Australian Meteorological and Oceanographic Journal 59 (2010) 25-34

Analysis differences and error variance estimates from multi-centre analysis data

Mozheng Wei, Zoltan Toth and Yuejian Zhu
NOAA/NWS/NCEP, Environmental Modeling Center, Camp Springs, MD, USA

Analysis fields produced by modern data assimilation (DA) systems are considered to be the best estimate of the state of nature and are used as initial conditions for numerical weather prediction (NWP) models. However, all analysis data have errors which come from the errors in the background fields, observation data and the model itself, in addition to the errors generated by the DA techniques used. These analysis errors ultimately limit the prediction skill of NWP weather forecasts. One important development in extending NWP forecasting capability is the development and implementation of ensemble forecasting systems. To generate an efficient ensemble system with a limited number of ensemble members, one needs to construct the initial perturbations from the initial analysis error covariance. Thus, how to estimate the analysis error variance is an important and challenging issue in variational DA systems such as 3D/4D-Var.

This paper presents one of our efforts at estimating analysis error variance. In this method, we use analysis data-sets from several NWP centres. It can be shown that the squared centre mean (CM) analysis error with respect to the unknown truth is smaller than the mean squared error of all individual analysis fields from different centres. In general, the CM analysis is closer to the truth than the individual analysis, especially when the number of centres is large. Our results show that the long-term averaged differences and standard deviations between the individual analysis and the CM analysis indicate less uncertainty over regions with a large number of conventional observations, such as the North American and Eurasian regions. Larger uncertainties are found mostly over oceanic regions where conventional observations are sparse.

However, there are systematic errors or biases in analyses from different centres due to the differences in models, observation errors, methods of quality control and DA methodologies. These systematic errors do not necessarily represent the true analysis errors. We introduce a method that will remove the systematic errors from the different centre analyses before estimating the analysis error variance. It is found that the time-averaged differences between the different centre anomalies and the CM anomaly represent the uncertainties over different regions according to the observation densities over land and ocean after systematic errors are removed from the raw data. The spread over the average anomaly from the different centres represents the analysis error variance better. Our results demonstrate that this quantity could provide a more accurate estimate of the true analysis error variance that we are seeking.

Introduction

NWP forecast performance has made great progress during the past decade due to a few important factors. First, the numerical forecast models at major numerical weather prediction (NWP) centres have improved tremendously due to more accurate physics parametrisation schemes and increased computing power, which permits the use of higher resolution forecast models. Second, more observations, more accurate observing systems and improved data assimilation (DA) methods have been developed, such as 4D-Var (Rabier et al. 2000) and ensemble Kalman filters (Whitaker and Hamill 2002; Tippett et al. 2003; Whitaker et al. 2007). More accurate DA systems have played a key role in providing more accurate initial conditions for the NWP models, which have improved weather forecasts, particularly over the short and medium ranges.

Corresponding author address: Mozheng Wei, NOAA/NWS/NCEP, Environmental Modeling Center, 5200 Auth Road, Rm. 207, Camp Springs, MD 20746, USA.
Email: Mozheng.Wei@noaa.gov

On another front of weather forecasting, advances in the development and implementation of ensemble forecast systems at some major NWP centres (Toth and Kalnay 1993, 1997; Molteni et al. 1996; Houtekamer et al. 1996) have provided an opportunity to generate state-dependent estimates of forecast uncertainty. The forecast capability has been improved to a new level compared with the traditional deterministic single forecast. Centres that produce operational global ensemble weather prediction include the National Centers for Environmental Prediction (NCEP), the European Centre for Medium-Range Weather Forecasts (ECMWF), the Canadian Meteorological Center (CMC), the United Kingdom Meteorological Office (UKMO), the Fleet Numerical Meteorological and Oceanography Center (FNMOC), Meteo France, the Japanese Meteorological Agency (JMA), the China Meteorological Agency (CMA) and the Korea Meteorological Agency (KMA). Different ensemble systems and their performance have been evaluated and reviewed by, e.g., Hamill et al. (2000), Wei and Toth (2003), Buizza et al. (2005), Bowler (2006), Wei et al. (2006, 2008), Leutbecher and Palmer (2008), Park et al. (2008) and Bougeault et al. (2010). Detailed descriptions and evaluations are beyond the scope of this paper, so interested readers should consult these references.

In ensemble forecasting, a number of different numerical forecasts are generated to estimate the range of possible future states of a dynamical system. It has been widely accepted that the initial ensemble perturbations should sample the probability density function (PDF) that is our best knowledge about the initial state of the dynamical system. Thus, in the operational environment at an NWP centre, the analysis error covariance of the DA system that produces the initial analysis fields for the forecast is important in generating the initial perturbations. So far, the analysis error variance information has been used by only a few different ensemble methods in NWP centres to only a certain extent. A recent description and comparison on how analysis error variance information is being used in ensemble initial perturbation techniques are given in Tables 1 and 2 in Wei et al. (2008).

In the NCEP global ensemble forecast system, an ensemble transform with rescaling (ETR) has been used to generate the initial perturbations, as described in detail in Wei et al. (2005, 2008). In the ETR method, the initial perturbations depend on the accuracy of the analysis error variance information. This analysis error variance is provided in operations by the DA system. At NCEP, the operational DA system is the Gridpoint Statistical Interpolation (GSI) which is based on a three-dimensional variational analysis (3D-Var) (Derber et al. 1991; Parrish and Derber 1992; Wu et al. 2002). In the variational analysis system, the analysis is solved by minimising the cost function based on the background fields and observations and their respective error covariance matrices. The analysis error covariance matrix in 3D/4D-Var is determined by the background and observation error covariance matrices, and it cannot be computed directly due to its huge memory demand.

Fisher and Courtier (1995) proposed three approximate methods to estimate the analysis error variance in a 3D/4D-Var system. Among these, the Lanczos method was implemented for the ECMWF DA system. This method produces analysis error variance estimates by computing the leading singular vectors of the Hessian matrix. It takes advantage of the close link between the Lanczos method and the conjugate gradient method which was implemented in DA for the minimisation of the 4D-Var cost function. Since only a limited number of singular vectors can be estimated and included in the computation due to computing resource limits, a kind of calibration has to be performed to compensate for the missing trailing singular vectors. In general, the leading eigenvectors correspond to the most data-dense areas, so that the effects of aircraft observations over the US and Europe are well represented. However, this method is not very good at representing the reduction in variance due to satellite observations. This is because the reduction is spread over the globe, and would require many eigenvectors to be represented (Fisher 2007, personal communication). This method has been tested in the NCEP GSI and is described in a paper by two of the present authors (Wei et al. 2010).

In recent years, as high quality analysis data became available from more NWP centres, the analysis data from different centres have been exploited as another way of estimating the analysis error variance. Buizza et al. (2005) indicated that the spread of three centres' (NCEP, ECMWF and CMC) initial states could be considered as a crude lower-bound estimate of the analysis error variance. Swanson and Roebber (2008) studied the differences between the NCEP and ECMWF reanalysis data and showed that the reanalysis difference could be considered as a 'shadow' of the analysis error. They showed that the analysis difference contains certain aspects of the true flow-dependent analysis error and has a significant impact on the short-term forecast skill in downstream regions.

In another related recent work, Langland et al. (2008) studied the differences between the NCEP and FNMOC analyses from 1 January to 30 June 2007. They found that the differences and root mean of the squared daily differences in 500 hPa temperature are closely related to the distribution of radiosonde observations. The large differences between the two analyses are found to be associated with the regions having mostly satellite observations. Park et al. (2008) compared the ensemble performance from the THORPEX Interactive Grand Global Ensemble (TIGGE) data. They found that there are large variations between the different analyses from different centres. The performance scores of an ensemble depend on the verifying analysis used. They argued that the mean analysis from the different centres will probably be best as a reference analysis in comparing the performance of ensembles from each centre. The quality of an analysis could be estimated from the deviation between it and the centres' mean. Bowler et al. (2008) also argued that the mean of analyses from multiple centres is generally better than the analysis from any one centre.

In this study we will use the analysis data from NCEP, ECMWF, UKMO, CMC and FNMOC to estimate the analysis error variance information which will be used in generating the initial perturbations in the ETR based ensemble forecast system at NCEP. Firstly, we will study the differences between each centre's analysis and the mean of all centres. The standard deviations of each centre's analysis from the centre mean are also studied. Further, we will propose a new method to remove the systematic bias from the different centres and estimate a more sensible analysis error variance. We then provide descriptions and results of the analysis differences and variance among all the NWP centres. The analysis anomalies and error variance estimation are then introduced followed by discussion and conclusions.

Mean and variance of analyses from different centres

Suppose that $a_i(t)$ is the analysis data field at time t from any NWP centre $i = 1, 2, \ldots k$. k is the number of data centres such as NCEP, ECMWF, UKMO (or UKM for simplicity), CMC and FNMOC (or FNO). The centre mean (CM) analysis field at time t for all different NWP centres is simply

$$\bar{a}(t) = \frac{1}{k}\sum_{i=1}^{k} a_i(t) \qquad \ldots 1$$

Since the data assimilation (DA) systems and observations used are different among these centres, the analyses produced by these centres are bound to be different. Reasons for these differences include different DA methods (i.e. 3D or 4D-Var or EnKF), different forecast models that provide the backgrounds, different observation numbers and types with different procedures or methods for observation quality control and bias correction. Due to these differences the analyses produced by the NWP centres mentioned above could be considered as independent estimates of the nature which is not known. One would expect that the CM is better than any of the individual estimates. This can be seen from the mean squared error ($e^2(t)$) of all centres' analyses with respect to the true analysis, $q(t)$. It can be shown that if the assumption of independence is correct, then

$$e^2(t) = \frac{1}{k}\sum_{i=1}^{k}[a_i(t) - q(t)]^2 = \frac{1}{k}\sum_{i=1}^{k}[a_i(t) - \bar{a}(t) + \bar{a}(t) - q(t)]^2$$

$$= s^2(t) + [\bar{a}(t) - q(t)]^2 \qquad \ldots 2$$

where $s(t)$ is the spread of different analyses around the CM, i.e.

$$s(t) = \sqrt{\frac{1}{k}\sum_{i=1}^{k}[a_i(t) - \bar{a}(t)]^2} \qquad \ldots 3$$

and

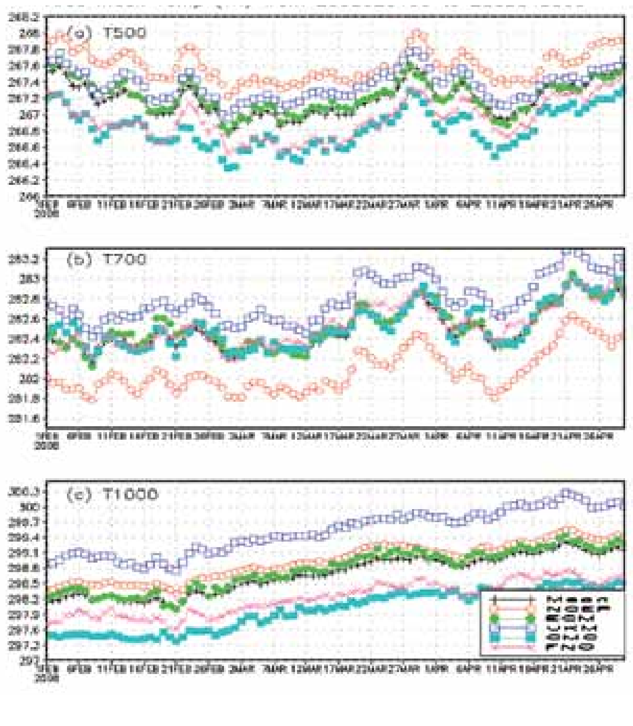
$$[\bar{a}(t) - q(t)]^2 = e^2(t) - s^2(t) \qquad \ldots 4$$

It is clear from Eqn 4 that the squared mean analysis error with respect to the true analysis $q(t)$ is generally smaller than the mean squared error of all individual analysis fields from the different centres. This is particularly true when the number of centres is very large. This conclusion is consistent with the principle of ensemble forecasting where the ensemble mean is generally better than any one of the perturbed ensemble members (Toth and Kalnay 1997). Based on this philosophy, Bowler et al. (2008) argued that the mean of analyses from multi-centres is generally better than the analysis from any one centre. Generally speaking, if different centres' analysis errors are of comparable magnitudes and relatively independent, then the CM will be closer to the truth than the individual analysis from each centre. In this case, the analysis error (AE) in a different centre's analysis can be estimated by the differences between that centre's analysis and the CM, i.e.

$$e_i(t) = a_i(t) - \bar{a}(t) \qquad \ldots 5$$

To see the analysis differences among the different centres, the area-averaged temperatures (T) of all five centres over the tropics (latitudes 20°S to 20°N) at three different pressure levels (500 hPa, 700 hPa and 1000 hPa) are shown as a function of time from 0000 UTC 1 February 2008 to 30 April 2008 in Figs 1(a), (b) and (c). The CM analysis of temperature is also shown in black at all three levels. There are large differences over the tropics at all three pressure levels. For example, the NCEP 500 hPa temperature (T500) analysis tends to be the highest with CMC the lowest. The difference between NCEP and CMC is about one degree. If the CM (indicated in black) is considered

Fig. 1 Area-averaged temperature analyses for all centres over the tropics as a function of time from 0000 UTC 1 February 2008 to 0000 UTC 30 April 2008 for three levels, (a) 500 hPa, (b) 700 hPa and (c) 1000 hPa. The centre mean analysis is shown in black.



to be the best estimate of truth, as we argued above, then the AE from the different centres can estimated from the distance between that centres analysis and the CM from Eqn 5. For T500, the ECM analysis is the closest to the CM with UKM second. The CMC and FNO analyses are colder, while NCEP and UKM are warmer than the CM. However, T at 700 hPa (Fig. 1(b)) shows that the NCEP analysis is coldest while UKM is warmest compared with the CM. Again, ECM is the closest to the CM and FNO is second. At 1000 hPa, the UKM and CMC analyses are the warmest and coldest respectively compared with the CM, while ECM and NCEP are closest to the CM. The temperature differences among the centres over the northern hemisphere (NH) and southern hemisphere (SH) are less than over the tropics (not shown).

Figures 2(a) to (f) show the CM of temperature and the differences between the CM and NCEP, ECM, UKM, CMC and FNO at the 500 hPa level, respectively. These values are averaged over a 90-day period from 0000 UTC on 1 February 2008

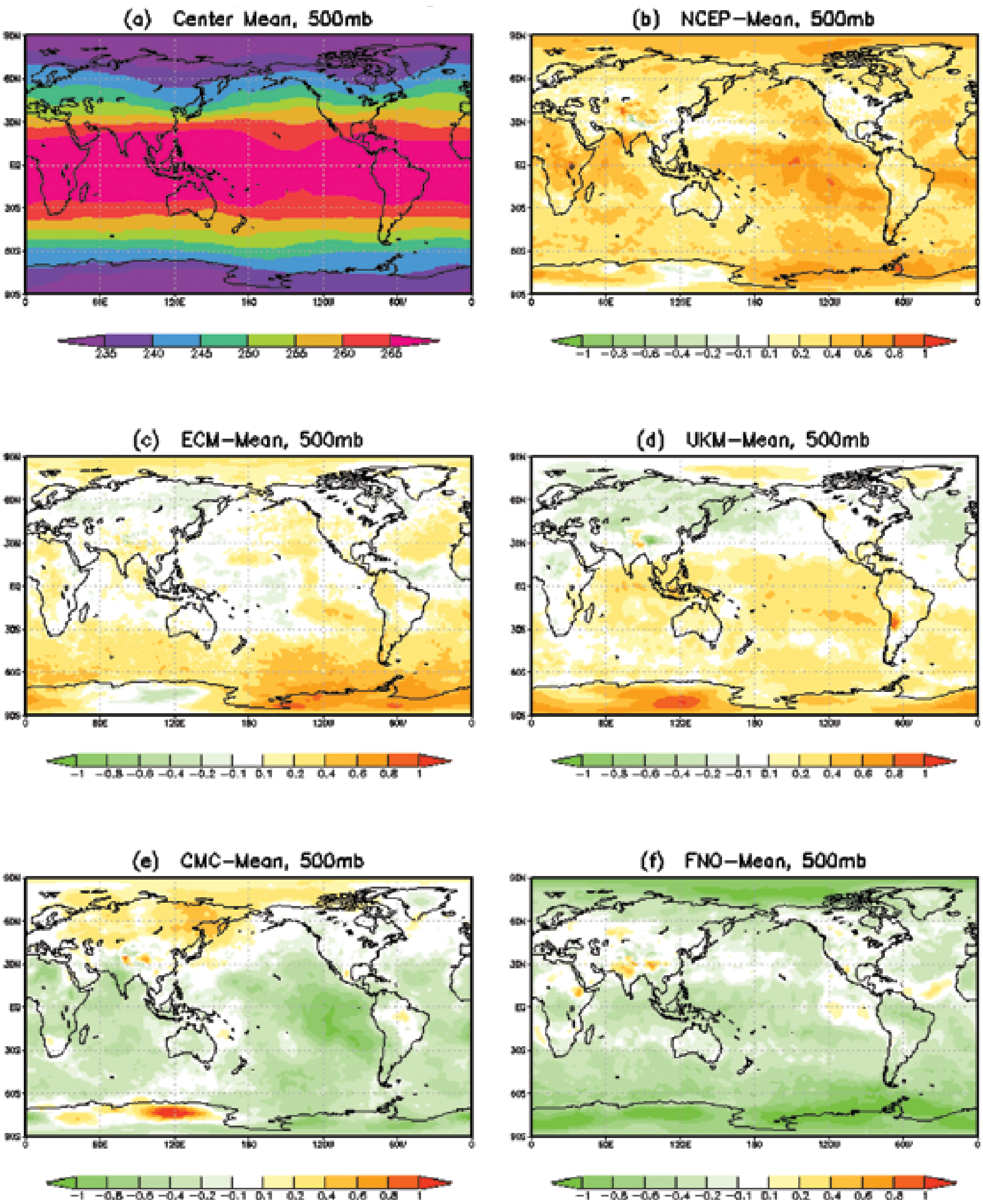


Fig. 2 Average values over the period from 0000 UTC 1 February 2008 to 0000 UTC 30 April 2008 of (a) CM temperature, and the temperature differences between CM and (b) NCEP, (c) ECMWF, (d) UKM, (e) CMC, (f) FNO.

to 0000 UTC on 30 April 2008. It is clear that all the analysis differences with the CM, which is supposed to be closer to the truth, are generally small in the traditionally data-dense regions such as North America, Eurasia and Australia. In the other regions with mostly satellite observation data, the differences are larger. This indicates that there are larger variations among these NWP centres in the areas with mostly satellite observation data. Handling satellite data is a more complicated issue than dealing with conventional data. Different centres may have different methods for bias correction, observation error specification and quality control. Langland et al. (2008) studied the differences between the NCEP and FNO analyses. They pointed out that the large differences between the two analyses are associated with the regions having mostly satellite observations. Swanson and Roebber (2008) looked at differences between the NCEP and ECMWF reanalysis data, and showed that these reanalysis differences are linked to the analysis errors which can impact forecast skill in the downstream regions.

The NCEP analysis temperature shown in Fig. 2(b) is slightly warmer than the CM in most regions except for North America, Eurasia and Australia. The ECM analysis (Fig. 2(c)) is slightly warmer in the southern parts of the SH. The UKM analysis is slightly cooler in the NH and warmer in the SH, while the CMC analysis is almost the opposite of the UKM analysis. Figure 2(f) shows that FNO temperature is generally cooler than the CM almost everywhere except for the traditionally data-dense regions.

Figure 3 is the same as Fig. 2, but is for 500 hPa geopotential height (Z500). In general, all centres show smaller deviations from the CM in the conventional data-dense regions. NCEP (Fig. 3(b)) shows lower values around the tropics and slightly larger values in the NH and SH, while the UKM analysis is larger in the tropics and smaller in the NH. The ECM and UKM analyses show mixed positive and negative differences with the CM in different regions. Again, the FNO analysis (Fig. 3(f)) is smaller than the CM in most regions.

If the CM is the closest to the truth, as in the case where the number of centres is large, i.e., $\bar{a}(t) \approx q(t)$, Eqn 4 shows that

$$s(t) \approx e(t) \qquad \text{...6}$$

This means that the spread around the CM (SCM) is equivalent to the root mean squared differences with respect to the truth, $q(t)$, which is the standard deviation of different centres' analyses with respect to the truth. Therefore, at a particular time t, the standard derivation of the analysis error of these centres can be estimated by computing the SCM $s(t)$, as defined in Eqn 3. The time-averaged SCM over the same 90-day period for temperature at 500 hPa (T500), wind component in the west-east direction at 500 hPa (U500) , 500 hPa geopotential height (Z500), and relative humidity (RH) at 850 hPa (RH850) are shown in Figs 4(a) to (d). The SCM for T500 (Fig. 4(a)) is clearly smaller over the conventional data-dense regions and larger over the oceans, particularly near the South Pole, indicating the analyses from different centres are very close over these data-rich regions but that there are large variations near the Pole.

Figure 4(b) shows that the SCM for U500 is relatively larger near the tropics over the oceans. The largest spread is located in a small area near Mount Everest in south Asia. For Z500 in Fig. 4(c), the largest variations are mostly located to the south of 60°S. The variance of the RH analysis at 850 hPa shows the largest spatial variation among the different centres' analyses. The spreads near the mountainous regions of South Asia, Greenland and the South Pole are about seven times larger than in most continental regions. In fact, the surface pressures in these high altitude regions are lower than 850 hPa. The analyses for these regions received from the different centres are not real analysis values; they are obtained by the different centres from extrapolating the RH values above the ground. The average RH analysis at 850 hPa during this period over these regions shows relatively large differences among different centres (not shown). This is consistent with the large SCM values we have seen in these regions. These larger SCM values are mainly due to the different methods or procedures used by different centres; these include different methods of interpolating to 850 hPa, different vertical coordinate systems and topographies specified in their models. These differences are the typical systematic differences from different NWP centres, but they are not necessarily the analysis error variance we will need to generate initial perturbations for an ensemble forecast system. These systematic differences can be removed by using the method we will propose in the next section.

Error variance estimate based on anomaly analysis

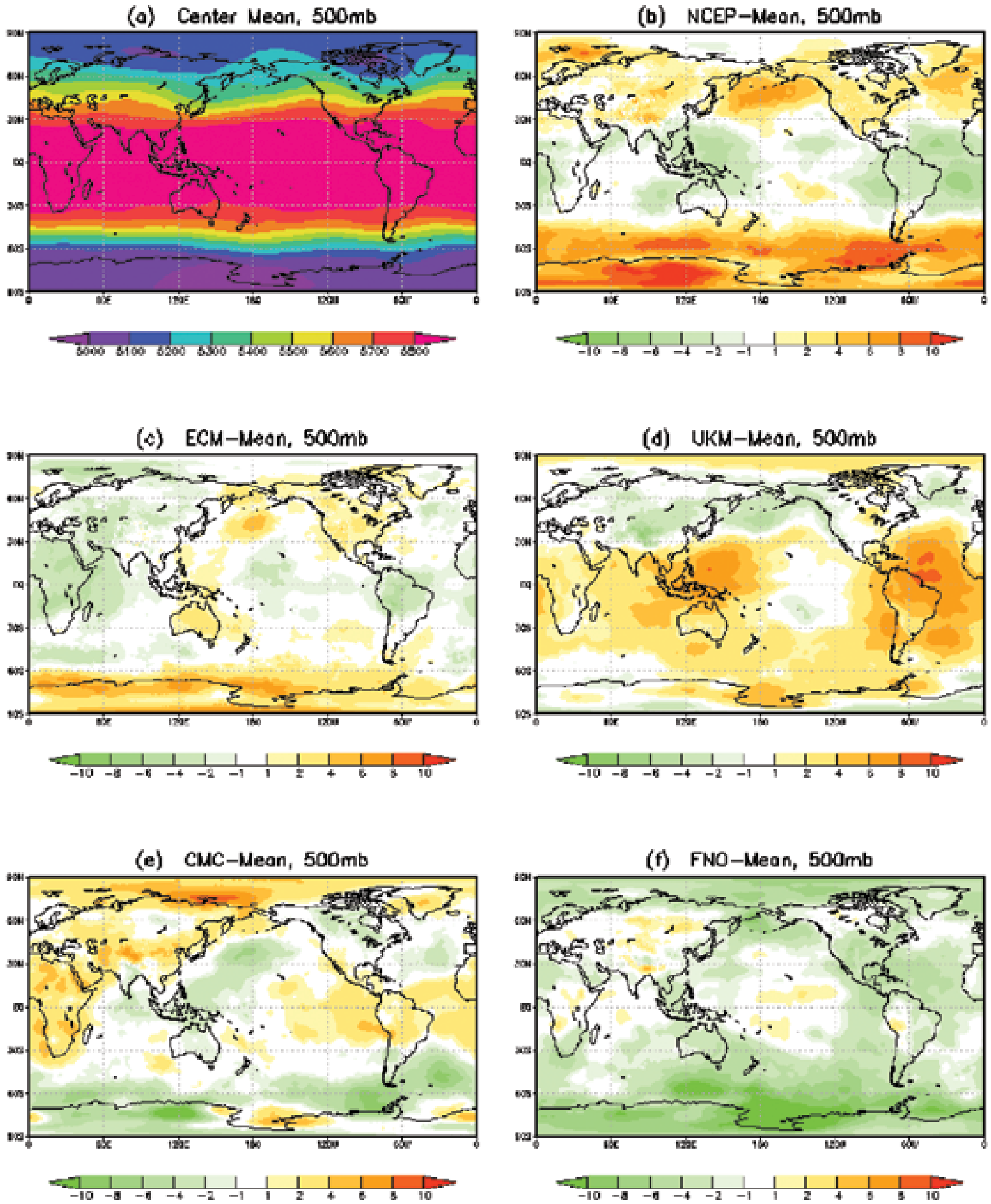
The SCM computed in the previous section describes the variations in analyses generated by different NWP centres. As we described before, different centres use different models that provide different backgrounds and different observation errors are assigned for similar types of observations. They also use different methods of quality control and bias correction. In addition, most data assimilation systems at these centres are based on 4D-Var, except for NCEP and FNO where the GSI is based on 3D-Var. All these differences at different centres produce systematic differences in the analysis data. But these systematic differences do not necessarily represent the real analysis error variance information which is needed to generate ensemble initial perturbations. To estimate a more accurate analysis error variance for generating ensemble perturbations, we calculate the spread of anomalies from different centres by removing the systematic behaviours.

Let $\bar{a}(t)$ denote the long-time averaged analysis field at time t for centre i, which can be computed by using a recursive filter as in

$$\tilde{a}_i(t) = (1.0 - \alpha)\, \tilde{a}_i(t-1) + \alpha\, a_i(t) \qquad \text{...7}$$

where t is the time step. The units of t are normally days or the cycling length of the DA system. α is the weighting factor assigned to the current analysis field. We have chosen α =

Fig. 3 The same as in Fig. 2, but for Z500.



0.05 in the following. The anomaly of analysis for each centre is defined as

$$d_i(t) = a_i(t) - \tilde{a}_i(t) \qquad ...8$$

The average anomaly of different centres is

$$\bar{d}(t) = \frac{1}{k}\sum_{i=1}^{k} d_i(t) = \frac{1}{k}\sum_{i=1}^{k} [a_i(t) - \tilde{a}_i(t)] \qquad ...9$$

The difference between the different centre anomaly and the centre-mean anomaly can be computed as

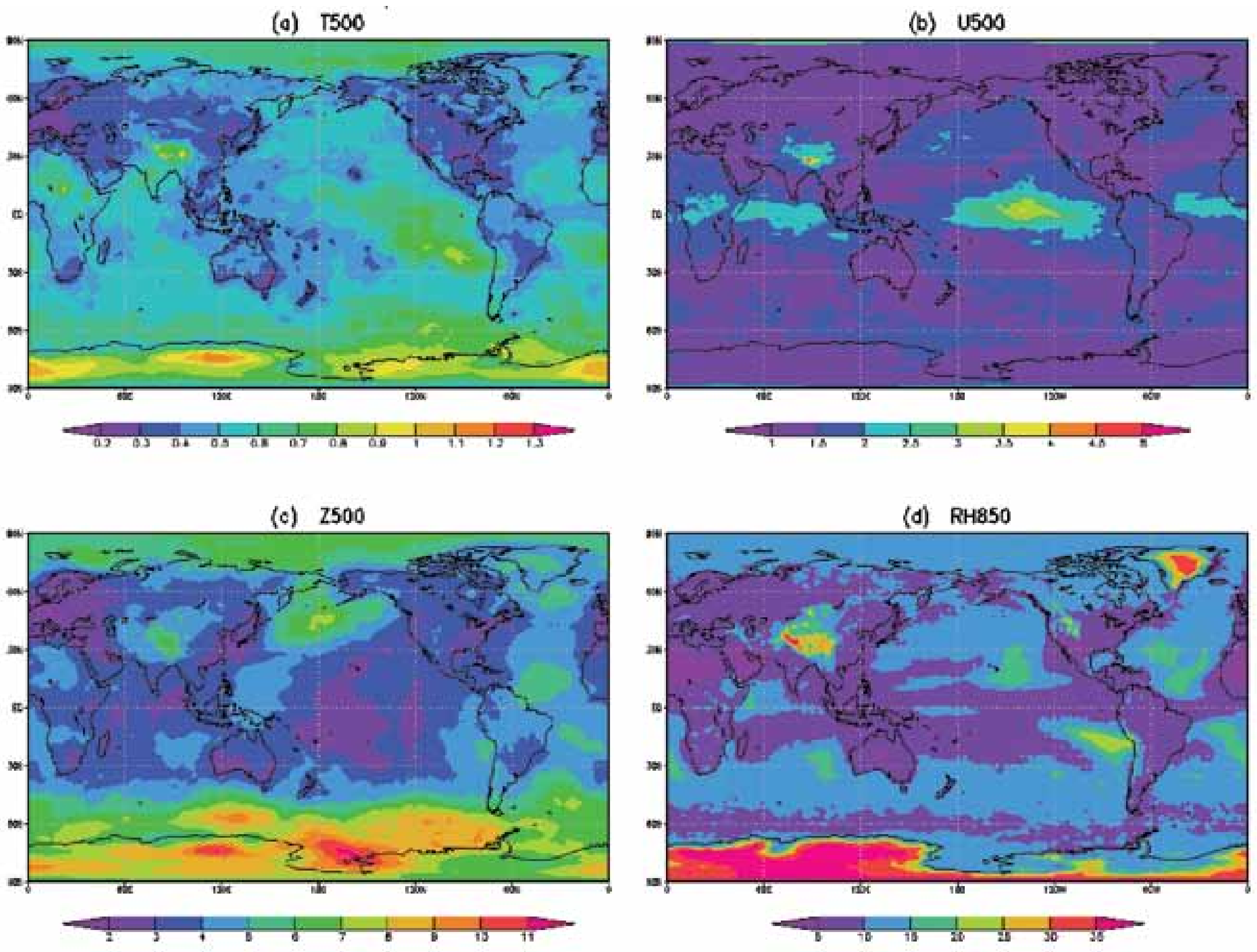
$$c_i(t) = d_i(t) - \bar{d}(t) \qquad ...10$$

Thus, the standard deviation of error variance or the spread over the average anomaly (SPA) from different centres is computed as

$$v(t) = \sqrt{\frac{1}{k}\sum_{i=1}^{k} c_i^2(t)} = \sqrt{\frac{1}{k}\sum_{i=1}^{k} [d_i(t) - \bar{d}(t)]^2} \qquad ...11$$

The experiment of the recursive filter based on Eqn 7 is run

Fig. 4 Time-averaged spread around the centre mean from 0000 UTC 1 February 2008 to 0000 UTC 30 April 2008 of (a) T500, (b) U500, (c) Z500 and (d) RH850.



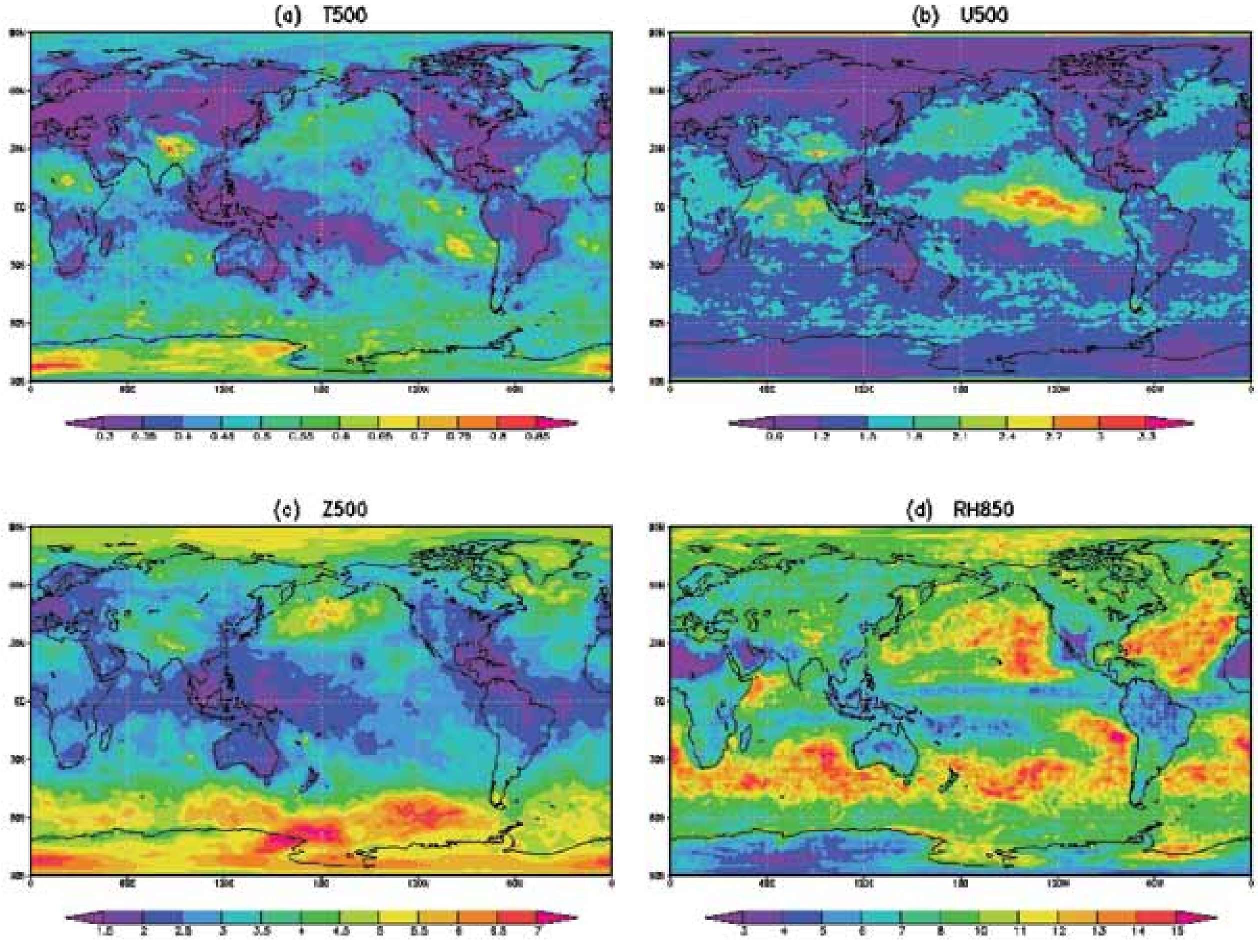
for a 90-day period starting from 0000 UTC 1 February 2008 to 0000 UTC 30 April 2008. By choosing a small value of α, we want to put more weight on the past analysis and less on the current data in computing the long-term average. Based on this small value of α, the recursive mean will not depend on when it started if the filter has run for more than two months.

Figures 5(a) to (d) show the SPA over the same 90-day time period for T500, U500, Z500 and for RH850 as in Fig. 4. These results provide a good comparison with those in Fig. 4. The SPA distributions of T500, U500 and Z500 in Figs 5(a) to (c) are similar to those of SCM shown in Figs 4(a) to (c), except that the SPA values are reduced by about one-third. But the low values of SPA are more organised and localised in the conventional data-dense regions and the larger SPA values are more focused over the oceans and polar regions than those of SCM. A larger difference between SPA and SCM can be seen in the RH850 shown in Figs 5(d) and 4(d). The SPA has larger values over the northern Pacific and northern Atlantic and all the southern oceans. The large values of SCM (due to the systematic differences between different centres) near the south Asian mountainous area, South Pole and Greenland seen in Fig. 4(d) do not appear in the SPA in Fig. 5(d). As we have discussed earlier, these systematic differences at different centres can be removed by using spread about the mean of the centres' anomalies based on Eqn 11.

To see SPA change as a function of time we choose two locations in the northern Pacific and Atlantic Oceans at (175°W, 45°N) and (30°W, 45°N). Figures 6(a) and 6(b) show the SPA for T500 and Z500 as functions of time from 0000 UTC 1 February 2008 to 0000 UTC 30 April 2008. As expected, the SPA for both variables over the North Pacific is larger than over the North Atlantic for most of the time during this period.

The SPA defined in Eqn 11 is the spread of the anomalies after the systematic behaviours are removed, it should be closer to the standard deviation of the analysis error. Figures 7(a) to (d) show the SPA distributions for U500, V500 (north-south wind component), T850 (temperature at 850 hPa) and RH850 at 0000 UTC 15 April 2008. For U500 and V500, low SPA values occur over North America, Eurasia and Australia. Areas near South Africa and South America also show lower

Fig. 5 The same as in Fig.4, but for the spread over the average anomaly.



SPA values. These regions with low values of SPA probably correspond to a higher density of wind speed observations. For T850 and RH850, the low values of SPA are located in Europe, the southern part of Asia and other regions closer to the tropics, such as Africa, Australia and Central America.

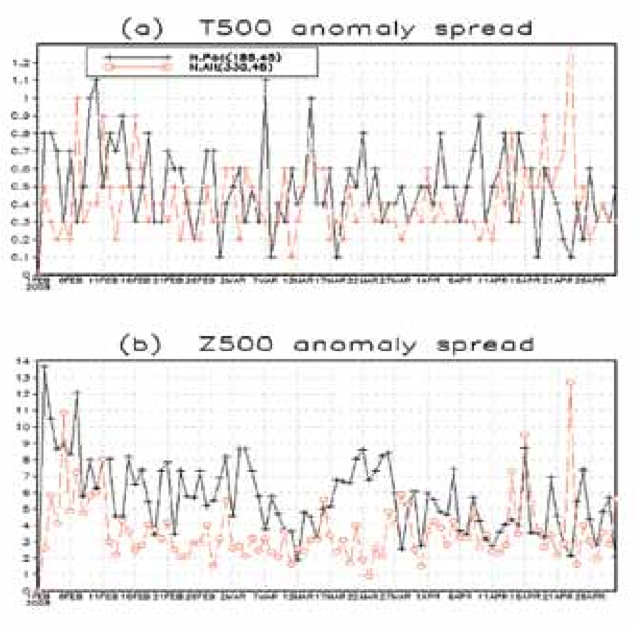
Discussion and conclusions

A good estimate of analysis error variance is essential for generating the initial perturbations in some methods, such as ET or ETR, which was first implemented in 2006 at NCEP. Unlike the ensemble-based Kalman filter DA system, as in Whitaker et al. (2007) where the analysis error variance is easily available, estimating the analysis error variance information from the 3D/4D-Var DA systems that have been widely implemented at all major NWP centres is not straightforward (Fisher and Courtier 1995).

In this paper, we have attempted to use the analysis data from five different NWP centres to estimate the analysis error variance information. Since the analysis from each centre can be considered to be an independent estimate of the unknown true state, there is no definite answer as to which centre's analysis is the closest to the truth. However, it can be shown that the mean squared analysis error with respect to the truth is smaller than the mean squared error of all individual analysis fields from the different centres. One has reason to believe that the CM analysis is generally better than any individual estimate when the number of centres is very large. The results show that the long-term averaged deviation of each centre's analysis from the CM is strongly associated with the observation network density and systematic errors of each centre.

The spread in each centre's analysis around the CM is also strongly related to the observation network density, with larger values seen over the oceans and the South Pole. There are large spreads of values for the interpolated relative humidity at 850 hPa over the mountain regions in south Asia, Greenland and the South Pole due to the systematic differences from different centres. These include the different interpolation methods, vertical coordinates and topographies

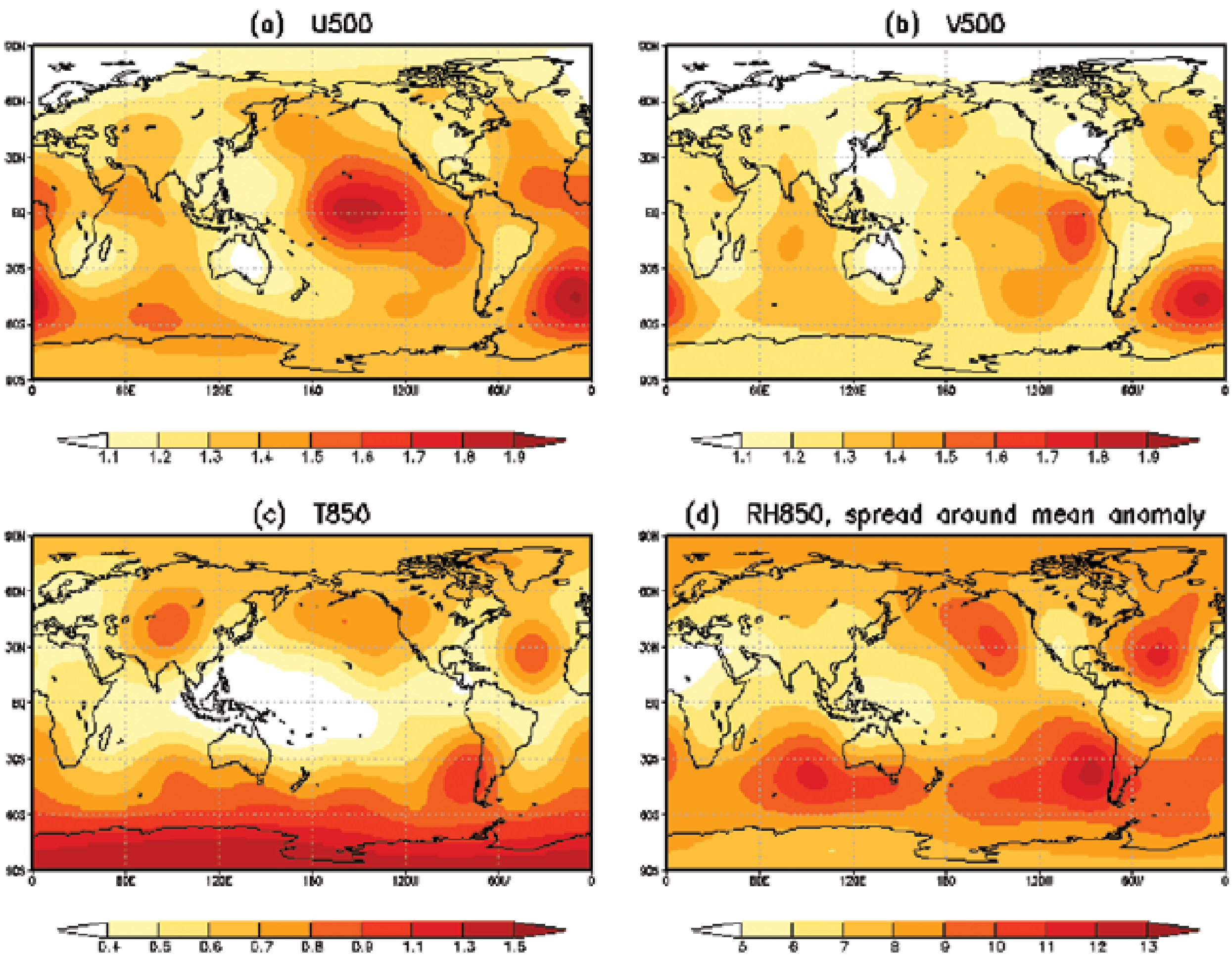
Fig. 6 The spread over the average anomaly for temperature at (a) 500 hPa and (b) Z500 as a function of time for grid-points (175°W, 45°N) (black) and (30°W, 45°N) (red).



specified at the different centres. Apart from the differences in DA methods used, systematic errors can also come from the different models, the observations and how their corresponding error variances are specified, and different methods of quality control and bias correction. All of these sources can contribute to the systematic error; thus, the spread over the CM is not necessarily the real analysis error variance we need.

In the new method proposed in this paper, we compute the anomaly of each centre's analysis by removing the long-term mean using a recursive filter. The spread over the average anomaly (SPA) from different centres can then be computed. It is found that the time-averaged distribution of SPA is even more related to the observation network density, compared with the spread around the CM analysis. More importantly, the typical systematic errors that appear in the spread around the CM over high altitude regions in south Asia, Greenland and regions near the South Pole are completely removed. The instantaneous values of SPA at any cycle for various variables bear a strong resemblance to the unknown analysis error variance. We believe that the spread of anomalies from different centres, after removing the systematic errors, is closer to the standard deviation of the analysis error we need for initialising ensemble perturbations.

Fig. 7 The instantaneous spread over the average anomaly at 0000 UTC 15 April 2008 for (a) U500, (b) V500, (c) T850 and (d) RH850.



However, we recognise that the number of analysis centres tested in this work is still very small. We do not claim that this spread in anomalies from the different centres is the true analysis error variance. But the idea and method proposed in this paper offer a useful and practical way of estimating the analysis error variance information, particularly for those centres where resources are limited and that must rely on the analysis data from other centres. As time goes by and real-time data exchange becomes more widespread among the NWP centres and analysis data from more centres becomes available, the estimate of analysis error variance derived from this method is expected to become more accurate.

Another disadvantage of this method is that the estimate of analysis error variance depends on the data flow between different centres, and this process is vulnerable to any unexpected technical problems in communications.

Acknowledgments

We are grateful to colleagues at NOAA/NCEP/EMC for their help during this work. Particularly, we thank Drs Tom Hamill and Glen White for their critical comments and insightful internal reviews, Mary Hart for editorial check up and Fanglin Yang and Bo Cui for helpful discussions. We are very thankful to Dr Carolyn Reynolds at NRL for providing the FNMOC analysis error estimate for us to compare, and to two anonymous reviewers for their comments and suggestions.

References

Bougeault, P., Toth, Z., Bishop, C., Brown, B., Burridge, D., Chen, D.H., Ebert, B., Fuentes, M., Hamill, T.M., Mylne, K., Nicolau, J., Paccagnella, T., Park, Y.-Y, Parsons, D., Raoult, B., Schuster, D., Dias, P.S., Swinbank, R., Takeuchi, Y., Tennant, W., Wilson, L. and Worley, S. 2010. The THORPEX Interactive Grand Global Ensemble (TIGGE). *Bull. Am. Met Soc.*, in press.

Bowler, N.E. 2006. Comparison of error breeding, singular vectors, random perturbations and ensemble Kalman filter perturbation strategies on a simple model. *Tellus, 58A*, 538-48.

Bowler, N.E., Arribas, A. and Mylne, K. 2008. The benefits of multi-analysis and poor-man's ensemble. *Mon. Weath. Rev., 136*, 4113-29.

Buizza, R., Houtekamer, P.L., Toth, Z., Pellerin, P., Wei, M. and Zhu, Y. 2005. A comparison of the ECMWF, MSC and NCEP global ensemble prediction systems. *Mon. Weath. Rev., 133*, 1076-97.

Derber, J.C., Parish, D.F. and Lord, S.J. 1991. The new global operational analysis system at the National Meteorological Center. *Weath. forecasting, 6*, 538-47.

Fisher, M. and Courtier, P. 1995. Estimating the covariance matrices of analysis and forecast error in variational data assimilation. *ECMWF Tech. Memo., No. 220*, August 1995.

Hamill, T.M., Snyder, C. and Morss, R.E. 2000. A comparison of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles. *Mon. Weath. Rev., 128*, 1835-51.

Houtekamer, P.L., Lefaivrem, L., Derome, J., Ritchie, J. and Mitchell, H.L. 1996. A system simulation approach to ensemble prediction. *Mon. Weath. Rev., 124*, 1225-42.

Langland, R.H., Maue, R.N. and Bishop, C.H. 2008. Uncertainty in atmospheric temperature analysis. *Tellus, 60A*, 598-603.

Leutbecher, M. and Palmer, T.N. 2008. Ensemble forecasting. *J. Comp. Phys., 227*, 3515-39.

Molteni, F., Buizza, R., Palmer, T. and Petroliagis, T. 1996. The ECMWF ensemble prediction system: Methodology and validation. *Q. Jl R. Met. Soc., 122*, 73-119.

Park, Y.-Y., Buizza, R. and Leutbecher, M. 2008. TIGGE: Preliminary results on comparing and combining ensembles. *Q. Jl R. Met. Soc., 134*, 2029-50.

Parrish, D.F. and Derber, J. 1992. The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Weath. Rev., 120*, 1747-63.

Rabier, F., Jarvine, H., Klinker, E. Mahfouf, J.-F. and Simmons, A. 2000. The ECMWF operational implementation of four-dimensional variational assimilation. I: experimental results with simplified physics. *Q. Jl R. Met. Soc., 126*, 1143-70.

Swanson, K.L. and Roebber, P. 2008. The impact of analysis error on medium-range weather forecasts. *Mon. Weath. Rev., 136*, 3425-31.

Tippett, M.K., Anderson, J.L., Bishop, C.H., Hamill, T. and Whitaker, J.S. 2003. Ensemble square root filters. *Mon. Weath. Rev., 131*, 1485-90.

Toth, Z. and Kalnay, E. 1993. Ensemble forecasting at NMC: the generation of perturbations. *Bull. Am. Met. Soc., 174*, 2317-30.

Toth, Z. and Kalnay, E. 1997. Ensemble forecasting at NCEP and the breeding method. *Mon. Weath. Rev., 125*, 3297-319.

Wei, M., De Pondeca, M.S.F.V., Toth, Z. and Parish, D. 2010. Estimating observation impact signals in NCEP GSI using the Lanczos method. NCEP/EMC Report, NCEP, Camp Springs, USA.

Wei, M. and Toth, Z. 2003. A new measure of ensemble performance: Perturbations versus Error Correlation Analysis (PECA). *Mon. Weath. Rev., 131*, 1549-65.

Wei, M., Toth, Z., Wobus, R., Zhu, Y. and Bishop, C.H. 2005. Initial Perturbations for NCEP Ensemble Forecast System. Thorpex Symposium Proceedings for the First THORPEX Internal Science Symposium 6-10 December 2004, Montreal, Canada. The Symposium Proceedings in a WMO Publication 2005, *WMO TD No. 1237*, *WWRP THORPEX No. 6*, 2005, 227-30.

Wei, M., Toth, Z., Wobus, R., Zhu, Y., Bishop, C.H. and Wang, X. 2006. Ensemble Transform Kalman Filter-based ensemble perturbations in an operational global prediction system at NCEP. *Tellus, 58A*, 28-44.

Wei, M., Toth, Z., Wobus, R. and Zhu, Y. 2008. Initial perturbations based on the Ensemble Transform (ET) technique in the NCEP global operational forecast system. *Tellus, 60A*, 62-79.

Whitaker, J.S. and Hamill, T.M. 2002. Ensemble data assimilation without perturbed observations. *Mon. Weath. Rev., 130*, 1913-24.

Whitaker, J.S., Hamill, T.M., Wei, X., Song, Y. and Toth, Z. 2007. Ensemble data assimilation with the NCEP global forecast system. *Mon. Weath. Rev., 136*, 463-82.

Wu, W.-S., Purser, R.J. and Parrish, D.F. 2002. Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Weath. Rev., 130*, 2905-16.