigned to vectors with progressively smaller Lyapunov exponents. Note that in the following discussion, characteristic Lyapunov vectors are referred to as growing/unstable or decaying/stable based on their corresponding Lyapunov exponent value, regardless of actual filter dynamics along those directions. Also, for the remainder of this paper, ‘Lyapunov vectors’ or ‘Lyapunov modes’ will refer to characteristic vectors, and not backward or forward Lyapunov vectors.

Assessing errors in the projected space is useful because error along a particular (non-orthogonal) characteristic Lyapunov mode evolves according to the corresponding local Lyapunov exponent from time t to $t + \tau$ (λ^{local})

$$\tilde{\varepsilon}_j(t + \tau) = \tilde{\varepsilon}_j(t) * \exp(\lambda_j^{\text{local}} \tau) + h.o.t.(\tilde{\varepsilon}_1(t), \tilde{\varepsilon}_2(t), \dots, \tilde{\varepsilon}_n(t)), \quad (19)$$

where *h.o.t.* refers to contributions from higher order terms neglected in the linearized Lyapunov analysis. The first (linearized) part of eq. (19) follows from the fact that errors along characteristic Lyapunov modes are intransient under linearized flow (eq. 13), i.e. the tangent linear model propagates the i th characteristic mode at time t to the i th characteristic mode at time $t + 1$. Importantly, (19) indicates that error along a particular mode can be affected by other modes only through non-linear contributions. Furthermore, it underscores the role of the local Lyapunov exponent for actual linear growth over a particular time period, which can differ from the long time average global value. The combined effect of non-linear movement and intermittent linear growth may make error along stable modes important. This may become problematic for a filter if the ensemble has insufficient degrees of freedom to represent those stable modes. Note that the separation of individual modes under linearized dynamics in equation (19) occurs only with characteristic Lyapunov vectors.

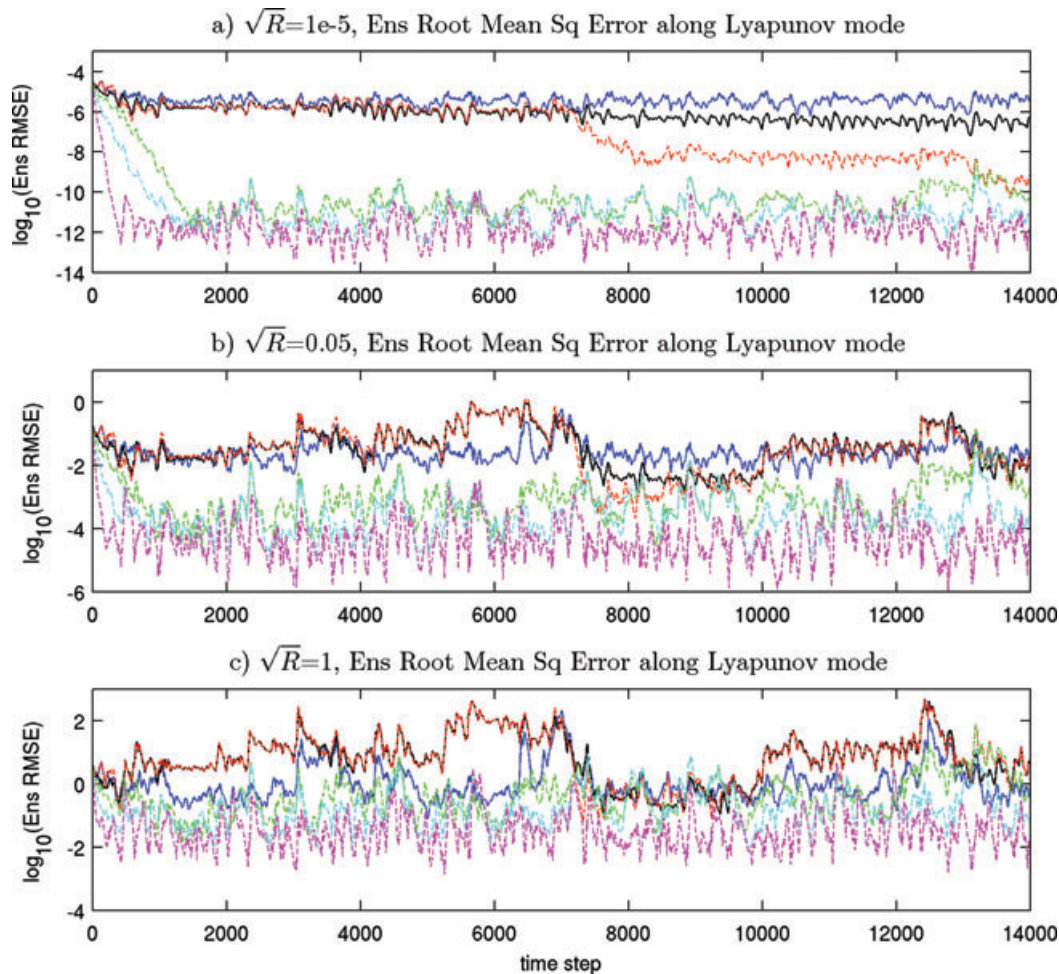


Fig. 7. Time series of ensemble root mean squared error along characteristic Lyapunov vectors for 61-member runs of the 6-D Lorenz-98 model, with different \sqrt{R} values. Blue shows ensemble root mean squared error along the 1st Lyapunov vector, black along the 2nd, red along the 3rd, green along the 4th, cyan along the 5th and magenta along the 6th. Unstable directions (those corresponding to non-negative Lyapunov exponents) are shown with solid lines (blue and black), and decaying directions are shown with dashed lines (red, green, cyan and magenta). The greatest separation between decaying and growing modes occurs for very small \sqrt{R} .

It does not occur with backward Lyapunov vectors, which generally undergo an upward propagation of energy from more stable to more unstable modes during linearized flow. Thus, although backward Lyapunov vectors are easier to compute and are adequate for identifying the unstable subspace important for EnKF updates, an analysis based on characteristic Lyapunov vectors has the crucial advantage of making it possible to assess how linear versus nonlinear dynamics affect various modes of the system and ultimately impact EnKF performance.

To test the role of different growth modes, we show the time series of errors in the projected space for three runs with different \sqrt{R} values in Fig. 7. To simplify the visualization, we examine the ensemble root mean square of the transformed error replicates rather than individual replicates. For each run, an ensemble size of 61 and a full observation network are used to insure that the filter is full rank and does not diverge over

the plotted period of 14 000 time steps. Although the filter runs shown in Figs. 7a and b do not diverge over the total period investigated (5×10^5 time steps), the run shown in Fig. 7c, where $\sqrt{R} = 1$, does eventually diverge (after the time period shown in Fig. 7). As suggested earlier, \sqrt{R} is a surrogate measure of non-linearity because larger \sqrt{R} values are associated with larger estimation errors and greater departures from linearity assumptions.

It can be seen that for all \sqrt{R} values considered, errors project more on linearized growth modes with lower Lyapunov indices (i.e. growth modes with larger Lyapunov exponents). We note that if an ensemble does not mostly follow linear characteristics, the *h.o.t.* is the most significant component in (19), and evaluating the ensemble relative to Lyapunov modes is not very informative. Figure 7 shows that for our examples, the linearized component of error in (19) generally plays the dominant role,

and thus assessing the errors relative to Lyapunov modes can be useful. However, we are also interested in analyzing how the ensemble changes as the magnitude of the smaller higher order component of error increases. In particular, we want to determine how this secondary component may contribute to filter divergence.

Noting the log-scale of the y-axis in Fig. 7, it can readily be seen that for $\sqrt{R} = 1 \times 10^{-5}$, the magnitude of errors along the decaying Lyapunov vectors (the four dashed lines) becomes negligible. This is to be expected, because for such small \sqrt{R} , the errors are kept very small by frequent updates and the system behaviour is adequately captured by the linearized Lyapunov analysis. In this case, the higher-order terms in (19) are small and the Lyapunov exponents provide a very accurate representation of actual non-linear dynamics. After some time, state errors in the long-term stable directions reflect their negative global Lyapunov exponents and eventually decay, leaving non-negligible errors only in the linearized long-term unstable subspace. In this case, an ensemble size greater than or equal to the unstable subspace dimension (corresponding to non-negative Lyapunov exponents) is sufficient for the filter to avoid divergence. This result is consistent with the results summarized in Figs. 1–6 as well as the arguments provided by investigators such as Trevisan et al. (2010) and Toth and Kalnay (1997). It is also in line with the motivation for many reduced rank filter formulations.

It should be noted, however, that the magnitudes of errors along stable characteristic Lyapunov vectors in Fig. 7a do not decay asymptotically to zero. During early times (<2000 time steps), root mean squared error in the three most decaying modes shows an overall decrease, as expected from the negative Lyapunov exponents. After that, these magnitudes oscillate around a baseline value $[O(10^{-12})]$. The most weakly decaying mode also approaches this baseline by about 14 000 time steps. For exact linearized dynamics, Pires et al. (1996) showed (theoretically and experimentally) that stable mode error components in 4D-Var vanish to zero in the limit of infinite observation times. Pham et al. (1998) similarly demonstrated that the EKF analysis covariance converges exclusively to the unstable subspace over time with the tangent linear model if there is no system noise. Tests confirmed that these results extend to stable Lyapunov mode errors for large EnKF ensembles in the linearized and perfect model case (not shown here). However, the non-zero stable mode errors in the full non-linear EnKF case shown in Fig. 7a are too small to have an adverse effect if missed by the filter.

In contrast, Figs. 7b–c show that energy along the decaying Lyapunov vectors becomes more prominent as non-linearity (\sqrt{R}) increases. For $\sqrt{R} = 0.05$, the root mean-squared error along the most decaying Lyapunov vectors is still considerably smaller than in the linearized long-term unstable subspace (confirming dominance of linear dynamics), but the mean-squared error magnitude along the 3rd Lyapunov vector (the most weakly decaying), and occasionally along the 4th and 5th Lyapunov vectors, can at times be comparable to those along the unstable

subspace. For $\sqrt{R} = 1$, there is still generally greater ensemble error along the modes with larger Lyapunov exponents, but there is no longer such clear separation of the mode magnitudes.

The increasing presence of ensemble errors along the linearized stable modes as non-linearity (or \sqrt{R}) increases provides insight into why more replicates are required to achieve good filter performance for these conditions. Due to the limited number of degrees of freedom in the ensemble, the ensemble covariance can no longer fully represent the actual magnitude of notable errors along all modes. Once the ensemble starts to underrepresent uncertainty in the linearized long-term unstable subspace, fast growing errors may be inadequately corrected, leading to divergence. The need for sufficient ensemble sizes is indeed noted in our tests with greater non-linearity. For $\sqrt{R} = 0.05$, an $N - 1$ of at least 4 or 5 is needed to avoid divergence. For $\sqrt{R} = 1$, any reduced rank ensemble will eventually diverge.

To understand more specifically how non-linearity affects errors along different modes and ultimately impacts the performance of a reduced rank filter, it is useful to compare short-term error growth predicted from a linearized analysis with actual short-term error growth obtained with the fully non-linear system. Figure 8 shows the intermittent growth rate of errors over a single forecast time step along each Lyapunov vector for 3500 successive time steps (note that error changes due to the update step are not shown). Figure 8a shows local Lyapunov exponents, which represent the short-term linear growth of infinitesimal errors. Figure 8b shows the actual growth of errors (also expressed as exponents) that occurred in a large ensemble filter run for $\sqrt{R} = 0.05$. The results in Fig. 8b account also for the higher-order non-linear contribution in (19) while the results in Fig. 8a do not. While all Lyapunov modes exhibit some differences in growth between the two plots, the discrepancy is most dramatic for the decaying modes. The local Lyapunov exponents plotted in Fig. 8a predict very few periods of intermittent linear growth along the decaying Lyapunov modes, yet Fig. 8b shows that the actual error grows significantly at numerous times after an initial period. The differences between Figs. 8a and b reflect the impact of the higher-order non-linear terms in (19). Note that Fig. 8 depicts multiplicative growth, and not additive growth, which accentuates non-linear effects along stable modes. This is because errors along stable modes are relatively small for all \sqrt{R} values tested (see Fig. 7) due to predominantly linear decay, allowing small absolute non-linear contributions to result in significant relative growth for these modes. Unstable modes also undergo non-linear influences, but the similarities shown in Figs. 8a and b for these directions indicate that they are relatively minor compared to their error magnitudes.

How errors arise due to higher-order terms is illustrated in Fig. 9, which shows 2-D phase plots of the filter ensemble error replicates along the first (growing) and last (decaying) Lyapunov vectors. Note that although the axes are shown at right angles, actual characteristic Lyapunov vectors are non-orthogonal for the system. The filter ensemble is shown using black asterisks

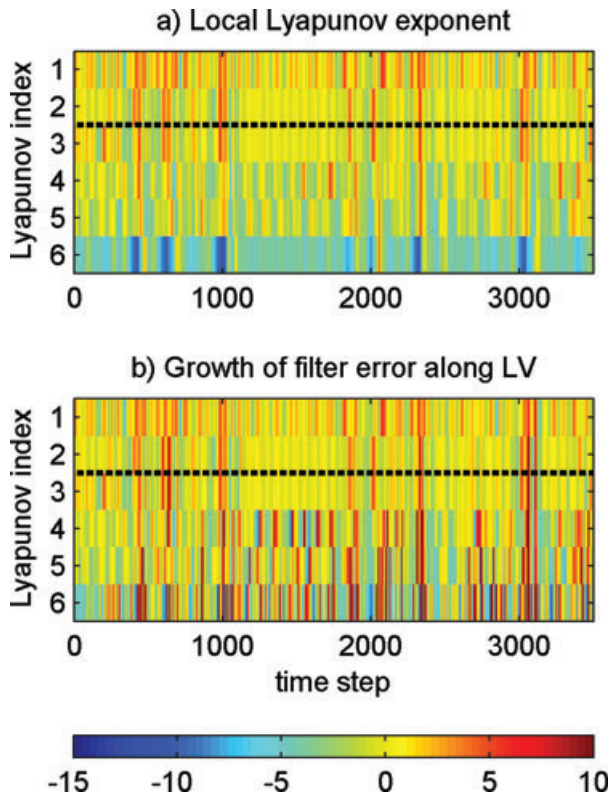


Fig. 8. Times series of intermittent growth along characteristic Lyapunov vector directions for the 6-D Lorenz-98 model. Growth is expressed as exponents, or the natural log of the multiplicative growth over one time step, divided by the units of time. Above the dotted black line indicates long-term growing modes. The local Lyapunov exponents plotted in (a) represent linear growth of very small perturbations. The growth of actual filter errors plotted in (b) are obtained from a 61-member ensemble with $\sqrt{R} = 0.05$. The difference between the actual growth exponents and linear local Lyapunov exponents represent non-linear error growth not accounted for by the Lyapunov vectors. Significant non-linear growth is apparent at some times in the long-term decaying modes in the filter.

at the start of a forecast period (after a transient spin-up period) and again 10 steps later, for a filter run with $N = 61$ and $\sqrt{R} = 0.05$. The grey points show forecast errors derived from a tangent linear model applied over the same 10 steps. The linearized model predicts that errors shrink along the decaying Lyapunov direction to align more strongly with the leading Lyapunov vector. In contrast, the actual errors obtained from a full non-linear model bend from the first toward the last Lyapunov mode. This movement of growing errors towards stable modes is clearly not captured by the tangent linear model. Figures 8 and 9 suggest that the primary impact of non-linearity, as conveyed by the higher-order terms in (19), is to move growing errors aligned with unstable linear growth modes into the directions corresponding to stable linear growth modes.

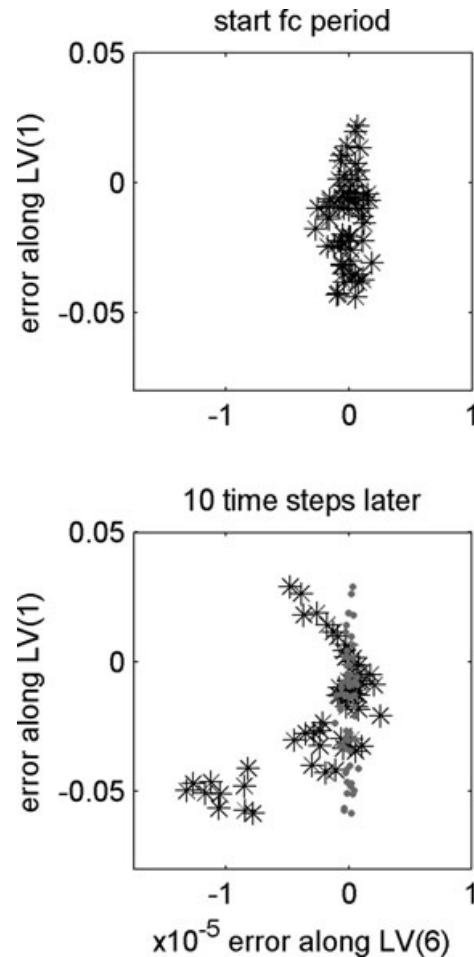


Fig. 9. 2-D phase plots of filter error replicates in black asterisks, along the 1st (growing) and 6th (decaying) characteristic Lyapunov vectors, for the 6-D Lorenz-98 model, with a 61-member ensemble and $\sqrt{R} = 0.05$. Grey points show forecast results from a tangent linear model, starting with the same ensemble at the start of the forecast period. The tangent linear model predicts that errors decrease in the stable direction. In reality, the actual errors grow in the stable direction. Non-linearities not captured by the tangent linear model have the effect of moving the actual replicate errors from the 1st Lyapunov vector direction into the 6th Lyapunov vector direction.

Once energy is pumped from unstable to stable Lyapunov modes due to non-linear dynamics, it frequently decays due to the dominant linear dynamics. However, decaying Lyapunov modes occasionally exhibit temporary but rapid linear growth (Fig. 8a). The phase plots in Fig. 10 show how an insufficiently large ensemble can fail to capture all the important error dimensions, due to combined non-linear and intermittent linear growth along stable modes. These plots show error replicates along the 2nd characteristic Lyapunov vector (corresponding to the 0-value Lyapunov exponent) and the 4th characteristic Lyapunov vector (a decaying mode), at every 15 time steps after time step 3370, for the filter run of Fig. 8, with $\sqrt{R} = 0.05$. The grey

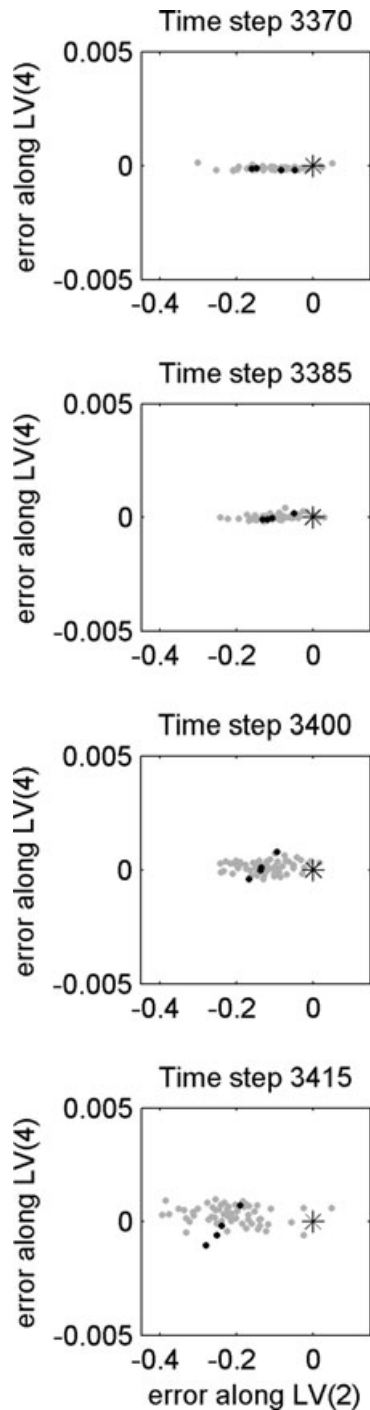


Fig. 10. 2-D phase plots of filter error replicates, along the 2nd (corresponding to 0-value Lyapunov exponent) and 4th (corresponding to negative Lyapunov exponent) characteristic Lyapunov vectors, for the 6-D Lorenz-98 model, and $\sqrt{R} = 0.05$. Grey markers show a nearly optimal 61-member ensemble, and the black markers show a 4-member ensemble that begins to diverge from the true state (0 error is shown with an asterisk). As the 4-member ensemble moves into the direction of the 4th Lyapunov vector, underdispersion becomes significant along the 2nd Lyapunov vector.

points are the error replicates for the $N - 1 = 60$ (non-divergent) case and represent a near-optimal ensemble. The black points are the error replicates for a run with $N - 1 = 3$, which begins to diverge by the end of the period depicted. Inadequate ensemble spread along strongly decaying modes is inevitable under full non-linear dynamics when using $N - 1 < n$ due to insufficient degrees of freedom in the ensemble. The small ensemble at early times in Fig. 10 indeed underrepresents spread along the stable mode, yet the magnitude is relatively unimportant, compared to error in the unstable mode (it is barely visible in Fig. 10). Thus, overall, the small ensemble adequately represents the analysis error up to time step 3370.

The progression of phase plots in Fig. 10 illustrates how the small ensemble (black dots), with only three degrees of freedom, eventually moves into and stretches along the 4th Lyapunov mode after time step 3370, leaving a strong underrepresentation of error along the 2nd Lyapunov vector. This underdispersion is apparent by comparing the small ensemble with the large ensemble, which approximates the optimal EnKF result. A detailed comparison of results in Figs. 8a versus b reveals that non-linear growth occurs along the 4th Lyapunov mode during time steps 3370–3385, originating from uncontained growth along the 1st Lyapunov vector. At the same time, error along the 2nd Lyapunov vector slightly decays under linearized dynamics. As a result, the small ensemble, with insufficient degrees of freedom, tilts away from the 2nd mode in order to increase its alignment with the 4th Lyapunov vector instead. This alone may not necessarily be problematic because the 4th mode corresponds to a negative Lyapunov exponent and will tend to shrink as a first order approximation. However, during the following time steps 3385–3415, the 4th Lyapunov vector also undergoes strong intermittent linear growth, causing the small ensemble to continue tilting rapidly into that direction. By time step 3400, error along the 2nd Lyapunov vector also grows, exacerbating the underdispersion of the small ensemble that has developed along the 2nd mode due to insufficient degrees of freedom in the ensemble. After this, the small ensemble no longer adequately represents spread along the 2nd Lyapunov vector, and the filter cannot correctly use observations to update error in that direction. Because the 2nd Lyapunov mode is not a long-term decaying mode, the filter diverges.

Together, Figs. 7–10 provide insight into the dynamic properties that affect EnKF performance. When perturbations are small, and dynamics are nearly linear, an ensemble with $N - 1$ equal to the number of non-negative Lyapunov exponents is sufficient to avoid divergence. This is because errors along decaying Lyapunov modes will rapidly shrink and become irrelevant (even though non-zero), and the ensemble perturbation replicates will self-align with the relevant linearized long-term unstable subspace. Error growth in the unstable subspace is then well-represented by the ensemble and thus controlled by the filter.

In the mildly non-linear case, greater ensemble sizes become necessary, but reduced rank EnKF formulations may still be

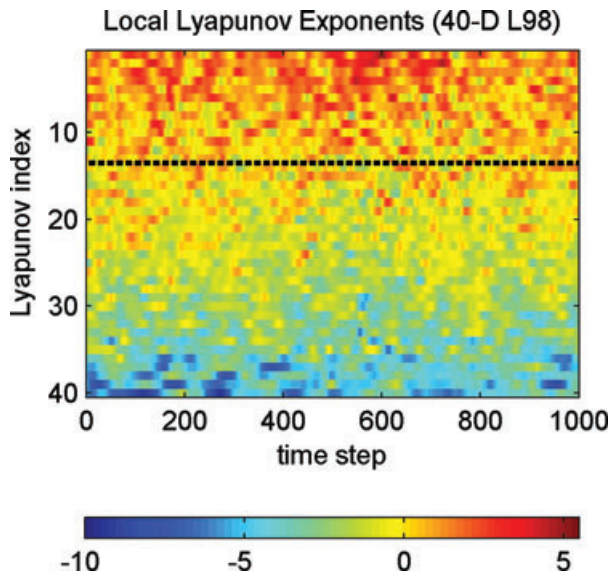


Fig. 11. Local Lyapunov exponents for the 40-D Lorenz-98 model indicate intermittent linear growth along Lyapunov vectors. Above the black dotted line indicate indices corresponding to long-term growing modes. Significant intermittent linear growth occurs occasionally for some long-term decaying modes.

possible. In such cases, the non-linear movement from growing Lyapunov modes to decaying modes may still be small, but they can no longer be ignored. Although those errors along decaying modes tend to later shrink, errors along modes corresponding to less negative Lyapunov exponents can be boosted at certain times due to intermittent linear growth. The combined effects of non-linear error movement from unstable to stable modes and intermittent growth along stable modes produce errors along the decaying Lyapunov vectors that can be comparable at times to errors along growing vectors. Intermittent linear growth along a stable mode will be later reversed and further attenuated by intermittent linear decay, yet an unstable mode may start to be underrepresented by a small ensemble (with too few degrees of freedom) during times when stable mode error magnitudes and growth compete. As a result, to avoid divergence, a sufficiently large ensemble size is needed to represent not only the linearized long-term unstable subspace, but also those decaying Lyapunov modes that can at times exhibit significant error magnitudes. Otherwise, unrepresented spread can develop along growing Lyapunov modes, eventually leading to filter divergence.

Generally, the decaying Lyapunov modes most prone to significant errors are those that more commonly undergo intermittent linear growth. This is evident in Fig. 7, which shows weakly decaying modes to be more likely than strongly decaying modes to have error magnitudes close to those of unstable modes at times. Local Lyapunov exponent time series, such as that shown in Fig. 8a, thus provide a useful tool for determining necessary ensemble sizes for robust, reduced rank EnKF. The analogous time series of local Lyapunov exponents shown in Fig. 11 for

the 40-D Lorenz-98 model clearly reveals the number of modes affected by notable intermittent linear growth (~ 25 – 30). This value, which is significantly larger than the number of unstable modes (14), corresponds well to the minimum ensemble size threshold needed for the mildly non-linear case of $\sqrt{R} = 0.1$ in Fig. 2.

When a system becomes more strongly non-linear, ensemble sizes significantly larger than $n + 1$ become necessary. This can occur for large \sqrt{R} or sparse observations, both of which cause error magnitudes to become significant. In these cases, even though ensemble spread may still be greatest along long-term growing modes, the non-linear portion of (19) becomes significant for all modes. The effect of sampling error on the ensemble forecast covariance then becomes critical, even for ensemble sizes greater than $n + 1$. In such cases, fixes such as inflation and localization will likely be inevitable.

6.3. Implications for larger-scale models

Our work with the Lorenz-98 model has shown that EnKF performance depends on both long-term and intermittent short-term dynamics. The respective roles of these two types of linearized error growth depend on the degree of non-linearity in the system. In particular, for very accurate and dense observations, filter performance depends only on linearized long-term dynamics, and an ensemble size corresponding to the dimension of the linearized long-term unstable subspace is adequate to avoid divergence. For cases of greater non-linearity, larger ensembles are needed, but reduced ensemble sizes less than the state size may still be adequate as long as they also accommodate decaying Lyapunov modes that become important due to intermittent short-term linearized growth. This result can have significant implications for larger scale problems if the reduced dimension size of the critical subspace approaches feasible computational loads. Additional tests conducted with the Lorenz-98 model (not shown here) showed that the number of Lyapunov modes affected by intermittent short-term linear growth is a smaller fraction of the state size as model dimension is increased.

To assess a more realistic model, Fig. 12 shows the first 1000 Lyapunov exponents for an idealized primitive equation general circulation model adapted from that used in Whitaker and Hamill (2002). It is a 2-level model with the prognostic variables of baroclinic and barotropic vorticity, baroclinic divergence, and barotropic potential temperature (barotropic divergence is assumed to be zero). It uses a spectral solution with truncation at wavenumber 15 (T15), which results in a little over 1000 degrees of freedom. The same physical parameters as Whitaker and Hamill's (2002) T31 version are adopted, except for the quad-harmonic diffusion e-folding time scale, which we set to 18 h so that the spectral distribution of the kinetic energy spectrum appropriately follows the classic power law (to the power -3). Vannitsem and Nicolis (1997) showed Lyapunov exponent sensitivity to this parameter. It can be seen from Fig. 12 that

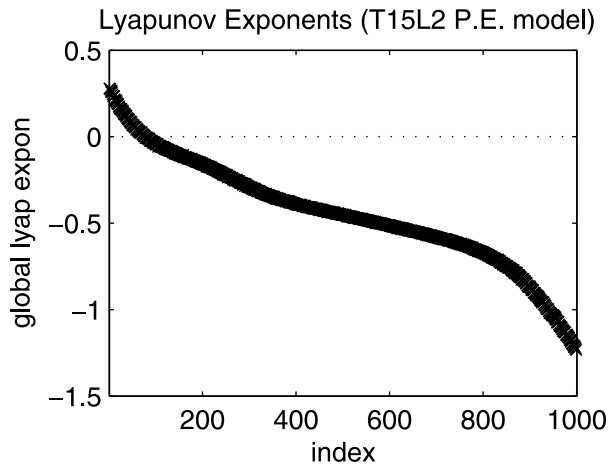


Fig. 12. The first 1000 Lyapunov exponents for a simplified 2-level primitive equation model with a little over 1000 degrees of freedom. There are 78 positive Lyapunov exponents.

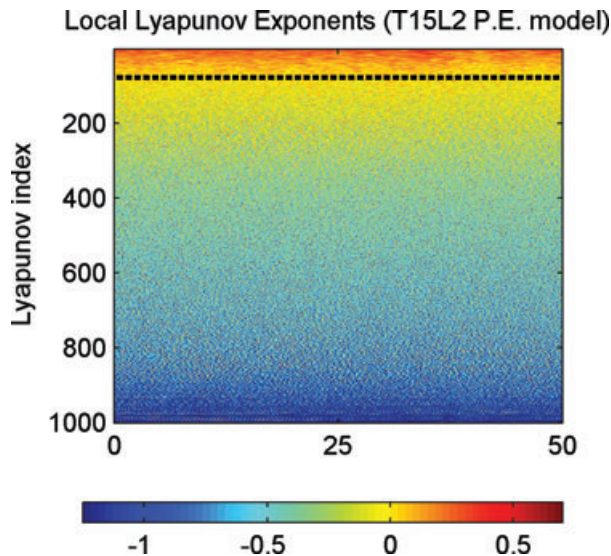


Fig. 13. Local Lyapunov exponents for the 2-level primitive equation model indicate intermittent linear growth along Lyapunov vectors. Above the black dotted line indicate indices corresponding to long-term growing modes. Significant intermittent linear growth occurs occasionally for some long-term barely decaying modes. The number of modes affected by strong intermittent linear growth is small relative to the full model size.

there are only 78 positive Lyapunov exponents, an impressive decrease from the model size. To assess the dimensionality of the subspace susceptible to intermittent linear growth, which becomes important for systems with greater uncertainty, we look at the time series of local Lyapunov exponents for the 2-level T15 P.E. model in Fig. 13. Intermittent linear growth mostly affects the first 100–150 modes, still a small fraction of the total state dimension.

Small linearized long-term unstable subspace dimensions have also been observed in other larger-scale geophysical models. Vannitsem and Nicolis (1997) calculated about 100 positive Lyapunov exponents for a 3-level T21 quasi-geostrophic model that has about 1500 degrees of freedom, and Snyder and Hamill (2003) found only 20 positive Lyapunov exponents for a quasi-geostrophic model with $o(10^5)$ degrees of freedom. These significantly reduced effective dimensionalities, determined by non-negative Lyapunov exponents, suggests that our findings on the relationship between the linearized long-term unstable subspace and EnKF performance could be useful with more realistic models.

Although the greatly reduced sizes of linearized long-term unstable subspaces indicate that our findings have implications for practical filtering, the ensemble sizes required to prevent divergence may still exceed current computational limitations for realistic, large-scale models. However, understanding that dynamics in these smaller numbers of modes drive EnKF performance is important for robust filtering. These insights can help improve variance inflation and localization. For example, in a number of EnKF tests carried out with variance inflation, we found that the dimension of the linearized long-term unstable subspace still acts as a minimal threshold on ensemble sizes to avoid divergence. Additionally, as observation error is increased (e.g. $\sqrt{R} = 1$), the ensemble size needed to be increased, even with inflation, to cover decaying Lyapunov modes that exhibit occasional strong intermittent linear growth. Long-term, intermittent and non-linear error growth are all important components for robust EnKF performance.

7. Conclusion

The influence of model dynamics has often been overlooked when assessing EnKF performance in geophysical applications. Studies that have considered its role in reduced rank filtering have varied in the types and time scales of dynamic properties examined. This work uses a simplified model of a chaotic attractor, the Lorenz-98 model (Lorenz and Emanuel, 1998), to clarify the dynamic features that affect EnKF performance.

Our results show that, even in the face of sampling error, reduced rank ensembles ($N - 1$ less than the model dimension) can produce non-divergent filter estimates that track the true state, without the use of variance inflation, localization, or any other modifications, as long as enough high quality measurements are available to justify the use of linearized Lyapunov theory. However, even in this idealized case, there is a minimum threshold for the ensemble size that corresponds approximately to the dimension of the linearized long-term unstable subspace. We find that the non-divergent filter ensemble only aligns weakly with singular vectors (representing short-term error dynamics), while it aligns strongly with leading characteristic or backward Lyapunov vectors (representing long-term error dynamics). Furthermore, divergence is avoided only when the ensemble fully

spans the entire linearized long-term unstable subspace, which is the space spanned by the characteristic (or backward) Lyapunov vectors corresponding to non-negative Lyapunov exponents. Ensemble alignment with the long-term unstable subspace is important for EnKF performance because errors in that space will eventually grow, and they can be corrected only if those directions are represented by the filter ensemble. The ensemble automatically moves to that linearized long-term subspace after some time, because most other errors decay. Sampling error is then mostly restricted to an irrelevant decaying subspace that has a minimal impact on EnKF performance.

Our results for nearly linear cases are consistent with the 4D-Var work of Pires et al. (1996) and Trevisan et al. (2010) and the new EKF work of Trevisan and Palatella (2011). However, we find that as measurement errors and the degree of non-linearity in the system increases, this subspace is no longer able to capture all important filter errors. As long as linear dynamics are still relevant, though, linearized error properties can continue to provide a framework for managing filter performance. We showed that for increasing non-linearity, some linearized long-term decaying modes, defined using characteristic Lyapunov vectors, must also be represented in the ensemble. This is because non-linear movement of errors from long-term linearized growing modes to long-term decaying modes prevents errors along decaying modes from shrinking indefinitely. Although this non-linear movement is generally smaller than growth in the long-term unstable subspace, some decaying modes occasionally undergo strong intermittent linear growth. When the non-linear movement into the long-term decaying mode is combined with intermittent linear growth, errors along these decaying modes may reach magnitudes that cannot be ignored. In such cases, it is critical for the ensemble to properly represent parts of the system's stable subspace as well as the unstable subspace. If the ensemble size is too small to adequately cover both the unstable subspace and decaying Lyapunov modes with intermittently large errors, it may underestimate spread in some of the unstable space. The filter then becomes highly prone to divergence because it is unable to correct larger-than-expected errors in the subspace spanned by the ensemble.

While a reduced rank EnKF is possible for some filter configurations without fixes like variance inflation and localization, such schemes are indispensable as non-linearity becomes more important. Even with ensemble sizes one or two orders of magnitude larger than the state dimension, the filter diverges with the Lorenz-98 model when observations are very noisy or very sparse. However, even in our highest \sqrt{R} value tests, the ensemble aligns preferentially with the linearized long-term growth modes corresponding to larger Lyapunov exponents. Furthermore, even with inflation, we found it is still necessary to have ensemble sizes large enough to cover the entire linearized long-term unstable subspace and additional decaying modes that undergo strong intermittent linear growth. This indicates that although sampling error in the forecast ensemble covariance

becomes significant with greater non-linearity, error growth dynamics considered here can still impact EnKF performance. Of course, dynamics based on linearized flow descriptors such as Lyapunov vectors may become meaningless if errors become too large. As a result, we caution that the relative importance of the linear and higher order term component of EnKF dynamics should be evaluated in any particular application before applying concepts from linearized dynamics. Fortunately, operational applications in meteorology and oceanography may have sufficient observations to make linearized analyses relevant, as long as proper attention is given to the synergistic effects of intermittent errors in decaying modes and non-linear and linear propagation of errors from unstable to stable modes.

Overall, our results have shown that the linearized long-term unstable subspace and the subset of stable modes that experience notable intermittent linear growth must both be captured to avoid divergence and insure robust EnKF performance with chaotic systems. The relative importance of the unstable and stable subspaces depends on the degree of non-linearity in the filter, which is influenced by the observation properties. The size of the linearized long-term unstable subspace and the number of additional attractor modes susceptible to intermittent linear growth appears to be a relatively small fraction of the full model dimension, even in more realistic geophysical models. This has promising implications for the development of accurate and robust reduced rank filtering approaches. It is likely that practical limitations on computational loads and observation capabilities may still make fixes such as variance inflation and localization necessary. However, the work described here can help guide future developments in robust filtering that properly account for the impact of chaotic dynamics on EnKF performance.

8. Acknowledgments

Funding for this research was provided by the following National Science Foundation grants: 0530851 (CMF program), 0540259 (DDAS program), and 0121182 (ITR/AP program). We thank Jeff Whitaker for providing the code for the 2-level primitive equation model used in part of this work. We are also grateful for helpful discussions with Pavel Sakov, Jim Hansen and Paola Malanotte-Rizzoli. An anonymous reviewer contributed useful comments and suggestions that significantly improved this paper.

References

- Anderson, J. L. and Anderson, S. L. 1999. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.* **127**, 2884–2903.
- Anderson, J. L. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.* **129**, 2884–2903.
- Annan, J. D. 2004. On the orthogonality of bred vectors. *Mon. Wea. Rev.* **132**, 843–849.

- Bowler, N. 2006. Comparison of error breeding, singular vectors, random perturbations and ensemble Kalman filter perturbation strategies on a simple model. *Tellus* **58A**, 538–548.
- Carne, S., Pham, D.-T. and Verron, J. 2001. Improving the singular evolutive extended Kalman filter for strongly nonlinear models for use in ocean data assimilation. *Inverse Probl.* **17**, 1535–1559.
- Carrassi, A., Vannitsem, S., Zupanski, D. and Zupanski, M. 2009. The maximum likelihood ensemble filter performances in chaotic systems. *Tellus* **61A**, 587–600.
- Cohn, S. E. and Todling, R., 1996. Approximate data assimilation schemes for stable and unstable dynamics. *J. Meteor. Soc. Jpn.* **74**, 63–75.
- Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**, 10143–10162.
- Evensen, G. 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.* **53**, 343–367.
- Farrell, B. F. and Ioannou, P. J. 2001. State estimation using a reduced-order Kalman filter. *J. Atmos. Sci.* **58**, 3666–3680.
- Fukumori, I. and Malanotte-Rizzoli, P. 1995. An approximate Kalman filter for ocean data assimilation: an example with an idealized Gulf Stream model. *J. Geophys. Res.* **100**, 6777–6793.
- Furrer, R. and Bengtsson, T. 2007. Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *J. Multivariate Anal.* **98**, 227–255.
- Hamill, T. M., Whitaker, J. S. and Snyder, C. 2001. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.* **129**, 2776–2790.
- Hansen, J. A. and Smith, L. A. 2000. The role of operational constraints in selecting supplementary observations. *J. Atmos. Sci.* **57**, 2859–2871.
- Heemink, A. W., Verlaan, M. and Segers, A. J. 2001. Variance reduced ensemble Kalman filtering. *Mon. Wea. Rev.* **129**, 1718–1728.
- Houtekamer, P. L. and Mitchell, H. L. 1998. Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.* **126**, 796–811.
- Houtekamer, P. L. and Mitchell, H. L. 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.* **129**, 123–137.
- Lawson, W. G. and Hansen, J. A. 2004. Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon. Wea. Rev.* **132**, 1966–1981.
- Legras, B. and Vautard, R. 1996. A guide to Lyapunov vectors. In: *Seminar on Predictability*. ECMWF, Reading, United Kingdom, 143–156.
- Lermusiaux, P. F. J. and Robinson, A. R. 1999. Data assimilation via error subspace statistical estimation. Part I: Theory and schemes. *Mon. Wea. Rev.* **127**, 1385–1407.
- Lorenz, E. 1996. Predictability: A problem partly solved. In: *Proc. Seminar on Predictability*. Vol. 1, ECMWF, Reading, United Kingdom, 1–18.
- Lorenz, E. and Emanuel, K. A. 1998. Optimal sites for supplementary weather observations: simulation with a small model. *J. Atmos. Sci.* **55**, 399–414.
- Molteni, F., Buizza, R., Palmer, T. N., and Petroliagis, T. 1996. The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.* **122**, 73–119.
- Oseledec, V. I. 1968. A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.* **19**, 197–231.
- Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J. and Yorke, J. A. 2004. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus* **56A**, 415–428.
- 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. doi:10.1111/j.1600-0870.2009.00419.x.
- Pham, D. T., Verron, J. and Roubaud, M. C. 1998. A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Syst.* **16**, 323–340.
- Pires, C., Vautard, R. and Talagrand, O. 1996. On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus* **48A**, 96–121.
- Sacher, W. and Bartello, P. 2008. Sampling errors in ensemble Kalman filtering, Part I: Theory. *Mon. Wea. Rev.* **136**, 3035–3049.
- Sacher, W. and Bartello, P. 2009. Sampling errors in ensemble Kalman filtering, Part II: Application to a barotropic model. *Mon. Wea. Rev.* **137**, 1640–1654.
- Sakov, P. and Oke, P. R. 2008. Implications of the form of the ensemble transformation in the ensemble square root filter. *Mon. Wea. Rev.* **136**, 1042–1053.
- Snyder, C. and Hamill, T. M. 2003. Leading Lyapunov vectors of a turbulent baroclinic jet in a quasigeostrophic model. *J. Atmos. Sci.* **60**, 683–688.
- Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. 2003. Ensemble square root filters. *Mon. Wea. Rev.* **131**, 1485–1490.
- Toth, Z. and Kalnay, E. 1993. Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.* **74**, 2317–2330.
- Toth, Z. and Kalnay, E. 1997. Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.* **125**, 3297–3319.
- Trevisan, A., D’Isidoro, M. and Talagrand, O. 2010. Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension. *Q. J. R. Meteor. Soc.* **136**, 487–496.
- Trevisan, A. and Palatella, L. 2011. On the Kalman Filter error covariance collapse into the unstable subspace. *Nonlin. Proc. Geophys.* **18**, 243–250.
- Trevisan, A. and Pancotti, F. 1998. Periodic orbits, Lyapunov vectors, and singular vectors in the Lorenz system. *J. Atmos. Sci.* **55**, 390–398.
- Trevisan, A. and Uboldi, F. 2004. Assimilation of standard and targeted observations within the unstable subspace of the observation-analysis-forecast cycle system. *J. Atmos. Sci.* **61**, 103–113.
- van Leeuwen, P. J. 1999. Comment on “Data assimilation using an ensemble Kalman filter technique.” *Mon. Wea. Rev.* **127**, 1374–1377.
- Vannitsem, S. and Nicolis, C. 1997. Lyapunov vectors and error growth patterns in a T21L3 quasigeostrophic model. *J. Atmos. Sci.* **54**, 347–361.
- Verlaan, M. and Heemink, A. W. 1995. Data assimilation schemes for non-linear shallow water flow models. *Stoch. Hydro. Hydraul.* **11**, 349–368.
- Whitaker, J. S. and Hamill, T. M. 2002. Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.* **130**, 1913–1924.
- Wolfe, C. L. and Samelson, R. S. 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.