Weight interpolation for efficient data assimilation with the Local Ensemble Transform Kalman Filter

Shu-Chih Yang\textsuperscript{1,2}, Eugenia Kalnay\textsuperscript{1,3}, Brian Hunt\textsuperscript{1,3} and Neill E. Bowler\textsuperscript{4}

\textsuperscript{1} Department of Atmospheric and Oceanic Science, University of Maryland
\textsuperscript{2} Global Modeling and Assimilation Office, NASA/GSFC
\textsuperscript{3} Institute for Physical Science and Technology, University of Maryland
\textsuperscript{4} Met Office, FitzRoy Road, Exeter, UK

Corresponding author:
Shu-Chih Yang
Email: scyang@atmos.umd.edu

Abstract

We investigate a method to substantially reduce the analysis computations within the Local Ensemble Transform Kalman Filter (LETKF) framework. Instead of computing the LETKF analysis at every model grid point, we compute the analysis on a very coarse grid and interpolate onto a high-resolution grid by interpolating the analysis weights of the ensemble forecast members derived from the LETKF. Because the weights vary on larger scales than the analysis fields or analysis increments, there is little degradation in the quality of the weight-interpolated analyses compared to the analyses derived with the high-resolution grid, and the results from the weight-interpolated analyses are more accurate than the ones derived by interpolating the analysis increments.
1. **Background**

The resolution of modern numerical models and observation density have greatly increased in recent years in order to resolve the dynamic processes in smaller scales such as convective scales. This burdens the computational cost of the data assimilation (DA) procedure, so that for variational analysis (3DVar and 4DVar), methods to reduce the heavy computational cost focus on reducing the computation during the minimization process of the cost function (Courier et al. 1994). In 4D-Var, the so-called inner loop is conducted by running the adjoint model at a low resolution and/or with simplified physics. For the ensemble methods as Ensemble Kalman Filter, the computational cost is alleviated by allowing the analysis to be computed in parallel in local regions (Keppenne et al. 2002, Ott et al. 2004 and Hunt et al., 2007). However, the computational burden for such local analyses is still constrained by the ensemble size and the number of total local regions in a high-resolution model. It is possible to reduce the computation further by carrying out the ensemble analyses on a coarse resolution, as done with variational analyses, and then interpolating to the finer resolution. Unfortunately, in both methods this interpolation step degrades the accuracy of the analysis, compared to a full-resolution analysis.

In this study, we investigate the feasibility of reducing the computational cost of the assimilation by using the output of an ensemble of full-resolution forecasts without sacrificing the accuracy of the analysis. This method is developed following a suggestion
of Bowler (2006, see section 6.4), who computed the transform matrix derived from the solution of Ensemble Transform Kalman Filter (Bishop et al. 2001) on a coarse grid and then interpolated it onto the high-resolution grid. A similar idea of weight interpolation is used to speed up the analysis computation of an ocean ensemble Kalman filter (EnKF, Keppenne et al. 2007). However, in these studies the properties of the weights and how the weight-interpolation ensures the analysis accuracy are not discussed.

In this study, the weight-interpolation method is adapted to the framework of the Local Ensemble Transform Kalman Filter (LETKF, Ott et al., 2004 and Hunt et al. 2007), which has the advantage of being configured for local analyses, and where the analysis ensemble is expressed with weights that linearly combine the forecast ensemble. The interpolation is done through these weights for the background ensemble in order to create an analysis at full resolution. As we will show, this method allows us to substantially reduce the required number of local analyses for LETKF assimilation without sacrificing analysis accuracy.

This paper is organized as follows: The implementation of the LETKF is briefly described in Section 2, including the setup of the assimilation experiments in a Quasi-Geostrophic model. In Section 3, we explain how the sparse analysis is done within the LETKF framework. The results of the interpolated data assimilation experiments are discussed in Section 4. A summary and discussion is given in Section 5.

2. Local Ensemble Transform Kalman Filter (LETKF)
The LETKF scheme, described in detail in Hunt et al. (2007), is an ensemble-based Kalman Filter that performs analysis within a local region, with the background error covariance in each region described by the corresponding local background ensemble. This scheme is a more efficient version of the Local Ensemble Kalman Filter proposed by Ott et al. (2004). The LETKF has been shown to have similar accuracy as other sequential ensemble-based Kalman Filters implemented in the same numerical weather prediction model (Whitaker et al. 2007), but its parallel implementation, made possible because the analysis at each grid point is independent of other grid points, becomes more efficient as the number of processors grows.

We briefly discuss the LETKF algorithm, where the data assimilation is performed at the analysis central grid point within a local region. The local error statistics are estimated based on the background states and the available observations within this local region. The LETKF determines a transform matrix that converts the local background ensemble perturbations into the analysis ensemble perturbations. The local analysis error covariance can be written as (1), where $K$ is the ensemble size, $X_f$ is the matrix whose columns are the background (forecast) ensemble deviations from the background ensemble mean, $X_a$ is the corresponding matrix of the analysis ensemble perturbations, $P_a$ is the analysis error covariance and $\tilde{P}_a$, the analysis error covariance in ensemble space.

$$\mathbf{P}_a = \frac{1}{K-1} \mathbf{X}_a \mathbf{X}_a^T = \mathbf{X}_f \tilde{\mathbf{P}}_a \mathbf{X}_f^T,$$  (1)

and the transform matrix is computed as:

$$\mathbf{P}_a = \frac{1}{K-1} \mathbf{X}_a \mathbf{X}_a^T = \mathbf{X}_f \tilde{\mathbf{P}}_a \mathbf{X}_f^T,$$
\[
\tilde{P}_a = [(HX_f)^T R^{-1} (HX_f) + (K-1)I]^{-1}.
\] (2)

In (2), \(H\) is observation operator that converts variables from model space to observation space and \(R\) is the observation error covariance. Both \(H\) and \(R\) depend on the region because only local observations are used. The matrix \(\tilde{P}_a\) is efficiently computed within the ensemble space. After \(\tilde{P}_a\) is obtained, the mean analysis, \(\bar{x}_a\), at the central grid point of the local region is computed from the background ensemble mean, \(\bar{x}_f\), according to (3).

\[
\bar{x}_a = \bar{x}_f + X_f \tilde{P}_a (HX_f)^T R^{-1} (y_o - H(\bar{x}_f)) = \bar{x}_f + X_f \bar{w}_a
\] (3)

In (3), the \(K \times 1\) vector of weight \(\bar{w}_a\), carries the information about observational increments.

In the final step, the analysis ensemble perturbations at the central grid point are derived by multiplying the background ensemble perturbations by the symmetric square root of \((K-1)\tilde{P}_a\):

\[
X_a = X_f ((K-1)\tilde{P}_a)^{1/2} = X_f W_a.
\] (4)

In (4), \(W_a\) is a multiple of the symmetric square root of the local analysis error covariance in the ensemble space. The use of a symmetric square root ensures that the sum of the analysis ensemble perturbations is zero and depends continuously on \(\tilde{P}_a\).

Therefore, adjacent points with slightly different \(\tilde{P}_a\) will yield similar analysis ensemble perturbations, necessary to ensure the smoothness of the analysis (Hunt et al. 2007). This property of the symmetric square root also ensures that the analysis ensemble
perturbations are consistent with the background ensemble perturbation since the $W_a$ is the square root matrix closest to the identity matrix, given the constraint of the analysis error covariance (Ott et al. 2004). Harlim (2006) demonstrated that the symmetric solution has better performance than one obtained with a non-symmetric square root, given the same ensemble size.

Eqs. (3) and (4) show that the analysis ensemble at each grid point is simply a linear combination of the background ensemble, with weighting coefficients given by $\tilde{W}_a$ (a $K \times 1$ vector) for the mean analysis, and $W_a$ (a $K \times K$ matrix) for the perturbation. So, the $k_{th}$ analysis ensemble member is given by

$$x_{a,k} = \bar{x}_f + X_{f,k} \left[ \tilde{W}_a + W_{a,k} \right]$$

(5)

where $W_{a,k}$ is the $k_{th}$ column of $W_a$. Let $v$ denotes a vector of $K$ ones, which is an eigenvector of the computed $W_a$ at each analysis grid point (Hunt et al. 2007). The zero mean of the background ensemble perturbations, $X_f v = 0$, implies that

$$X_a v = X_f W_a v = 0,$$

(6)

which means that the analysis perturbations also have a zero mean.

In this study, the LETKF is implemented on a quasi-geostrophic (QG) channel model. There are 64 grid-points in zonal direction, 33 grid-points in meridional direction and a total of 7 variables in the vertical. The model physics include processes of advection, diffusion, and relaxation at all levels and Ekman pumping at the bottom level, with a
zonally periodic solution (Rotunno and Bao 1996). The model variables are potential vorticity, in the five internal levels, and potential temperatures at the bottom and top levels. They are the analysis variables in the DA experiments. Observations are generated by adding random Gaussian errors to the true profiles of zonal and meridional winds and potential temperatures.

Yang et al. (2007) compared the performance of LETKF with other DA schemes in this model. They showed that with a perfect model configuration (no model errors), the analysis derived from the LETKF was more accurate than the analyses from the 3D-Var or from the 4D-Var with a short assimilation window (12-hour), but with 24-hour windows 4D-Var was more accurate. A similar DA setup as in Yang et al. (2007) is used here to test the method of interpolated LETKF analyses. There are 128 “rawinsonde” observations uniformly distributed within the model domain. The observation error is 0.8 ms$^{-1}$ for the zonal wind, 0.5 ms$^{-1}$ for the meridional wind and 0.8°C for potential temperature. In this study, we allow the model to be imperfect by changing the vertical mixing parameter used for the Ekman pumping from 5 for the truth run to 4.75 for the data assimilation cycle.

For the LETKF analysis, 20 ensemble members are used in this study and the local patch is chosen to be 7×7 grid-points in the horizontal and to include the whole column (7 levels). A Gaussian function with a decorrelation length of 5 grid points is applied to the observation error covariance to reduce the impact of more distant observations on the
analysis (Miyoshi, 2005). In addition to a multiplicative covariance inflation of 8%, an additive perturbation is used to optimize the LETKF performance by adding a very small amount of random perturbations onto the analysis ensemble perturbations (Corazza et al., 2007).

3. **Interpolation of the analysis weights**

The local analysis error covariance in the LETKF is estimated by combining the contributions from each available observation in each local region. Such contributions are represented in the local ensemble space by the weighting coefficients ($\tilde{w}_a$, $W_a$ in Eqs. (3-4)). The same information from observations and background states is used over several regions due to the overlapping between local regions. This ensures that the weights vary slowly and allows us to perform a local analysis on a limited number of grid points (a coarse-resolution grid), and to spread out the information of the error statistics to the higher-resolution grid through the interpolation of the weights.

Figure 1 illustrates how the sparse analysis is done within the LETKF configuration. The background ensemble is available at all the grid points of high resolution denoted as dots and crosses in Figure 1. The dots, arranged on a coarse grid, denote the grid points where the LETKF analyses and the weighting coefficients ($\tilde{w}_a$ vector and $W_a$ matrix) are actually computed. In this example, only one analysis is computed for every $3 \times 3$ grid box so that the coverage of the grid (i.e., the number of analysis grid points divided by the
total number of grid points, except for modifications at the northern and southern boundaries) is 11%. After the weight coefficients are collected on the coarse grids, we interpolate $\mathbf{\tilde{w}}_a$ and $\mathbf{W}_a$ onto the high-resolution grid points where no analysis has been computed, in order to generate fields of weights. Corresponding to each forecast ensemble member, we will have one map of weight coefficients associated with the mean analysis increment, and $K$ maps of weights associated with the analysis ensemble perturbations. As discussed in Section 2, the first map of weights represents the observational impact in correcting the mean state, and the latter $K$ maps apply “errors of the day” structures based on the dynamic evolution of the ensemble perturbations to generate the analysis perturbations. Yang et al. (2007) showed that both these two sources of information are important in improving the accuracy of the LETKF analysis. The unstable space of growing perturbations dominates the forecast errors, and the ensemble perturbations provide information that makes the analysis increments project well onto this unstable space.

For the interpolation of the weights, we apply a smooth bi-variate interpolation scheme (Akima, 1978) based on locally fitting quintic polynomials as the function of the zonal and meridional positions of the analysis grid. The chosen interpolation scheme is linear in the interpolated values. This ensures that the vector $\mathbf{v}$ of $K$ ones is still an eigenvector of the interpolated $\mathbf{W}_a$, so that the analysis ensemble perturbation calculated from the interpolated weights will maintain the property of zero mean as in Eq (6).

We also conducted assimilation experiments with even coarser analysis grids than the one shown in Figure 1, so that there is one LETKF analysis computed in every $3\times3$, $5\times5$, or $7\times7$ horizontal grid-box. The corresponding analysis coverage is 11%, 4% and 2%.
respectively. We will show in the next section that even though the number of total local analyses is substantially reduced, the analyses derived by interpolating the weights to full-resolution show little degradation.

Bowler (2006) indicated that the spatial consistency of the weighting coefficients is what makes advantageous the interpolation of weight to perform analyses on a coarse grid. With the use of the symmetric square root matrix, the weighting coefficients derived at adjacent points are typically close to each other, as illustrated in Figure 2(a)-(d) showing the weight coefficients for the observational correction of the background ensemble mean. We choose to show as an example the contribution from the 4th ensemble member (an element of the $\bar{w}_a$ vector of maps) to the analysis mean. Figure 2(a) is the weight map obtained from the full analysis at an arbitrary time. The empty spots in Figure 2(a) are areas where there are no observations available in the corresponding local patch and the background ensemble is therefore not updated in these regions. Figure 2(b)-(d) are the interpolated maps of this weighting coefficient with coarser analysis grids. Because the weights tend to be consistent at adjacent points, the interpolated weight structures can represent the original structures reasonably well.

We now examine the weights applied to form the analysis ensemble perturbations, based on a diagonal element of $W_a$ and an off-diagonal element. Figure 2(e)-(h) show the maps of the weight coefficients of the first element in the first column of the matrix $W_a$, representing the contribution from the first background ensemble perturbation to the first
analysis ensemble perturbation. The values derived from the full analysis coverage are shown in Figure 2(e). Figure 2(f)-(h) are the same weights after interpolation on coarser analysis grids. As could be expected from (2) and (4), the first ensemble forecast perturbation has the most influence in determining the first analysis perturbation, with weights ranging from 0.7 to 1.0. Regions with lower values in these maps indicate that the contributions from other ensemble become also important. The main features of the weight map derived from full analysis coverage are well recovered in the interpolated maps, except for the 2% case, in which the patterns are smoothed out. As an example of the off-diagonal elements, Figure 2(i)-(l) show the maps of the weight coefficients of the fourth element in the first column of the matrix $W_a$, representing the contribution from the fourth background ensemble perturbation to the first analysis ensemble perturbation. The large values in Figure 2(i)-(l) correspond to small values appearing in Figure 2(e)-(h). This again indicates that at this location, the other ensemble perturbations contribute more. Overall, their amplitude is much smaller compared to Figure 2(e)-(h), due to the property that $W_a$ is close to the identity matrix. The main features shown in Figure 2(i) are well recovered in Figure 2(j)–(l) after interpolation.

Besides reducing the analysis computation, the interpolated weights also provide an additional benefit in obtaining reasonable weights for those local patches without available observations (the empty points in the interior domain of Figure 2 (a), (e) and (i)). For these points, $\tilde{w}_a$ is simply a vector of zeros and $W_a$ is equal to the identity matrix in the original analysis setting without weight interpolation, and these values may not be optimal. This interpolation provides a feasible way to update the background ensemble in
the regions that have not been observed, or that are under-observed compared to neighboring regions.

4. Analysis Results

In this section, we compare the LETKF sparse analyses constructed from weight interpolation with the analysis derived at full resolution. For comparison, we also conduct sparse analyses using interpolated analysis increments, after running the full-resolution LETKF for 20 days. The increment interpolation is a traditional method used to convert a coarse resolution analysis into a full-resolution analysis. The time series of the root mean square of analysis error in terms of the potential vorticity is shown in Figure 3. The LETKF analysis derived from full resolution or weights interpolation from sparse grid-point outperform the 3D-Var analysis derived at the high (full) resolution. Results show that the sparse analyses with the interpolated weights at different analysis coverage (blue lines) have a quality similar or even slightly better to the one obtained at full resolution. Such result suggests that the interpolated weights are very useful in retaining the quality of the analysis. The analysis accuracy from the interpolated weights becomes poorer only when the analysis grid is so coarse (2% or less) that the local patches of each analysis do not overlap. In addition, the sparse analyses with the interpolated increments (shown as red lines) are more sensitive to the sparseness of the analysis grid-point and have much lower analysis accuracy than the ones derived from the interpolated weights.

By contrast, the analysis accuracy from the interpolated increments remains satisfactory only with a high analysis coverage of 50%, where the analysis grid-point is arranged as a
staggered grid available every other grid-point. Once the coverage decreases, the accuracy degrades quickly. With the 25% analysis coverage, the analysis increments are smoothed out and stretch isotropically and thus the analysis has accuracy similar to the 3D-Var analysis. With a lower analysis coverage of 11%, the LETKF with the interpolated increments diverges (the dashed red line in Figure 3). These results show that interpolating the analysis increments or the full fields (not shown) leads to a serious degradation of the analysis, whereas by interpolating the weights, we retain the analysis accuracy and the advantages of the LETKF and the efficiency of low-resolution analysis.

The results shown in Figure 3 can be understood by the structures of the analysis increments (the differences between analysis and background) from the sparse analyses constructed with interpolated weights. The analysis increments, as shown for potential vorticity at the mid-level, in Figure 4, represent forecast errors stretched by the flow, and they have elongated structures and scales similar to that of bred vectors (Corazza et al. 2002).

We found that the analysis increments obtained from interpolated weights at different analysis coverage (11%, 4% and 2%) are very similar, as shown in an example in Figure 4(a)-(d). It is clear that interpolating the weights succeeds in recovering quite well the full analysis increments, and because of the large scales of variations in the weighting maps, the obtained analysis increments are insensitive to the analysis coverage. By contrast, if the interpolation is done on the analysis increments, like Figure 4(e)-(h), the analysis increment structures are quickly smoothed out as the analysis grid becomes coarser. Only
with a high analysis-coverage of 50%, can the local characteristics in the analysis increment at full resolution be retained and thus maintain an advantage over 3D-Var. In Figure 4(h), the pattern of interpolated analysis increment with 2% analysis coverage has an unrealistic large-scale feature. As could be expected, this will impose false corrections to the background state and lose the advantage of using the time-dependent error statistics in the LETKF.

5. Summary

In this study, we investigated an efficient method to reduce the analysis computational cost within the framework of the LETKF following a suggestion by Bowler (2006). The LETKF analyses are computed on coarse grids, but the weights used to update the background ensemble are interpolated onto the high-resolution grids. Instead of repeatedly using the observations and the background ensemble to perform the LETKF, the analysis at the high-resolution is now derived through estimating the interpolated weights. In LETKF, the weights of the analysis ensemble represents two sources of information: one is associated with the observations contribution to the mean background state and the other is associated with the dynamically evolving error structures obtained from background ensemble perturbations. Interpolations are done separately on these two components of weights by taking the advantage of the symmetric square root solution of the transform matrix used in LETKF.
We showed that the weights derived from LETKF are smooth and consistent among nearby points so the interpolated weight maps for a coarse analysis-grid can represent evolving features very well. Furthermore, interpolation results in some smoothing of the weights, which may provide additional balance in the analysis increments. Because the weights vary on larger scales than the analysis or analysis increments, there is little degradation in the quality of the weight-interpolated analyses compared to the analyses derived with the high-resolution grid. Therefore, the corresponding analysis accuracy is still high even when we reduce the analysis grid to just 2% of the grid points of the full resolution. In addition, the results are insensitive to the analysis coverage in this study (but they are sensitive to the size of the local patch and the ensemble as discussed in Yang et al. 2007). The results also show that interpolating the weights gives an analysis that is much more accurate than the analysis constructed from interpolating the analysis increments (or the full fields, not shown). The quality of interpolation of the analysis increments is much more sensitive to the sparseness of the analysis-grid due to their characteristic stretched dynamical scales.

Besides the purpose of performing sparse analyses, the weight-interpolation can also be used to provide analysis weights for regions without local observations, so that instead of returning the background ensemble values without updating them, as in the original LETKF procedure, the analysis ensemble can still be computed for these under-observed regions. This advantage of smoothing the weights in handling under-observed regions may explain why the interpolated weight results are not just comparable but even slightly better than the full resolution analysis, so that smoothing the weights may be
advantageous even when using the full resolution analysis grid. We also point out that any conserved quantity that is a linear function of the model state will be equally conserved in the original LETKF analysis and in the analysis performed with the interpolated weights. With the method of weights interpolation, the analysis remains in the subspace of the forecast ensemble, so that properties such as conservation of total mass and balance, satisfied by each ensemble member, are also satisfied in the analysis.

Although the goal of this study was to test whether sparse analyses with interpolated weights is a suitable method for operational use, the results obtained in this study may be over-optimistic since the quasi-geostrophic model used in the experiments was only slightly imperfect.

6. References


Soc.*, **120**, 1367-1387.

the breeding technique in the estimation of the background error covariance matrix
for a quasigeostrophic model. AMS Symposium on Observations, Data Assimilation

----------------, E. Kalnay, and S.-C. Yang, 2007: An implementation of the Local
Ensemble Kalman Filter for a simple quasi-geostrophic model: Results and
comparison with a 3D-Var data assimilation system. *Nonlinear Processes in


Harlim, H., 2006: Errors in the initial conditions for numerical weather prediction: A
study of error growth patterns and error reduction with ensemble filtering. *PhD
dissertation*, University of Maryland, 86 pages.

Keppenne, C.L., and M.M. Rienecker, 2002: Development and initial testing of a parallel
Ensemble Kalman filter for the Poseidon isopycnal ocean general circulation model.

-------------------, M. M. Rienecker, J. P. Jacob and R. Kovach, 2007: Error covariance

Miyoshi, T., 2005: Ensemble Kalman filter experiments with a Primitive-Equation global
at http://hdl.handle.net/1903/3046.


7. Acknowledgements

We are grateful for helpful interactions with Jeff Whitaker, Istvan Szunyogh, Eric Kostelich, and Milija Zupanski. Rebecca Morss and Matteo Corazza provided the original 3D-Var code. S.-C. Yang was supported by NASA grants NNG004GK78A and NNG06GB77G.
**Figure Captions**

**Figure 1** The grid arrangement for an 11% coarse analysis. The dots are the grids where the LETKF analysis is actually performed. The crosses indicate high-resolution grid points whose analysis will be derived by weight-interpolation or increment interpolation (see the explanation in the text).

**Figure 2** (a) The weighting coefficients corresponding to the fourth element of the LETKF mean weighting vector for updating the background ensemble mean, derived with 100% analysis grid-point coverage (full resolution); (b)-(d) the same as (a), except that the weight is interpolated at the 11%, 4% and 2% analysis grid coverage. (e)-(h) are the first element of the first column of the LETKF weighting matrix for constructing the first analysis ensemble perturbation derived at the same resolutions used in (a)-(d). (i)-(l) are the fourth element of the first column of the LETKF weighting matrix for constructing the first analysis ensemble perturbation derived at the same resolutions used in (a)-(d). The empty spots denote the local regions with no available observations. Their weight values are therefore zero for (a) and (i) and one for (e).

**Figure 3** The time series of the RMS analysis error in terms of the potential vorticity from different DA experiments. The LETKF analysis from the full-resolution is denoted as the black line and the 3D-Var derived at the same resolution is denoted as the grey line. The LETKF analyses derived from weight-interpolation with different analysis coverage are indicated with blue lines. The LETKF analyses derived after the first 20 days from increment-interpolation with different analysis coverage are indicated with the red lines.
Figure 4 The analysis increment for potential vorticity at the 3rd level at an arbitrary time from the LETKF with (a) a full resolution, (b) obtained through interpolated weights with 11% analysis coverage, (c) the same as (b) except for 4% analysis coverage, (d) the same as (b), except for 2% analysis-grid coverage. (e)-(h) are the interpolated analysis increments computed by taking 50%, 11%, 4% and 2% analysis coverage of analysis increment obtained at the full-resolution and interpolating back to the full-resolution.
Figure 1 The grid arrangement for an 11% coarse analysis. The dots are the grids where the LETKF analysis is actually performed. The crosses indicate high-resolution grid points whose analysis will be derived by weight-interpolation or increment interpolation (see the explanation in the text).
Figure 2 (a) The weighting coefficients corresponding to the fourth element of the LETKF mean weighting vector for updating the background ensemble mean, derived with 100% analysis grid-point coverage (full resolution); (b)-(d) the same as (a), except that the weight is interpolated at the 11%, 4% and 2% analysis grid coverage. (e)-(h) are the first element of the first column of the LETKF weighting matrix for constructing the first analysis ensemble perturbation derived at the same resolutions used in (a)-(d). (i)-(l) are the fourth element of the first column of the LETKF weighting matrix for constructing the first analysis ensemble perturbation derived at the same resolutions used in (a)-(d). The empty spots denote the local regions with no available observations. Their weight values are therefore zero for (a) and (i) and one for (e).
**Figure 3** The time series of the RMS analysis error in terms of the potential vorticity from different DA experiments. The LETKF analysis from the full-resolution is denoted as the black line and the 3D-Var derived at the same resolution is denoted as the grey line. The LETKF analyses derived from weight-interpolation with different analysis coverage are indicated with blue lines. The LETKF analyses derived after the first 20 days from increment-interpolation with different analysis coverage are indicated with the red lines.
Figure 4 The analysis increment for potential vorticity at the 3rd level at an arbitrary time from the LETKF with (a) a full resolution, (b) obtained through interpolated weights with 11% analysis coverage, (c) the same as (b) except for 4% analysis coverage, (d) the same as (b), except for 2% analysis-grid coverage. (e)-(h) are the interpolated analysis increments computed by taking 50%, 11%, 4% and 2% analysis coverage of analysis increment obtained at the full-resolution and interpolating back to the full-resolution.