Comparison of the Ensemble Kalman Filter and the Determinstic Ensemble Kalman Filter for a QG Model

By: Kieran Bhatia

Predictability

5/10/2012

I. Introduction

Due to imperfect dynamical models, sparse observational networks, and errors in observations, data assimilation has emerged as a vital field in the atmospheric sciences. The process of accurately determining the initial conditions of a model based on a background state and inadequate observations is a critical part of all forecasts. When Evensen (1994) updated the traditional Kalman filter (KF) with the Ensemble Kalman Filter (EnKF), it emerged as one of the premiere data assimilation schemes. As a Monte Carlo approximation to the traditional KF, the EnKF utilizes a collection of forecast runs to determine the background error covariance. The error covariance matrix P (superscript can be added depending on analysis or forecast) is computed with the equation

$$P=1/(m-1) * \sum [i=1 \text{ to } m] \ (X_i-x)(X_i-x)^T \quad (1)$$

where X_i represents a model run and x is the ensemble mean,

$$x =1/m \ \sum [i=1 \text{ to } m] \ X_i \ . \quad (2)$$

Like the traditional KF, the EnKF is based on the analysis equation

$$x^a = x^f + K(d-Hx^f) \quad \text{where } K= P^fH^T(HP^fH^T + R)^{-1} \quad (3)$$

where K is the Kalman gain, x^a is the analysis, x^f is the forecast, d is the vector of observations, H is the matrix that maps the first guess vector onto the observations, P^f is the background error covariance matrix, R is the observation covariance matrix, and the superscripts f and a represent the forecast and analysis respectively (Sakov and Oke 2008, hereafter SO2008).

The EnKF was such a significant improvement over previous data assimialtion techniques that used background error covariances based on mainly climatology because the EnKF background error covariance is allowed to change based on the current atmospheric pattern and the geography of a data point. As a result, updates to the first guess vector are more grounded dynamically. However, Houtemaker and Mitchell (1998) discovered that the EnKF, due to small ensemble sizes in operations, was often underdispersive and added random perturbations (added noise to d in equation 3) to the observations. As a result, the term EnKF is now currently associated with an EnKF involving perturbed observations (SO2008).

Although adding noise to observations helped with the EnKF's inherent underestimation of analysis error, there are problems with this technique (Tippett et al. 2003). The addition of noise to observations lowers the accuracy of the analysis error covariance estimate by increasing the sampling error and thereby raising the likelihood that the analysis error covariance will be underestimated by the ensemble (Whitaker and Hamill 2002). Deterministic methods, such as the Ensemble Square Root Filters

(ESRF), were proposed to avoid the sampling issues associated with the use of perturbed observations (Tippett et al. 2003). There are a number of different ESRF algorithms in the literature, and it is not clear if any of these techniques are superior for every situation. One such ESRF is proposed by SO2008 and was named the deterministic EnKF (DEnKF). The DEnKF uses the expression for the analyzed error covariance without perturbations

$$P^a = P^f - P^f H^T K^T - KHP^f + KHP^f + KHP^f H^T K^T \quad (4)$$

and the fact that $P^f H^T K^T = KHP^f$ (5)

to get $P^a = P^f - 2KHP^f + KHP^f H^T K^T$. (6)

Then, Sakov and Oke assume that the KH product is sufficiently small so the second term is much greater than the third term (considered negligible). Finally, with this assumption, the theoretical covariance of the traditional Kalman filter can be matched by halving the Kalman gain K. The analyzed error covariance then becomes equal to

$$P^a = (I-KH)P^f + 1/4 * KHP^f H^T K^T. \quad (7)$$

When $HP^f H^T << R$, the EnSRF analyzed covariance proposed by Whitaker and Hamill (2002) becomes equivalent to (7). Therefore, when the analysis corrections are small (observations have more error than the background), the DEnKF can be considered a linear approximation to the EnSRF. After formulating the update equations for the DEnKF, SO2008 tested the DEnKF by applying it to three different dynamical models and comparing its performance with the EnKF and serial EnSRF. Evensen's (2004) Linear Advection model, Lorenz and Emanuel's nonlinear Lorenz-40 model (1998), and a reduced-gravity quasi-geostrophic (QG) model with double-gyre wind forcing and bi-harmonic friction were the three small models examined.

The experiments conducted in this study aim to quantify the skill of the DEnKF and EnKF when applied to the QG model (EnSRF omitted due to issues with Sakov's EnKF Matlab package) by duplicating figure 7 of SO2008. The QG model was selected because it is the most similar to an actual Global Circulation Model: its dimension (127 X 127) is significantly larger than the other two models in SO2008. The exact methodology of the numerical tests conducted is introduced in section 2. In section 3, a portion of the localization radius-inflation factor subspace in figure 7 of SO2008 is recreated as well as this same parameter space for a larger ensemble. In section 4, a justification for the differences between the presented results and SO2008 results are discussed. The discrepancy between the effects of increasing ensemble size on the skill of EnKF and DEnKF is examined in section 5, followed by the conclusions in section 6.

II. Methodology

The particular design of the QG model used for this study is detailed carefully in section 4.3 of SO2008. In order to accurately reproduce figure 7 of SO2008, the data assimilation parameters required to produce the series of experiments depicted in figure 7 of SO2008 are followed exactly. The data assimilation is conducted as follows: data assimilation occurs at every fourth time step; 300 observations of Ψ with observation variance of 4.0 are inputted at each step; the initial ensemble is formed by 25 random samples of 2000 fields; ensemble inflation factors range from 1.0 to 1.18; a Gaussian localization function is used with localization radii varying from 5 to 45. Since the ensemble size is significantly smaller than the dimension of the model's sub-space, both the EnKF and DEnKF require localization of the background covariance matrix to lower the probability of filter divergence. For a small ensemble, spurious correlations between distant grid points are possible and localization reduces the impact of remote grid points on the analysis (Oke and Sakov 2006). Also, inflation is necessary to help with the underestimation of the background error covariance. Covariance inflation amplifies the ensemble anomalies at the end of each assimilation step by some factor, and consequently increases the background error covariance (SO2008). Figure 1 is included as an example of a typical analysis and true field for the QG model at the final assimilation step. SO2008 selected parameters for the data assimilation systems with the intention of producing a variety of non-divergent runs for a range of inflation factors and localization radii.

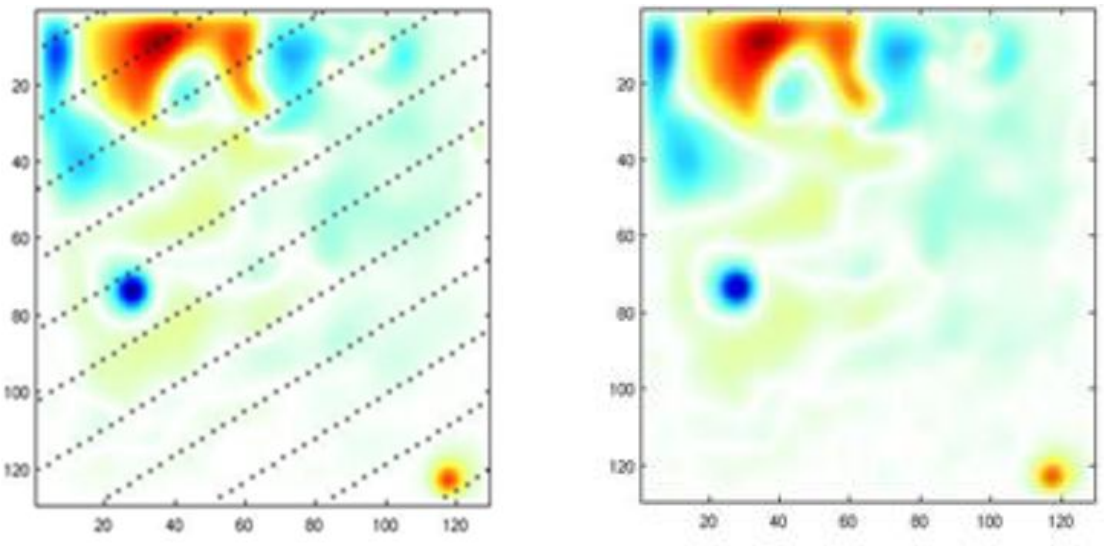

Figure 1. An example of a final state for an individual QG model experiment run to reproduce figure 7 of SO2008. The dots represent the observation locations for the particular run. True Ψ field is on the left, and the analysis Ψ field is on the right.

Figure 7 of SO2008 shows that the QG model is run for localization radii increments of 5 and inflation factor increments of 0.02. The area of localization radii-inflation factor subspace investigated here includes localization radii between 15 and 25 and inflation factors between 1.04 and 1.08 (nine bins). For each combination of the ensemble inflation and localization radius, 10 runs of 1200 steps each are carried out.

To quantify the performance of the DEnKF and DEnKF, the root mean squared error (RMSE) of the analysis of the streamfunction Ψ is computed. The RMSE values depicted in figure 7 of SO2008 represent an average over 10 realizations, for assimilation cycles 52 to 301 of each realization. As a reference, figure 7 is included with the localization radii and inflation factors reproduced in this investigation demarcated by a red box (Figure 2).

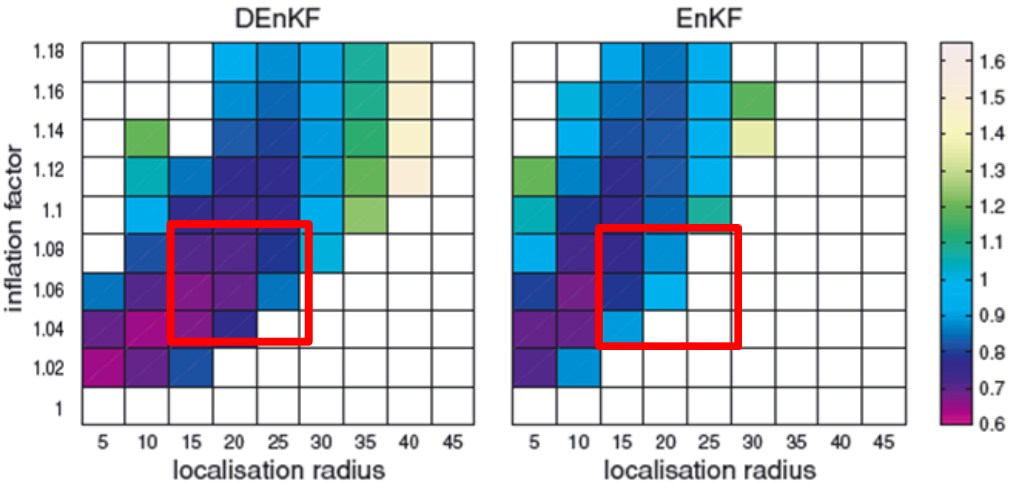

Figure 2. Figure 7 from SO2008 with a red box surrounding the localization radii and inflation factor combinations reproduced in this study. The shading refers to the size of the RMSE of the analysis of Ψ. White cells indicate experiments where at least one realization out of ten diverged or became unstable.

Figure 2 demonstrates that the DEnKF outperforms the EnKF considerably for the chosen data assimilation parameters. Not only are there fewer localization radii-inflation factor combinations that converge for the EnKF but the converging configurations have higher RMSE. The boxed area in Figure 2 appears qualitatively to be one of the regions of inflation factors and localization radii with the largest discrepancy in performance between the two data assimilation methods. The inferior results of the EnKF in this region of localization radii-inflation factor subspace is largely attributable to EnKF's higher sensitivity to the optimality of the parameters for the given ensemble size. Both data assimilations methods have their lowest RMSE values for moderate inflation factors with low localization radii. The location of this best performing region in the given parameter space indicates that the configuration of the data assimilation is certainly not optimal (background error covariance requires significant adjustments). In theory, the EnKF requires a larger inflation than any ESRF in these situations due to the small ensemble size's substantial impact on the EnKF's background error covariance (Whitaker and Hamill 2002).

Figure 2 agrees well with figures 3 and 4 from Whitaker and Hamill (2002), which shows a much smaller parameter space where the filter diverges for the EnSRF method compared to the EnKF. However, it is important to note that the RMSE values displayed in Figure 2 are not necessarily representative of the true performance of the two data assimilation systems. Each realization incorporates an initial ensemble that is formed by

25 random samples of 2000 fields. By only averaging over ten realizations, the RMSE in a square of Figure 2 might be quite different depending on which 10 realizations are averaged over. The results of section 3 highlight this important observation as well the impact of ensemble size on the performance of the EnKF and DEnKF.

III. Results

Results using the same QG model and the data assimilation parameters described in SO2008 to create Figure 2 are shown in Figure 3. The 3 X 3 parameter subspace depicted is the boxed region in Figure 2. The RMSE values are clearly different than SO2008's figure 7 so different realizations of the initial ensemble were clearly used which is not surprising. One of the more notable differences between Figure 3 and the boxed region of Figure 2 is the presence of five white squares EnKF's red box in Figure 2 but only one white square in the EnKF subplot of Figure 3. Again, this result is expected because the small sample size when randomly selecting ten realizations will naturally lead to differing amounts of diverging model runs.

Still, the overall trends in Figures 2 and 3 remain largely identical. The RMSE magnitudes for both the EnKF and DEnKF in Figure 3 agree very well with the corresponding bins in Figure 2. The DEnKF significantly outperforms the EnKF in both Figure 2 and 3. There are no combinations of parameters in the boxed region of Figure 2 or the entirety of Figure 3 that results in smaller RMSE values for the EnKF in comparison to the corresponding parameter combinations for the DEnKF. Also, the optimal parameter combinations are very similar in the two figures; smaller localization radii and larger inflation factors lead to smaller mean RMSEs in both Figure 2 and Figure 3.

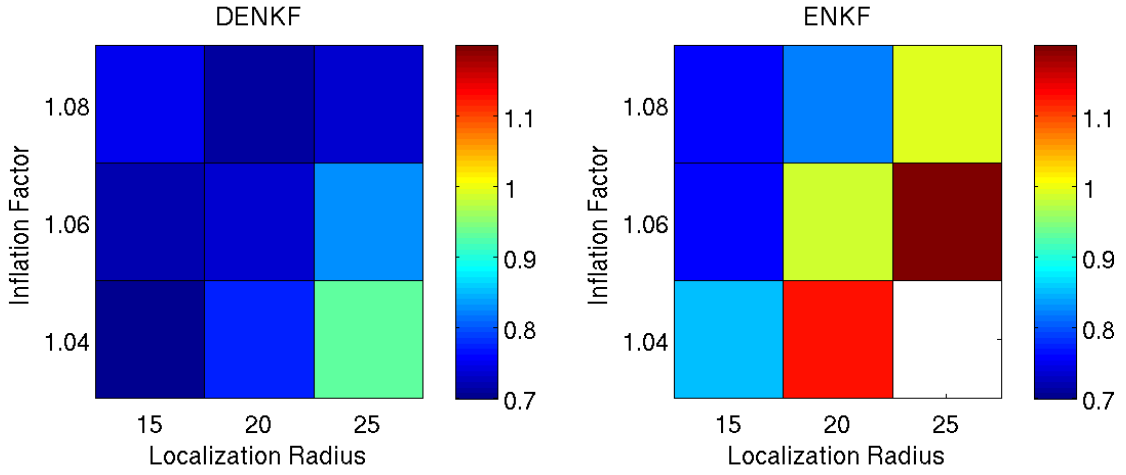

Figure 3. Comparison of the RMSE of the analysis of Ψ for the DEnKF and EnKF as a function of localization radius and inflation factor for the ensemble size of 25, averaged over 10 realizations and assimilation cycles 52-301 within each realization. White cells correspond to experiment where at least 1 realization out of 10 diverged. The color bar ranges from RMSE values of 0.7 to 1.1.

When evaluating Figure 3 and Figure 2, it is clear that with the data assimilation parameters and QG model employed in SO2008, the DEnKF is the better performing data assimilation system. However, one needs to consider how the DEnKF and EnKF improve when the data assimilation parameters becomes more optimal. In other words, what happens to each system's RMSE when improving the likelihood of a better analysis? There are a numbers of ways to theoretically achieve a better data assimilation system: lower the observational error, add more observations, increase the number of ensemble members, etc. In this study, the number of ensemble members was doubled to simulate the impact of bolstering the data assimilation configuration. Figure 4 applied the same data assimilation system to the SO2008 QG model as Figure 3 except for the increased amount of ensemble members.

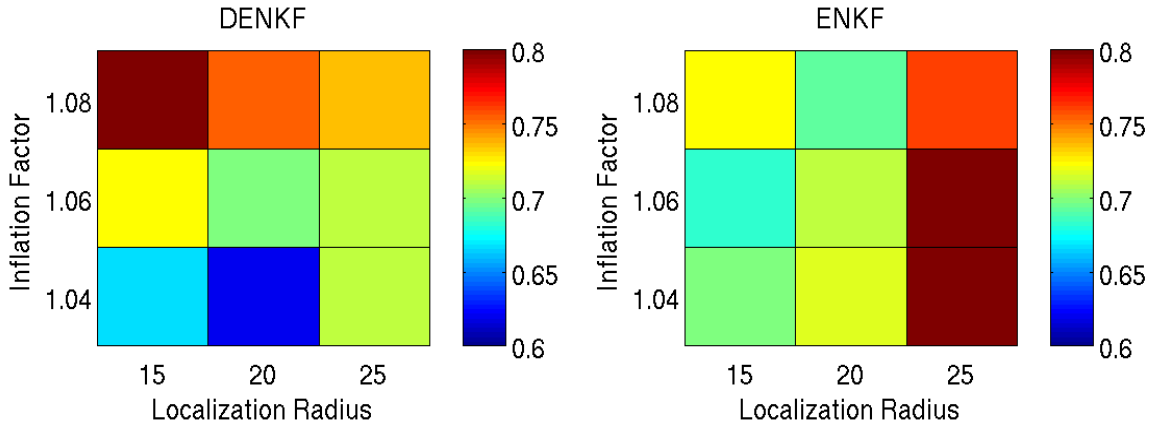

Figure 4. Comparison of the RMSE of the analysis of Ψ for the DEnKF and EnKF as a function of localization radius and inflation factor for the ensemble size of 50, averaged over 10 realizations and assimilation cycles 52-301 within each realization. The color bar ranges from RMSE values of 0.6 to 0.8.

Similar to Figure 3, Figure 4 includes the same localization radius and inflation factor ranges boxed in Figure 2 but there are several conspicuous differences between the two figures. The lower RMSE values for all bins are accentuated by the different color bar ranges in Figure 4 compared to Figure 3. Secondly, there are localization radii-inflation factor combinations where the EnKF RMSE is lower than the corresponding DEnKF bins. In other words, the EnKF is now performing better than the DEnKF for certain parameter configurations. The trends involving the localization radius and inflation factor are now significantly different for the two schemes. In general, the EnKF performs better for smaller localization radii and larger inflation factors but the gradient in RMSE across bins is not as dramatic as Figure 3. The DEnKF shows no noticeable patterns besides performing slightly better for lower inflation factors. However, smaller localization radii are not necessarily optimal for the DEnKF in this larger ensemble. Although the trends in the EnKF localization radii-inflation factor are predominantly the same for the different ensemble sizes, the DEnKF exhibits different trends between the two ensemble sizes. Based on Figure 4, it appears the DEnKF does not need inflation with this larger ensemble while EnKF still benefits considerably. The possible reasons for this conclusion will be discussed further in section 5. The results displayed in Figures

3 and 4 show only a small portion of the possible inflation factor-localization radius combinations illustrated in Figure 2 but the comparisons between the different figures provide several important conclusions about the merits of each data assimilation scheme.

IV. Discussion of the Discrepancies between Figures 2 and 3

In order to assess the performance of different data assimilation schemes on three relatively small models, SO2008 conducted a set of experiments where skill was quantified by the RMSE of the analysis Ψ. In this study, the performance of two data assimilation schemes, EnKF and DEnKF, is diagnosed for the QG model using the same experimental setup as SO2008. However, reproducing the SO2008 experiments exactly is impossible because the initial ensemble inputted into the two data assimilation systems is selected randomly from a large sample. Therefore, the RMSE values for each square in Figure 2 represent an average over ten numerical tests with randomly selected 25 initial ensemble members for each test (without knowledge of which ensemble members were used for each test, it is not possible to recreate Figure 2).

Additionally, the small number of samples creating each bin's average RMSE is responsible for the noticeably different results between Figures 2 and 3. In fact, the standard deviations of the ten realizations in each bin of Figure 3 suggest that if the experiment was reproduced again, different conclusions could be made. The same inference is appropriate for the bins in Figure 4.

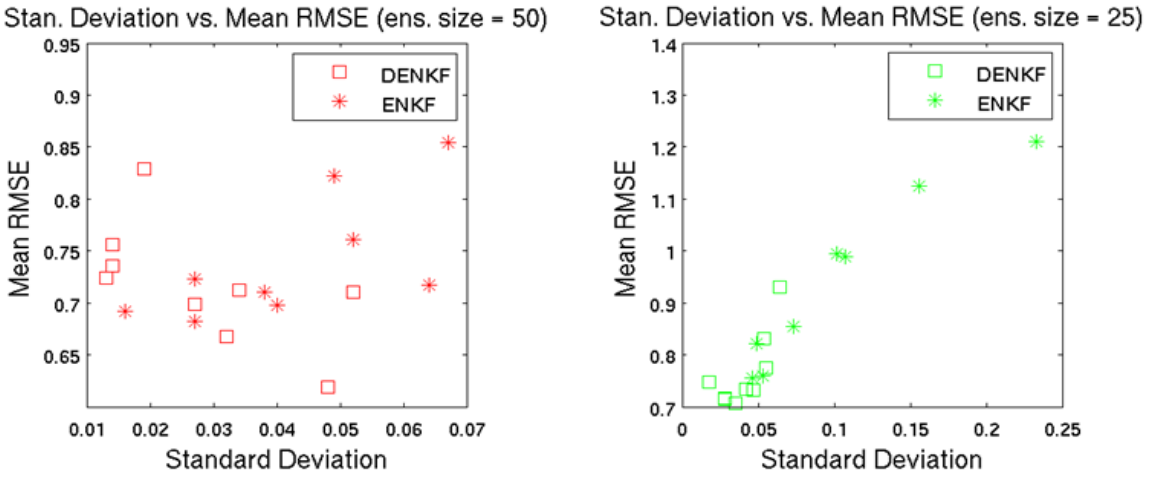

Figure 5. The mean RMSE of the ten realizations for each of the nine radius-inflation factor combinations is plotted against the standard deviation of the ten realizations. The DEnKF is denoted by squares and the EnKF is denoted by stars. The plot on the left represents the experiment conducted with an ensemble size of 50 while the plot on the right represents the experiment with the ensemble size of 25.

Figure 5 provides justification for the differences in mean RMSE between corresponding bins in Figures 2 and 3 as well as information on the performance of each data assimilation scheme. For each of the subplots in Figure 5, the standard deviation of the RMSE for each bin's ten realizations is on the x axis while the mean

RMSE of the ten realizations is on the y axis. Each of the nine squares represents one localization radius-inflation factor combination for the DEnKF; each star represents one localization radius-inflation factor combination for the EnKF. Data is included for both ensemble sizes.

The ratio of the standard deviation divided by the corresponding mean RMSE for each bin varies from approximately five percent to over twenty percent for the 25-member ensemble. For the 50-member ensemble, the standard deviations are much smaller in comparison to the RMSE mean, ranging from less than two percent to nine percent of the mean RMSE. These results show that realizations can vary significantly in individual bins and as a result, certain localization radius-inflation factor combinations will produce vastly different mean RMSEs depending on which initial ensemble members are selected for each realization. As a result, bins in Figure 2 should include many more realizations to lower the RMSE standard deviations of each bin and lead to more robust conclusions about data assimilation scheme performance and the effects of the size of the localization radius and inflation factor.

Figure 5 also demonstrates a relationship between the spread of the realizations' RMSE (standard deviation interpreted as proxy for spread) in a bin and the mean RMSE of the bin. For the 25-member ensemble, both data assimilation methods show a strong linear relationship between the standard deviation of the realizations' RMSE and the mean RMSE of the bin. As the variability between each realization's RMSE increases, the mean RMSE of the bin increases proportionally. This relationship can demonstrate the consistency of the data assimilation system; the system performs best when the ensemble-based estimate if the background error variance correlates well with the actual background error variance (SO2008). For the larger ensemble, there is much weaker correlation between mean RMSE and the RMSE standard deviation of a bin. EnKF still shows a positive correlation but DEnKF displays a weak negative correlation. This relationship between RMSE spread and mean RMSE further emphasizes the fact that a larger ensemble, or more optimal data assimilation parameterization, does not benefit the DEnKF in the same way as the EnKF.

V. Discussion of Ensemble Size Effects on DEnKF and EnKF

Comparing each bin's performance for the different ensembles sizes in the EnKF and DEnKF highlights the trends discussed in section 3. Figure 6 shows the difference in RMSE between each bin for the ensemble size of 25 and 50. This plot highlights the bins that display the greatest improvement when the ensemble members are increased. Figure 6 emphasizes the EnKF improves more than the DEnKF when the ensemble size is expanded. SO2008 provides a possible explanation for the lack of improvement in the DEnKF for the larger ensemble. Equation (7) demonstrates that DEnKF always overestimates the analyzed error covariance and the difference between the analysis of the error covariance of the DEnKF and the Kalman filter only differs considerably when the analysis correction is large (when KH assumption in introduction fails). For a larger ensemble, the background error covariance matrix formulated by the ensemble technique (part of both EnKF and DEnKF) is more realistic. As a result, the overestimation of the error covariance matrix by the DEnKF becomes more

unnecessary for the more optimal data assimilation configuration. The impact of the more accurate background error covariance is also visible in the EnKF results. The bins with the largest improvement in RMSE between the smaller and larger ensembles are the least optimal parameter combinations. Smaller adjustments from inflation and localization are required because the larger ensemble produces a more precise background error covariance matrix.

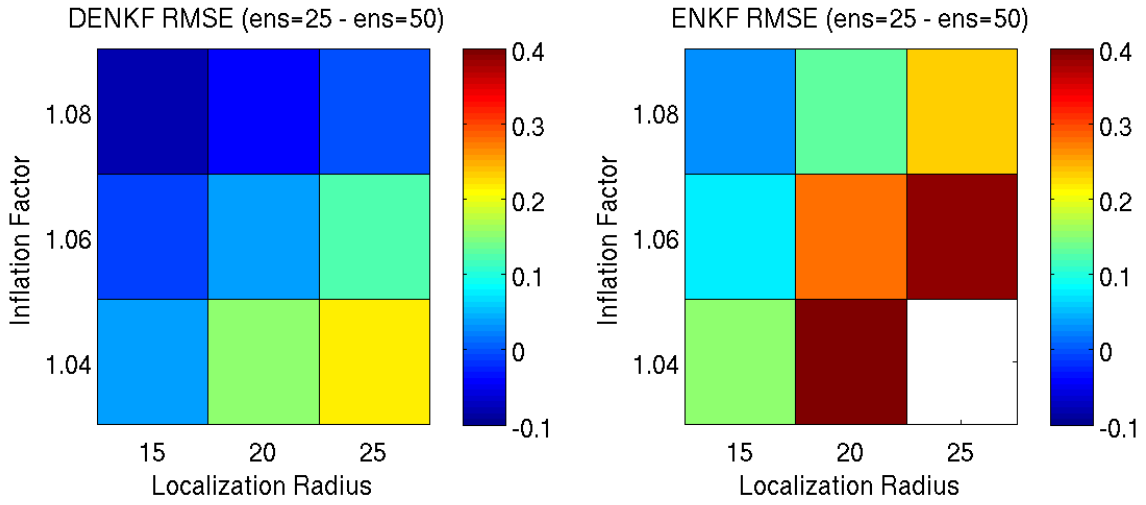

Figure 6. RMSE difference between 25-member ensemble experiment and 50-member ensemble experiment plotted against localization radius and inflation factor for DEnKF (left) and EnKF (right). White squares indicate one realization of either the 25-member experiment or 50-member experiment diverged.

VI. Conclusions

The primary goal of this study is to analyze and recreate the results presented in SO2008's figure 7. Additionally, the effects of upgrading the data assimilation schemes used in figure 7 by doubling the ensemble members are discussed. The QG model delineated in SO2008 provides a realistic, but small model to assess the performance of the EnKF and DEnKF data assimilation systems. However, to produce significant conclusions comparing the EnKF and DEnKF, SO2008 needed more thorough and comprehensive statistical analyses. The mean RMSE values of the bins in figure 7 of SO2008 were composed of 10 realizations. The typical standard deviations of the RMSE values in each bin indicate that more realizations are necessary to compare the performance of the EnKF and DEnKF and to even compare the different parameter configurations. Also, it is not mentioned how adjusting some of the QG equation's (eqn. 20 of SO2008) coefficients or data assimilation parameters would affect the conclusions drawn on the performance of the EnKF compared to the DEnKF. By doubling the ensemble size and repeating the experiments that produced figure 7, significant changes in EnKF and DEnKF performance are observed. If the data assimilation parameters become more optimal with another doubling of the ensemble size or lowering of the observation error, how will the DEnKF fair against other data

assimilation schemes? Future experimentation with the DEnKF is needed before considering it a legitimate data assimilation technique for operations.

References

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res., 99, 10 143–10 162.

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

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. Mon. Wea. Rev., 126, 796–811.

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

Oke, P. R. and P. Sakov, and S. P. Corney. Impacts of localization in the EnKF and EnOI: Experiments with a small model. Ocean Dynamics, 2006, 57, 32-45.

Sakov, P. and P. R. Oke. A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. Tellus, 2008, 60A, 361-371.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. Mon. Wea. Rev., 131, 1485–1490.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. Mon. Wea. Rev., 130, 1913–1924.