ECMWF Feature article

from Newsletter Number 144 – Summer 2015

METEOROLOGY

Promising results in hybrid data assimilation tests

cosmin4000/iStock/Thinkstock

www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/4n5hz0y5

This article appeared in the Meteorology section of ECMWF Newsletter No. 144 – Summer 2015, pp. 33–39.

Promising results in hybrid data assimilation tests

Mats Hamrud, Massimo Bonavita & Lars Isaksen

Data assimilation systems for global numerical weather prediction (NWP) combine short-range forecasts (the 'background') with the latest observations to arrive at the best possible representation of the current state of the atmosphere. They are designed to ensure that the resulting analysis of the atmosphere is as accurate and realistic as possible, while taking into account the errors associated both with the background and with observations.

ECMWF has been a pioneer in the development and operational implementation of a data assimilation method called 4DVAR. '4D' stands for the three spatial dimensions plus time, as this method uses observations as they come in over a period of time, while 'VAR' refers to variational methods.

An alternative algorithm, called the Ensemble Kalman Filter (EnKF), is also suitable for operational use. In recent tests, an EnKF-based data assimilation system developed at the Centre has shown good forecast performance, and a hybrid 4DVAR/EnKF approach has been found to perform significantly better than the two systems in standard configuration individually. The EnKF algorithm is also highly scalable, making it particularly well-suited to future, massively parallel computer architectures.

The performance of the hybrid 4DVAR/EnKF system has been found to be comparable to that of a low-resolution version of the data assimilation system in operational use at ECMWF, which is a 4DVAR system combined with an ensemble of data assimilations.

The EnKF system

A Kalman Filter is an algorithm designed to estimate the state of a system based on predictions of the system's behaviour on the one hand and observations on the other. In NWP, the system is the atmosphere, the predictions are short-range forecasts, and the observations are the most recent set of atmospheric observations available. An Ensemble Kalman Filter makes direct use of an ensemble of short-range forecasts to estimate the errors associated with the background.

By contrast, ECMWF's operational data assimilation system is a 4DVAR system which uses error statistics provided by an Ensemble of Data Assimilations (EDA). The EDA is an ensemble of lower resolution 4DVAR assimilations using perturbed observations and perturbed model tendencies.

Despite these differences, the EnKF system developed at ECMWF uses the operational EDA component in a number of ways. The model propagating the ensemble is the same, and the operators calculating the observation equivalents are the same. Both use the ECMWF Integrated Forecasting System (IFS) code base. The observation Quality Control (QC) and data selection procedure are also similar, using the same kind of background checks and data thinning.

The EnKF system also uses large parts of the technical infrastructure of the operational system at ECMWF. The observation equivalents, H(x), are evaluated at the observation time by running forecasts over the analysis window and applying the observation operator to the evolving state. The post-processing of the analysis, including producing pressure level data, is also performed by the IFS.

Two main variants of EnKF have been implemented: the Ensemble Square Root Filter (EnSRF) (*Whitaker & Hamill*, 2002) and the Local Ensemble Transform Kalman Filter (LETKF) (*Hunt et al.*, 2007). Both these EnKF implementations do not require perturbing the observations, as in the EDA approach, which has the theoretical advantage of not introducing additional sampling errors. Early experimentation showed that there was no statistically significant difference in performance between the two EnKF schemes. The focus in this article will be on the LETKF (see Box A for more details on the LETKF).

The analysed atmospheric variables in the ECMWF EnKF are temperature, wind vector components (u, v), specific humidity, surface pressure and surface pressure tendency.

 doi:10.21957/4n5hz0y5

The LETKF algorithm

A

The version of the EnKF used in the experiments is the Local Ensemble Transform Kalman Filter (LETKF). The LETKF algorithm can be thought of as a way of minimizing the standard 4DVAR cost function, i.e.:

$$J(\boldsymbol{x}) = \tfrac{1}{2}\left(\boldsymbol{x}-\boldsymbol{x}^{\boldsymbol{b}}\right)^{T}\mathbf{B}^{-1}\left(\boldsymbol{x}-\boldsymbol{x}^{\boldsymbol{b}}\right) + \tfrac{1}{2}\left(\mathbf{y}-\mathbf{H}(\boldsymbol{x})\right)^{T}\mathbf{R}^{-1}\left(\mathbf{y}-\mathbf{H}(\boldsymbol{x})\right) \quad (1)$$

in the space spanned by the ensemble. This means that in Equation 1 the background error covariance matrix is directly sampled from the ensemble background forecasts: $\boldsymbol{x}_i^b$

$$\mathbf{B} = \tfrac{1}{N_e-1}\textstyle\sum_{i=1}^{N_e}\left(\boldsymbol{x}_i^b-\overline{\boldsymbol{x}^b}\right)^{T}\left(\boldsymbol{x}_i^b-\overline{\boldsymbol{x}^b}\right) = \tfrac{1}{N_e-1}\mathbf{X}^{\boldsymbol{b}}\left(\mathbf{X}^{\boldsymbol{b}}\right)^{T} \quad (2)$$

For linear observation operators, the analysis is the minimum of the cost function (1) in the space spanned by the ensemble background forecasts $\boldsymbol{x}_i^b$ and it can be expressed in closed form as:

$$\boldsymbol{x}^{\boldsymbol{a}} = \overline{\boldsymbol{x}^{\boldsymbol{b}}} + \mathbf{X}^{\boldsymbol{b}}\boldsymbol{w}^{\boldsymbol{a}} \quad (3)$$

$$\boldsymbol{w}^{\boldsymbol{a}} = \left((N_e-1)\boldsymbol{I} + \left(\mathbf{Y}^{\boldsymbol{b}}\right)^{T}\mathbf{R}^{-1}\left(\mathbf{Y}^{\boldsymbol{b}}\right)\right)^{-1}\left(\mathbf{Y}^{\boldsymbol{b}}\right)^{T}\mathbf{R}^{-1}\left(\mathbf{y}-\overline{\mathbf{H}(\boldsymbol{x}_i^b)}\right) \quad (4)$$

where $\mathbf{Y}^{\boldsymbol{b}}$ are matrices whose columns are formed by $\mathbf{H}\left(\boldsymbol{x}_i^b\right) - \overline{\mathbf{H}(\boldsymbol{x}_i^b)}$. From Equations (3) and (4), it is apparent that the analysis essentially determines which linear combination of the ensemble members is the best estimate of the current state, given the current batch of observations. The error covariance of the analysed state is given by:

$$\mathbf{P}^{\boldsymbol{a}} = \mathbf{X}^{\boldsymbol{b}}\left((N_e-1)\boldsymbol{I} + \left(\mathbf{Y}^{\boldsymbol{b}}\right)^{T}\mathbf{R}^{-1}\left(\mathbf{Y}^{\boldsymbol{b}}\right)\right)^{-1}\left(\mathbf{X}^{\boldsymbol{b}}\right)^{T} \quad (5)$$

Taking the symmetric square root of Equation (5) provides the perturbations to add to the analysis (3) to form the full analysed ensemble, which can then be integrated forward in time.

Cap on observations

In the LETKF context, it has been found useful to limit the number of observations affecting the analysis of any grid point. This number is currently limited to 30 for each combination of different report type and observation variable, with the observations closest to the analysed grid point being selected. This was found to produce a significant advantage in terms of forecast scores in the northern hemisphere compared to using all available observations in the analysis update. At the same time it enables massive computational savings. Results are remarkably insensitive to the precise number of selected observations (for example, doubling or halving the number of locally selected observations has a marginal impact on analysis quality).

Covariance inflation

As in most EnKF implementations, two complementary types of covariance inflation have been implemented, namely multiplicative and additive inflation (*Whitaker & Hamill*, 2012). The multiplicative inflation relaxes the analysis variance towards the background variance to prevent it from dropping to unrealistically low values in densely observed regions. Experiments have shown that the EnKF skill is not very sensitive to the relaxation factor a; an a of 0.9 has been used in the experimentation. The additive inflation is based on the difference between the 48-hour and 24-hour IFS forecasts verifying at the same time, sampled over a 6-year period. A random sample of these forecast differences is added to each ensemble member at each analysis cycle. These are scaled by a factor of 0.25, which converts the forecast difference to a nominal 6-hour difference.

Bias correction

In the ECMWF 4DVAR implementation, a variational bias correction method is used for satellite radiance data. The computational cost of the variational bias correction is insignificant in terms of the whole 4DVAR cost. In the EnKF context, an equivalent scheme can be implemented but at a non-negligible computational cost. We have not implemented such a scheme; instead we have relied on the stored bias corrections produced by the 4DVAR system. In the context of the hybrid system presented below, the bias correction can be performed by the 4DVAR component of the system.

Controlling analysis noise

It is well documented that EnKF analyses tend to have problems in representing the model dynamical balance constraints, leading to the excitation of spurious inertia-gravity waves in the short-range forecast. This is normally attributed to the effect of covariance localization disrupting the mass-wind field geostrophic balance. Diagnostics within the ECMWF EnKF system do show evidence of significant small-scale noise associated with inertia-gravity waves in forecasts based on the EnKF analysis. To counter this problem, EnKF systems often apply some form of Digital Filter Initialization (DFI).

Experimentation with different versions of DFI applied to the EnKF analysis did in fact reduce the imbalances, but always degraded forecast scores. We have therefore developed a new method which adjusts the column horizontal wind divergence to match the analysed surface pressure tendency. This 'divergence adjustment' (see *Hamrud et al.*, 2014) is performed for all ensemble members. The wind increments caused by this procedure are normally very small (much smaller than the analysis increments) and the details of their distribution in the vertical do not seem to be very important. The significant change in the initial pressure tendency achieved by this procedure is illustrated in Figure 1. This method effectively imposes an additional dynamical balance constraint on the analysis and thus removes a large portion of the excess inertia-gravity waves from the short-range forecast.

The impact of this 'divergence adjustment' on the quality of the analysis is significant and positive. The 6-hour forecast fit for land-based surface pressure observations improves by around 5% and by around 10% for marine surface pressure observations. The effect on the observational fit of other variables is negligible.

Scalability and computational cost

There are no fundamental algorithmic constraints on the scalability of an EnKF data assimilation system. The forecasts and the computation of the observation equivalents can be done independently for all ensemble members. The EnKF analysis itself is local. In principle, the limit on dividing the computations into independent pieces of work is given by the size of the forecast state.

There are, however, a number of practical limits to the scalability of the EnKF system. One issue is the Input/Output (IO) bottleneck. In the analysis stage, all forecasts are needed to describe the background state, thus they have to be read in and distributed. At the end of the analysis, all the analysis states have to be written out to provide the initial states for the next analysis cycle's forecasts. In addition, all the observational data, including the observation departures for all ensemble members, have to be read in. This IO load is distributed over tasks in the ECMWF EnKF implementation, but for large applications the overall throughput of the IO subsystem still limits the scalability.

The other practical scalability issues for the analysis stage are quite different for the EnSRF and LETKF analyses. Limiting our analysis to the LETKF (the interested reader can find a broader discussion in *Hamrud et al.*, 2014), where the state is distributed in contiguous areas and the observations are distributed so that each task has local access to the observations it needs, the problem becomes that of load balancing. The cost of analysing a grid point is dependent on the number of nearby observations, and as the observation density is not homogeneous, the cost varies for different parts of the globe. Thus a distribution with an equal number of state points on each processor would lead to a poor load balance.

To prevent this, we measure the cost of performing the analysis of each state column as a function of the number of close observations. Then, in subsequent analysis steps, we find out the number of near by observations for each state profile in the global view and combine this information with the empirical relation between observation density and computational cost to compute weights for the distribution of the state space variables. The grid points are then distributed so that each processor ends up with approximately the same computational weight, rather than with the same number of grid points. This method normally gives a reasonable load balance, with a ratio of the fastest to the slowest analysis time on an individual processor of typically around 0.85.

In terms of computational cost, the natural comparison for an EnKF implementation is ECMWF's EDA system, based on multiple independent copies of 4DVAR (*Isaksen et al.*, 2010). Here an EnKF implementation has an obvious advantage: the cost per member of an EnKF system is considerably lower than for the EDA, mainly due to the fact that the Kalman Gain is only computed once in the EnKF.

Sensitivity experiments

EnKF systems are relatively easy to set up and tend to produce acceptable results with a broad range of 'reasonable' choices of the main filter parameters. A few examples of parameters that affect LETKF performance will be discussed here.

A larger ensemble size is desirable as it reduces the sampling errors in the background error covariance estimates. A set of experiments was carried out in which the ensemble size was set to 60, 120 and 240. The experiments were run with a triangular spectral truncation of 159 for the IFS model (~120 km linear grid spacing) and 91 vertical levels. All available observations were used, with the exception of radiances from satellite imagers and scatterometer winds, which, at the time of the experiments, had not yet been implemented in the EnKF. The sample ensemble covariances were localized in order to reduce sampling errors, with a localization radius of 2,800 km in the horizontal and 2.5 scale heights in the vertical (scale height is here the vertical distance over which pressure falls by a factor of 1/e).

 doi:10.21957/4n5hz0y5

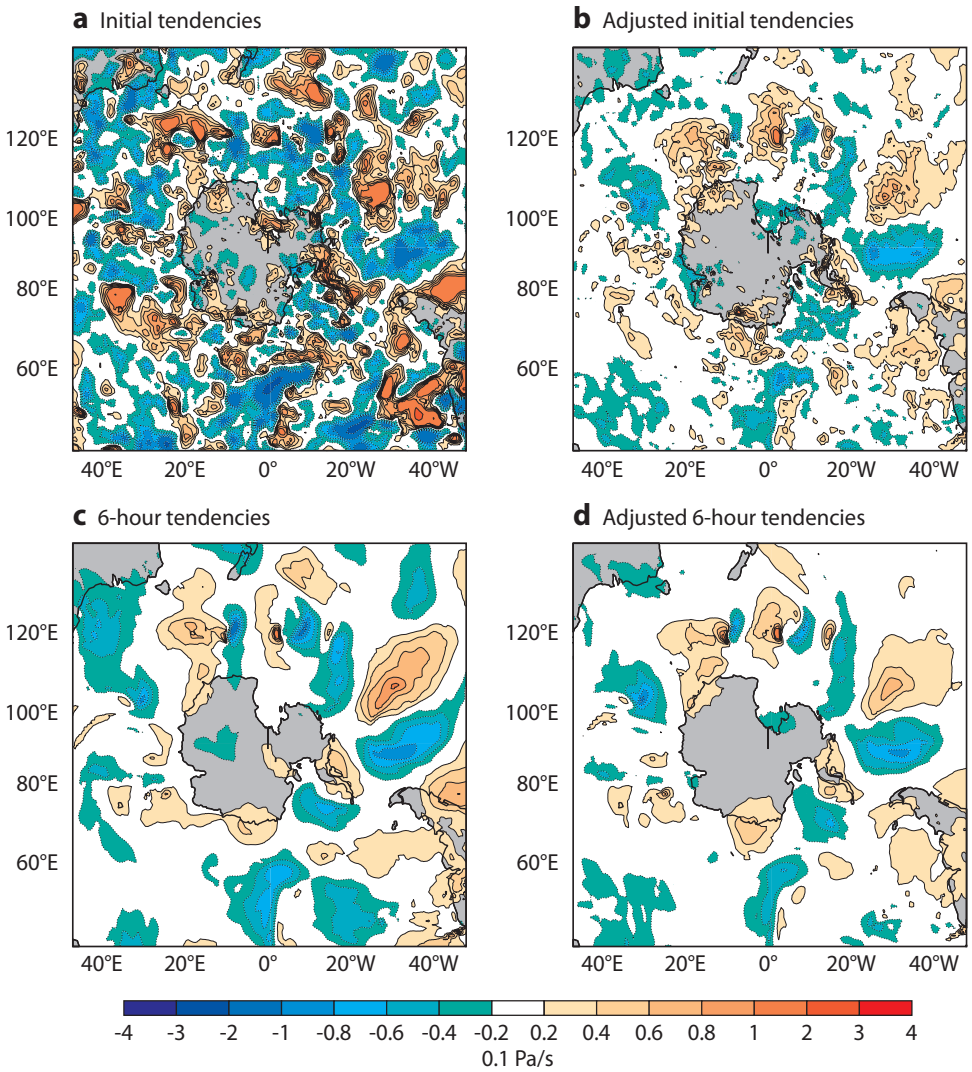

Figure 1 Surface pressure tendency fields over the southern hemisphere in forecasts based on the ensemble mean analysis for (a) initial tendencies without divergence adjustment, (b) initial tendencies with divergence adjustment, (c) tendencies after 6 hours without divergence adjustment, and (d) tendencies after 6 hours with divergence adjustment.

Figure 2 shows how the increase in ensemble size helps to better characterize sampled error correlations from the EnKF. Figures 2a–b show, as a function of horizontal and vertical distance, the average of the sample absolute correlations between 3-hour forecasts of 500 hPa temperature and observations, in the form of AMSU-A (Advanced Microwave Sounding Unit A) brightness temperature model equivalents, for the 60-member EnKF (Figure 2a) and the 120-member EnKF (Figure 2b). Figures 2c–d show the average of the sample absolute correlations between 500 hPa zonal wind 3-hour forecasts and AMSU-A brightness temperature model equivalents for the 60-member EnKF (Figure 2c) and the 120-member EnKF (Figure 2d). In both instances it is apparent how the structure of the spatial correlations is much better defined in the 120-member ensemble, which presents lower values of spurious long-range correlations.

Figure 3 shows forecast skill scores, averaged over the month of February 2011, of forecasts based on the EnKF ensemble mean analyses for a 60-member EnKF, a 120-member EnKF and a 240-member EnKF in the northern extra-tropics (Figure 3a) and the southern extra-tropics (Figure 3b). The improvement from a 60-member to a 120-member ensemble is consistent and statistically significant throughout the forecast range. On the other hand, gains are more limited when moving from a 120-member ensemble to 240 members. Part of the reason is the fact that covariance localization parameters were not changed in these experiments. Re-running the 240-member EnKF experiment with double the vertical covariance localization factor shows a clear improvement.

Higher spatial resolution is beneficial for the EnKF as the short-range forecast fields behave more realistically in the high wavenumber part of the spectrum, and they benefit from a more accurate

representation of the physiographic constraints. This is shown in Figure 4, where we present the deterministic skill scores for forecasts, over the months of February and March 2011, based on ensemble mean analyses in 60-member EnKF assimilation experiments run at three different resolutions: spectral truncations TL159 (approximate grid-point spacing 120 km), TL319 (approximate grid-point spacing 60 km) and TL639 (approximate grid-point spacing 30 km). The positive impact of increasing horizontal resolution is evident going from TL159 to TL319, while diminishing returns are seen from TL319 to TL639. The increase in horizontal resolution also has the important side effect of generating more active ensemble background fields with a larger spread. This in turn helps to reduce the reliance of the EnKF on the adaptive covariance inflation algorithm that is used to prevent the filter collapse.

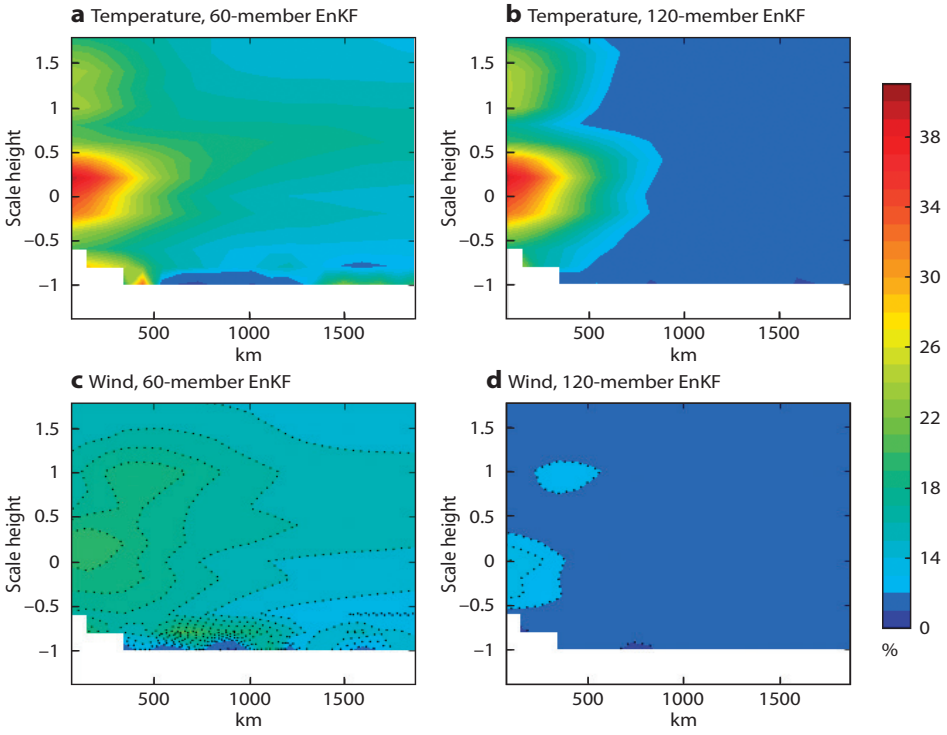

Figure 2 Azimuthally-averaged absolute correlation between short-range (3-hour) forecasts of (a) temperature at 500 hPa and AMSU-A brightness temperature model equivalents for a 60-member EnKF; (b) temperature at 500 hPa and AMSU-A brightness temperature model equivalents for a 120-member EnKF; (c) zonal wind at 500 hPa and AMSU-A brightness temperature model equivalents for a 60-member EnKF; and (d) zonal wind at 500 hPa and AMSU-A brightness temperature model equivalents for a 120-member EnKF.

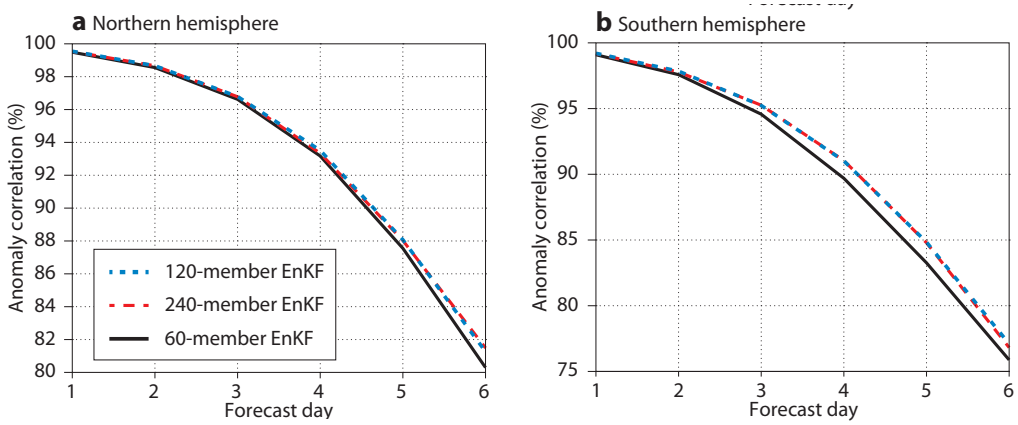

Figure 3 500-hPa geopotential forecast anomaly correlation scores for the 60-member EnKF, the 120-member EnKF, and the 240-member EnKF computed with respect to ECMWF's operational analysis and averaged over the month of February 2011 in (a) the northern hemisphere and (b) the southern hemisphere.

doi:10.21957/4n5hz0y5

EnKF performance

The development activities described in the previous sections have led to the availability of a stable EnKF-based data assimilation system at ECMWF. This system is able to assimilate all the available observation types, it has good scalability properties and it is computationally efficient. It is thus possible to perform a detailed comparison with the 4DVAR system available at ECMWF under controlled conditions.

An extended assimilation experiment has been run to evaluate the skill of the ECMWF EnKF analysis in a deterministic sense, i.e. when the ensemble mean is used as the initial condition for a single, deterministic forecast. This experiment has been run with triangular spectral truncation TL399 (corresponding to approximately 50-km grid spacing) and 137 model levels (near surface to 0.01 hPa), which is the current resolution of the EDA, using IFS model cycle 40r1.

An example of the results is given in Figure 5, where the forecast skill scores (geopotential height anomaly correlation at 500 hPa) of the EnKF system and a 4DVAR system using climatological background errors and covariances are presented. It is apparent that the two lines are hardly separable. The general conclusion that can be drawn is that, at this stage of development, the ECMWF EnKF-based data assimilation system has similar deterministic forecast performance to a 4DVAR system using static background error covariance estimates. The EnKF is at TL399 resolution and 4DVAR at TL399 outer loop and TL95/TL159 inner loop resolution.

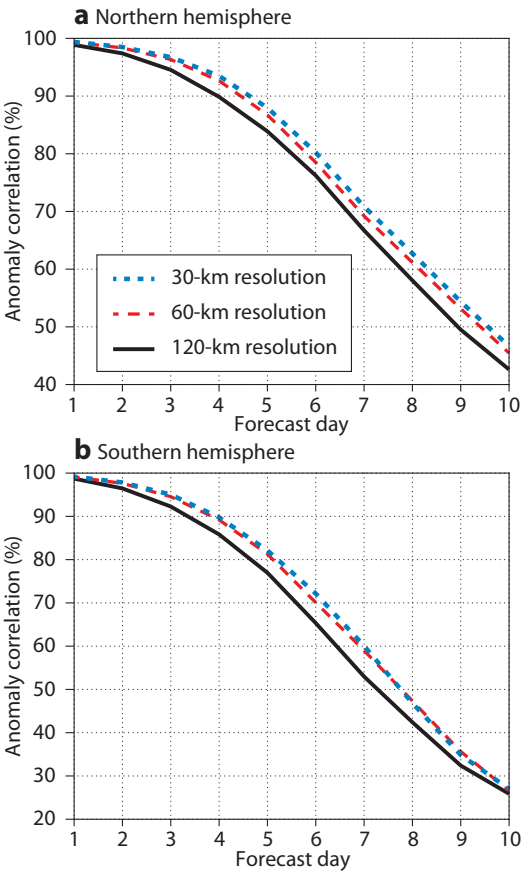

Figure 4 1,000-hPa geopotential forecast anomaly correlation scores for ensemble mean forecasts with 120-km resolution (T159) EnKF, 60-km resolution (T319) EnKF, and 30-km resolution (T639) EnKF for (a) the northern hemisphere and (b) the southern hemisphere. Scores computed with respect to ECMWF operational analysis and averaged over the months of February and March 2011.

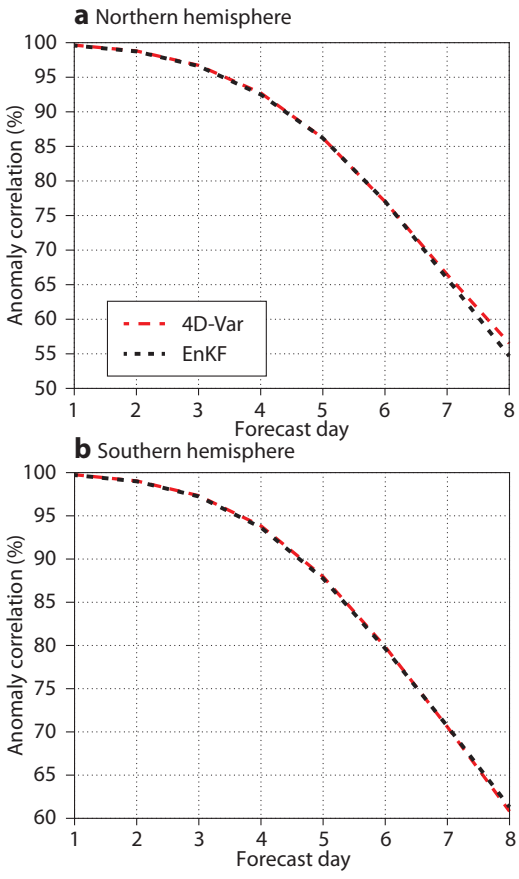

Figure 5 500-hPa geopotential forecast anomaly correlation for a 100-member EnKF and 4DVAR using climatological background errors and covariances, both at 50-km resolution (TL399), in (a) the northern hemisphere and (b) the southern hemisphere. All scores computed with respect to ECMWF's operational analysis and averaged over the period from 1 May to 31 August 2012.

doi:10.21957/4n5hz0y5

Hybrid Gain EnDA

Recent work by *Penny* (2014) proposed, in an idealized setup, a new hybrid variational/EnKF formulation that involves taking a weighted mean of the EnKF and variational analysis and using this 'control' state to re-centre the EnKF analysis ensemble (Box B). This is equivalent to blending the Kalman gains of the EnKF and 4DVAR systems and is thus called Hybrid Gain Ensemble Data Assimilation (HG-EnDA). At ECMWF, such a system produced a considerable gain in forecast skill compared to that of the EnKF and a 4DVAR system using climatological background errors and covariances (Figure 6, similar results are obtained for most variables and vertical levels). Where does this improvement originate?

The Hybrid Gain EnDA

B

In the Hybrid Gain EnDA scheme, in addition to a standard EnKF analysis update as described in Box A, an incremental 4DVAR analysis update is also performed. The 4DVAR analysis uses the background forecast from the previous ensemble mean analysis as background and starting linearization trajectory. Formally, the incremental 4DVAR cost function is minimised, i.e. in standard notation:

$$J(\delta \boldsymbol{x}) = \tfrac{1}{2}\delta \boldsymbol{x}^T \mathbf{B}^{-1}\delta \boldsymbol{x} + \tfrac{1}{2}(\mathbf{H}\delta \boldsymbol{x} - \mathbf{d})^T \mathbf{R}^{-1}(\mathbf{H}\delta \boldsymbol{x} - \mathbf{d}) \quad (1)$$

where $\mathbf{d} = \mathbf{y}^o - H(\mathbf{x}^b)$ is the innovation vector and $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^b$ is the incremental control vector, with the background provided by a short (t+3h) nonlinear model integration of the Hybrid Gain mean analysis valid at the previous analysis update time:

$$\mathbf{x}^b = M\left(\overline{\mathbf{x}^a_{\mathbf{tEnKF}}}(t_{k-1})\right) \quad (2)$$

The EnKF and 4DVAR analysis fields valid in the middle of the assimilation window (i.e., 0000, 0600, 1200, 1800 UTC in the 6-hour cycling setup used in the experiments) are then linearly combined to produce a 'hybrid gain' analysis around which the EnKF analysis ensemble is re-centred:

$$\mathbf{x}^a_{Hyb} = \alpha\overline{\mathbf{x}^a_{tEnKF}} + (1-\alpha)\mathbf{x}^a_{4DVar} \quad (3)$$

The weight given to each of the contributing analyses (denoted as α) is a free tuning parameter, reflecting the expected accuracy of the EnKF and 4DVAR analyses. A value of α = 0.5 has been used in the experiments reported in this work. This is thought to be not far from the optimal value, as later experimentation with α = 0.75 and α = 0.25 has generally produced inferior results.

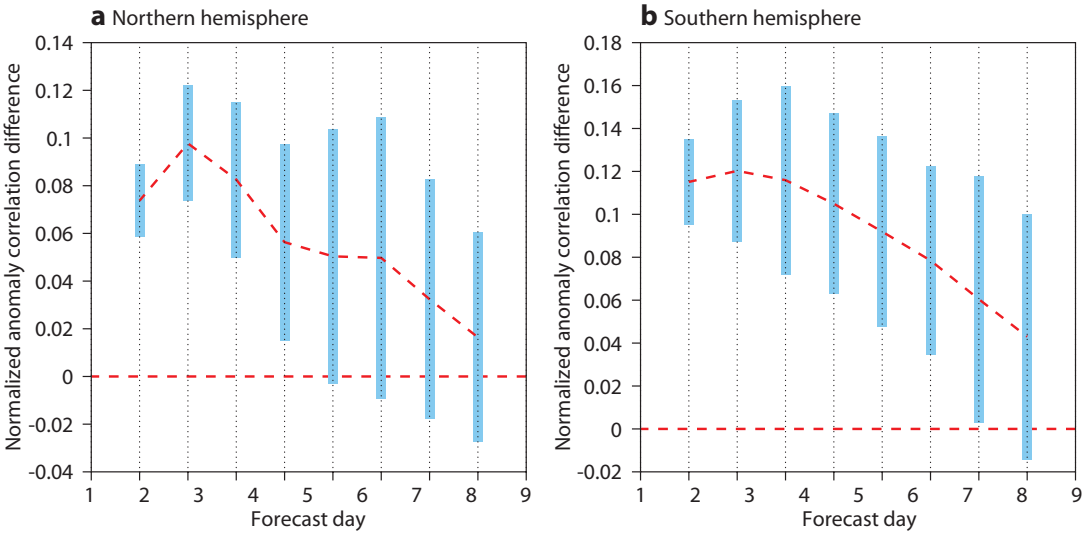

Figure 6 Normalised difference of the 500-hPa geopotential height forecast anomaly correlation between Hybrid Gain EnDA and 4DVAR using climatological background errors and covariances in (a) the northern hemisphere and (b) the southern hemisphere. Positive values indicate superior skill of the Hybrid Gain EnDA. Error bars show 95% confidence intervals. All scores computed with respect to ECMWF's operational analysis and averaged over the period from 1 May to 31 August 2012.

doi:10.21957/4n5hz0y5

An indication is given by Figure 7, where we show the mean sea level (MSL) pressure control background forecast and its standard deviation for an HG-EnDA experiment valid on 25 February 2014 at 0000 UTC (Figure 7a); the EnKF analysis increment to a single surface pressure observation located at (58.5°N, 30.3°W) and 1 hPa lower than the background forecast (Figure 7b); and the corresponding analysis increment of the 4DVAR analysis (Figure 7c).

It is apparent that the EnKF analysis increments have a smaller footprint due to the applied horizontal covariance localization; 4DVAR analysis increments have a considerably larger spatial extension as they are based on climatological correlations which have not been localized. On the other hand the EnKF increment shows interesting, flow-dependent, short-range covariance structures while the corresponding 4DVAR increments present a more isotropic structure (due to the climatological sampling of the 4DVAR covariances and the locally isotropic structure of the background error model used in 4DVAR).

The HG-EnDA, as a mixture of these two increments, is thus able to retain some of the flow-dependent aspects of the EnKF analysis update but also to partially correct for its two main apparent drawbacks, i.e. the underestimation of longer-range correlations and the underdispersiveness of the ensemble background in specific meteorological situations.

The use of a static 4DVAR in the re-centring step of the Hybrid Gain EnDA can be seen as a regularization procedure of the EnKF analysis which has two main benefits: a) it reduces the effect of sampling noise and localization; and b) it introduces climatological information into the background error covariance estimates, which is important to reduce the effect of model biases in the analysis and is not easy to do in a standard EnKF. The results show that, similar to the operational hybrid 4DVAR/EDA system, the Hybrid Gain algorithm clearly improves on the accuracy of its individual EnKF and 4DVAR components while its computational costs are only marginally higher than those of a standard EnKF algorithm. But it does require the development and maintenance of both the EnKF and the 4DVAR analysis systems.

Further tests not detailed here suggest that the hybrid 4DVAR/EnKF system is comparable in performance to a low-resolution version of the operational hybrid 4DVAR/EDA system (see *Hamrud et al.*, 2014).

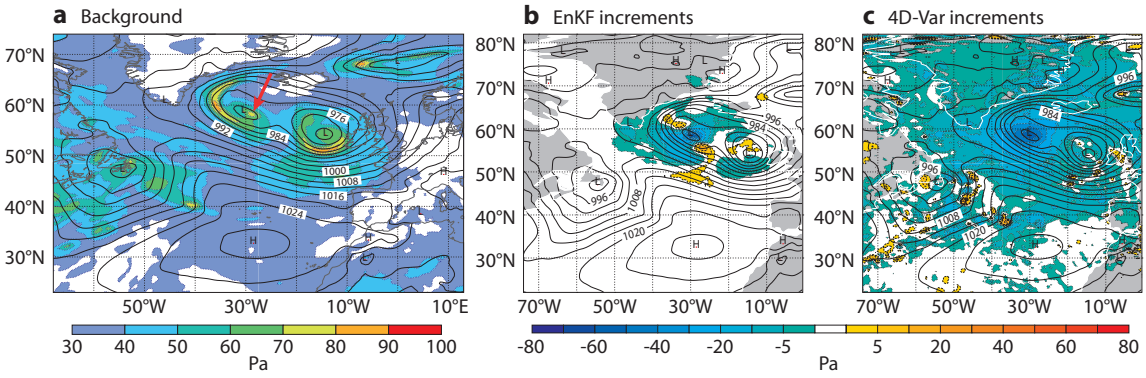

Figure 7 (a) Mean sea level (MSL) pressure control background forecast (black isolines, units: hPa) and standard deviation of MSL pressure ensemble background forecasts (shading) from a Hybrid Gain EnDA experiment valid on 25 February 2014 at 0000 UTC, (b) analysis increments for the single observation experiment with a synthetic surface pressure observation located at 58.5°N, 30.3°W (red arrow in left-hand panel) and 1hPa lower than the background forecast, in the EnKF mean analysis, and (c) analysis increments for the same observation experiment in the 4DVAR analysis.

doi:10.21957/4n5hz0y5

Outlook

The wish to carry out detailed comparisons between variational and more scalable ensemble-based assimilation systems has led to the development of a state-of-the-art EnKF system at ECMWF. The results of these comparisons are promising in terms of both forecast accuracy and computational cost, especially in the Hybrid Gain configuration. Research is continuing to further evaluate the system and also to verify whether these results, obtained at the resolution of the currently operational EDA, can be confirmed at the higher horizontal resolutions that are planned to be introduced later this year. Another important aspect to evaluate is the impact of using the Hybrid Gain analysis ensemble to initialize ECMWF's ensemble forecasts. The Hybrid Gain is one of several hybrid 4DVAR configurations that will be explored at ECMWF over the coming years as potential alternatives to the current hybrid 4DVAR/EDA system.

Further reading

Hamrud, M., M. Bonavita & **L. Isaksen,** 2014: EnKF and Hybrid Gain Ensemble Data Assimilation. *ECMWF Tech. Memorandum* **No 733.**

Hunt, B. R., E. J. Kostelich & **I. Szunyogh,** 2007: Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D,* **230,** 112–126.

Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher & **L. Raynaud,** 2010: Ensemble of data assimilations at ECMWF. *ECMWF Tech. Memorandum* **No 636.**

Penny, G. S., 2014: The Hybrid Local Ensemble Transform Kalman Filter. *Mon. Wea. Rev.,* **142,** 2139–2149.

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

Whitaker, J. S. & **T. M. Hamill**, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.,* **140,** 3078–3089.

 doi:10.21957/4n5hz0y5