Ocean Dynamics (2007) 57: 32–45
DOI 10.1007/s10236-006-0088-8

Peter R. Oke · Pavel Sakov · Stuart P. Corney

Impacts of localisation in the EnKF and EnOI: experiments with a small model

Received: 23 June 2005 / Accepted: 27 July 2006 / Published online: 12 September 2006

Abstract The performance of an inexpensive, ensemble-based optimal interpolation (EnOI) scheme that uses a stationary ensemble of model anomalies to approximate forecast error covariances, is compared with that of an ensemble Kalman filter (EnKF). The model to which the methods are applied is a pair of "perfect", one-dimensional, linear advection equations for two related variables. While EnOI is sub-optimal, it can give results that are comparable to those of the EnKF. The computational cost of EnOI is typically about N times less than that of EnKF, where N is the ensemble size. We suggest that EnOI may provide a practical and cost-effective alternative to the EnKF for some applications where computational cost is a limiting factor. We demonstrate that when the ensemble size is smaller than the dimension of the model's sub-space, both the EnKF and EnOI may require localisation around each observation to eliminate effects of sampling error and to increase the effective number of independent ensemble members used to construct an analysis. However, localisation can degrade an analysis if the length-scales of the localising function are too short. We demonstrate that, as the length-scale of the localising function is decreased, localisation can significantly compromise the model's dynamical balances. We also find that localisation artificially amplifies high frequencies for applications of the EnKF. Based on our experiments, for applications where localisation is necessary, the length-scales of the localisation should be larger than the decorrelation length-scales of the variables being updated.

Keywords Data assimilation · Ensemble Kalman filter

Responsible editor: Jörg-Olaf Wolff

P. R. Oke (✉) · P. Sakov · S. P. Corney
CSIRO Marine and Atmospheric Research
and Wealth from Oceans Flagship Program,
Hobart, Tasmania, Australia
e-mail: peter.oke@csiro.au

1 Introduction

For many oceanographic and meteorological centres, the long-term objective is a global high-resolution forecast system, coupled to an optimal data assimilation system that synthesises all available observations to produce realistic, dynamically consistent analyses. A practical limitation of such a system is the computational requirements of data assimilation systems like the ensemble Kalman filter (EnKF) and four-dimensional variational assimilation, which, in the case of EnKF, often limit the possible size of the ensemble.

The EnKF and other related ensemble-based assimilation systems use the Kalman filter equations to blend observed and modelled fields to produce an ensemble of analyses. The EnKF uses the statistics from an ensemble of forecasts to weight and distribute information from observations onto a model grid.

The sensitivity of the EnKF to ensemble size has been extensively investigated in the literature, with varied results (e.g. Houtekamer 1995; Houtekamer and Mitchell 1998; Keppenne 2000; Mitchell et al. 2002; Keppenne and Rienecker 2002); it is clear, however, that because the increments derived from ensemble-based schemes, without localisation, are a linear combination of ensemble perturbations (Evensen 2003), it is essential for a successful ensemble-based assimilation that the ensemble adequately spans the model sub-space. A possible solution to situations when only a small ensemble is feasible is the use of a technique called localisation. Localisation is typically achieved either by masking of covariances between distant elements of the model state vector (Hamill et al. 2001; Houtekamer and Mitchell 2001) or by applying filters locally in physical space (Ott et al. 2004).

In the extreme case, where the computational resources only afford a single forecast, one could employ an ensemble-based data assimilation system that uses a stationary ensemble, such as that described by Oke et al. (2002), Evensen (2003) and Oke et al. (2005). Evensen (2003) called this approach ensemble optimal interpolation (EnOI). Although EnOI typically

allows the use of a bigger ensemble than an EnKF, it is often not nearly enough to span the full sub-space of the model, making it necessary to use localisation.

This manuscript compares the performances of EnKF and EnOI for a small model application and investigates the advantages and disadvantages of localisation, including an analysis of the impacts of localisation on a model's dynamical balances. Because of the simplicity of the model used here, we are able to explicitly assess the level of imbalance introduced by localisation and investigate the influence of localisation on the modal composition of analyses. This sets the current study apart from previous papers that have examined the impacts of localisation on a model's dynamical balance (e.g. Mitchell et al. 2002).

This paper is organised as follows: The assimilation methodology is described in Section 2, followed by a description of the idealised model and the experiment design in Section 3. The results are presented in Section 4, followed by a summary and conclusion in Section 5.

2 Assimilation methodology

2.1 Analysis equations

We define w^f as an m-dimensional forecast state, w^o is a p-dimensional set of observations, H is an operator that interpolates from the model space to the observation space, P^f is an $m \times m$ forecast-error covariance matrix and R is a $p \times p$ observation-error covariance matrix. The minimum error-variance estimate of the analysed state w^a is given by the analysis equation:

$$w^a = w^f + K\left(w^o - Hw^f\right), \tag{1}$$

$$\text{where } K = P^f H^T \left(HP^f H^T + R\right)^{-1} \tag{2}$$

is the gain, or weight, matrix. The matrix K determines the relative weights of the forecast and the observations in an analysis and performs the interpolation of the forecast innovation $\left(w^o - Hw^f\right)$, on to all model variables at all model grid points. The main challenge for a successful application of Eqs. 1 and 2 is the estimation of P^f.

2.2 The ensemble Kalman filter (EnKF)

The EnKF approximates P^f in Eq. 2 as

$$P^f = A' A'^T/(N-1), \tag{3}$$

where N is the number of ensemble members and A' is an $m \times N$ matrix of ensemble perturbations that are defined as

$$A' = A^f - \overline{A^f}, \tag{4}$$

$$\text{where } \mathrm{A^f} = [\mathrm{w_1^f}, \mathrm{w_2^f}, \cdots, \mathrm{w_N^f}] \tag{5}$$

is an ensemble of model forecasts, w_i^f is the ith ensemble member in the forecast ensemble and $\overline{A^f}$ is the ensemble mean. During an assimilation step, each ensemble member is compared to the observations and is updated according to the analysis equation (Eq. 1) after substituting an approximation for the covariance matrix (Eq. 3) into the gain matrix (Eq. 2), yielding

$$K = A'(HA')^T \left(HA'\left(HA'\right)^T + (N-1)R\right)^{-1}. \tag{6}$$

For applications to an imperfect model, the covariances of the model error also need to be included in the evolution of A'. This can be achieved in a number of different ways; for example, by inflating the ensemble perturbations (e.g. Anderson and Anderson 1999; Whitaker and Hamill 2002) or through the addition of spatially correlated noise (Evensen 2003). When the ensemble members in an EnKF are appropriately initialised, and evolved in time, the EnKF arguably offers a nearly optimal method for assimilation. However, the EnKF is expensive, requiring N versions of the model to be run simultaneously.

2.3 Ensemble optimal interpolation (EnOI)

An approximation to the EnKF is EnOI. This involves using a stationary ensemble to define P^f in Eq. 3, using

$$A' = \alpha\left[w_1', w_2', \cdots, w_N'\right], \tag{7}$$

where N is, again, the number of ensemble members; A' is an ensemble of scaled model anomalies; α is a constant that scales the magnitude of the model anomalies, and therefore, the magnitude of the assumed forecast errors and w_i' is the *ith* model anomaly. As an example, a model anomaly could be an anomaly from a long-term mean, or in the case of Oke et al. (2005), intraseasonal anomalies in a long model integration. EnOI doesn't require each ensemble member to be integrated with time; the ensemble is stationary.

The rationale for the EnOI-approach is threefold. Firstly, one might expect that the magnitude of a model's forecast

errors, represented by the diagonal elements of P^f, are probably related to the variance that the model simulates. So, in regions of strong variability, where the model is most likely to produce large errors, the assumed forecast-error variance is also large. Secondly, if two elements of the model state are typically correlated in a long model integration, then one might reasonably expect that the errors of those state elements are also correlated. So, under the above-stated assumption, if two state elements typically vary together, then an adjustment to one will result in an adjustment to the other, and the relative magnitude of these adjustments will be directly related to the covariance between these elements. Thirdly, in the same way as the EnKF, EnOI produces analyses that lie within the model's sub-space.

2.4 Localisation

To understand further how ensemble-based assimilation works, it is helpful to recognise that the increments derived from these systems can be expressed as

$$w^a - w^f = A'c, \tag{8}$$

where c is an N -dimensional column vector. Evensen (2003) refers to this as a weakly non-linear combination because the coefficients c depend on the combined vectors in A. By contrast, Nerger et al. (2005) refer to this as a weighted average; however, we note that this implicitly assumes that the coefficients sum to one, which is generally not the case. When expressed in this form, it is clear why the success of any ensemble-based data assimilation system critically depends on the representativeness of its ensemble members. If the ensemble doesn't span the same space as the forecast errors, then it cannot reliably produce increments that bring the model in agreement with the true state. This is particularly a problem when a small ensemble size is used; and unfortunately, for most realistic applications, the computational resources limit the number of ensemble members that can be used to a modest size.

For ensemble-based assimilation, the rank of the estimated forecast error covariance matrix is, at most, $(N-1)$. This rank can be increased significantly using localisation (e.g. Hamill et al. 2001; Houtekamer and Mitchell 2001). A typical implementation of localisation may involve the multiplication of the ensemble-based covariances in P^f by a correlation function so that the gain matrix in Eq. 2 is re-expressed as

$$K = (C \circ P^f)H^T\left(H^T(C \circ P^f)H^T + (n-1)R\right)^{-1}, \tag{9}$$

where C is an $m \times m$ correlation matrix and the operation denoted by the open circle is an element-by-element matrix multiplication (also called a Schur or Hadamard product); so if $B = C \circ P^f$, then $B_{ij} = C_{ij}P^f_{ij}$, where C_{ij} is a correlation between the ith and jth state elements of the model that depends only on the Euclidean distance between these elements in physical space. Here, we use a Gaussian function with an $e-$folding scale of L. Consider an example where a single observation is assimilated. Then, the innovation covariance matrix, the matrix to be inverted in Eq. 9, is a scalar. The distribution of the forecast innovation, $(w^o - Hw^f)$, onto the rest of the model state is determined by the covariances in P^f. When localisation is used, these covariances are artificially reduced to near zero with an $e-$folding length-scale of L.

The primary benefit of localisation is often considered to be the reduction of spurious long-range covariances occurring when a small ensemble is used (Hamill et al. 2001; Houtekamer and Mitchell 2001). Importantly, localisation also effectively increases the rank of the system. To understand this, consider an example where two observations are assimilated. Suppose these observations are far enough apart that the off-diagonal elements of $H(C \circ P^f)H^T$ are zero. Then, the combination of ensemble anomalies that comprise the increment, represented in Eq. 8, around each observation is effectively independent. So, the ensemble anomalies can be effectively split into two separate ensembles, one for each observation, and the column vector c , in Eq. 8, is effectively doubled in size. That is, the ensemble size is effectively doubled. More generally, suppose we have p observations that are greater than $4L$ apart; then, the effective ensemble size could be expected to be about pN.

For an example with many observations that roughly span the full model domain, the effective ensemble size is $N \times N_x$, where N is the ensemble size and N_x is the number of independent sub-domains. For a one-dimensional domain, of length X, we might suppose that the covariances are near enough to zero at a distance of $2 \times L$ from an observation. So, there are $X/(4L)$ independent domains in a domain of length X. If $X = 1{,}000$ and $L = 50$, the rank, or effective ensemble size, could be expected to be about $5N$.

By increasing the rank of the system, an analysis can be computed that fits the observations better. This explains why Houtekamer and Mitchell (1998), for example, obtain satisfactory results using only 16 ensemble members for a global application. Without appreciating the effects of localisation, these results seem incredible, given that the atmospheric model used has many more degrees of freedom than the size of their ensemble.

The benefits of localising have been demonstrated by several authors (Hamill et al. 2001; Houtekamer and Mitchell 2001); however, Evensen (2003) and others have suggested that it will result in non-dynamical modes being introduced into the model. This issue has been explored in the context of an atmospheric general circulation model by Mitchell et al. (2002). To gain a better understanding on the effects of localisation on dynamical balance, we present analyses from a series of experiments with a simple, linear

model where the model's dynamical balances are known exactly.

3 Experiment design

The model used here is a variation of that employed by Evensen (2004). It is a one-dimensional linear advection model, solving

$$\frac{da}{dt} + u\frac{da}{dx} = 0, \tag{10}$$

where $u = 1$ is the speed of advection, a is a model variable, t is time and x is space, ranging from 1 to 1,000, with grid spacings of 1. The computational domain is periodic in x. The main difference between this and Evensen's (2004) application is that we use two related model variables, variable a and variable b, where b is initialised as

$$b = 0.5 + 10\frac{da}{dx}, \tag{11}$$

where the multiple 10 was chosen to get the range for the variable b of the same order as for the variable a for fields generated by the procedure described below. The relationship between a and b in Eq. 11 is motivated by the relationship between pressure and velocity in oceanic and atmospheric applications. That is, geostrophic velocities are proportional to pressure gradients. This experiment design facilitates an assessment of multivariate ensemble-based data assimilation when only one variable is observed. This is somewhat analogous to the case where sea-level anomalies, from satellite altimetry, are assimilated into an ocean model to adjust all state variables including ocean currents. We seek to evaluate the extent to which observations of variable a can be used to adjust variable b, and additionally, whether the initial dynamical balance between a and b, represented by Eq. 11, is maintained during the assimilation process.

Both the ensemble and the true field are generated from random samples. For each sample, variable a is set as a linear combination of 25 sine curves with random uniformly distributed amplitude and phase and a random offset (generated by the harmonic with $k = 0$):

$$a_i^0 = \sum_{k=0}^{25} A_k \sin\left(\frac{2\pi k}{1000} i + \varphi_k\right), \quad 1 \le i \le 1000,$$

$0 < A_k < 1$, $0 < \varphi_k < 2\pi$, where a_i^n is the value of variable a in the *ith* grid node at the *nth* time step. Subsequently, the vector $a^0 = (a_1^0, \cdots, a_{1000}^0)^T$ is normalised so that it has a variance of 1. Variable b is initialised as a centred difference approximation to the second term in the right-hand side of Eq. 11:

$$b_i^0 = 5(a_{i+1}^0 - a_{i-1}^0),$$

The first generated sample is offset so that variable a has a mean of 6 and variable b has a mean of 0.5; it represents a reference field. The true field is defined as a sum of the reference field and the next sample. The remaining N samples are generated to form the ensemble. For each ensemble sample, the ensemble mean is subtracted and variable a is normalised to have a variance of 1. The reference field is then added to each ensemble sample. Because we subtract the ensemble mean after the initial normalisation, each ensemble perturbation for a will have a variance that is slightly different from 1.

This procedure is similar to that described by Evensen (2004); however, for simplicity, both the penalty for shorter waves and the improved sampling strategy described by Evensen are not used here. The decorrelation length-scales for variables a and b are both around 20, which is equal to the half-period of the harmonic with the smallest period.

For the experiments that use EnOI, the model is initialised by the ensemble mean (that is, the reference field), and the ensemble members do not change with time. For applications of the EnKF, where each ensemble member evolves in time, the time-evolving ensemble mean is used as the best estimate from the system for comparison with the true field.

To advance the model in time, instead of integrating Eq. 10 numerically, we use the general solution

$$a(x, t) = f(x - ut),$$

where $f(x)$ is an arbitrary function; we then choose $u = 1$ and simply shift at each time step the current values of variables a and b by one grid node: $a_i^{n+1} = a_{i-1}^n$, $i = 1, \cdots, 999$, $a_1^n = a_{1000}^{n+1}$; $b_i^{n+1} = b_{i-1}^n$, $i = 1, \cdots, 999$, $b_1^n = b_{1000}^{n+1}$.

For each assimilation experiment, four equally spaced observations of a are periodically obtained from the true state and are contaminated with uncorrelated, normally distributed noise with a variance of 0.01, after Evensen (2004). Therefore, an observation error variance of 0.01 is used in all experiments considered in this study. For EnKF, the normally distributed noise with the same variance is also independently added to these values for each ensemble member. These observations are obtained and assimilated every five time steps.

Because we use 25 orthogonal sine functions with random phase and amplitude as well as a random offset to initialise the ensemble fields, the dimension of the model state space is 51. Therefore, in the following experiments, ensembles of size $N = 10$ or $N = 20$ can be considered small, ensembles of size $N = 50$ intermediate and ensembles of size $N = 100$ large.

4 Results

4.1 Ensemble spread as a tuning parameter

The parameter α in Eq. 7 controls the spread of the stationary ensemble used in EnOI experiments. Larger values of α are found to give faster initial convergence rates; however, they also lead to larger asymptotic values of the root-mean-squared error (RMSE) between the analysed and true fields. We note that we use the term "asymptotic value" for the RMSE value reached by both EnOI-based and EnKF-based systems at the time when it stops changing in a practical sense.

Overall, using a smaller value of α gives a more conservative approach in relation to the convergence of an EnOI system, although this may change for a non-stationary system when a faster initial convergence rate may be required for a better performance. For example, Oke et al. (2005) use values of $\alpha = 0.5$ for an application to a global eddy-resolving ocean model. In this study, we do not focus on the problem of optimal α because it may depend on many factors, and we use the value $\alpha = 0.05$ unless stated otherwise.

4.2 General behaviour of the system

Figure 1 shows the convergence in assimilation experiments conducted with EnKF and EnOI using 100 ensemble members. This figure demonstrates that both variable a and variable b approach the true field when the EnKF is used. Similarly, both variables become well aligned with the true fields using EnOI; however, for small times ($t \leq 300$), the EnKF clearly outperforms EnOI. The reason for this is in the gradual reduction of the ensemble spread in the EnKF, which leads to a more accurate estimate of the forecast error covariances. By contrast, the ensemble spread is unchanged when EnOI is used, and the static covariances calculated with this static ensemble do not improve in time. Therefore, the analysis converges to the true field more gradually.

The same example is repeated using 100 ensemble members, but with the inclusion of localisation using an e-folding scale of 50 (Fig. 2). A comparison between Figs. 1 and 2 shows that localisation improves the performance of both EnOI and EnKF, with the EnKF again having a faster rate of conversion.

The experiments presented in Figs. 1 and 2 are repeated using only 20 ensemble members without localisation. These experiments initially fail to produce reasonable results for both the EnKF and EnOI (not shown); however, after reducing α to 0.03, the EnOI-based system robustly, albeit slowly, converges. This is an example when an EnOI outperforms EnKF. The reason for this surprising result is that EnOI does not advect the ensemble, while EnKF does. Because of that, for EnOI at each assimilation event, the increment is expanded in an effectively different basis. While the ensemble itself has insufficient range to reproduce the true field, being taken with different offsets over a period of time, it effectively spans the full model space. To take advantage of this feature of EnOI here, one needs to use a small ensemble spread to avoid instabilities

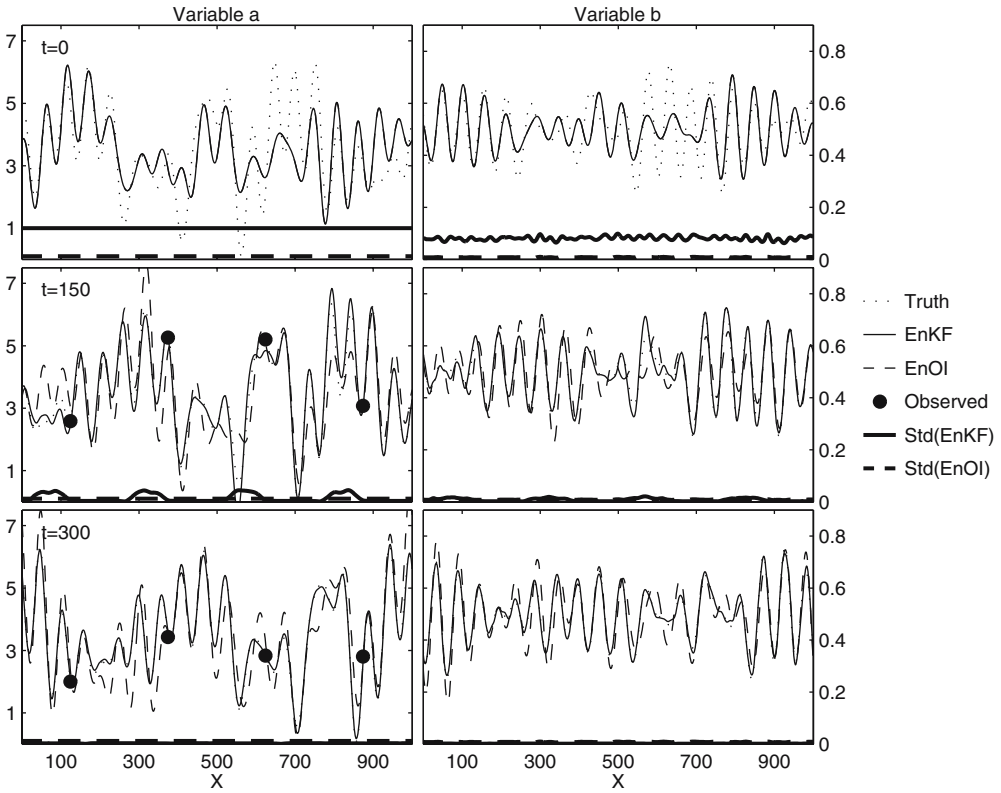

Fig. 1 Example of an EnKF and EnOI ($\alpha = 0.05$) experiment showing the true field, EnKF mean analysis, EnOI analysis, observations, EnKF standard deviation and EnOI standard deviations for variable a (*left*) and variable b (*right*) for time $t = 0$, 150 and 300 (top to bottom) using 100 ensemble members. Where the truth field is not visible, the EnKF solution overlays the truth

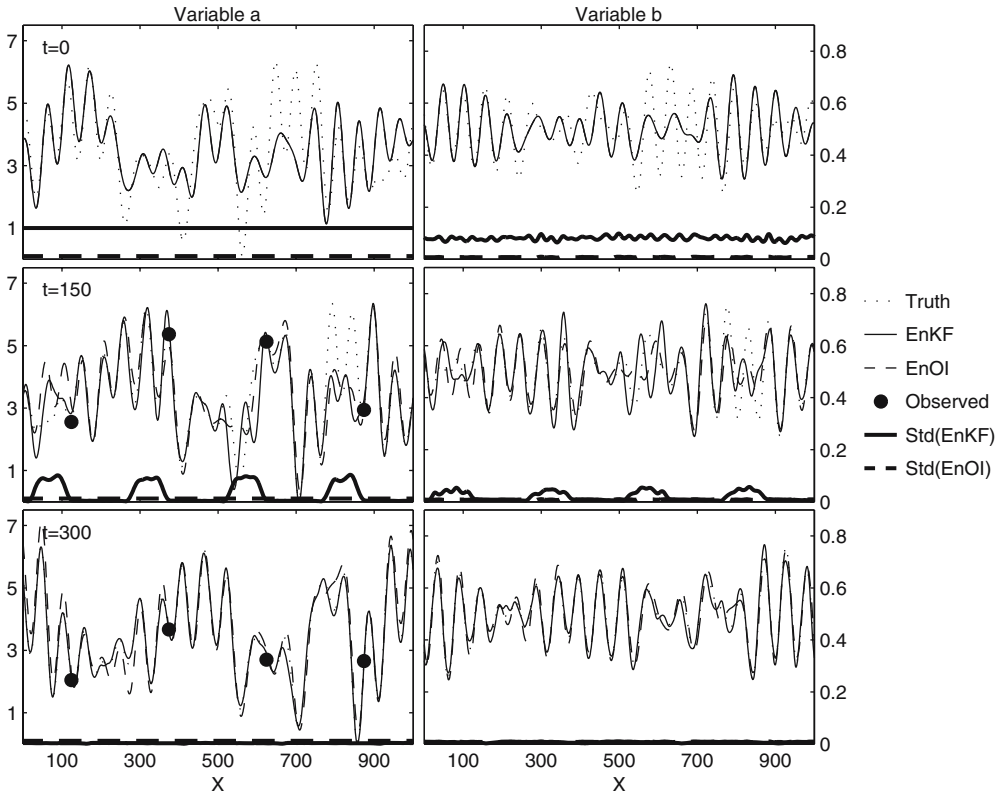

Fig. 2 As for Fig. 1, using 100 ensemble members, except also using localisation with a length-scale of 50

and integrate over a long period of time to allow statistical accumulation of the increments. In most practical situations, evolution of the model will prevent this accumulation, which is very slow; however, the stationary character of the model in hand does allow it.

When localisation is used with a 20-member ensemble and an e -folding length scale of 50, both the EnKF and EnOI produce analyses that converge quickly to the true field (Fig. 3). Again, when localisation is used, there are only small differences between the final analyses using EnOI and EnKF.

In the idealised model used here, the consecutive model states represent a series of identical shapes moving through the domain from left to right. In the experiments using the EnKF, each ensemble member aligns with the true field, resulting in the regions to the right of the observation points in an ensemble mean that agrees well with this field, and rendering the variance of the ensemble near zero. This can be clearly seen in Figs. 2 and 3 at $t = 150$, for example. By contrast, to the left of each observation, the ensemble spread is relatively large, and so, the forecast error variance is also large. When a new set of observations are assimilated, the regions where the ensemble spread is small are not significantly changed during the analysis step, while in the regions where the ensemble spread is larger, the assimilation of new observations can result in a significant change to each ensemble member, depending on the covariances and on the magnitude of the forecast innovations. For this idealised model, this is exactly how we expect the EnKF to work.

Now consider the examples with EnOI, where the ensemble spread is stationary. Initially, the ensemble spread is too small, so the forecast error variance is assumed to be smaller than it really is. As a result, observations are under-utilised. Despite this, the corrections are, on average, in the right direction, and after several assimilation cycles, the EnOI-based analysis becomes nicely aligned with the true field. However, because the ensemble spread does not decrease as the observations move through the domain, at some moment, the estimated forecast error variance becomes larger than it really is. In this case, observations are over-utilised, and this can degrade the analysis. The performance of EnOI in this respect can be improved by gradually reducing the scaling constant α in Eq. 7. For an application with an imperfect model, this issue is not likely to be such a problem.

4.3 Sensitivity experiments

We now turn to a statistical comparison of the performance of the EnKF and EnOI for different configurations of the assimilation systems. To gain statistically meaningful results, we generate 50 independent realisations of the true field and the 100-member ensemble following the procedure described in Section 3. In the analysis that follows, we present the RMSE and the anomaly correlation coefficient (CC) between the true and analysed fields (the ensemble mean for EnKF applications), averaged over these 50 independent realisations. The RMSE and CC are presented as functions of time. With a speed of advection of

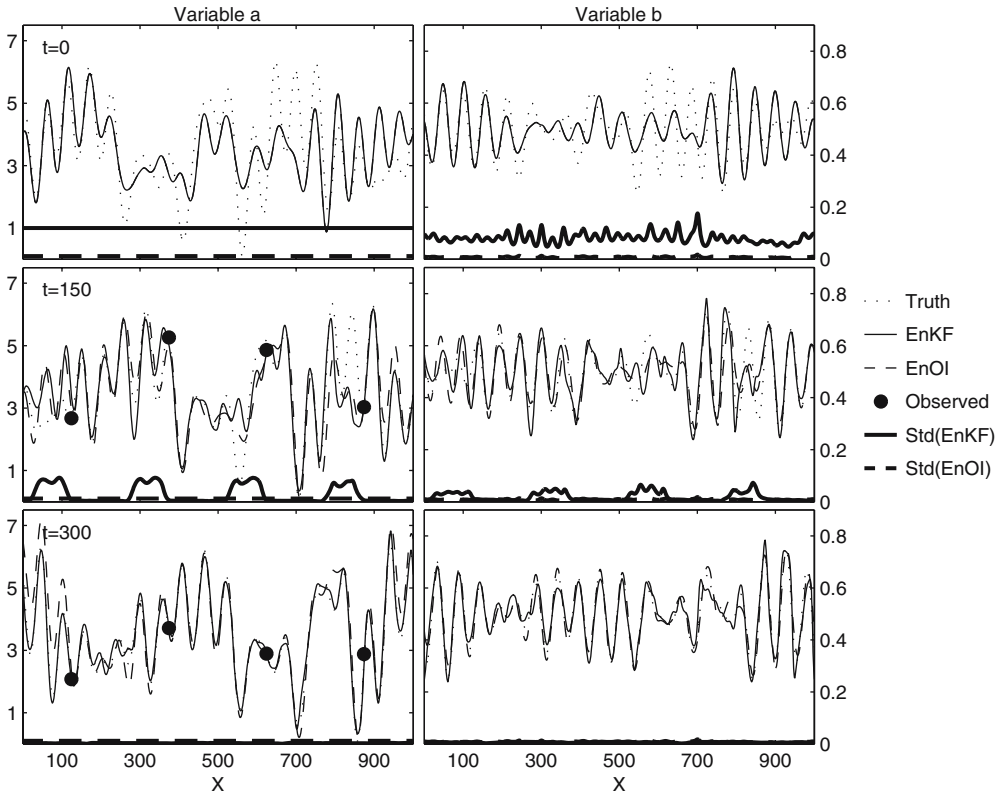

Fig. 3 As for Fig. 1 except using 20 ensemble members and localisation with a length-scale of 50. The standard deviation of the ensemble differs from Figs. 1 and 2 because a different ensemble size is used here

$u = 1$, and a separation between observation sites of 250, it takes 250 time steps for the fields to advect from one observation site to the next. We might therefore expect that it takes less than 250 time steps for the analysis from an optimal assimilation system to converge on the true field.

4.3.1 Ensemble size without localisation

Figure 4 presents RMSE and CC as functions of time for the EnKF and EnOI with 10, 20, 50, and 100 ensemble members. It clearly demonstrates that ensemble size is of first-order importance for ensemble data assimilation. Results range from failures, when only 10 ensemble members are used for EnOI, to examples where the RMSE decreases quickly to its asymptotic value and correlations increase quickly to near one over about 230 time steps.

Applications of both EnKF and EnOI with a 10- and 20-member ensemble are unable to substantially reduce RMSE from the initial level. This can be explained by the fact that the rank of the ensemble is smaller than that dimension of the model state space (i.e. 51), and the range of the ensemble does not change over time. Because of that, the true field cannot be represented as a linear combination of the ensemble fields, although some minor improvements in fitting the analyses to the true field can still be made. This would not necessarily be true for applications with a non-linear or imperfect model, when the sub-space spanned by the EnKF ensemble can change over time. The divergence shown in Fig. 4 for assimilation with EnOI can be eliminated by reducing the spread parameter α. It can also be eliminated or reduced by using observations at random positions rather than fixed positions.

Note that the divergence shown by statistically averaged performance criteria for an assimilation system does not necessarily mean that it is present in all or most of the 50 realisations; even one realisation with exponentially unstable analyses can result in a major deterioration in the average results.

The experiments with 10- and 20-member ensembles show that, in the case of a perfect linear model, some realisations diverge when a small ensemble is used without localisation. We also find that, with a 100-member ensemble, the EnKF system outperforms the EnOI system. With 50 members, the final states at $t = 500$ for EnOI and EnKF are almost the same, which occurs because of the premature collapse of the 50-member EnKF. Interestingly, the 100-member EnOI outperforms the 50-member EnKF. For this comparison, the 100-member EnOI involves the computation of one model integration and one analysis; by contrast, the 50-member EnKF involves the computation of 50 model integrations and 50 analyses. That is, in this case, EnOI is about 50 times less expensive than the EnKF if we neglect the differences in the computational cost of computing an analysis with a 50- or 100-member ensemble. In general, these results support the contention that EnOI may offer a reasonable alternative to the EnKF for some applications. For example, suppose for the application considered here that only a 10- or, even, a 50-member ensemble could be afforded for the EnKF, but a 100-member ensemble could be afforded for EnOI. For this case, EnOI offers a practical alternative to the EnKF.

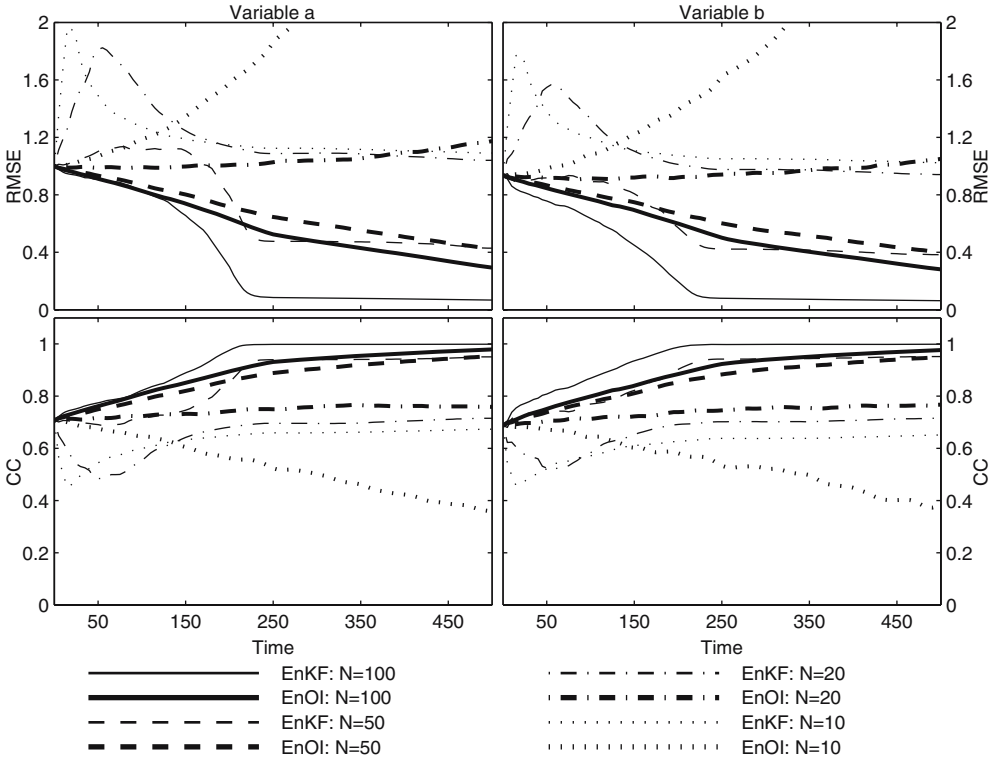

Fig. 4 Time series of the RMSE (*top*) and anomaly CC (*bottom*) between the true and analysed fields for variables *a* (*left*) and *b* (*right*) for EnKF and EnOI (*bold*) for a range of ensemble sizes

4.3.2 Ensemble size with localisation

We repeat the series of experiments presented in Fig. 4, but with localisation with a localising length-scale of 50 (Fig. 5). When localisation is used, all experiments successfully reproduce the true field, with low RMSE and correlations near 1.0. For both the EnKF and EnOI, the performance starts to deteriorate when a 10-member ensemble is used, but even in this case the systems with localisation outperform systems without localisation and an ensemble size of 50. Interestingly, Fig. 5 shows worse performance for variable *b* compared with variable *a* . This

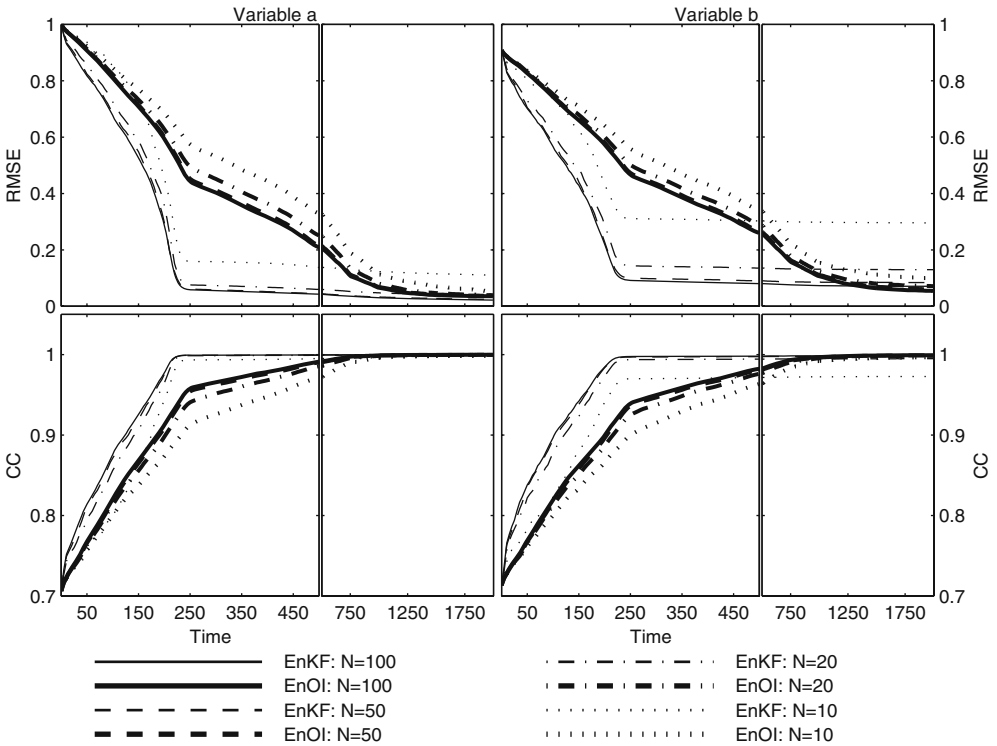

Fig. 5 As for Fig. 4, except using localisation with a length-scale of 50; and for results out until $t = 2000$ to show the asymptotic values of the EnOI experiments

may indicate a violation of the dynamic balance; we will investigate this issue in more detail later. Similar to the experiments without localisation, the performance of EnOI is inferior to that of the EnKF, in particular concerning the convergence rate. This indicates that the EnKF is more optimal than EnOI, as we expect. However, EnOI can achieve equivalent and sometimes smaller asymptotic values of the RMSE compared to the EnKF. For example, for both variables a and b, the 10-member EnOI has a smaller asymptotic RMSE value than the 10-member EnKF. For experiments with 20, 50 and 100 ensemble members the EnKF has smaller asymptotic RMSE values for variable a. By contrast, these experiments show that EnOI has smaller asymptotic RMSE values for variable b.

Based on the estimates in Section 2.4, we argue that, for the experiments with $N = 20$ and $L = 50$, for a domain of length 1,000 with $p = 4$ equally spaced observations, the effective rank of the ensemble is roughly $pN = 80$. We therefore expect this experiment to outperform the experiment with $N = 50$, with no localisation, and to perform marginally worse than the experiment with $N = 100$, with no localisation. The results in Figs. 4 and 5 support this contention.

4.3.3 Localisation length-scale

We perform a series of 50 experiments using a 20-member ensemble with a range of different length-scales for the localisation (Fig. 6). Recall that a 20-member ensemble can be considered small because the dimension of the model state space is 51. For a 20-member ensemble, the experiments that perform the worst are the ones that did not use localisation. For both variable a and variable b, the configuration that performs the best is the EnKF with localisation using an $e-$folding length-scale of 50. For EnOI, the analyses generally improve as the localising length-scale decreases.

As the localising length-scale reduces, the assimilation starts to approach data insertion. This increasingly results in producing analyses outside the model state space and violating constraints between the model variables (dynamical balance). This aspect of localisation is examined below.

Figure 6 shows the existence of an optimal localisation length, with the experiments using $L = 50$ and 100, outperforming those using $L = 25$ and ∞. If the localisation length-scale is too short, the model's dynamical balances may be violated and the effective range of the ensemble is increased; whereas if the localisation length-scale is too long, the effective range of the ensemble may be less than that of the model's sub-space.

4.4 Impact of localisation on dynamical balances

Variable b is defined as a derivative of variable a. To investigate violations of the dynamical balance in the data assimilation, one can calculate the discrepancy between the analyses of b and the values given by a centred difference approximation to Eq. 11, given an analysis of variable a; we will refer to the former as b^{assim} and the latter as $b^{balanced}$. To assess how localisation impacts the dynamical balance of the system, we calculate time series of the RMS difference between b^{assim} and $b^{balanced}$ for a number of different localising length-scales using both the EnKF and EnOI. Analysis of the results of these experiments presented in

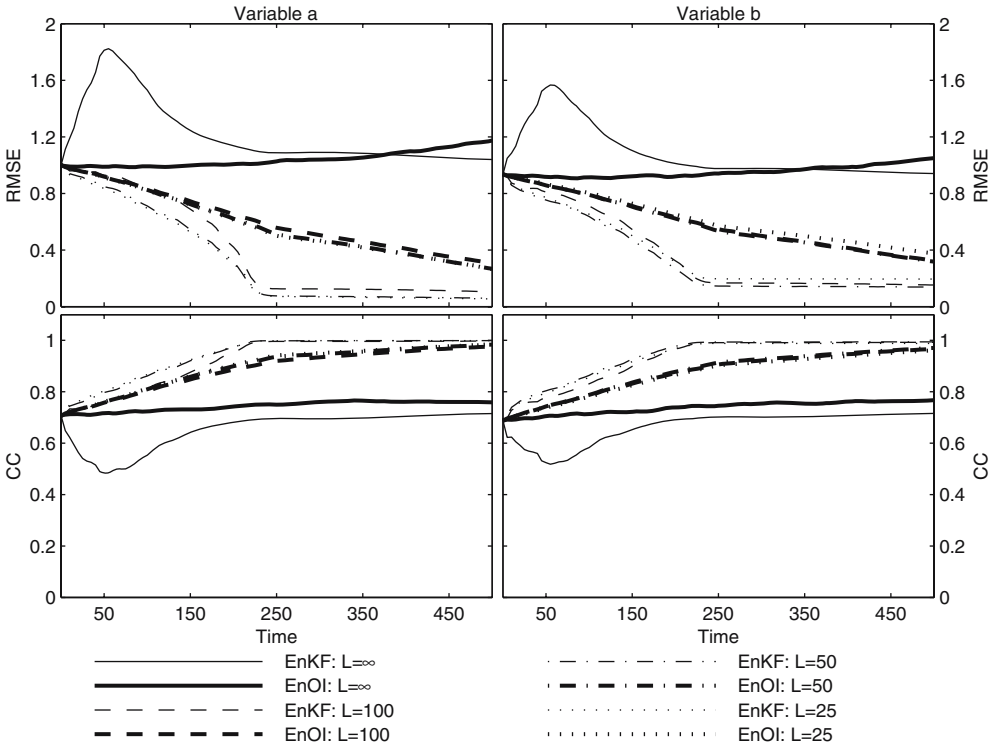

Fig. 6 As for Fig. 4, except using 20 ensemble members with a range of different localisation length-scales

Fig. 7, which demonstrates two important points. Firstly, the degree of imbalance increases with a decrease of localisation length-scale. Secondly, in the limit of no localisation, $L = \infty$, variables a and b are perfectly balanced, as we expect.

Another result in Fig. 7 is that the degree of model imbalance is greater using the EnKF than it is using EnOI. The reason for this is clear. In the EnKF, the dynamical balance of the ensemble is degraded at each assimilation step because of the localisation. By contrast, the ensemble for EnOI is stationary, and is therefore always dynamically balanced. Consequently, the degree of imbalance in the EnOI analysis is not as severe as for the EnKF. However, we note that, for more realistic applications, the dynamical imbalances may be restored by the model.

The impact of localisation on dynamical imbalance is further highlighted by contrasting the mean power spectral density (PSD) of variables a, b and $b^{balanced}$ for a series of 50 experiments using the EnKF with a 20-member ensemble and localisation with a length-scale of 100 (Fig. 8) and 50 (Fig. 9). We present the PSD of the true field and the ensemble mean at $t = 0$ and $t = 500$. For the experiments with $L = 100$ (Fig. 8), the PSD of the ensemble mean becomes well aligned with the PSD of the true field after 500 time steps for both variables a and b. Similarly, the PSD of b and $b^{balanced}$ are also in good agreement. We also find that the mean PSD of the ensemble perturbations of b and $b^{balanced}$ become well aligned by $t = 500$. This demonstrates that, for $L = 100$, the dynamical imbalance introduced by localisation is relatively insignificant. However, consideration of the mean PSD of the ensemble perturbations for the experiments when $L = 50$ (Fig. 9) renders quite a different story. We find that the agreement between b and $b^{balanced}$ becomes very poor, particularly for the higher frequencies. The ramifications of this for a linear model are not severe; however, for a non-linear model, such imbalance will likely cause serious initialisation problems that are likely to degrade future forecasts. This demonstrates that the length-scales used with an EnKF should be chosen with great care.

We also note that the power in the PSD of the mean in Figs. 8 and 9 is virtually zero for modes 26–50, for which the amplitude of the true field is zero. By contrast, the mean PSD of the ensemble perturbations shows considerable power in higher frequencies that are not spanned by the true field (mode 26–40) or the initial ensemble. The absence of any significant power at these frequencies in the ensemble mean indicates that the components of the ensemble perturbations at these frequencies must cancel out in the ensemble mean.

4.5 Analysis of the gain

Estimates of elements of the true gain matrix obtained for a single observation of a at an isolated location and variables a and b are shown in the top panels of Figs. 10 and 11, respectively. These estimates are calculated from the average of 50 independent realisations of the gain from 1,000-member ensembles. We compare this estimate of the true gain to estimates of the gain for a range of different ensemble sizes, both with and without localisation. We define the error of the gain to be the difference between the estimated true gain and the gain calculated from an individual ensemble. For both variables a and b, the RMSE of the gain increases as the ensemble size is reduced (second panels of Figs. 10 and 11). This is what we expect because sampling error is more significant for a smaller

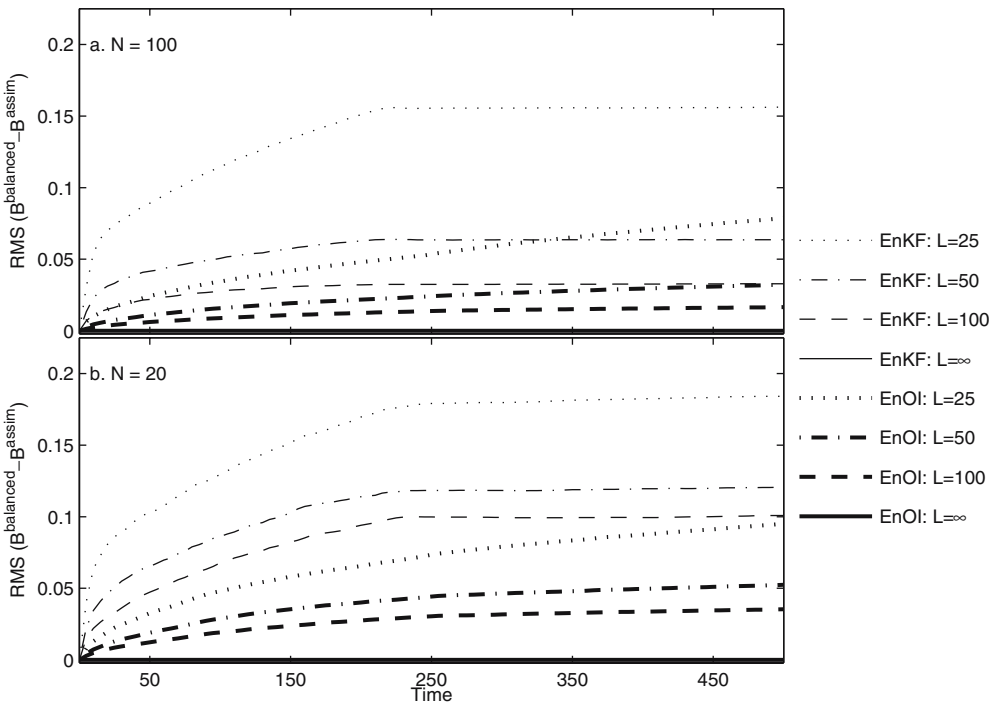

Fig. 7 Time evolution of the root-mean-squared difference between $b^{balanced}$ and b^{assim} using 100 (*top*) and 20 (*bottom*) ensemble members, where $b^{balanced}$ is determined using the analysed estimate of a and the balance constraint (Eq. 11); and b^{assim} is determined statistically using the analysis Eqs. 1–2. For comparison, the spatial RMS of b is typically 0.13, so an imbalance of 0.013 represents an imbalance of about 10%

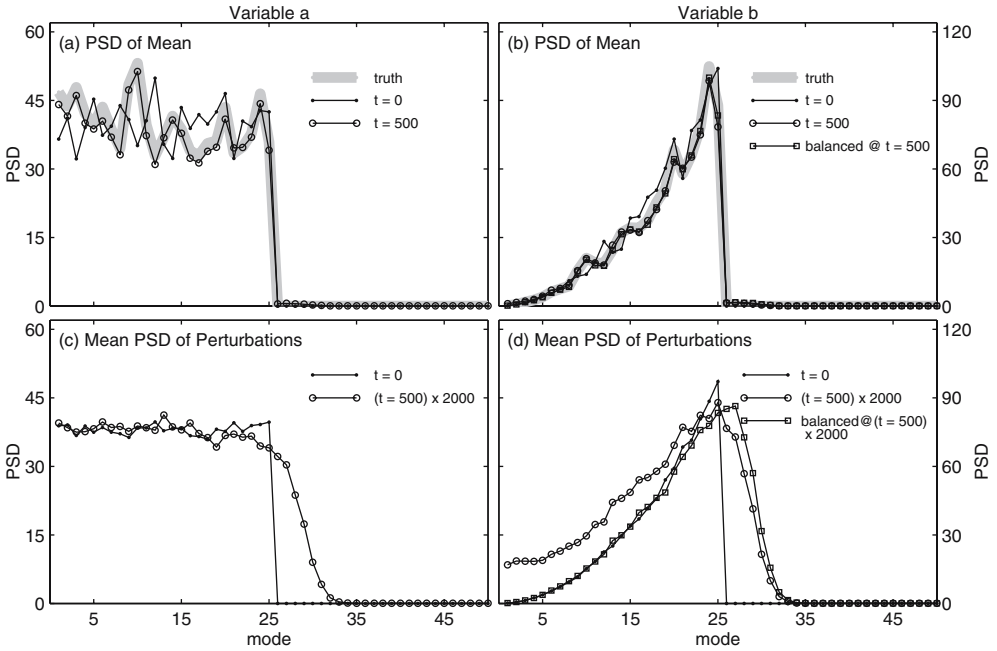

Fig. 8 PSD of the true field and the ensemble mean at the start ($t = 0$) and end ($t = 500$) for **a** variable a and **b** variable b for an EnKF using 20 ensemble members and localisation with a length-scale of 100, and the PSD of the corresponding ensemble perturbations at the start and end (multiplied by 2000) for **c** variable a and **d** variable b. The PSD of the balanced b is also shown in panels **b** and **d**. The PSD is averaged over 50 realisations to give statistically robust results

ensemble. For both variables, sampling error is smallest at the observation location and largest adjacent to the observation.

Finally, we consider the RMSE in the gain when a 20-member ensemble is used with a range of localising length-scales. When $L = 100$, the RMSE of the gain is almost unchanged from when localisation is not used. When $L = 50$, the RMSE is reduced for both variables a and b almost everywhere. There is a region at about $X - X^o = 50$–70, where localising with $L = 50$ increases the RMSE of the gain, but in general, the RMSE of the gain is significantly reduced, and particularly for variable a at $|X - X^o| > 100$. This reduction will clearly have a positive impact on analyses calculated using this gain. When $L = 25$, the RMSE of the gain is reduced almost everywhere; however, it is significantly increased at about $X - X^o = 50$–70 for both variables a and b. In this case, the localisation is too severe, and is acting to reduce the gain to near zero in regions where it should be non-zero. Recall that the decorrelation length-scales of a and b are about 20. It appears as though there is a window of length-scales over which localisation is beneficial. If the localisation length-scale is too small, then the gain can be degraded. If the localisation length-scale is too large, then the sampling error is not significantly influenced. For this application, a localising length-scale of 50 appears to be most appro-

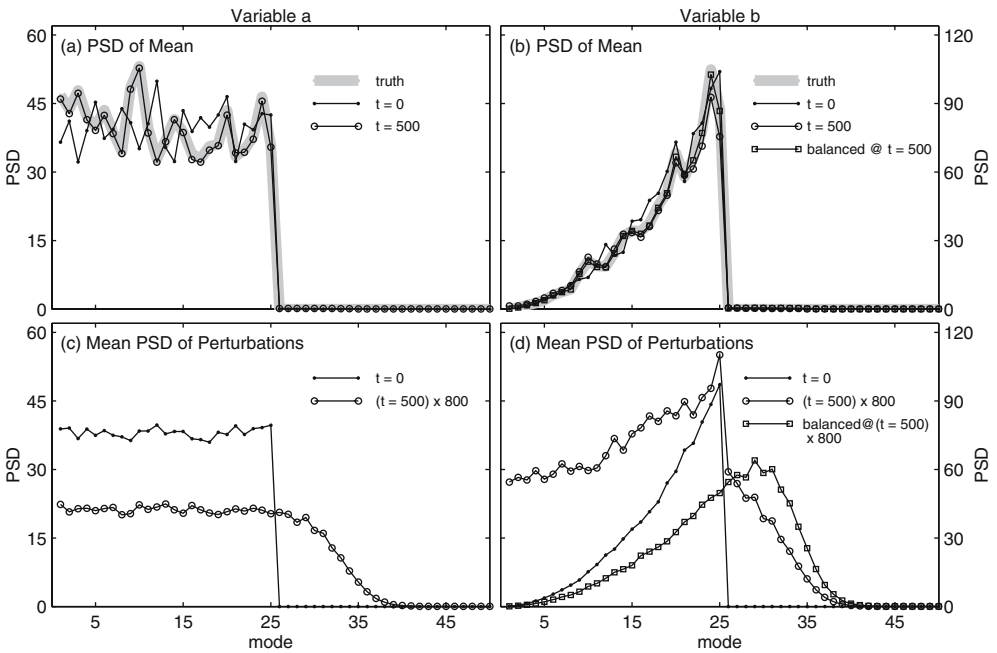

Fig. 9 As for Fig. 8 (EnKF, $N = 20$), except with a localising length-scale of $L = 50$

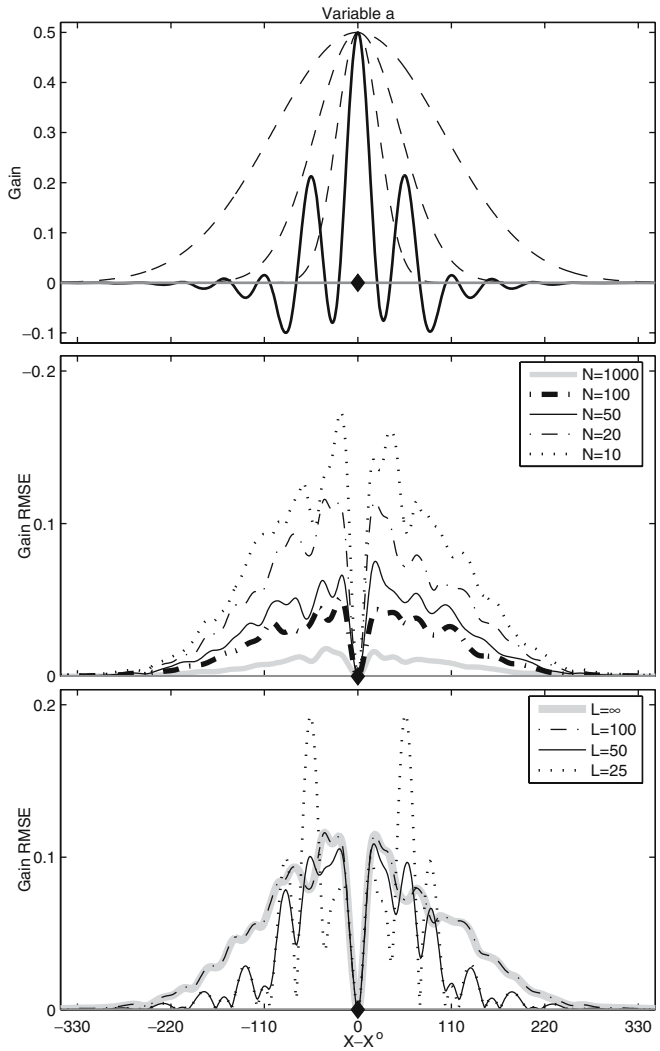

Fig. 10 The *top panel* shows elements of the true gain matrix, from Eq. 2, between an observation at the location denoted by the diamond, and the surrounding region for variable *a*; the *dashed lines* in the *top panel* show the localising correlation function, multiplied by a half, for $L = 25$, 50 and 100. The *middle panel* shows the RMSE of the gain with no localisation using a range of ensemble sizes. The *bottom panel* shows the RMSE of the gain for a 20-member ensemble with different localising length-scales

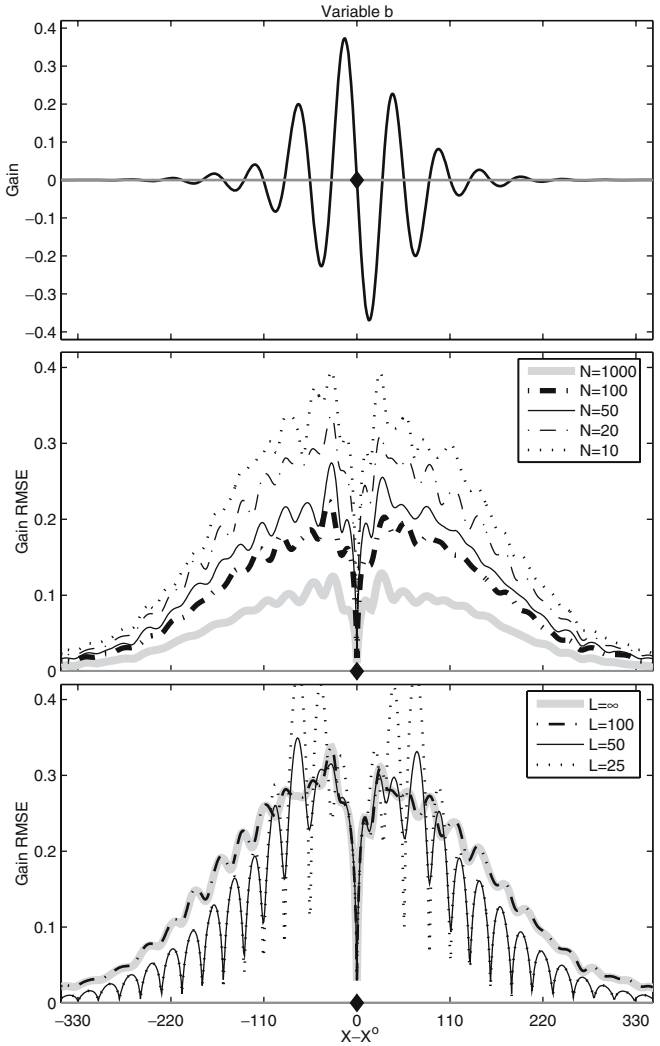

Fig. 11 As for Fig. 10, except for variable *b*

priate. This aspect of localisation clearly needs to be tuned for specific applications, but it seems obvious that the localising length-scale must be greater than the decorrelation length-scales of the fields being adjusted in a model.

4.6 Can localisation enrich an ensembles spectrum?

We compare the PSD of the mean and ensemble for a series of 50 experiments with an EnKF using 20-ensemble members and localisation with a length-scale of 100, except where the ensemble is initialised with zero amplitudes for a range of harmonics that are spanned by the true field (Fig. 12). Therefore, the initial ensemble does not span the full sub-space of the true field. We seek to determine whether localisation can enrich the spectrum of an ensemble by expanding the effective sub-space that is spanned by the ensemble. We find that by $t = 500$, the PSD of the ensemble mean is in close agreement with the true field (Fig. 12a,b), despite the fact that the ensemble did not initially span sufficient space to represent all the modes of the true field (Fig. 12c,d). The mean PSD of the perturbations at $t = 500$ shows that the dimension of the final ensemble is greater than that of the initial ensemble (Fig. 12c,d) and that it spans roughly the same space as the true field. This demonstrates that localisation has the effect of enriching the spectra so that the effective sub-space of an ensemble is increased.

We repeat the analysis described above for a series of 50 experiments using EnOI with a 20-member ensemble and localisation with a length-scale of 100. Again, the ensemble is initialised with zero amplitudes for a range of harmonics that are spanned by the true field (Fig. 13). Obviously, the ensemble perturbations for EnOI remain unchanged throughout these experiments. However, again we see that one effect of localisation is to enrich the spectra, and the analysis gradually comes into agreement with the

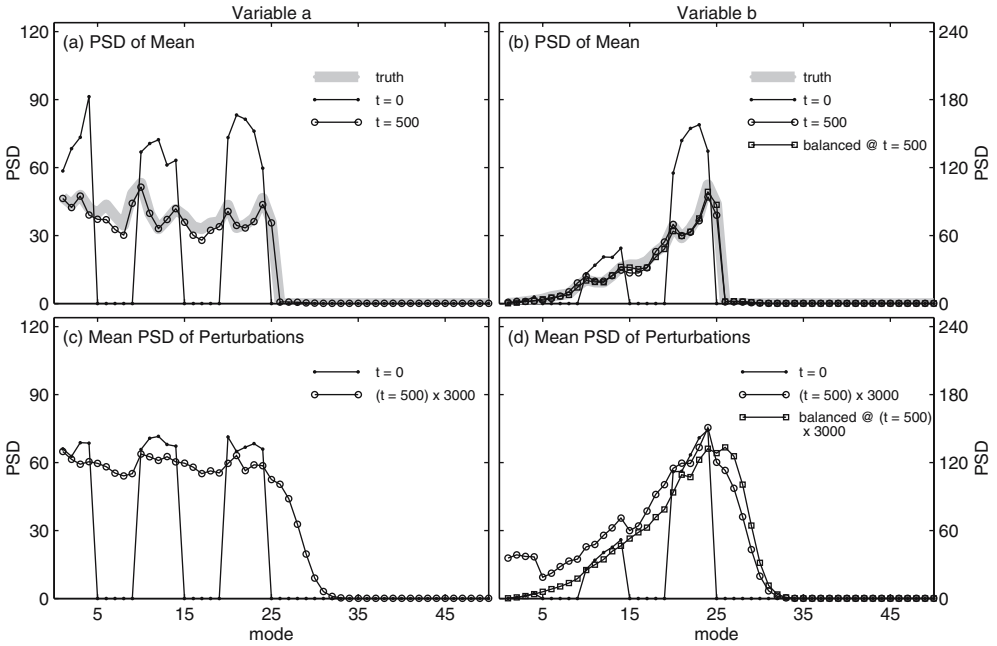

Fig. 12 As for Fig. 8 (EnKF, $N = 20$, $L = 100$), except for a case where the initial ensemble does not span the complete model sub-space (i.e. some modes are initialised with zero amplitude)

true field for all frequencies that are represented in the true field.

Therefore, this series of experiments demonstrates that localisation can effectively expand the sub-space spanned by an ensemble, even for EnOI, where the ensemble remains unchanged.

5 Summary and conclusions

We have demonstrated on a simple linear model that, while EnOI is less optimal than the EnKF, for some cases the performance of EnOI may be comparable (e.g. 50-member ensembles without localisation), or even superior, to that of the EnKF. For example, we find that without localisation, a 100-member EnOI outperforms a 50-member EnKF. For this case, EnOI is significantly less computationally expensive than EnKF. In practice, for a large-scale application, using comparable computational resources, one might be able to afford either EnOI with a large ensemble or EnKF with a small ensemble. Based on the results of this study, we suggest that in such cases EnOI may offer a reasonable alternative to the EnKF. An example of such an application where EnOI was successfully applied to an application for which an EnKF was not feasible is described by Oke et al. (2005). For both the EnKF and EnOI, when a small ensemble is used (e.g. $n = 20$), the estimated forecast error covariances may require localisation around each observation to achieve reasonable results. We find that the performance of either the EnKF or EnOI is sensitive to the length-scales used in the localisation. If the localisation is too strong, or too weak, the analyses can be degraded. This is consistent with the conclusions of Mitchell et al. (2002) and others. There are two reasons why localisation improves the performance: sampling error is reduced and the effective rank of the system is increased. This allows the ensemble to give a good fit to the forecast innovations, thus making the system more optimal.

The effect of localisation in reducing the influence of noise in the covariance matrix at large distances caused by the limited ensemble size has been underlined in previous studies by Hamill et al. (2001) and Houtekamer and Mitchell (2001). In our opinion, the effect of the increase in the effective rank of the ensemble is more important for the performance of the EnKF. Figure 14 shows the best possible RMSE with a given ensemble and true field as a function of ensemble size, averaged over 50 realisations.

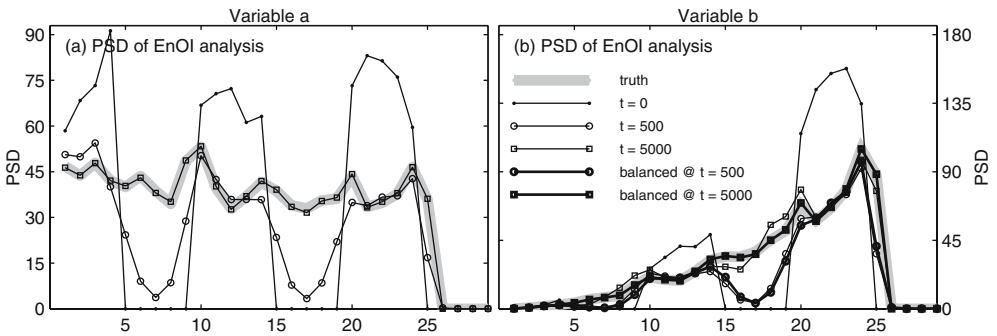

Fig. 13 As for Fig. 12a,b, except for a series of experiments using EnOI ($N = 20$, $L = 100$)

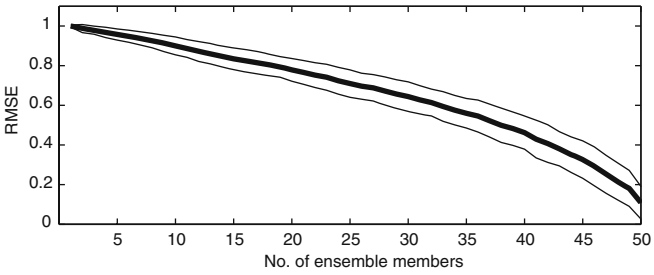

Fig. 14 The best achievable RMSE for a given ensemble and true field realisation, averaged over 50 random realisations. The *thick line* shows the average, *thin lines* show the average plus/minus one standard deviation

For a given ensemble A and true field x_t, the best possible RMSE σ_{min} may be calculated as the residual of the solution of the system

$$As = x_t$$

in the least squares sense:

$$s = (A^T A)^{-1} A^T x_t, \sigma_{min} = ||As - x_t||. \quad (12)$$

This value remains quite large even for relatively large ensembles, with ensemble size close to the dimension of the model state space of 51. For example, the expected best possible RMSE for an ensemble of 40 members (calculated by Eq. 12 and averaged over 50 realisations) is approximately 0.46. This means that even with an optimal system, one cannot achieve successful data assimilation with an ensemble of insufficient rank without localisation. In contrast, with localisation, it is possible to obtain an RMSE that is close to zero with an ensemble with as few as 20 or 10 members, or even fewer. The effective size of the model state space depends on a particular application, but one can perhaps expect dimensions of the order of 10^3 –10^4 or greater in high-resolution operational oceanographic and meteorological models. This means that, in practical ensemble-based assimilation in oceanography or meteorology, localisation may be considered a necessary attribute of a data assimilation system.

We find that localisation degrades the dynamical balances of model fields, and as the localisation becomes more severe, the imbalance in these fields increases. We find that the length scales of the localisation should be larger than the decorrelation length-scales of the variables being updated. The problem of dynamical balance is more severe for the EnKF, where dynamical imbalance can accumulate over time. However, we note that this effect is likely to be more serious in the linear model that which we used for our experiments. We expect that this will not be as serious for applications to a non-linear model, which can gradually restore the underlying dynamical balances.

Acknowledgements This research is funded by Australia's Commonwealth Scientific & Industrial Research Organisation through appropriation funding and by the US Office of Naval Research Ocean Modelling Program through grant N000140410345. The authors thank the anonymous reviewers for constructive comments that led to substantial improvements in this manuscript.

References

Anderson JL, Anderson SL (1999) A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Mon Weather Rev 127:2741–2758

Evensen G (2003) The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn 53:343–367

Evensen G (2004) Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn 54:539–560

Hamill TM, Whitaker JS, Snyder C (2001) Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. Mon Weather Rev 129:2776–2790

Houtekamer PL (1995) The construction of optimal perturbations. Mon Weather Rev 123:2888–2898

Houtekamer PL, Mitchell HL (1998) Data assimilation using an ensemble Kalman filter technique. Mon Weather Rev 126:796–811

Houtekamer PL, Mitchell HL (2001) A sequential ensemble Kalman filter for atmospheric data assimilation. Mon Weather Rev 129:123–137

Keppenne CL (2000) Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter. Mon Weather Rev 128:1971–1981

Keppenne CL, Rienecker M (2002) Initial testing of a massively parallel ensemble Kalman filter with the poseidon isopycnal ocean general circulation model. Mon Weather Rev 130:2951–2965

Mitchell HL, Houtekamer PL, Pellerin G (2002) Ensemble size, balance, and model-error representation in an ensemble Kalman filter. Mon Weather Rev 130:2791–2808

Nerger L, Hiller W, Schroter J (2005) A comparison of error subspace Kalman filters. Tellus 57A:715–735

Oke PR, Allen JS, Miller RN, Egbert GD, Kosro PM (2002) Assimilation of surface velocity data into a primitive equation coastal ocean model. J Geophys Res 107, DOI 10.1029/200JC000511

Oke PR, Schiller A, Griffin DA, Brassington GB (2005) Ensemble data assimilation for an eddy-resolving ocean model of the Australian region. Q J R Meteorol Soc 131:3301–3311

Ott E, Hunt BR, Szunyogh I, Zimin AV, Kostelich EJ, Corazza M, Kalnay E, Patil DJ, Yorke JA (2004) A local ensemble Kalman filter for atmospheric data assimilation. Tellus 56A:415–428

Whitaker JS, Hamill TM (2002) Ensemble data assimilation without perturbed observations. Mon Weather Rev 130:1913–1924