J5.1

A SCHEME FOR THE CHARACTERIZATION AND SYNTHESIS OF ANISOTROPIC BACKGROUND ERROR COVARIANCES SUITABLE FOR ADAPTIVE VARIATIONAL ASSIMILATION

R. James Purser

Science Applications International Corporation
NOAA/NCEP/EMC, Camp Springs, MD

1. INTRODUCTION

The simplest nontrivial covariance models for use in 3D or 4D variational assimilation tend to be based on the horizontally isotropic Gaussian form. The reason for this is that this particular shape, or approximations to it, can be efficiently synthesized numerically by a variety of convenient methods. Examples include using the compact support approximations to Gaussians proposed by Gaspari and Cohn (1999), the simulated diffusion methods of Derber and Rosati (1989) and Weaver and Courtier (2001), and the use of spatial digital filters designed to mimic a Gaussian, as discussed by Purser and McQuigg (1982); Hayden and Purser (1995); Huang (2000); Dévényi and Benjamin (2003). However, studies of the actual structure of forecast errors (e.g., Baker et al. 1987, Thiébaux et al. 1990) indicate that this simple choice is deficient, not only in its restriction to an isotropic form, but also because the Gaussian shape itself gives insufficient weight to both the smallest and the largest resolved scales in comparison to those intermediate scales close to the Gaussian's defining scale parameter.

We address these deficiencies by presenting a broader parametric family of distributions, of which the Gaussians are special members, but which also accommodates a very general tensorial prescription of anisotropy, together with adaptive control over the degree of spatial "kurtosis" of the shape of the covariance distribution. The broader family of distributions generalizes to two or three dimensions what Purser et al. (2003) refer to as the "Hyper-Gaussian" family of distributions. These may be thought of as a particular class of positive mixtures of Gaussians of different scales. Our new family of covariance models enjoys several algebraically convenient attributes, including the fact that the set of shapes of the implied power-spectra are exactly of the kind that are accommodated by the same parameteric model applied in the Fourier domain. Because each member of the proposed family can be formed as an additive mixture of anisotropic Gaussians, the efficient numerical methods that facilitate the practical application to variational assimilation of the simpler Gaussian covariances can also be extended without difficulty to the applications involving discrete approximations to these more appropriate and more general hyperGaussian covariance forms.

The conceptual basis for the generalization we are proposing involves an essentially geometrical viewpoint, which we shall elaborate in the next two sections. In section 4 consideration is given to the practical implementation of approximations of these covariances to variational assimilation. The final section lists some of the techniques by which the parameters of hyperGaussians may be estimated objectively using the information available from observations or from forecast ensembles.

Corresponding author address: R. J. Purser, W/NP2 RM 207, WWBG, 5200 Auth Road, Camp Springs, MD 20746

2. A GEOMETRICAL CHARACTERIZATION OF GAUSSIAN SHAPES

The geometrical viewpoint we adopt regarding the shapes of Gaussian covariances is that each "aspect tensor" (Purser et al., 2003) of normalized and centered spatial second moments of the given distribution is identified with a particular "point" in "aspect space". The space has dimensionality equal to the number of independent components of the symmetric aspect tensor (three components for two physical dimensions, six components for three physical dimensions). Furthermore, aspect space is assumed to be endowed with a continuous Riemann metric in terms of which the concept of a "distance" between any given pair of Gaussian shapes becomes meaningful and definite.

In two physical dimensions, an aspect tensor $\boldsymbol{A}$ defining the shape of a Gaussian profile,

$$G(\boldsymbol{x}) = G(\boldsymbol{0})\exp(-\boldsymbol{x}^T\boldsymbol{A}^{-1}\boldsymbol{x}/2), \tag{1}$$

and its three independent components may be formed into a 3-vector:

$$A_1 = (A_{xx} - A_{yy})/2, \tag{2a}$$

$$A_2 = A_{xy}, \tag{2b}$$

$$A_3 = (A_{xx} + A_{yy})/2, \tag{2c}$$

and we deduce that, since the principal components of $\boldsymbol{A}$,

$$\lambda_{\pm} = A_3 \pm (A_1^2 + A_2^2)^{1/2}, \tag{3}$$

most both be nonnegative for a valid Gaussian, the feasible region of aspect space comprises the right circular cone:

$$(A_1^2 + A_2^2) \leq A_3^2, \tag{4a}$$

$$A_3 \geq 0. \tag{4b}$$

The hyperboloids,

$$A_3^2 - (A_1^2 + A_2^2) = D, \tag{5}$$

define the surfaces of constant determinant, $D = |\boldsymbol{A}| \equiv \lambda_+\lambda_-$ and therefore group together the Gaussians of equal effective areal coverage.

A natural measure of separation for aspect tensor $\boldsymbol{A}$ and an infinitesimal perturbation to it, $\boldsymbol{A}+\boldsymbol{dA}$, is obtained by taking $\sqrt{2}$ times the smallest Frobenius norm (e.g., Golub and Van Loan 1989) among infinitesial spatial deformation operators $\boldsymbol{dT}$ that transform $\boldsymbol{A}$ into $\boldsymbol{A} + \boldsymbol{dA}$ according to the conjugacy:

$$\boldsymbol{A} + \boldsymbol{dA} \equiv (\boldsymbol{I} + \boldsymbol{dT})\boldsymbol{A}(\boldsymbol{I} + \boldsymbol{dT})^T. \tag{6}$$

The square of this natural distance measure is then:

$$\|\boldsymbol{A}, \boldsymbol{A} + \boldsymbol{dA}\|^2 = \frac{1}{2}\mathrm{trace}[(\boldsymbol{A}^{-1}\boldsymbol{dA})^2]. \tag{7}$$

It can be shown that for *any* pair of valid aspect tensors $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, (7) implies a distance of separation satifying:

$$\|\boldsymbol{A}_1, \boldsymbol{A}_2\|^2 = \frac{1}{2}\sum_{i=1}^{3}(\log\Lambda_i)^2, \tag{8}$$

where Λ_i are the eigenvalues of $(\boldsymbol{A}_1^{-1}\boldsymbol{A}_2)$. Equivalently,

$$\|\boldsymbol{A}_1, \boldsymbol{A}_2\|^2 = \frac{1}{2}\mathrm{trace}\{[\log(\boldsymbol{A}_1^{-1}\boldsymbol{A}_2)]^2\}, \tag{9}$$

where the logarithm function has been extended in the natural way to tensor arguments.

It can be shown that, under this metric, the hyperboloidal surfaces of (5) each form a geodesic subspace of constant negative-unit intrinsic ("Gaussian") curvature (neighboring, apparently parallel geodesics in such a subspace diverge exponentially from each other, regardless of their initial orientation). A sample of the geodesics in another section, the plane containing the cone's axis (which corresponds to the set of *isotropic* aspect tensors) is shown in Fig. 1. Here, the Gaussian curvature vanishes, so, despite the distorted appearance of the geodesics, the intrinsic geometry actually remains Euclidean. However, at no orientation in aspect space is the intrinsic curvature positive — a property shared by the aspect spaces of $N = d(d+1)/2$ dimensions associated with physical spaces of d dimensions for all $d > 1$. The natural metric defined by (8) or (9) interprets each associated aspect space infinite and unbounded with an intrinsic non-positive curvature (when $d > 1$) that is homogeneous (symmetrical to translations), but not fully isotropic (symmetrical to rotations), throughout the aspect space.

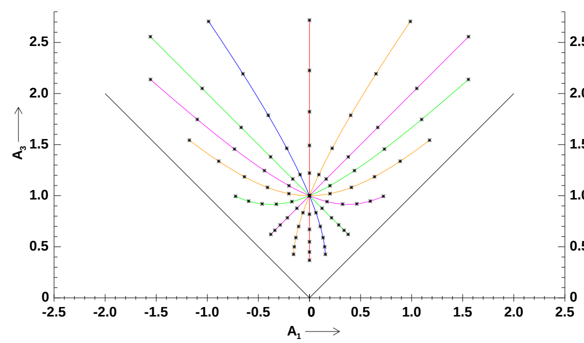

Figure 1. Geodesics through a central point (corresponding to an isotropic Gaussian) in the (A_1, A_3) plane of the cone of feasible aspect tensors for two-dimensional Gaussian profiles. The marks along each geodesic curve are uniformly spaced in the natural metric defined in section 2.

3. CHARACTERIZATION OF HYPERGAUSSIAN SHAPES

While a point in aspect space identifies only a Gaussian's spatial shape, a delta-function impulsive weight distribution at a point can prescribe both this shape and the amplitude (the variance). We propose to model each fat-tailed distribution as the continuous superposition of Gaussian components. This immediately suggests that fat-tailed, symmetric covariance shapes may be identified by more general weight distributions in aspect space than simply a delta function. With a continuous weight function, its value at each point $\boldsymbol{A}$ in aspect space quantifies the additive contribution of the corresponding Gaussian component. Clearly, there are very many way of carrying out such a superposition. However, for the purposes of objective data analysis, we wish to define a *family* of parameterized covariance shapes, the parameters of which provide direct control over the degree of generalized "kurtosis" in different orientations, but are not too numerous. We also ask that the Gaussians themselves belong to our family and that at least some of their algebraically convenient attributes are inherited by the expanded family.

For example, we shall require closure of our family under linear transformations of physical space – that is, uniform stretching, rotation and shearing operations. These are operations that, geometrically, correspond to the isometric translation and rotation symmetries of aspect space under the Riemann metric defined in section 2. Therefore, in examining the anatomy of a representative member of the proposed family, it is sufficient primarily to consider the "standard" forms of weight distributions that are centered on the location corresponding to the identity aspect tensor – the same point at which the geodesics intersect in Fig. 1. The isotropic Gaussian covariance of unit width corresponding to this point is shown in Fig. 2.

We shall also ask that the family be closed under Fourier transformation, in the sense that the power spectrum (Fourier transform) of each standardized synthetic covariance possesses a formally identical composition, except in the reciprocal space of Fourier wave-vectors. Purser et al. (2003) show how a family of hyperGaussian profiles, synthesized as a Gaussian mixture (in log-scale space) of Gaussians possesses the algebraic properties listed above. One such member of this family is depicted in Fig. 3. Here we show how this class of hyperGaussian distributions can be further generalized to a more versatile family by replacing a Gaussian mixture over a single parameter (log-scale) by a "diffusive" mixture over the full aspect space.

In order to maximize the formal symmetry between a generic hyperGaussian and the Fourier transform of one, we adopt the convention that a mixing weight comprising a unit delta-function at aspect, $\boldsymbol{A}$, corresponds to the Gaussian covariance with the amplitude, $G(\mathbf{0})$, of (1) given by:

$$G(\mathbf{0}) = |\boldsymbol{A}|^{-1/4}. \tag{10}$$

In d dimensions, we may obtain the Fourier transform in wavevector $\boldsymbol{k}$ according to:

$$\tilde{G}(\boldsymbol{k}) = (2\pi)^{-d/2} \int G(\boldsymbol{x}) \exp(-i\boldsymbol{k}^T \boldsymbol{x})\, dx_1 \ldots dx_d. \tag{11}$$

Whence,

$$\tilde{G}(\boldsymbol{k}) = |\boldsymbol{A}|^{1/4} \exp(-\boldsymbol{k}^T \boldsymbol{A}\boldsymbol{k}/2), \tag{12}$$

emerges as a new Gaussian obeying the same amplitude convention (10), except with $\boldsymbol{A}^{-1}$ replacing the original $\boldsymbol{A}$. Note that $\boldsymbol{A}$ and $\boldsymbol{A}^{-1}$ lie on the same aspect space geodesic passing through the identity point, $\boldsymbol{I}$, which they straddle symmetrically. Thus, when a continuous mixing weight distribution W for

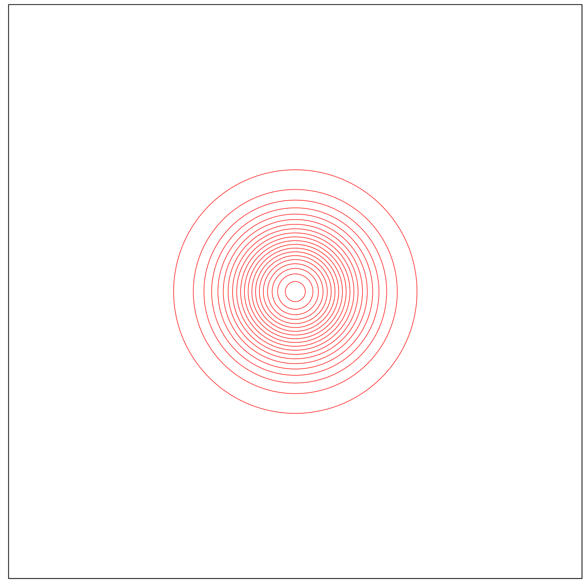

Figure 2. Contours of a Gaussian distribution with identity aspect tensor

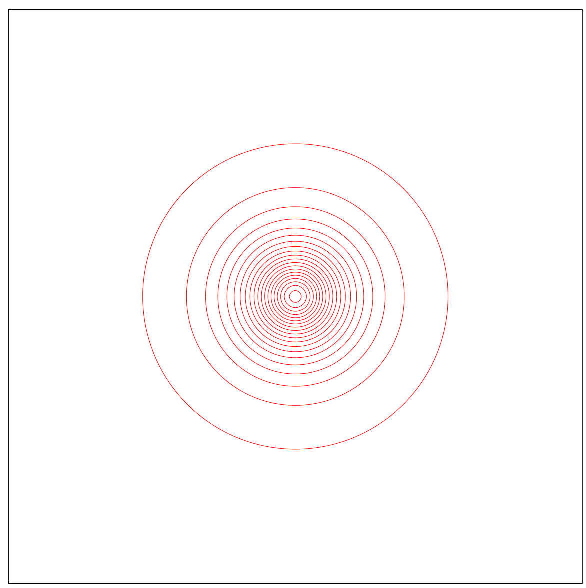

Figure 3. An isotropic fat-tail hyperGaussian distribution obtained by a weighting function W resulting from diffusion of an initial impulse at $\boldsymbol{A} = \boldsymbol{I}$ along the axis of the aspect cone

the Gaussian contributions to a covariance model,

$$H(\boldsymbol{x}) = \int W(\boldsymbol{A})|\boldsymbol{A}|^{-1/4} \exp(-\boldsymbol{x}^T \boldsymbol{A}^{-1} \boldsymbol{x}/2) d\alpha, \tag{13a}$$

$$d\alpha \equiv dA_1 \dots dA_N, \tag{13b}$$

obeys the symmetry,

$$W(\boldsymbol{A}) = W(\boldsymbol{A}^{-1}), \tag{14}$$

then H and its Fourier transform, $\tilde{H}$, are identical. One family of such symmetrical weight functions, including the Gaussian as a special limiting case, is obtained by subjecting an initial delta function at $\boldsymbol{A} = \boldsymbol{I}$ to a diffusive process with specified diffusivity and duration. This is analogous to the way Derber and Rosati (1989) and Weaver and Courtier (2001) employ diffusion to generate Gaussian covariances. However, in the present case it is a weight function $W(\boldsymbol{A})$ that is the end result of the diffusion process, and the space in which it operates is the

Like Weaver and Courtier, we need not restrict the effective diffusivity to being isotropic but can allow it to assume any prescribed symmetric non-negative second-rank tensorial form. However, since the N aspect space position vector components are themselves components of a tensor of second rank in physical space, we may legitimately identify a second-rank aspect space diffusivity tensor $\hat{\boldsymbol{K}}$ with a tensor $\boldsymbol{K}$ of *Fourth* rank in physical space. The symmetries of $\boldsymbol{K}$ all follow from the rules:

$$K_{i,j;k,m} = K_{j,i;k,m} = K_{k,m;i,j}, \tag{15}$$

that reduce the number of independent components to six in two spatial dimensions, and to 21 in three spatial dimensions.

There is a geometrical subtlety related to the non-Euclidean nature of aspect space; except in special isotropic cases it makes no sense to speak of a tensor being "constant" over the space. Therefore, in formally constructing the standard (identity-centered) weight functions for members of the hyperGaussian family, we must be careful to specify that the aspect space diffusivity $\hat{\boldsymbol{K}}(\boldsymbol{A})$ at each $\boldsymbol{A}$ to be given by "parallel transport" of $\hat{\boldsymbol{K}}(\boldsymbol{I})$ along the geodesic joining $\boldsymbol{I}$ to $\boldsymbol{A}$. The weight function $W(\boldsymbol{A})$ is then obtained by "diffusing", for an imagined duration of one half unit of "time", an initial delta function at $\boldsymbol{A} = \boldsymbol{I}$ of specified amplitude C:

$$W(\boldsymbol{A}) = W'(\boldsymbol{A}, 1/2), \tag{16a}$$

where,

$$\frac{\partial}{\partial \tau} W'(\boldsymbol{A}, \tau) = \nabla \cdot \hat{\boldsymbol{K}} \nabla W'(\boldsymbol{A}, \tau), \tag{16b}$$

$$W'(\boldsymbol{A}, 0) \equiv C\delta(\boldsymbol{A} - \boldsymbol{I}). \tag{16c}$$

This will give a distribution for W, obeying (14), and of approximately Gaussian form with a dispersion tensor (in the usual second-moment sense) of approximately $\hat{\boldsymbol{K}}$. (Again, it is the non-Euclidean nature of aspect space that makes these "approximately" qualifications necessary; these approximations are excellent when the dispersion implied by $\hat{\boldsymbol{K}}$ is small in comparison to the space's unit intrinsic curvature.)

For a generic hyperGaussian centered on some general aspect tensor, it is sufficient to invoke the translation isometries of the aspect space and apply a parallel transport operation to the appropriate member of the family of standardized hyperGaussian weight funstions that are centered about the identity aspect. In this most general form, the hyperGaussian in a physical space of d dimensions is defined by a single amplitude parameter, C, the $N = d(d+1)/2$ independent parameters locating the central aspect tensor, and the $N(N+1)/2$ independent dispersion parameter defining $\boldsymbol{K}$.

A convenient and simpler subset of this family occurs when we restrict the dispersion model $\hat{\boldsymbol{K}}$ to a tensor in aspect space whose matrix representation has rank one. That is,

$$\hat{\boldsymbol{K}} \equiv \hat{\boldsymbol{V}} \hat{\boldsymbol{V}}^T, \tag{17}$$

for some aspect space vector, $\hat{\boldsymbol{V}}$, (or, equivalently, some corresponding second-rank symmetric tensor

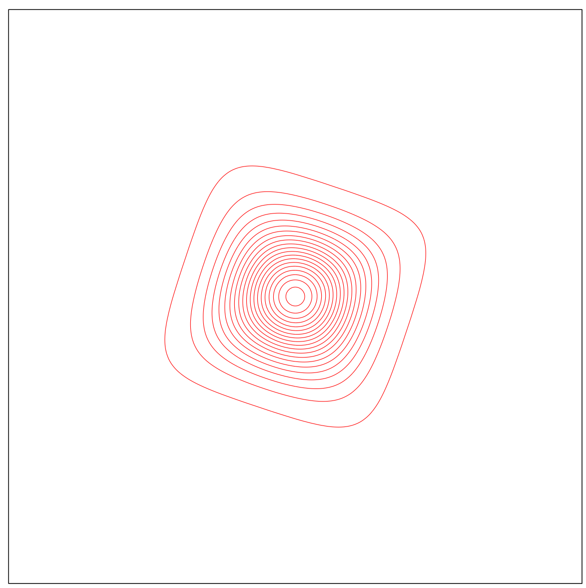

Figure 4. A hyperGaussian whose weighting function W is the result of diffusing an initial impulse at $\boldsymbol{A} = \boldsymbol{I}$ along an oblique geodesic contained in the constant $|\boldsymbol{A}|$ subspace

$\boldsymbol{V}$ in physical space). In this case, the diffusion of W' in (16) is confined along one particular geodesic through the center. Even with this restriction, a surprising variety of covariance forms may be produced. abstract parameter space of aspect components, not physical space.

Fig. 4 shows how a slightly square-looking covariance emerges from a weighting function formed by diffusion of an impulse along a geodesic contained in the subspace (5) of constant $|\boldsymbol{A}|$. (Note that the orientation of the covariance shape changes at half the rate of any change in the orientation of this geodesic.) In this case, the contributing Gaussians all share the same effective geographical area. In Fig. 5, this is no longer the case. The diffusion that results in weighting function W is here along a geodesic oriented obliquely to both the aspect cone's axis and to the geodesic subspaces defined by (5). As a result, we find that the elongation of covariance at small scale (see contours close to the center) is orthogonal to that of the larger scales (see outer contours).

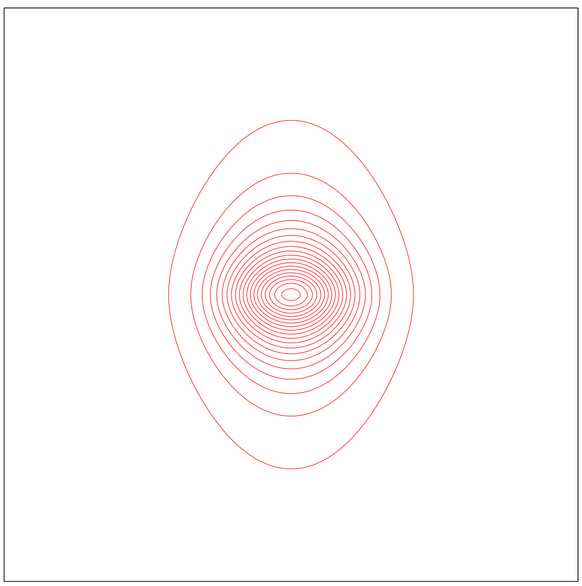

Figure 5. A hyperGaussian whose weight function W is the result of diffusing an initial impulse at $\boldsymbol{A} = \boldsymbol{I}$ along a geodesic angled obliquely to the constant $|\boldsymbol{A}|$ subspaces

4. PRACTICAL APPROXIMATIONS TO IDEAL HYPERGAUSSIANS

In the introduction we discussed several of the practical methods by which a Gaussian covariance (or strictly, the convolution operation implied by a kernel of Gaussian form) may be incorporated into a variational assimilation. The property of a Gaussian convolution kernel that makes it a uniquely attractive choice from the point of view of computational efficiency is that its various multi-dimensional forms can all be synthesized from a short sequence of simple one-dimensional Gaussian smoothers. While we formally lose this property in the extension to the hyperGaussian family, we can still exploit it indirectly provided we approximate the continuous superposition defining the true hyperGaussian by a discrete superposition of a modest tally of carefully selected contributing Gaussians. For the rank-one subset of hyperGaussians we have considered, the weighting distribution W is itself a Gaussian (in s, the metrical distance along the geodesic from the initial center) and a very neat solution to the discretization is then suggested by taking the nodes and discrete quadrature weights of the Gauss-Hermite quadrature for which the associated orthogonal polynomials are the Hermite polynomials orthogonal with respect to weight W. Press et al. (1992) provide a convenient algorithm for these quantities.

For the full-rank hyperGaussians, the weighting function W remains approximately Gaussian providing its degree of dispersion does not greatly exceed unity, the aspect space's characteristic scale of intrinsic curvature. When this is the case, W can be approximately factored (in the convolution sense) into one-dimensional Gaussians, each of which may be discretely sampled at their Gauss-Hermite quadrature nodes. The cartesian product of these quadrature points form a reasonably efficient sampling array for the reconstruction of our full rank hyperGaussians. Obviously, the cost is substantially greater than in the equivalent rank-one discretization, especially when a relatively high-order of quadrature is selected. However, in most cases, the main benefits of using hyperGaussian covariances in preference to Gaussians are attained with quadrature discretizations of as few as two points in each dimension.

5. PARAMETER ESTIMATION

We have provided a detailed description of the construction of a useful family of fat-tailed distributions that generalize the commonly-used Gaussian family in multi-dimensions, we have noted some of their algebraic properties and illustrated a few examples. In order to apply these methods to data assimilation, we need to be able to make some intelligent estimates of the various parameters that now define not just amplitude (variance) and scale, but those attributes of shape that we have bundled together as "generalized kurtosis". While direct curve- or surface-fitting procedures (e.g., see Thiebaux et al. 1990; Baker et al. 1987) provide a statistically robust approach to estimating covariance parameters when validating data are plentiful, there are other objective methods that also have attractive features. Maximum likelihood (Dee and da Silva 1999) and its Bayesian generalization (Purser and Parrish 2003) offer options that are statistically efficient, though not necessarily robust; Wahba's method of "Generalized Cross-Validation" (e.g., Craven and Wahba 1979; Wahba 1990) is a possibly more robust method, but has not been extensively applied to simultaneous estimation of such large sets of covariance parameters as would be required for hyperGaussians.

Finally, we should not omit to include the possibility that forecast ensembles might provide some of the required information. It is not too difficult to extract estimates of local spatial moments of background error implied by the sample covariances of appropriate gradients of ensemble members, provided the latter are collectively consistent with the error statistics desired. While there is not a clean one-to-one correspondence between the fourth spatial moments of the hyperGaussian covariance and its $\boldsymbol{K}$ parameters, there is at least a partial corresponence which could conceivably be exploited by gleaning the relevant sample-moment information from ensembles.

6. ACKNOWLEDGMENTS

This work was partially supported by the NSF/NOAA Joint Grants Program of the US Weather Research Program. This research is also in response to requirements and funding by the Federal Aviation Administration (FAA). The views expressed are those of the authors and do not necessarily represent the official policy or position of the FAA.

REFERENCES

Baker, W. E., S. C. Bloom, J. S. Woollen, M.S. Nestler, E. Brin, T. W. Schlatter, and G. W. Branstator, 1987: Experiments with a three-dimensional statistical objective analysis scheme using FGGE data. *Mon. Wea. Rev.,* **115,** 272–296.

Craven, P., and G. Wahba, 1979: Smoothing noisy data with spline functions. *Numer. Math.,* **31,** 377–403.

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

Dee, D. P., and A. M. da Silva, 1999: Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Mon. Wea. Rev.,* **127**, 1822–1834.

Dévényi , D., and S. G. Benjamin, 2003: A variational assimilation technique in a hybrid isentropic-sigma coordinate. *Meteor. Atmos. Phys.,* **82,** 245–257.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.,* **125,** 723–757.

Golub, G. H., and C. F. Van Loan, 1989: *Matrix Computations.* Johns Hopkins, 642 pp.

Hayden, C. M., and R. J. Purser, 1995: Recursive filter objective analysis of meteorological fields: applications to NESDIS operational processing. *J. Appl. Meteor.,* **34,** 3–15.

Huang, X-Y, 2000: Variational analysis using spatial filters. *Mon. Wea. Rev.,* **128,** 2588–2600.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1992: *Numerical Recipes in Fortran 77, Second Ed.* Cambridge University Press, Cambridge, U. K. 934 pp.

Purser, R. J., and R. McQuigg, 1982: A successive correction analysis scheme using recursive numerical filters. Met. O 11 Tech. Note, No. 154, British Meteorological Office. 17 pp.

Purser, R. J., and D. F. Parrish, 2003: A Bayesian technique for estimating continuously varying statistical parameters of a variational assimilation. *Meteor. Atmos. Phys.,* **82,** 209–226.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.,* **131,** 1526–1548.

Thiébaux, H. J., L. L. Morone, and R. L. Wobus, 1990: Global forecast error correlation. Part I: Isobaric wind and geopotential. *Mon. Wea. Rev.,* **118,** 2117–2137.

Wahba, G., 1990: *Spline Models for Observational Data*, SIAM. CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59, 165 pp.

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