P1.88

A COMPARISON OF PRESCRIPTIONS FOR USING BACKGROUND FIELD DIAGNOSTICS TO ADAPT COVARIANCES TO THE AMBIENT FLOW IN A 3D VARIATIONAL ASSIMILATION

Manuel S. F. V. De Pondeca*, R. James Purser*, David F. Parrish†, John C. Derber†

National Centers for Environmental Prediction

1. INTRODUCTION

A crucial component of a three-dimensional variational (3DVar) scheme for the assimilation of data for initializing a weather prediction model is the assumed covariance model for the background errors (e.g., Daley 1991). Traditionally, the covariance model of an operational 3DVar scheme has been rather simple in spatial structure, being horizontally isotropic and homogeneous for each of the separated scalar components into which the full dynamical fields are typically resolved. The balanced (quasi-geostrophic) component, having larger energy in its modes and in the errors of these modes (as measured by an 'energy norm') naturally has the covariance with the dominant amplitudes; the errors of the unbalanced divergent and rotational components are normally considered to be of lesser amplitude and significance. However, as we migrate to progressively finer scales in our forecast models and their accompanying assimilations, arguably it becomes less excusable to ignore the evident strong horizontal anisotropies and the vertically tilted meteorological structures, such as fronts, rainbands, and quasi-linear features of organized convection that are so dominant at the scales now resolved by these mesoscale models. Directional dependencies in the errors in the background are intuitively expected to mirror the manifest horizontal and vertical 'stretching' of the mesoscale features themselves to some degree, so the assumption that simple horizontally isotropic and untilted covariances will suffice seems less easy to justify.

Recently developed numerical tools, including explicit models of quasi-Gaussian covariances of finite support (Gaspari and Cohn 1998, 1999), various flavors of wavelet analysis (Fisher 2003, Auger and Tangborn 2004, Deckmyn and Berre 2005), synthetic diffusion (Weaver and Courtier 2001) and the 'Hexad' filtering algorithm (Purser et al. 2003b), now make it feasible to construct a background covariance operator that incorporates spatial inhomogeneity and a controlled degree of general anisotropy in a variational assimilation. Topographic alignments (coasts, valleys, mountain ranges, and so on) suggest an obvious influence on the anisotropy at low levels. Another reasonable input for choosing parameters that control the background error covariance might be some measure of the data quantity and quality at the *previous* assimilation time. However, a particularly important source of information about anisotropy in the free atmosphere is the latent dynamical information carried by the background field itself.

We consider a number of possible diagnostic measures available from the background. In principle, such a diagnostic can be extended to the analysis itself in an iterative formulation of the assimilation, which is clearly a desirable option if convergence can be guaranteed. Desroziers (1997) discusses some potential difficulties with the convergence of such an iterative procedure if care is not taken. In this preliminary study, we do not attempt an iterative refinement. Neither do we attempt to seek a dependency of the *amplitude* of the covariance on diagnostics of the background, although we recognize that this is another potentially valuable control mechanism for a fully adaptive covariance model.

The covariances we employ for this study have a quasi-Gaussian form – essentially what would be obtained as an appropriate amplitude-modulated response function representing a general diffusion operator acting over a finite duration, for example. The diffusion analogy has been made explicit in the work of Derber and Rosati (1989) and, for more generally horizontally anisotropic covariances, in Weaver and Courtier (2001). The spatial recursive filters by which we synthesize our own quasi-Gaussian covariances also act as accelerated diffusion operators (Purser et al. 2003a) and, by combining them with the sequential line-filtering 'hexad' algorithm (Purser et al. 2003b), any arbitrary anisotropy, as prescribed by the centered and normalized second-moment 'aspect' tensor of spatial dispersion can be obtained, provided it is a sufficiently smooth function of space.

---

* SAIC, Beltsville, MD; † NOAA/NCEP/EMC, Camp Springs, MD

Corresponding author address: Manuel De Pondeca, W/NP2 RM 207, WWBG, 5200 Auth Road, Camp Springs, MD 20746

2. NOTATION

We shall adopt the convention that, unless declared otherwise, vectors are oriented as columns and the gradient operator, defined:

$$\boldsymbol{\nabla} \equiv (\partial/\partial x, \partial/\partial y, \partial/\partial z)^T$$

acts upon either scalars or row-vectors. The orientation of any Jacobian matrix $\partial\hat{\boldsymbol{x}}/\partial\boldsymbol{x}$, is that implied by:

$$\partial\hat{\boldsymbol{x}}/\partial\boldsymbol{x} \equiv (\boldsymbol{\nabla}\hat{\boldsymbol{x}}^T)^T$$

so that

$$d\hat{\boldsymbol{x}} = (\partial\hat{\boldsymbol{x}}/\partial\boldsymbol{x})d\boldsymbol{x}.$$

The independent physical coordinates $\boldsymbol{x}$ are assumed orthogonal and may be taken to be local cartesians when, as we shall normally assume, the scales of correlations are much smaller than the radius of the earth. The notation for other variables is standard.

3. THE METHOD OF RIISHØJGAARD

Riishøjgaard (1998) suggests that a passively advected field would constitute a valuable indication of the recent flow distortions that would, in turn, have played a part in stretching the error covariances. In this sense, the selected scalar field acts as a proxy for a Lagrangian coordinate – an especially good analogy in the case of a tracer field evolving almost entirely by passive advection. Such ideal tracers do not exist in the atmosphere, however, and we opportunistically rely on convenient scalar quantities that only partially behave like tracer variables.

We start with the current Gaussian horizontally isotropic model in use at NCEP:

$$C_0(\Delta\boldsymbol{x}) = \exp\left(-\frac{1}{2}\Delta\boldsymbol{x}^T\boldsymbol{S}_0^{-1}\Delta\boldsymbol{x}\right), \qquad (1)$$

where Δx is the local cartesian vector of separation between particle pairs. The inverse aspect tensor is the diagonal matrix:

$$\boldsymbol{S}_0^{-1} = \mathrm{diag}\{L_h^{-2}, L_h^{-2}, L_v^{-2}\} \qquad (2)$$

where L_h is the horizontal correlation length, L_v the vertical correlation length, and both may be considered functions of latitude and altitude. (The correlation's aspect ratio, from which the tensor derives its name, is the ratio of the square-roots of these diagonal elements). The inverse aspect tensor clearly acts like a metric, producing an effective separation, Δs defined by:

$$\Delta s^2 = \Delta\boldsymbol{x}^T\boldsymbol{S}_0{}^{-1}\Delta\boldsymbol{x} \qquad (3)$$

such that the correlation,

$$C_0 = \exp\left(-\frac{1}{2}\Delta s^2\right) \qquad (4)$$

now becomes effectively *homogeneous* in the imagined space whose distances are measured by the metric definition (3).

For a chosen scalar function, q, assumed to be evolving predominantly under the influence of advection, the Riishøjgaard method would modify the default correlation (1) according to:

$$C(\Delta\boldsymbol{x}) \approx \exp\left(-\frac{1}{2}\Delta\boldsymbol{x}^T\boldsymbol{S}_0^{-1}\Delta\boldsymbol{x}\right) \times \exp\left(\frac{-(\Delta q)^2}{2L_q^2}\right), \qquad (5)$$

where L_q, with the same units as q, represents a correlation scale in the 'direction' of variations of q. Thus, in Riishøjgaard's method, the new metric in which (4) holds is that obtained as if the three spatial dimensions are mapped to a 3-manifold imbedded in a larger space whose orthogonal coordinates are those of $\boldsymbol{x}$, augmented with q and where the new metric is therefore,

$$\hat{\boldsymbol{S}}^{-1} = \mathrm{diag}\{L_h^{-2}, L_h^{-2}, L_v^{-2}, L_q^{-2}\}. \qquad (6)$$

To a first-order approximation, the local 'displacement' in the value of q is related to its gradient and to the real displacement:

$$\Delta q \approx \boldsymbol{\nabla} q \cdot \Delta\boldsymbol{x} \qquad (7)$$

so we may write the modified correlation:

$$C(\Delta\boldsymbol{x}) = \exp\left(-\frac{1}{2}\Delta\boldsymbol{x}^T\boldsymbol{S}^{-1}\Delta\boldsymbol{x}\right), \qquad (8)$$

where

$$\boldsymbol{S}^{-1} = \boldsymbol{S}_0^{-1} + \frac{1}{L_q^2}(\boldsymbol{\nabla} q)(\boldsymbol{\nabla} q)^T. \qquad (9)$$

Examples of suitable choices of q might be humidity (in units scaled appropriately with altitude), potential temperature, θ, or, in order to repond more directly to dynamics, the Ertel potential vorticity (PV),

$$q = \frac{1}{\rho}\boldsymbol{\eta} \cdot \boldsymbol{\nabla}\theta, \qquad (10)$$

where ρ is the density, θ the potential temperature, and

$$\boldsymbol{\eta} = f\boldsymbol{k} + \boldsymbol{\nabla} \times \boldsymbol{v},$$

is the three-dimensional absolute vorticity. If q is characterized by sufficient variability, and L_q^{-2} is large enough, the modification will have a general tendency to reduce the typical correlation scales of C relative to the default scale L_h and L_v, so the implementation of Riishøjgaard's method is typically accompanied by a compensating adjustment to larger values of these defaults.

The construction (9) is readily augmented to admit the influence of more than one scalar, since we may copy and add the form of the second term on the right with as many additional variables as one wishes to consider. However, in this case it is algebraically neater and potentially more general to gather the controlling scalars into a *vector* $\boldsymbol{q} = (q_1, \ldots, q_n)$ to whose n components one associates the corresponding correlation 'scales', L_i with $i = 1, \ldots, n$ in the same respective units as the q_i themselves. Then, assuming these q_i all act as independent fields, a diagonal matrix is formed:

$$\boldsymbol{S}_q^{-1} = \mathrm{diag}\{L_1^{-2}, \ldots, L_n^{-2}\} \qquad (11)$$

and the generalization of (9) written as:

$$\boldsymbol{S}^{-1} = \boldsymbol{S}_0^{-1} + (\nabla \boldsymbol{q}^T)\boldsymbol{S}_q^{-1}(\nabla \boldsymbol{q}^T)^T. \qquad (12)$$

A further generalization that this form suggests when the q_i are *not* mutually independent is to allow the matrix $\boldsymbol{S}_q^{-1}$ to be nondiagonal (but still symmetric and non-negative). While we have segregated the terms on the right of (12) into 'geometrical' and 'functional' contributions (first and second terms respectively), even this distinction becomes somewhat artificial when the possible choices for the components of $\boldsymbol{q}$ can include some that are themselves most naturally interpreted in geometrical terms (as we show below). As a consequence of the trivial Jacobian identity, $\nabla(\boldsymbol{x}^T) = \boldsymbol{I}$, a further algebraic consolidation of the formalism is achieved by absorbing the geometrical components of $\boldsymbol{x}$ into an augmented vector $\boldsymbol{q}$, and making room on the diagonal of (augmented) $\boldsymbol{S}_q^{-1}$ for the elements of $\boldsymbol{S}_0^{-1}$. Thus, the first term on the right of (12) may be dropped from the general formalism:

$$\boldsymbol{S}^{-1} = (\nabla \boldsymbol{q}^T)\boldsymbol{S}_q^{-1}(\nabla \boldsymbol{q}^T)^T, \qquad (13)$$

where it is understood that, in general, the vector $\boldsymbol{q}$ includes the geometric, or coordinate components, $\boldsymbol{x}$.

In this form, both geometrical and functional variables have essentially the same status and it is therefore not mandatory that all of the geometrical components $\boldsymbol{x}$ of $\boldsymbol{q}$ actively participate in the metric formula, as long as the remaining metric components corresponding to $\boldsymbol{q}$ still suffice to ensure that the embedded image of the real domain in $\boldsymbol{q}$-space stays three-dimensional. A useful example of this occurs when L_v^{-1} vanishes but L_θ^{-1}, corresponding to a potential temperature component, θ, of $\boldsymbol{q}$, does not. The correlation implied by this choice should lead to an assimilation virtually equivalent to that obtained using undeformed correlations in an isentropic coordinate; but here the accomplishment requires no formal transformation to an actual isentropic coordinate grid.

We shall see below that, by an extension of this judicious activation and deactivation of different metric coefficients, in a suitably augmented $\boldsymbol{q}$, a variety of implied deformations of an otherwise geometrically simple correlation are conveniently indicated. However, before we proceed to describe these, we introduce convenient alternative parameters $\kappa \equiv L^{-1}$ for each of the variables we might use in the metric definitions, so that infinities are avoided in the cases where some of the given variable happen not to participate actively in the metric (when the corresponding κs go to zero). Each κ has dimensions inverse to its respective variable and may be thought of as a sort of 'concentration' of correlation lengths per unit of the variable in question.

4. ADAPTATION OF THE METHOD OF DESROZIERS

A special class of prescriptions for adaptively deforming the correlations occurs when the correlation is envisaged to have a simple form in a variably-displaced copy of the physical coordinates. The generalized form of the Riishøjgaard method above accommodates this class of methods by simply nullifying the horizontal and vertical metric terms, say κ_h and κ_v, that corresponds to the geometrical variables $\boldsymbol{x}$, and activating *only* the κ terms associated with the appropriate horizontal and vertical transformed copies, $\boldsymbol{X}$, of the original $\boldsymbol{x}$. The archetype of such methods is that of Desroziers (1997), who proposed such a transformation to the 'geostrophic momentum' coordinates of the semi-geostrophic (SG) theory of Hoskins and Bretherton (1972), Hoskins (1975). This is a horizontal coordinate transformation which tends to expand cyclonic regions, shrink anticyclonic ones and broaden frontal zones. Since the background error covariance is expected to respond qualitatively in much the same way, Desroziers suggested the use of this coordinate change, suitably adapted to spherical geometry and modified within the tropics (where geostrophic control becomes invalid), as a way to enable spectral representations of background covariance to indirectly acquire some useful adaptivity to the environmental flow. (Spatial adaptivity in a spectral representation is not directly practical owing to the loss of diagonalization implied).

His implementation explicitly used a transformed computational grid, requiring interpolations between this grid and the regular Gaussian grid used for spectral transformation. Since the natural extension of the SG transformation employs potential temperature as the vertical transformed coordinate, we shall adopt this more complete form here. For locally Cartesian coordinates, the horizontal part of the transformation will give geostrophic momentum coordinates X_g and Y_g transversally displaced with respect to the geostrophic wind (u_g, v_g) according to:

$$X_g = x + \frac{1}{f} v_g, \qquad (14)$$

$$Y_g = y - \frac{1}{f} u_g. \qquad (15)$$

As noted by Desroziers, the singular behavior of the reciprocal-Coriolis parameter at the equator needs to be regularized. The effective regularized 'Coriolis time scale' that we denote τ_g and that approximates, and substitutes for $1/f$ in the transformation above is defined as a function of latitude ϕ, following Desroziers (1997), by

$$\tau_g(\phi) = [1 - \exp(-\phi^2/2\phi_0^2)] f^{-1},$$

where ϕ_0 is 15^o. For our purposes, it is acceptable to replace the geostrophic wind components by the true wind, so the approximation to geostrophic momentum coordinates $\boldsymbol{X}_g$ that we consider is:

$$X_g = x + \tau_g v, \qquad (16)$$

$$Y_g = y - \tau_g u, \qquad (17)$$

together with the substitution of θ for z to complete the set.

The optimal metric parameter, κ_g, that we associate with displacements of $\boldsymbol{X}_g = (X_g, Y_g)$ will not, in general, be equal to the optimal κ_h used in the default case of horizontally isotropic statistics. Neither should we expect the best choice of κ_θ to be the same when the horizontal coordinates are $\boldsymbol{X}_g$ as when they are $\boldsymbol{x}$. Thus, each new combination of the variables that are active in $\boldsymbol{q}$ requires a fresh optimization of the corresponding parameters κ *jointly.*

5. KINEMATIC DEFORMATION METHOD

In this construction, the idealized assumption is that the actual covariance is equivalent to the final result of a kinematic deformation acting over a finite duration, τ_k, upon a passively evolving covariance that *was* horizontally isotropic and homogeneous standard covariance at the beginning of this period. Since we do not have the luxury of being able to carry actual Lagrangian coordinates forward with the model's advection operator for this period, we make the drastic first-order simplification of the horizontal kinematics that attributes the location of each advected particle at the beginning of the period to the points $\boldsymbol{X}_k = (X_k, Y_k)$ where, by analogy to the prescription of the last subsection,

$$X_k = x - \tau_k u, \qquad (18)$$

$$Y_k = y - \tau_k v. \qquad (19)$$

For the vertical direction, we can again take θ as a proxy for a true Lagrangian coordinate, so that the active components defining the metric in this case are a κ_k (associated with $\boldsymbol{X}_k$) and κ_θ only. Although, unlike the Desroziers method, the τ in this case can be taken to be a constant, there is no reason why it must be, and it is quite possible that best results with such a purely kinematic approach might require this kinematic time scale τ_k to be a function of altitude, latitude, or both.

6. SOME PRACTICAL CONSIDERATIONS

Some of the methods we have described in this section can, under extreme conditions of the background field, yield embedded images of the mapping of real space to the augmented $\boldsymbol{q}$-space that self-intersect, so that pairs of points actually distant from one another seem unduly close or even coincident in terms of the metric in $\boldsymbol{q}$-space. In practical terms, the implied aspect tensor then becomes highly elongated or even singular. In such extreme conditions it is desirable to limit the aspect ratios of the shapes that might otherwise occur. We do this by expressing the unmodified inverse aspect tensor inferred by each tested diagnostic method in units that make the typical horizontal and vertical scale dimensions equitable, and we then artificially restrict the allowable range of the eigenvalues of this S^{-1}. This is easily done if we diagonalize this matrix through the similarity transformation defined by its matrix of eigenvectors, trim the eigenvalues to within their allotted bounds, and reconstruct our new S^{-1} via the corresponding inverse similarity transformation. However, it is clearly undesirable to *have to* carry out such remedial procedures and far better to choose a combination of the parameters κ that tends automatically to avoid the offensive occurrences.

7. DESCRIPTION OF PLANNED EXPERIMENTS

The generalized framework of the Riishøjgaard method provides us with opportunities to deviate from the 'pure Desroziers method' or the 'pure kinematic method' and consider various hybrids of these methods together with at least some active contributions from the geometrical coordinates $\boldsymbol{x}$. It is beyond the scope of this study to comprehensively explore the vast parameter space of possibilities available by taking different vertical and meridional

Exper.	κ_h	κ_v	κ_q	κ_θ	κ_g	κ_k
1	X	X				
2	X	X	X			
3	X			X		
4				X	X	
5				X		X

Table 1. Summary of the combinations of metric terms (marked with an 'X') actively involved in each of the five families of experiments.

profiles in different combinations from even the limited set of parameters κ that we have suggested, namely the set:

$$\boldsymbol{\kappa} \equiv (\kappa_h, \kappa_v, \kappa_q, \kappa_\theta, \kappa_g, \kappa_k) \tag{20}$$

Instead we primarily examine only a few simple combinations, summarized in Table 1, together with a few choices of the numerical values of the active κ parameters, designed to explore regions of the parameter space that would seem roughly to surround the particular restricted parameter selection that corresponds with NCEP's present horizontally isotropic and homogeneous operational configuration.

These expermeents are being conducted using the NCEP regional analysis system. Forecasts are being examined out to three days for a selection of cases throughout the year for the North American domain. The match of the forecasts of wind, and temperature at various elevations, and the precipitation, to the corresponding time's operational analyses are being examined for each of the experiental set-ups described in Table 1. A comparison of these experiments with the operational configuration (i.e, experiment 1) will allow us to assess the merits of the various strategies for inferring anisotropy from the background. Preliminary results will be provided at the conference.

8. CONCLUSION

We have shown that an adaptation of the method of Riishøjgarrd (1998) allows us to test a variety of strategies for deducing the anisotropies in the background error covariance from the background field itself. The general method implies an assumption that, in a suitably transformed space of possibly more than three dimensions, the image of the mapping of physical space causes the actual background error covariances to become isotropic and homogeneous. Included in this general approach are methods that make the covariance look homogenous in isentropic coordinates, or in geostrophic momentum coordinates, or in kinematically-lagged coordinates, or combinations of all these. The relative merits of these inputs can therefore be put to the test in this form of the generalisation of Riishøjgaard's method.

REFERENCES

Auger, L., and A. V. Tangborn, 2004: A wavelet-based reduced rank Kalman filter for assimilation of stratosperic chemical tracer observations. *Mon. Wea. Rev.,* **132,** 1220–1237.

Daley, R. A., 1991: *Atmospheric Data Assimilation.* Cambridge University Press, 457 pp.

Deckmyn, A., and L. Berre, 2005: A wavelet approach to representing background error covariances in a limited area model. *Mon. Wea. Rev.,* **133,** 1279–1294.

Derber, J. C., and A. Rosati, 1989: A global ocean data assimilation system.*J. Phys. Ocean.,* **19,** 1333–1347.

Desroziers, G.,1997: A coordinate change for data assimilation in spherical geometry of frontal structure. *Mon. Wea. Rev.,* **125,** 3030–3038.

Fisher, M., 2003: Background error covariance modelling. *Proceedings of the ECMWF Seminar on Recent developments in data assimilation for the atmosphere and ocean, 8–12 September 2003*, 45–64.

Gaspari, G., and S. E. Cohn, 1998: Construction of correlation functions in two and three dimensions. Office Note Series on Global Modeling and Data Assimilation, DAO Office Note 96-03R1, DAO, GSFC, 53 pp.

——, and ——, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.,* **125,** 723–757.

Hoskins, B. J., 1975: The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.,* **32,** 233–242.

——, and F. P. Bretherton, 1972: Atmospheric frontogenesis models: mathematical foundation and solution. *J. Atmos. Sci.,* **29,** 11–37.

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.

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

Riishøjgaard, L.-P., 1998: A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus,* **50A,** 42–57.

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