Tellus (1985), *37A*, 126–131

On approximations to geopotential and wind-field correlation structures

By H. JEAN THIÉBAUX, *Department of Mathematics, Statistics and Computing Science, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H8*

(Manuscript received August 2, 1983; in final form August 20, 1984)

ABSTRACT

The purpose of this note is to show that a correlation structure model which is rigorously derived from simple stochastic assumptions for geopotential anomalies, automatically satisfies the essential properties of correlation representations as they are used in multivariate optimal interpolation and diagnostic studies. The parameters of the illustrations are derived from observed covariances, with location-specific local-time-average values defining the anomalies of the covariance (correlation) computations. For this reference field, it is shown that the two-dimensional model in $(\Delta\phi, \Delta\lambda)$ arguments reproduces pronounced anisotropy in correlation structure of height and geostrophic–wind field anomalies. Comparisons are made with model-implicit isotropic correlation arrays which are similar to those computed directly from observed statistical arrays and (similarly) fail to provide accurate representation of true ensemble field relationships.

1. Introduction

The present note discusses some important properties of correlation structure models for constant-pressure-surface heights and displays the correlation structure of the geostrophic wind field implicit in a specific anisotropic model. These properties are relevant both to optimal interpolation of observation-minus-guess field differences and to diagnostic studies of geopotential and geostrophic winds, and apply regardless of the definition of the guess field. In recent years, the principal use of covariance models for geopotential fields has been in the objective analysis step of numerical weather prediction. However, this does not obviate consideration of reference of "guess" fields other than numerical forecasts in modeling stochastic structure of meteorological variables for diagnostic purposes. Here the parameters of the illustrations are not those appropriate to modeling statistical structure of observation-minus-forecast differences. Rather, they pertain to a diagnostic field analysis with location-specific local-time-average values defining the anomalies of the covariance (correlation) computations. For this reference field, it is shown that the two-dimensional model in $(\Delta\phi, \Delta\lambda)$ arguments reproduces pronounced anisotropy in correlation structure of height and geostrophic–wind field anomalies. Comparisons are made with model-implicit *interval-averaged correlation arrays* which are similar to those computed directly from observed statistical arrays and (similarly) fail to provide accurate representation of true ensemble field relationships.

The development and use of a direction-dependent correlation function for geopotential, in connection with multivariate analyses of geopotential and isobaric wind components (Thiébaux, 1976, 1977, 1981), has generated considerable interest in properties of correlation structure functions for the winds, which are derived from the geopotential correlation function via the geostrophic equations. The spatial-lag-correlation function:

$$R_{Z,Z}(\Delta\phi, \Delta\lambda) = \prod_{\tau_j = \Delta\phi, \Delta\lambda} \left[\cos(a_j|\tau_j|) + \frac{c_j}{a_j} \sin(\alpha_j|\tau_j|) \right] e^{-c_j|\tau_j|} \quad (1)$$

was originally developed in order to account for the evident anisotropy of second-order statistics of pressure-surface heights, and its derivatives have been used to model the auto- and cross-correlations of zonal and meridional wind components for multivariate optimal interpolation. Subsequently, enquiries have been made concerning inherent properties of the total modeled correlation tensor for heights and winds, and the large-scale structure of its elements. This note has been prepared in response to those enquiries. It will briefly discuss some useful properties of the model (1) and its family of derivatives—properties which pertain whether the guess field used to define anomalies by differencing with observations is a climatological field, a seasonal average, a local-time average, or a current forecast. The reference field used in constructing the figures shown here, is a 5-day local-time average. As discussed in Thièbaux (1974, 1977), the choice of this estimate of the (transient) ensemble mean state of the atmosphere minimizes bias in the corresponding estimates of true atmospheric covariance and, thus, correlation structure. Correspondingly the present illustrations are diagnostics of the ensemble structure of observed fields, rather than forecast error fields.

It is not relevant to include discussion of model parameterizations for forecast errors (f.e.), in conjunction with the model, eq. (1), because ensemble–average forecast-error statistics show little evidence of anisotropy. However it is of interest to note that the one-dimensional or isotropic analogue of the above model, namely,

$$R_{Z,Z}(\Delta s) = \left[\cos(a|\Delta s|) + \frac{c}{a}\sin(a|\Delta s|)\right] e^{-c|\Delta s|}, \tag{2}$$

does provide a reasonable fit to geopotential f.e. correlations. In work primarily designed to evaluate the impact of removing a Pacific-based ocean weather ship on accuracy of f.e. analyses (Thièbaux, 1980), comparison was made between this isotropic analogue and the traditional operational f.e. correlation model (Rutherford, 1973, 1976; Schlatter, 1975; Bergman, 1979; Lorenc, 1981), and the results suggest that its use in the data assimilation cycle of numerical forecasting could be of significant benefit, particularly in relatively data-sparse regions. Fitting the model, eq. (2), to forcast error statistics generally identifies $a \approx 0$, so that

$$R_{Z,Z}(\Delta s) = (1 + c|\Delta s|)\, e^{-c|\Delta s|}. \tag{3}$$

The latter special case of a *second-order autoregressive correlation function* for an isotropic field now has the further support of work reported by Balgovind et al. (1983) who arrive at precisely this form as a result of a dynamical approximation for f.e. correlation structure.

2. Properties of the autoregressive correlation models

Possibly the most important feature of eqs. (1) and (2) as models of R_{ZZ} correlation structure is that they are *true* correlation functions, i.e., they are *rigorously* derived from stochastic autoregressive representations for the anomaly field (Thièbaux, 1976). Accordingly, they and their families of derivatives (which, again, are rigorously the correlation functions of well-defined stochastic fields) automatically satisfy all the requisite properties of true correlation structure. An important example is the positive definiteness of the joint covariance matrix used to define the weighting scheme for multivariate statistical objective analysis—a property not necessarily satisfied by *ad hoc* representations of anomaly field covariances. For the models under discussion here, the "approximation" made in electing them lies in the initial representation of the anomaly field as a spatial (second-order) autoregressive process, not in the choice of a *function type* to represent an observed correlation array. Accordingly, the derived multivariate, spatial correlation structure used to represent geopotential and geostrophic wind components may fail to provide a good fit to a correlation array computed from observed anomalies due to the wrong choice of the *order* of the stochastic autoregressive model, or *inhomogeneity* in the observed anomaly fields, or *ageostrophic* components of the wind fields (as discussed by Buell and Seaman, 1983). However, it will not fail in fulfilling the properties of spatial correlation structure for variables truly related as:

$$Z(\lambda, \phi), \quad \frac{-\kappa}{\sin\phi}\frac{\delta Z}{\delta\phi} \quad \frac{\kappa}{\sin\phi\cos\phi}\frac{\delta Z}{\delta\lambda}, \tag{4}$$

provided the Z anomaly field is appropriately represented by the underlying autoregressive model. Thus, questions of *matching* relate back to the initial assumptions about stochastic and dynamic structures of the meteorological anomaly fields, and not to possible unseemly behavior of a function selected to represent geopotential, spatial-lag correlations on an *ad hoc* basis.

An additional benefit of this approach to correlation modelling is the ease with which the spectral properties of the fields so represented may be studied. Moreover, it is due entirely to the initial representation of the geopotential anomaly field as a stochastic autoregressive process.

We assume the anomaly fields to be differences between true or observed fields and appropriate first-guess or mean fields. We index values on constant-pressure-surfaces in (latitude, longitude) coordinates (ϕ, λ), so that the corresponding lag-correlations may be written as functions of separations in latitude and longitude lags, $\Delta\phi$ and $\Delta\lambda$. The latter choice of arguments makes the correlation functions independent of zonal circumference in distance units—evidently a critically important factor in fitting some observed large-scale correlation arrays for which those natural geographic units appear to be *functionally* natural units for stochastic field description (Thiébaux, 1980). In addition, correlation or covariance models in ($\Delta\phi, \Delta\lambda$) arguments are directly identifiable with spectral functions in two-dimensional wave-number or frequency arguments. This has significant implications for field diagnostic studies, particularly for spectral studies with data from highly irregular observing arrays (see Thiébaux, 1981, for discussion).

3. Implicit auto- and cross-correlations of geostrophic winds

The foregoing spatial geopotential correlation function, $R_{Z,Z}(\Delta\phi, \Delta\lambda)$, has been compared with other correlation models with regard to accuracy of analyses requiring a simple function of geographic separation as the basis for a multivariate "statistical guidance scheme" (Thiébaux, 1977), and has been examined with regard to its implications for the geostrophic eddy kinetic energy spectrum (Thiébaux, 1981). However, the geostrophic wind-field correlation functions which are implicit in this representation for a geopotential field, have not been presented explicitly, either there or elsewhere. Here, for a pair of locations for which (ϕ_i, λ_i) and (ϕ_j, λ_j) denote the (latitude, longitude) coordinates, we refer not only to the 4 distinct correlations among the wind components themselves, but to the height-wind cross-correlations as well. Thus, the full correlation array of interest is:

$$\mathfrak{X}_{i,j} = \begin{pmatrix} R_{Z_i,Z_j} & R_{Z_i,U_j} & R_{Z_i,V_j} \\ R_{U_i,Z_j} & R_{U_i,U_j} & R_{U_i,V_j} \\ R_{V_i,Z_j} & R_{V_i,U_j} & R_{V_i,V_j} \end{pmatrix}. \tag{5}$$

From the relationships between Z, U, V described by eq. (4), the correlations of this array are proportional to the corresponding components of:

$$\begin{pmatrix} R & -\dfrac{\partial R}{\partial \phi_j} & \dfrac{\partial R}{\partial \lambda_j} \\ -\dfrac{\partial R}{\partial \phi_i} & \dfrac{\partial^2 R}{\partial \phi_i \, \partial \phi_j} & \dfrac{\partial^2 R}{\partial \phi_i \, \partial \lambda_j} \\ \dfrac{\partial R}{\partial \lambda_i} & -\dfrac{\partial^2 R}{\partial \lambda_i \, \partial \phi_j} & \dfrac{\partial^2 R}{\partial \lambda_1 \, \partial \lambda_j} \end{pmatrix}, \tag{6}$$

where R without subscripts is simply an abbreviation for R_{Z_i,Z_j}. Rather than reproducing the specific algebraic forms for all 9 components of $\mathfrak{X}_{i,j}$, we simply note that the derivatives of the orthogonal factors of R, as given by the model, eq. (1), namely

$$\left[\cos(a\tau) + \frac{c}{a}\sin(a\tau)\right] e^{-c\tau} \qquad \text{for } \tau > 0, \tag{7}$$

all have the form

$$[k_n \cos(\alpha\tau) + l_n \sin(\alpha\tau)]\, e^{-c\tau}, \tag{8}$$

where, for all $n \geqslant 1$, the coefficients of derivative of order n are

$$k_n = (al_{n-1} - ck_{n-1}) \quad \text{and} \quad l_n = -(\alpha k_{n-1} + cl_{n-1}), \tag{9}$$

and we describe their large-scale structure in Fig. 1–4. V-V and Z-V correlation plots are analogous to those of U-U and Z-U, and thus are not shown. The parameter values which determine the scales and curvatures of the basic correlation structure were those evaluated to best-fit 500 mb, winter,

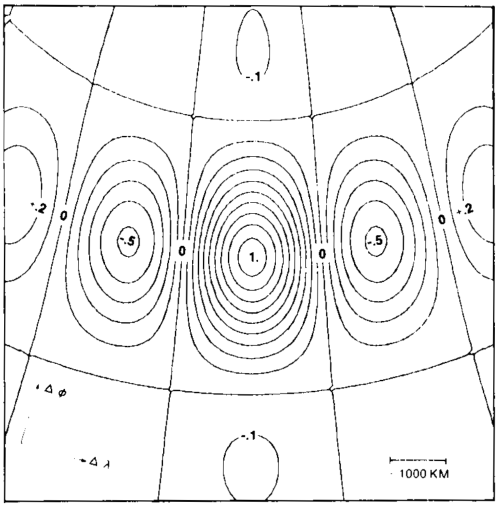

Fig. 1. Spatial-lag correlation function for geopotential in (latitude, longitude) separations: $R_{Z,Z}(\Delta\phi, \Delta\lambda)$, in polar conic projection with tangent circle at 40° N latitude. Contours are at intervals of 0.1.

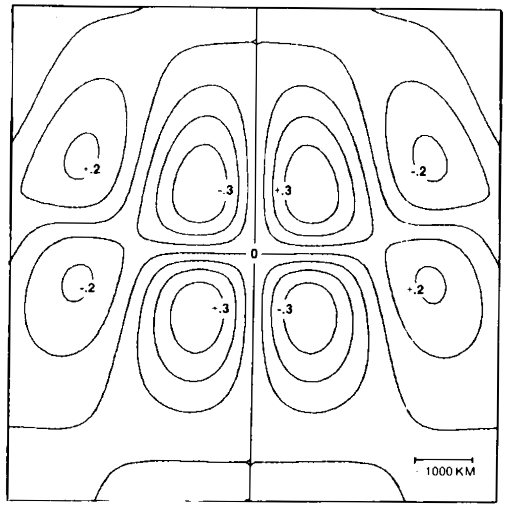

Fig. 3. Spatial-lag cross-correlation function for U and V wind components: $R_{U,V}(\Delta\phi,\Delta\lambda)$.

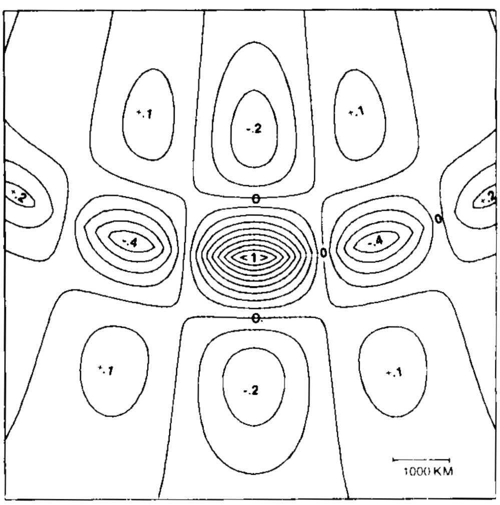

Fig. 2. Same as Fig. 1 for zonal wind components: $R_{U,U}(\Delta\phi,\Delta\lambda)$.

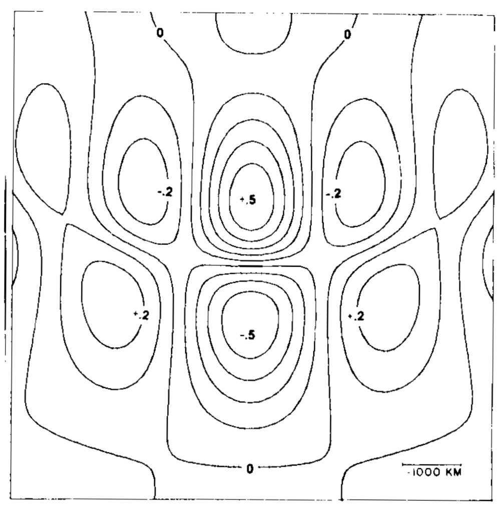

Fig. 4. Spatial-lag cross-correlation function for geopotential Z and the U-component of wind: $R_{Z,U}(\Delta\phi,\Delta\lambda)$.

geopotential statistics for 00Z observations from a 50 station North American radiosonde network, and are as follows, for latitude and longitude lags in radians:

$$\text{zonal factor:}\begin{cases} a = 7.76 \\ c = 1.51; \end{cases}$$

$$\text{meridional factor:}\begin{cases} a = 6.69 \\ c = 6.12. \end{cases}$$

4. Implicit interval-averaged correlation structure

Implications of the two-dimensional arrays, shown here in contour plots, for average-correlation behavior are illustrated in Fig. 5. The

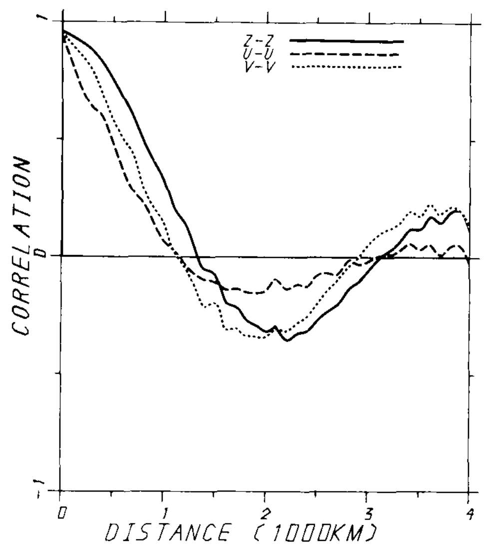

Fig. 5. Distance-interval-averaged autocorrelation of functions for geopotential and the *U*- and *V*-components of wind velocity, obtained by computing the anisotropic model values for all pairs of North American station locations and averaging their values within 50 great-circle-distance intervals.

technique used in the construction is analogous to the construction of any distance-interval-average array of observed correlations: anisotropic correlation *function* values for all pairs of radiosonde station coordinates and the great-circle distances between the stations are computed, and correlation function values averaged within successive 65 km distance intervals. Consequently, the curves of these figures are exactly what we would get if the true lag-correlations for the observed fields (as seen by this North American observing network) were as described by the two-dimensional correlation functions for *Z*, *U*, *V*. Thus, these would be the curves we would seek to represent if isotropic models of spatial correlations were desired as approximations to ensemble field structure. Comparison of Figs. 1 and 2 with Fig. 5 shows clearly that the "geographic structures" of the averaged correlation functions are inappropriate representations for the anisotropic function values, in the sense that they can be significantly dissimilar in magnitude *and* sign for many pairs of locations. Of course, the interval-averaged cross-correlations are near zero, and their plots, which have been omitted for economy of space, confirm this, except for variations due to non-uniform sampling intervals and the influence of latitudinal variation of the Coriolis force.

5. Acknowledgements

I wish to express my appreciation to A. Hollingsworth, R. Daley, and P. Lönnberg, for raising the questions addressed by this work and discussing their consequences with me, during a working visit to ECMWF. It is also a pleasure to acknowledge correspondence on this subject with L. S. Gandin from whose insight my work has benefited greatly. Finally I wish to acknowledge the useful suggestions of the reviewers and to thank P. R. Julian for helpful discussions during revision of the original manuscript. Research was supported by the Natural Sciences and Engineering Research Council of Canada under Grant No. A9182.

REFERENCES

Balgovind, R., Dalcher, A., Ghil, M. and Kalnay, E. 1983. A stochastic-dynamic model for the spectral structure of forecast error statistics. *Mon. Wea. Rev. 111*, 701–722.

Bergman, K. H. 1979. Multivariate analysis of temperatures and winds using optimum interpolation. *Mon. Wea. Rev. 107*, 1423–1444.

Buell, C. E. and Seaman, R. S. 1983. The "scissors" effect: anisotropic and ageostrophic influences on wind correlation coefficients. *Aust. Meteorol. Mag. 31*, 77–83.

Lorenc, A. 1981. A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev. 109*, 701–721.

Rutherford, I. D. 1973. Experiments on the updating of p.e. forecasts with real wind and geopotential data. Preprints. *Third Conf. on probability and statistics in atmospheric science*, Boulder, Colo., June 19–22, A.M.S. 198–201.

Rutherford, I. D. 1976. An operational three-dimensional multivariate statistical objective analysis scheme. Proc. JOC Study Group Conference on Four-dimensional data assimilation, Paris. Report No. 11, the GARP Programme on Numerical Experimentation.

Schlatter, T. W. 1975. Some experiments with a multivariate statistical objective analysis scheme. *Mon. Wea. Rev. 103*, 246–257.

Thiébaux, H. J. 1974. Estimation of covariances of meteorological parameters using local-time averages. *J. Appl. Meteorol. 13*, 596–600.

Thiébaux, H. J. 1976. Anisotropic correlation functions for objective analysis. *Mon. Wea. Rev. 104*, 994–1002.

Thiébaux, H. J. 1977. Extending estimation accuracy with anisotropic interpolation. *Mon. Wea. Rev. 105*, 691–699.

Thiébaux, H. J. 1980. Statistical modeling in short- and long-range weather forecasting. Preprints. *WMO Symposium on probabilistic and statistical methods in weather forecasting*, Nice, France, Sept. 8–12, 29–32.

Thiébaux, H. J. 1981. The kinetic energy spectrum *vis-à-vis* a statistical model for geopotential. *Tellus 33*, 417–427.