Flow Dependent Jb in a global grid-point 3D-var

J.C. Derber, R.J. Purser, W.-S. Wu, R. Treadon, M. Pondeca, D. Parrish, and D. Kleist

Environmental Modeling Center
NWS/NOAA
Camp Springs, MD USA

1. Introduction

The specification of the background error covariance has long been thought to be extremely important for determining the quality of analyses and resulting forecasts. While more complete and appropriate than many of the background error covariances used in OI systems, the current operational background error covariances still are fairly simple and inappropriate in many situations. For example, many current 3D and 4D implementations use spectrally defined homogeneous and isotropic background error covariances. For the analysis of hurricanes, frontal structures, precipitation, and other rapidly changing structures this definition of the background errors is probably inappropriate. NOAA/NWS/NCEP/EMC is attempting to improve our definition of the background errors by redefining the our analysis variables in grid space and using recursive filters to describe the background errors.

In spectral space, very reasonable homogeneous isotropic background error covariances can be specified using a simple diagonal matrix. When the forecast model is spectral and global, the spectral formulation of the analysis variables and background error has advantages with the formulation consistent with the model. Also, in spectral space the singularities at the poles are easy to handle. However, the extension of the spectral formulation to include inhomogeneous, anisotropic error covariances is not straightforward.

By defining the analysis variables in grid space and using the recursive filters to define the background error, we believe offers several advantages over the spectral definition of these quantities. These advantages include easier creation of inhomogeneous anisotropic background errors (still far from trivial), a local definition of the errors, easy discrimination between major features such as land-sea, tropics-midlatitudes, etc. and the ability to easily apply the same system to both regional and global systems.

Note that there are two major considerations with the background error covariance. First, one must be able to computationally describe the appropriate background error structures. We believe that we can do this using recursive filters. Second, one must determine the appropriate background error structures. Several ways of defining the background errors have been proposed (e.g., defining them along isentropic surfaces, defining them in semi-geostrophic coordinates). However, none of the proposed structures appear to be appropriate in all 3-D situations. The definition of the background errors is an ongoing (and will likely continue to be for many years) research problem at NOAA/NWS/NCEP/EMC and other organizations.

2. Recursive Filters

Recursive filters are the basic building blocks used to create background error covariance structures in the grid point analysis system. The general 3-D structures possible using recursive filters are created by application of a series of simple 1-D recursive filters. In this paper, simpler and more straightforward applications of the recursive filter will be described. More general applications to more complex problems and the underlying theory can be found in Purser et al., 2003a and b). We note that the recursive filters are closely related to the diffusion operator methods used in Derber and Rosati (1989), Weaver and Courtier (2001) and presented in a paper in this volume by Anthony Weaver.

The general 1-D form of the recursive filter has two steps. First the advancing step:

$$q_i = \beta p_i + \sum_{j=1}^{n} \alpha_j s_{i+j}$$

where p is the input and q is the output of the advancing step and n is the order of the recursive filter. Followed by the backing step:

$$s_i = \beta q_i + \sum_{j=1}^{n} \alpha_j s_{i+j}$$

where s is the final output from the recursive filter. Note that

$$\beta = 1 - \sum_{j=1}^{n} \alpha_j$$

and that the recursive filter is self-adjoint.

For a first order (n=1) recursive filter the equations simplify to:

$$q_i = (1-\alpha) p_i + \alpha q_{i-1}$$
$$s_i = (1-\alpha) q_i + \alpha s_{i-1}$$

From these equations, it can be seen that a single impulse in the input will quickly be propagated to the right (i increasing) by the advancing step and the propagated to the left (i decreasing) by the backing step. A single application of the recursive filter produces a quasi-Gaussian (see Fig. 1) response. As can be seen from Fig. 1, higher order filters can produce results closer to Gaussian as can repeated applications of lower order filters (with appropriate coefficients).

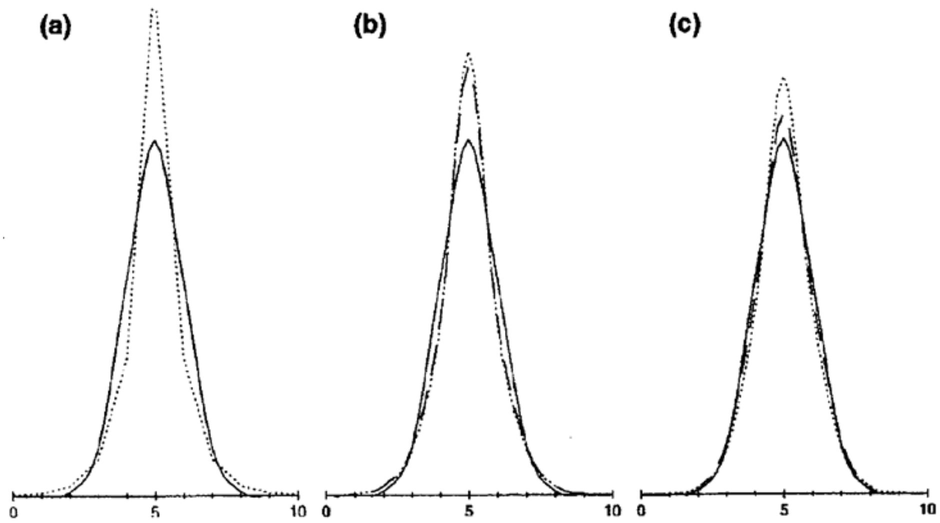

Figure 1. Comparison of one-dimensional applications of recursive filters approximating a Gaussian (solid). Long dashed curves show filter approximations: (a) order n=1, (b) n=2, and (c) n=4. From Purser et al. 2003a.

To create 2D isotropic fields, the recursive filter is applied first along one direction (e.g., along the x axis) and then in the other (e.g., along the y axis). Results for the creation of 2D isotropic fields are given in Fig. 2. Note that the higher order recursive filters (e.g., n=4) give a much better approximation to a Gaussian than lower order filters (e.g., n=1). Also note that multiple applications of the low order recursive filters (n=1x4 - four applications of a first order recursive filter) also greatly improves the approximation to a Gaussian.

Of course for many applications, the derivatives of the fields are very important. For that reason, it is important to ensure that the recursive filter not only approximates the Gaussian reasonably well, but also the derivatives are well approximated. In Fig. 3, the negative of the Laplacian of the fields after the application of a 2-D recursive filter approximating the Gaussian is shown. It is clear that the second order (*n*=2) recursive filter is not sufficient. However, 4th order and higher recursive filters reasonably approximate the Gaussian. Based on this result and others, we believe that one should use at least a 4th order recursive filter.

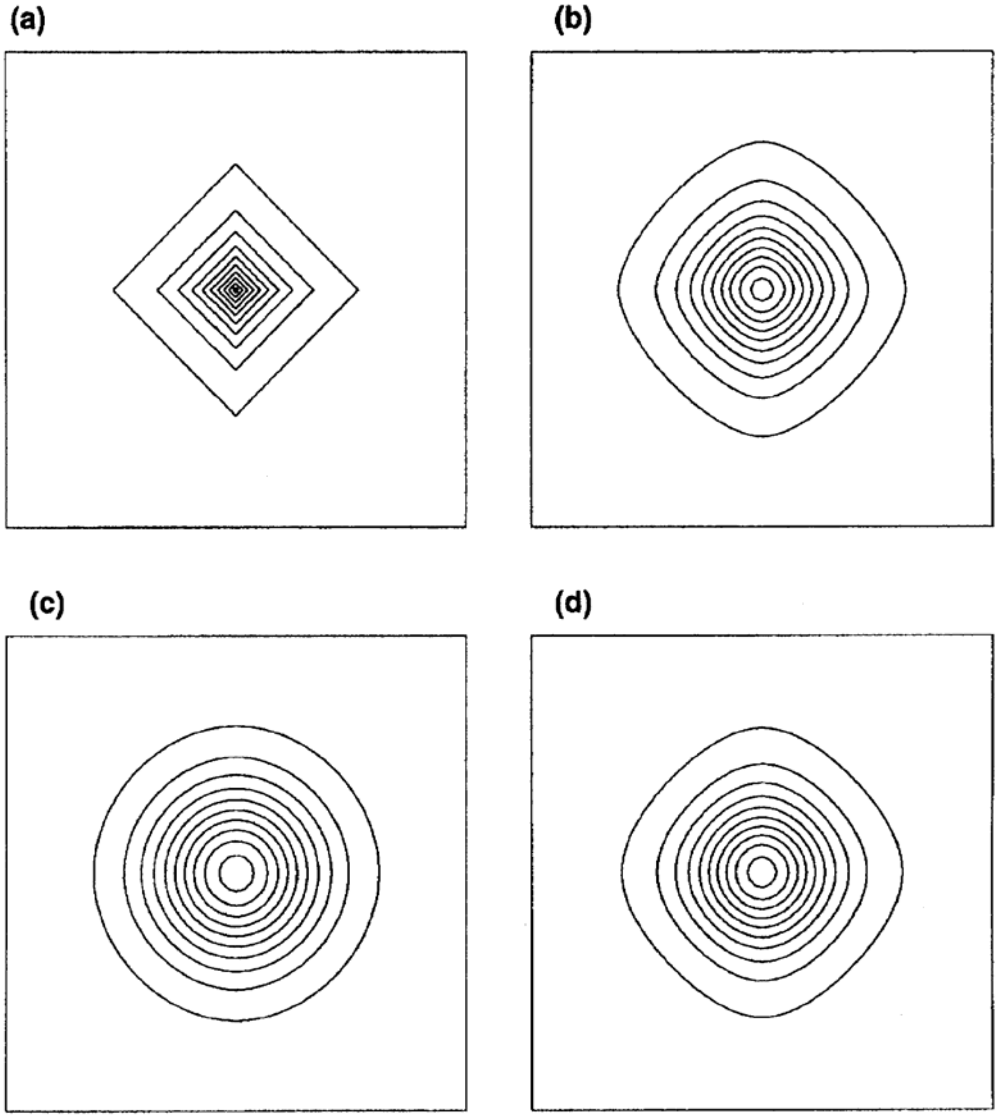

Figure 2: Sequential application of quasi-Gaussian recursive filters of order n in two dimensions (a) n = 1, (b) n = 2, (c) n = 4 and (d) four applications of filters with n = 1 with scale parameters adjusted to make the result comparable with the other single-pass filters. From Purser et al. (2003a).

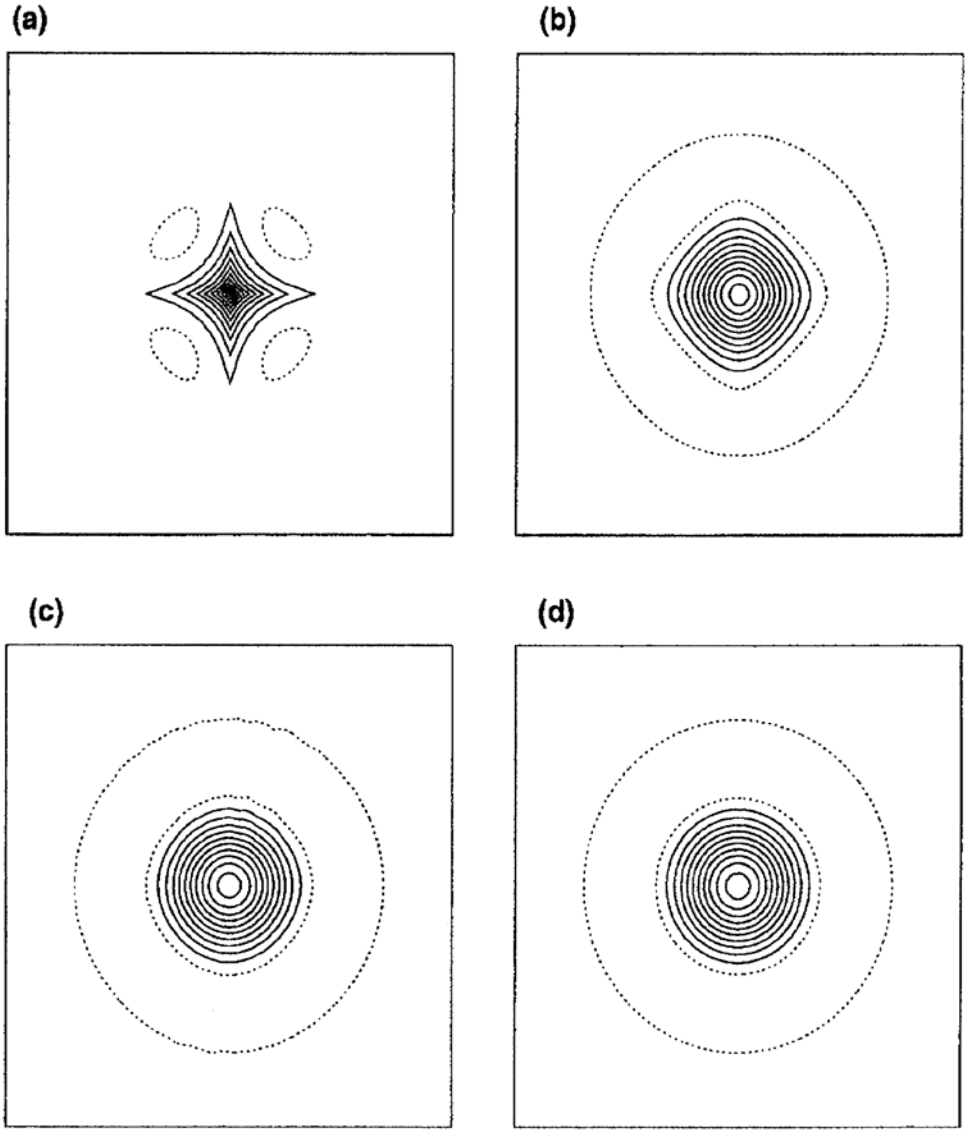

Figure 3: Negative Laplacian applied to quasi-Gaussian recursive filters with (a) n =2, (b) n = 4, (c) n =6, and (d) corresponding contours for the exact Gaussian. Contours are at multiples of odd numbers with negative contours shown as dotted curves. From Purser et al. (2003a).

Of course not every background error structure is Gaussian. Non-Gaussian isotropic structures can be constructed using by adding Gaussians (and Laplacians of Gaussians)of different structures. By doing this one can create fat-tailed error covariance which give greater weight to smaller scales than Gaussians. The fat-tailed nature of the covariances can be especially important in the transition regions between high density data and lower density data. By using fat-tailed covariances, it is less likely that strong gradients will be extrapolated long distances into the lower density data region.

The creation of inhomogeneous anisotropic covariances on the globe is a bit more complicated. For general covariances it is necessary to apply the 1D recursive filters not along the x, y or z axis but along directions determined by the structure of the covariances. This can be accomplished by use of the triad (2D) and hexad (3D) algorithms as discussed in Purser et al. (2003b). In these algorithms, the coefficients for the 1D recursive filters are determined by the local aspect tensors defined at each grid point. Some problems have been noted in the past in cases where the orientation of the covariances shifts by 90 degrees over short distances. This problem has been solved by defining a bridging function between the two orientations. An example of before and after is shown in Fig. 4.

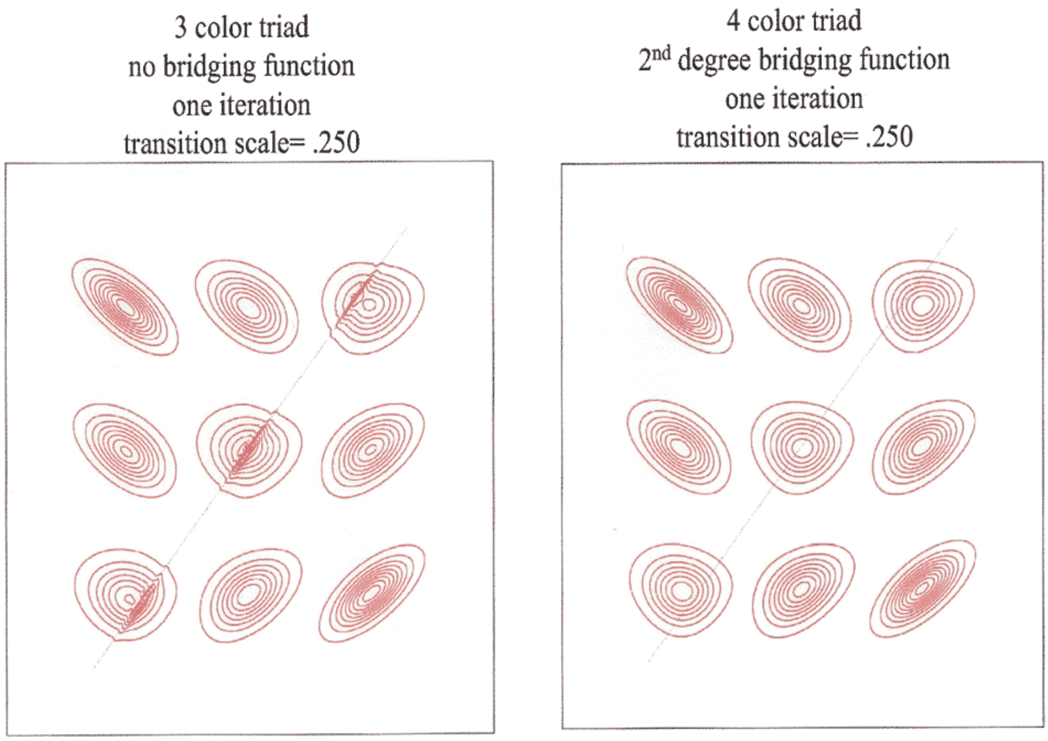

Figure 4: Correlation functions when strong (90o)transition in orientation over short distance. Left without bridging function. Right with bridging function.

The application of the recursive filters on the globe is not straightforward at the pole. There are several potential practical solutions to this problem. We have chosen to initially define three regular grids (tropical-midlatitude and two polar stereographic grids). The fields are interpolated to these three grids, the recursive filters are then applied on each grid and the grids recombined. Note that the interpolation and recombination must be done carefully to preserve the self-adjointness of the background error covariance matrix.

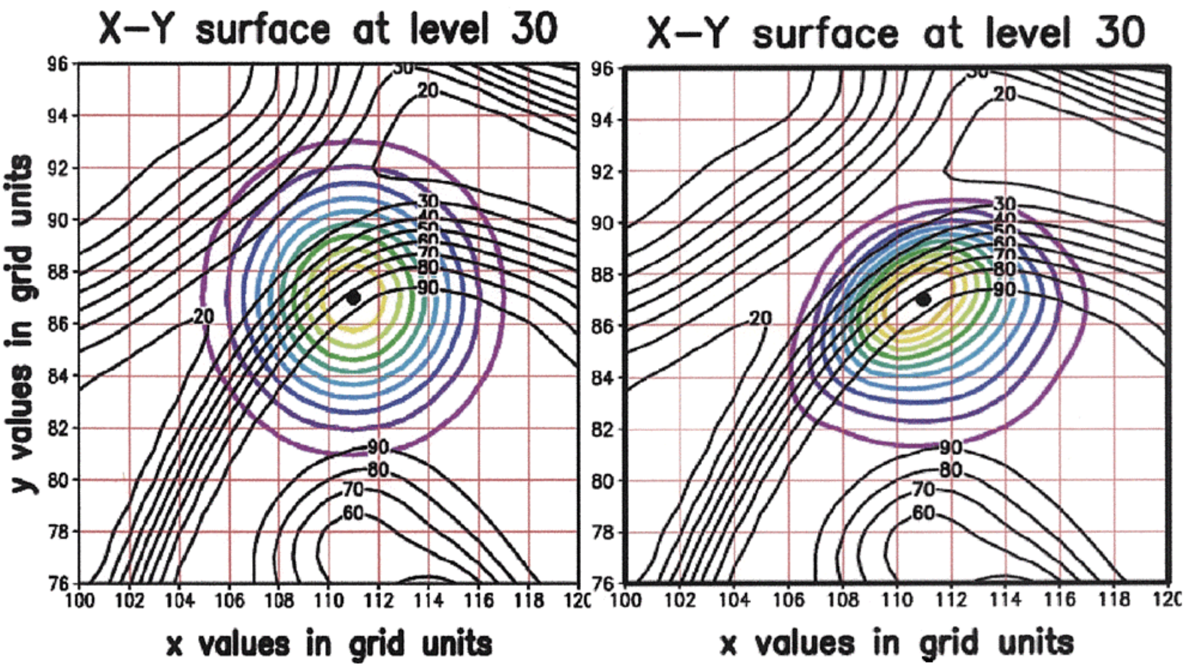

Figure 5: x-y section through 3-D moisture analysis with isotropic covariance functions (left) and anisotropic covariance functions (right)

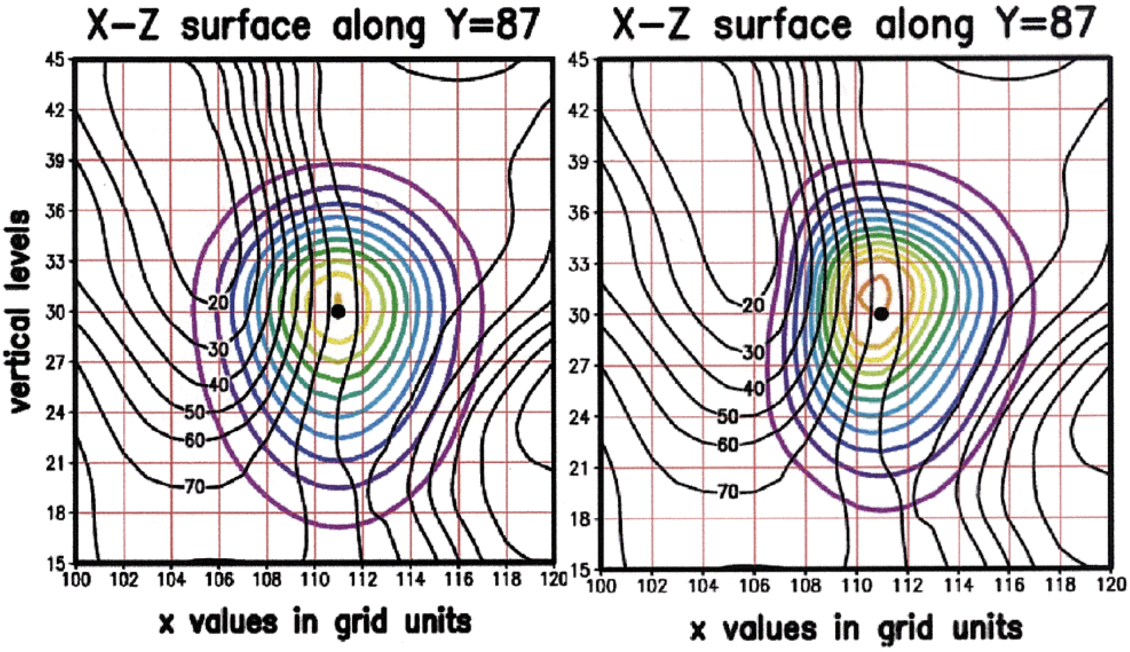

Figure 6: x-z cross section through 3-D moisture analysis with isotropic covariance functions (left) and anisotropic covariance functions (right).

By use of the recursive filters, general situation dependent background error covariance structures can be defined. For example, a moisture covariance matrix can be defined which is stretched along constant values of relative humidity. The resulting covariances for both the isotropic and anisotropic cases are shown in Fig. 5 for a horizontal section and in Fig. 6 for a vertical section. Note that the structures revert to isotropic in the regions of weak gradients in the relative humidity and that across strong gradients the correlation lengths are greatly reduced.

3. NCEP's initial global grid point version

The current pre-implementation version of the NCEP's global analysis system has been transformed to define the background term in grid space and to use recursive filters to define the background error covariances. We have attempted to get keep the versions as close as possible to each other, with the differences limited to the background error calculations, the modified error statistics, changes to the balance equation, a different minimization algorithm and the removal of the divergence tendency equation constraint (currently not used operationally). The version described below should not be interpreted as our final version, but rather a starting point in the development of our first operational grid point version. An earlier, but similar version of this code is described in Wu et al. (2002).

The background term initially assumes the form:

$$B = B_V^{T/2}(B_H^1 + B_H^2 + B_H^3)B_V^{1/2}$$

where $B_V^{1/2}$ includes the vertical component of the recursive filter and the balance relationships and $B_H^1 + B_H^2 + B_H^3$ represents three horizontal applications of the recursive filters. This form assumes the horizontal and vertical covariances are separable. Note that the vertical component $B_V^{1/2}$ can be incorporated into the definition of the analysis variables as is currently done for the complete background error covariance in the operational NCEP and ECMWF systems (e.g., Derber and Bouttier, 1999).

In the initial experiments we are attempting to produce a system which reasonably approximates the current spectral system. For this reason, we have produced horizontal error covariances which are homogeneous and isotropic in physical space. Note that since the system is solved on the model's Gaussian grid, the horizontal error covariances are not homogeneous in grid space and it was necessary to include the capability to create inhomogeneous covariances in the system. It is more difficult to duplicate the structures in the spectral version for the vertical terms because of the interaction of the balance equation and the vertical structures. Therefore for the vertical term, the vertical covariance terms (as are the balance equation terms) are defined as a smoothly varying function of latitude. The specific statistics used to define the vertical terms and B_H^1 are done using the NMC method. The lengths scales used in B_H^2 and B_H^3 are 1/2 and 1/4 those in B_H^1, respectively. This results in a fat-tail distribution for the covariances in the horizontal.

The balance equation used in the grid point version is fairly simple, but appears to be reasonably effective. The regression matrices are a function of latitude and are of the form:

$$T = A\psi$$
$$\chi = C\psi$$
$$p_s = D\psi$$

where A, C and D are empirically defined (using the NMC method) matrices. The global mean explained variances in the temperature and velocity potential field are shown in Fig. 7. For the surface pressure, 86% of the variance is explained by the balance equation. A significant percentage of the variance is explained by the balance equation throughout the atmosphere except near the model top. For the velocity potential, the balance equation only becomes important near the surface.

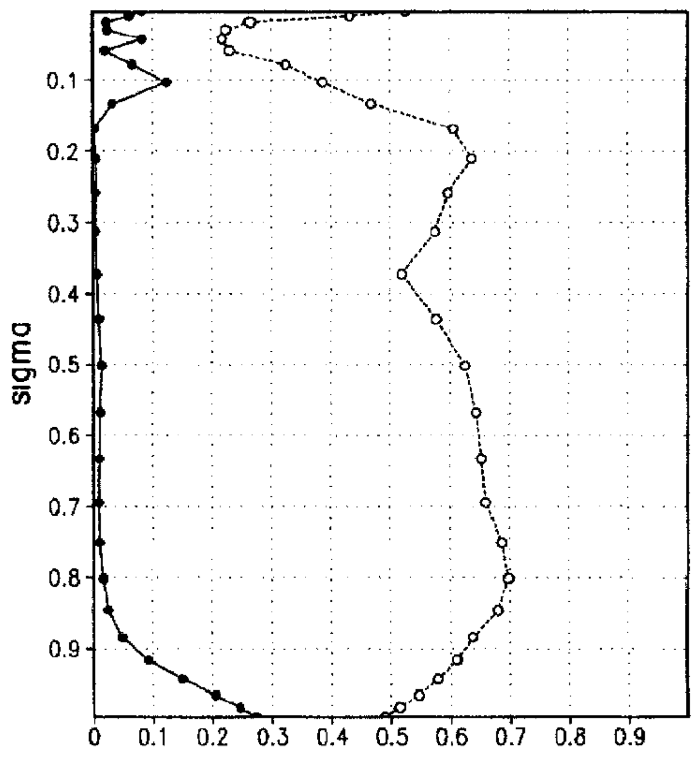

Figure 7: Global mean explained variance from balance equation for temperature (open circles) and velocity potential (filled circles). From Wu et al. (2002).

Since it is difficult to factor the $B_H^1 + B_H^2 + B_H^3$ matrix into square-root form, it is necessary to change the minimization algorithm to a form which is reasonably well conditioned and only requires multiplication by the B matrix. We have chosen to use the preconditioned conjugate gradient routine described in Derber and Rosati (1989). This minimization routine in combination with the grid point analysis produces convergence characteristics similar to the spectral analysis system.

The spectral and grid point versions of the analysis were compared initially by examining single point observations. Figs. 8-10 show various comparisons between the two analysis systems (run at T62) for a single temperature observation at 1000hPa at 45°N and 180°E. Note the similarity in the scale of the temperature increments and the similarity of the structure of the wind increments. This indicates that the horizontal component of the background errors and the balance equation are similar between the two systems. In the vertical, there are a few more differences. For the grid point version there is virtually no response above 600hPa and the response is more concentrated near the surface. Also, note the waviness away from the observation location in the horizontal plots from the grid point version This waviness is probably from the conversion back to spectral space at the end of the inner loop of the minimization and probably would be reduced at higher horizontal resolution.

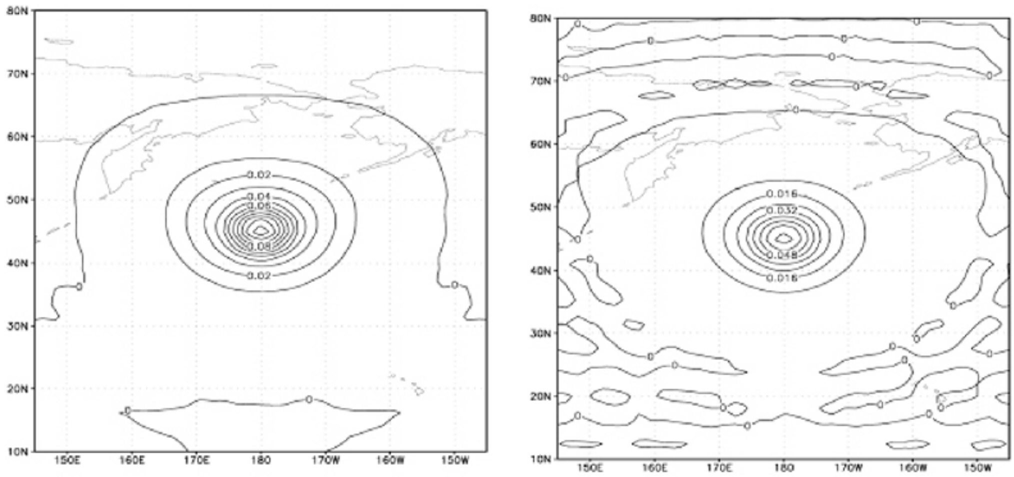

Figure 8: T62L28 analysis produced from a single 1000hPa temperature observaion. Difference between background and analysis for temperature field (left - spectral analysis, right - grid point analysis).

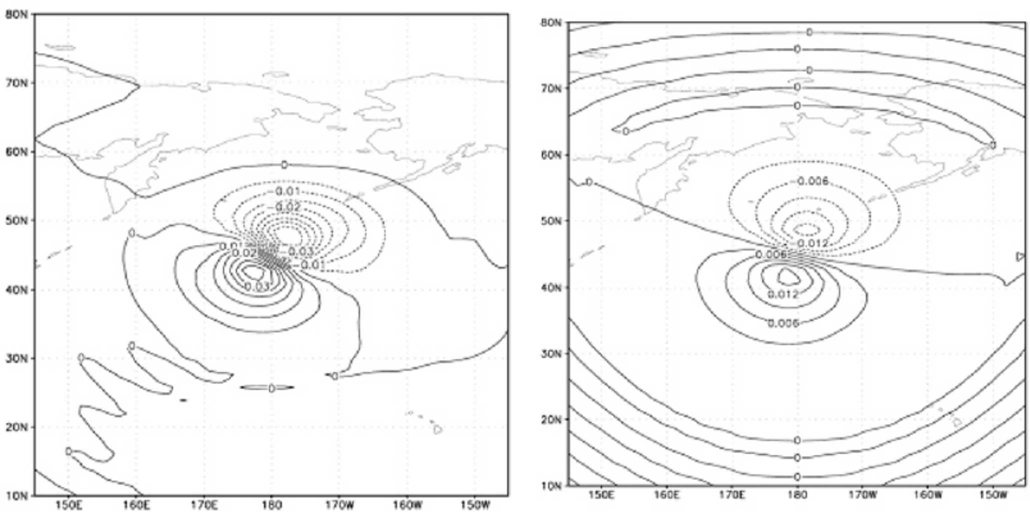

Figure 9: Same as Fig. 8 except u component of the wind at 1000hPa.

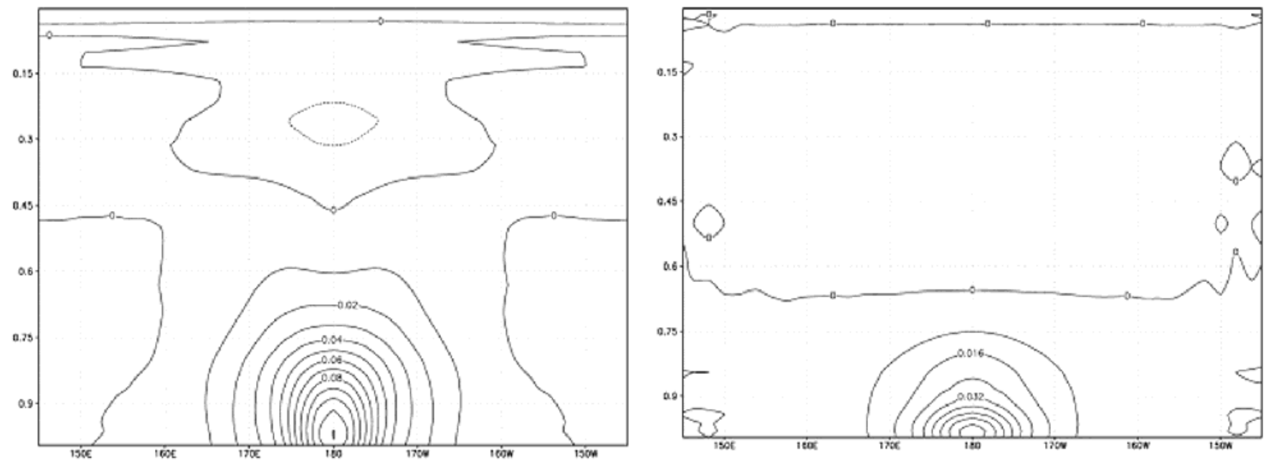

Figure 10: Same as Fig. 8 except west to east vertical cross section of temperature field.

Full resolution (T254L64) analyses have been run using the complete set of data being used in our pre-operational testing. This is similar to our current operational data set except with the inclusion of GOES-12 sounder radiances. The analysis increments produced by the spectral and grid point versions have similarities, but still have significant differences. We believe that most of the differences result from the different vertical covariances used in the grid-point system (which may be a good thing). The grid point version produces slightly better fits to the input observations (given exactly the same initial penalty). This difference probably results from either small difference in the variances or the different vertical correlation structures.

4. Final comments

NCEP is exploring a version of the global analysis system which uses a grid point form of the background error term. This work will be done in collaboration with the GMAO. Recursive filters are being used to allow local definition of the background error covariances. The initial test (some performed with bugs in the code) showed promising results especially in the tropics. Two important upgrades are currently underway prior to beginning tests to replace the spectral version. First, the code is being modified to improve the efficiency of the parallel processing. With the new grid point version some of the structuring of the data can be done in a much more efficient and flexible manner. These changes are expected to significantly improve the efficiency of the analysis system and the grid point version is expected to complete at least a quickly as the spectral version. Second, the code is being modified to work both globally and for our new regional system. This is intended to unify our global and regional analysis projects and allow improvements to be directly usable in both systems.

Eventually the goal is to enable the use of situation dependent background error covariances. The major difficulty with using the situation dependent background errors is determining how to define the structures. Several different possibilities exist including using the background state or using the ensembles. A program to estimate the appropriate error statistics based on the innovations (not the NMC method) is underway.

At NCEP, we have chosen to attempt to do this through a grid point version of the analysis system and by using recursive filters to define the background errors. However, this is not the only possibility. For example, Mike Fisher's presentation in this seminar presented an alternative using wavelets. Regardless of the particulars of the scheme, the incorporation of situation dependent background error covariances is thought, among experienced data assimilators, to be one of the most promising ways of significantly improving analyses and resulting forecasts.

References

Derber, J. and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333-1347.

Derber, J. C. and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195-221.

Purser, R. J., W.-S. Wu, D. F. Parrish and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.,* **131**,1524-1535.

Purser, R. J., W.-S. Wu, D. F. Parrish and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.,* **131**,1536-1548.

Weaver, A., and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion equation. *Quart. J. Roy. Meteor. Soc*., **127**, 1815-1846.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev*., **130**, 2905-2916.