Estimation of Three-Dimensional Error Covariances. Part I: Analysis of Height Innovation Vectors

QIN XU
NOAA/National Severe Storms Laboratory, Norman, Oklahoma

LI WEI
Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

ANDREW VAN TUYL AND EDWARD H. BARKER
Naval Research Laboratory, Monterey, California

(Manuscript received 21 August 2000, in final form 1 December 2000)

ABSTRACT

The statistical analysis of innovation (observation minus forecast) vectors is one of the most commonly used techniques for estimating observation and forecast error covariances in large-scale data assimilation. Building on the work of Hollingsworth and Lönnberg, the height innovation data over North America from the Navy Operational Global Atmospheric Prediction System (NOGAPS) are analyzed. The major products of the analysis include (i) observation error variances and vertical correlation functions, (ii) forecast error autocovariances as functions of height and horizontal distance, (iii) their spectra as functions of height and horizontal wavenumber. Applying a multilevel least squares fitting method, which is simpler and more rigorously constrained than that of Hollingsworth and Lönnberg, a full-space covariance function was determined. It was found that removal of the large-scale horizontal component, which has only small variation in the vertical, reduces the nonseparability. The results were compared with those of Hollingsworth and Lönnberg, and show a 20% overall reduction in forecast errors and a 10% overall reduction in observation errors for the NOGAPS data in comparison with the ECMWF global model data 16 yr ago.

1. Introduction

Optimal methods used in data assimilation require statistical descriptions of prediction and observation errors. The statistical analysis of innovation vectors (discrete arrays of observation minus forecast background values) is one of the most common techniques for estimating these error covariances for use in large-scale data assimilation. It is also the most accurate method; however, it can only be employed in regions of sufficiently dense data coverage. In data-sparse areas the National Meteorological Center (NMC, now known as the National Centers for Environmental Prediction) method (Parrish and Derber 1992; Rabier et al. 1998) is typically used instead. It estimates error covariances by analyzing the difference between two model forecasts verifying at the same time. The forecast differences at model grid points effectively substitute for the innovation values. A drawback of this technique is that it can only estimate the forecast error correlation structure, and perhaps the forecast error variance as well if further assumptions are made. It cannot estimate the observation error variance nor its vertical correlation structure, which must be approximated by other means. Thus use of the innovation method remains preferable whenever possible.

The innovation vector methods of Hollingsworth and Lönnberg (1986, henceforth referred to as HL86) and Lönnberg and Hollingsworth (1986, henceforth referred to as LH86), which use truncated spectral representations of forecast error covariance functions with assumed homogeneity and isotropy in the horizontal, form the basis for the current study. After some refinement, their technique is applied to North American data for the (geopotential) height field from the Navy Operational Global Atmospheric Prediction System (NOGAPS) (Hogan and Rosmond 1991). The data assimilation system description (section 2) and basic assumptions (section 3) are used to set up the methodology for isolating the computation of observation and prediction errors (section 4). The method is then illustrated for the horizontal dimension (section 5) and then extended to...
full space (section 6). Finally, the results from sections 4–6 are used to formulate the conclusions (section 7).

2. The NOGAPS assimilation system and statistical data

The radiosonde data used in this study came from the Fleet Numerical Meteorology and Oceanography Center’s NOGAPS operational posttime runs from 1 March to 31 May 1999. At that time NOGAPS was a spectral global model with T159 truncation and 24 levels. The data quality was controlled using the techniques described by Baker (1992). These methods include testing all values against climatological ranges and the complex algorithm developed by Collins and Gandin (1992), which tests lapse rates, hydrostatic consistency, and agreement with neighboring observations. Finally, suspicious values are tested within the multivariate optimal interpolation algorithm following Lorenc (1981), which determines the quality of the data by comparing the analysis at the observation locations with and without the use of the observation in question. The biases caused by solar radiation contamination of the radiosonde thermistor are removed using the International Radiosonde Intercomparison Tests (Nash and Schmidlin 1987) and the navy’s monitoring system, which fine tunes the bias corrections based on long-term seasonal averages of the radiosonde and 6-h forecast differences (Baker 1992). The observations that typically go into the data assimilation system posttime run are described by Goerss and Phoebus (1992), and include all of the meteorological data available 9 h after analysis time.

The background used in the NOGAPS data analysis is typically the 6-h forecast from an analysis 6 h earlier. The procedure involves computing the innovation (observation minus forecast) vector for the various variates; these become input to the analysis, which for NOGAPS is a multivariate optimal interpolation method (Lorenc 1981; Barker 1992; Goerss and Phoebus 1992). During the operational run, select portions of the innovation vector are archived for use in statistical analysis of the background and observation errors. The data selected for this study include the radiosonde observations and corresponding background values for the region between 25°–65°N latitude and 60°–130°W longitude. The reported values of temperature, wind, and dewpoint depression at both mandatory and significant pressure levels are stored, along with the corresponding forecast background values. The geopotential height field is then computed hydrostatically from temperature, moisture, surface height, and surface pressure.

3. Basic assumptions

We denote by \( z \) and \( z' \) the forecast and observation of height, respectively. The errors in \( z \) and \( z' \) are denoted by \( z'^{'} = z - z' \) and \( z'^{''} = z'^{'} - z'^{'} \), respectively, where \( z'^{'} \) represents the true height averaged in a four-dimen-

sional volume of space and time defined by the model’s resolution. The observation error includes both instrument error and sampling error (error of representativeness). In general, \( z'^{'} \) and \( z'^{''} \) are discrete arrays, and \( z' \) and \( z'^{'} \) can be considered as continuous fields in the pressure coordinate system \( (x, y, p) \). The height innovation vector, denoted by \( z'^{''} \), is a discrete array of observation minus forecast values at the observation locations, that is,

\[
z'^{''} = z'^{'} - z = z'^{'} - z'^{'}.
\]

In this paper, \( z' \) and \( z'^{''} \) are assumed to be unbiased or bias-removed. Thus, according to (3.1) we have

\[
\langle z'^{''} \rangle = \langle z'^{'} \rangle = \langle z'^{'} \rangle = 0,
\]

where \( \langle \cdot \rangle \) denotes the ensemble average of \( \langle \cdot \rangle \) over all possible realizations. Since there is only one realization for the observed atmospheric process, the ensemble average will be calculated by averaging in time under the ergodicity assumption, that is, \( \langle \cdot \rangle = \text{time mean of } \langle \cdot \rangle \).

In addition to the above assumed ergodicity and (3.2), the statistical analysis in this paper requires the following three assumptions:

1) The observation errors are independent of the forecast errors, so these two types of errors are not correlated, that is,

\[
\left\langle z'^{'}_{m} z'^{'}_{n} \right\rangle = 0 \quad \text{for any } i, j, m, \text{ and } n,
\]

where \( z'^{'}_{m} \) denotes the value of \( z'^{'} \) at the point \( (x_{i}, y_{j}, p_{m}) \) in pressure coordinates \( (x, y, p) \) and, similarly, \( z'^{'}_{n} \) denotes the value of \( z'^{'} \) at the point \( (x_{i}, y_{j}, p_{n}) \).

2) The (radiosonde) observation errors are not correlated between different stations, so their autocovariance has the following general form

\[
\left\langle z'^{'}_{m} z'^{'}_{n} \right\rangle = \begin{cases} 0 & \text{for } i \neq j, \text{ and any } m \text{ and } n, \\ C_{z}(p_{m}, p_{n}) & \text{for } i = j, \text{ and any } m \text{ and } n, \end{cases}
\]

where \( C_{z}(p_{m}, p_{n}) \) denotes the vertical covariance function for the observation errors.

3) The background error fields are statistically homogeneous and isotropic in the horizontal over a local area of synoptic scale, so their autocovariance is a function of only three independent variables and has the following general form

\[
\left\langle z'_{m} z'_{n} \right\rangle = C_{z}(r, p_{m}, p_{n}),
\]

where \( r = [(x_{i} - x_{j})^{2} + (y_{i} - y_{j})^{2}]^{1/2} \) is the horizontal distance between the two points \( (x_{i}, y_{i}) \) and \( (x_{j}, y_{j}) \), along a great circle on the earth’s surface. We will also denote \( C_{z}(r, p_{m}, p_{n}) \) simply by \( C_{z}(r) \) as long as \( p_{m} \) and \( p_{n} \) are not involved in the derivations and the meaning is clearly understood.

All the above assumptions will be used for the analysis, although the assumption (iii) is not as accurate as the other assumptions. The Hankel transformation of
When the innovation vector \( \mathbf{z}' \) is computed for each observation station according to (3.1), its autocovariance can be calculated for each pair of observation stations. The number of available station pairs within each bin of fixed \( r \) interval \((\Delta r = 100 \text{ km})\) is plotted in Fig. 1 for four different pressure levels over the range of \( 0 \leq r \leq 5000 \text{ km} \), where the total number of stations is also plotted at \( r = 0 \). The plotted numbers reach their maximum at 500 mb and remain maximal above 500 mb until the level becomes higher than 70 mb (not shown). Since most observation stations are above the 1000 mb level, the number of station pairs at 1000 mb is much smaller than at other levels. There are no station pairs in the first bin, only two pairs of stations in the second bin, and then the number of pairs increases to about 20 in the third bin and to about 40 in the fourth bin for the levels above 1000 mb. Although the radiosonde data had been subjected to and passed the routine quality control applied by the NOGAPS system, the available time levels of the qualified data for most of the pairs are very close to the total time levels (twice a day for 3 months). Thus, the sample size in each bin is approximately proportional to the number of pairs, the latter increases roughly linearly with \( r \) for \( r \leq 1000 \text{ km} \) as shown in Fig. 1.

Using (3.2)–(3.5), the autocovariance of \( \mathbf{z}' \) can be partitioned as follows:

\[
C_{zz}(r, p_m, p_n) = C_{zz}(0, p_m, p_n) + C_{zz}(r, p_m, p_n) \quad \text{for } i = j \quad (r \neq 0), \quad \text{and any } m \text{ and } n, \\
C_{zz}(r, p_m, p_n) \quad \text{for } i \neq j \quad (r \neq 0), \quad \text{and any } m \text{ and } n.
\]

Clearly, \( C_{zz}(r, p_m, p_n) = \langle z_i' z_j' \rangle \) can be computed for all \( i \neq j \), so its zero distance limit gives \( C_{zz}(0, p_m, p_n) \). Subtracting \( C_{zz}(0, p_m, p_n) \) from \( \langle z_i' z_j' \rangle \) with \( i = j \) gives \( C_{zz}(p_m, p_n) \). This basic idea and related technique have been widely used for error covariance estimates (Rutherford 1972; HL86; Thiebaux et al. 1986; Bartello and Mitchell 1992). In this paper, the technique is applied not only to single-level analysis but also to multilevel analysis, so the three-dimensional covariance of height forecast error and vertical covariance of height observation error can be estimated directly. The detailed technique and results are described in following sections.

5. Single-level analysis

When \( p_m = p_n \), \( C_{zz}(p_m, p_n) \) reduces to the observation error variance, \( C_{zz}(0, p_m, p_n) \) reduces to the forecast error variance, \( C_{zz}(r, p_m, p_n) \) gives the power spectrum of \( \mathbf{z}' \) in the following form:

\[
S_{zz}(k) = \int C_{zz}(r) J_0(k r) r \, dr \geq 0
\]

for any wavenumber \( k \),

\[
C_{zz}(r) = \int S_{zz}(k) J_0(k r) k \, dk.
\]

For a finite radial range, the integral in (3.7) is replaced by a summation of infinite spectral expansion.

A truncated spectral expansion will be used for the analysis, which assumes implicitly that the background error field is sufficiently smooth so that its covariance function is also smooth and can be adequately represented by the truncated spectral expansion [see (5.1)]. Since the analysis will be performed over a finite range \( r \leq D = 3000 \text{ km} \) constrained by the limited data coverage, an additional assumption is required for the asymptotic behavior of the background error covariance. Following HL86, it is assumed that the horizontal derivative of the background error covariance function becomes sufficiently flat as \( r \) approaches the boundary of the finite range so that its horizontal derivative is negligibly small at the boundary of the finite range. This assumption is supported by the behavior of the binned innovation covariance data in the vicinity of the boundary \( (r = D) \) of the finite range (see Fig. 1 of LH86 and Figs. 2 and 5 in this paper). Under this assumption, the contribution to the background error covariance from the large-scale components (not included within the finite range) can be represented by the constant part of the truncated spectral expansion [see (5.1)].
variance, and \( C_z(r, p_m, \rho_m) \) reduces to a single-level covariance, which can be denoted by \( C_z(r, \rho_m) \) or simply by \( C_z(r) \). The analysis is performed similarly to that in HL86. In particular, the forecast error covariance function is expressed over the finite radial range of \( r \leq D = 3000 \) km by the following truncated spectral expansion:

\[
C_z(r) = \sum S_z(k)J_0(kr),
\]

(5.1)

where \( k_0 = 0 \), and \( k_i \) (\( i = 1, 2, \ldots, M \)) are zeros of \( J_0(kD) \) or, equivalently, zeros of \( J_i(kD) \). Here, \( J_0(\cdot) = \partial J_0(\cdot)/\partial(\cdot) \), and \( J_i(\cdot) \) is the first-order Bessel function. The summation is from \( i = 0 \) to \( M \), where \( i = 0 \) [with \( k_0 = 0 \) and \( J_0(k_0) = 1 \)] corresponds to the constant part that represents the contribution from the large-scale components not included within the range of \( r \leq D = 3000 \) km (since pairs of innovation vectors are used only up to \( r = D \) in the analysis described below). The upper bound \( (M) \) of the summation is selected to ensure that \( S_z(k_i) \) is either automatically satisfied or satisfied with a small adjustment (relative to the total variance) for \( i \leq M \) so that \( C_z(r) \) is semidefinite positive over the range of \( r \leq D \).

The spectral coefficients \( S_z(k) \) are obtained by least squares fitting of (5.1) to the innovation covariances, binned for each interval of \( \Delta r = 100 \) km, over the range of \( r \leq D \) (with the variance at \( r = 0 \) excluded). The weight used in the least squares formulation is simply proportional to \( r \), similarly to that of Rutherford (1972) and HL86. Another possible choice is to let the weight be proportional to the sample size or number of station pairs in each bin (see Fig. 1), similar to that of Bartello and Mitchell (1992). Although this second choice is statistically based, the resulting spectral coefficients are found to be approximately the same as obtained with the first choice. The spectral coefficients \( S_z(k) \) can be also obtained by least squares fitting of (5.1) directly to the data cloud (covariances between all station pairs) without binning. This can be considered as the third choice, which is the limiting case of the above second choice with vanishing bin width \( \Delta r \). According to HL86, the result of the fitting changes only about 1% as \( \Delta r \) is reduced from 100 km to 10 km. This difference is even smaller than that between the first and second choices. Thus, the resulting spectral coefficients are not very sensitive to the above choices. The first choice is used in this paper due to its simplicity.

Examples of the above single-level fittings are shown for the 850-, 500-, and 200-mb levels in Fig. 2. The associated power spectra are shown in Fig. 3a with the truncation set to \( M = 8 \). As shown, \( S_z(k) \) is ensured for \( i \leq M \) at these vertical levels. The local wavenumber \( k_i \) can be related to the global wavenumber \( K_i \) by

\[
K_i(K_i + 1) = (ak_i/D)^2 \text{ for } i \geq 1,
\]

(5.2)

where \( a \) is radius of the earth (see Table 1 of HL86).

To compare with the global spectra in Fig. 7 of LH86, the local spectra (Fig. 3a) are multiplied by the square of \( J_0(kD) \) and plotted as functions of the global wavenumber \( K_i \) in Fig. 3b. As shown, the global spectra in Fig. 3b are similar to those in Fig. 7 of LH86.

The forecast error variance, denoted by \( [\sigma_r(p_m)]^2 \), is estimated by the spectral expansion (5.1) at the zero-distance limit; that is, \( C_z(r) \to [\sigma_r(p_m)]^2 \) as \( r \to 0 \). This is shown by the vertical intercept of each fitted curve at \( r = 0 \) in Fig. 2. Then, according to (4.1), the observation error variance, denoted by \( [\sigma_z(p_m)]^2 \), can be obtained by subtracting the forecast error variance from the total innovation variance, denoted by \( [\sigma_z(p_m)]^2 \). The resulting observation error variance is shown by the remaining part of the total variance above the interception point for each pressure level in Fig. 2. The estimated forecast error and observation error are plotted together with the total innovation in terms of their standard deviations, \( \sigma_z(p_m) \), \( \sigma_z(p_m) \), and \( \sigma_z(p_m) \), as functions of pressure level in Fig. 4.

The following results are achieved with the errors estimated in Fig. 2a of LH86 for the European Centre for Medium-Range Weather Forecast (ECMWF) data (for the period of 1 January–31 March 1983 over North America), we can see that the NOGAPS 6-h height forecast error over North America is smaller than that of ECMWF 16 yr ago by about 20%, and the (radiosonde) observation error is also slightly smaller than 16 yr ago.

The above reduction in the height forecast error is a reflection of the improved forecast skill of the 1999 NOGAPS with respect to the 1983 ECMWF system. This improvement may be explained largely by the enhanced resolution of the 1999 NOGAPS forecast model with respect to the 1983 ECMWF forecast model. As mentioned in section 2, the 1999 NOGAPS was a spec-
Fig. 3. As in Fig. 2 but for the associated power spectra \( \mathcal{S}_{zz}(k_i) \) plotted as functions of (a) the local wavenumber index \( i \) and (b) the global wavenumber \( K_i \).

Fig. 4. Forecast error (dotted), observation error (dashed), and total error (solid) in height measured by their standard deviations: \( \sigma(p_m) \), \( \sigma(p_m) \), and \( \sigma(p_m) \), respectively.

6. Multilevel analysis

When \( p_m \neq p_n \), the analysis can be done similarly to the above single-level analysis except that (5.1) is replaced by a truncated spectral expansion of \( \mathcal{C}_{zz}(r, p_m, p_n) \). Since \( p_m \neq p_n \), \( \mathcal{C}_{zz}(r, p_m, p_n) \) is not semidefinite positive and its associated cross-spectral coefficients \( \mathcal{S}_{zz}(k_i, p_m, p_n) \) are not subject to the same nonnegative condition as in (5.1). Instead, they have to satisfy the following condition:

\[
\frac{1}{2} \left| \mathcal{S}_{zz}(k, p_m, p_n) \right| \leq \left[ \mathcal{S}_{zz}(k, p_m, p_n) \mathcal{S}_{zz}(k, p_m, p_n) \right]^{1/2}.
\]

(6.1)

This condition is derived from the semidefinite positiveness of the autocovariances: \( \langle \xi_{im} + \xi_{jm} \rangle \xi_{jn} \) and \( \langle \xi_{im} - \xi_{jn} \rangle \xi_{jn} \), where \( \xi_{im} = \xi_{jm}^\beta \) and \( \xi_{jn} = \xi_{jm}^\beta \) for any positive number \( \beta \). The power spectra associated with these two autocovariances are nonnegative and have the following form:

\[
\mathcal{S}_{zz}(k_i, p_m, p_n) \mathcal{S}(k_i, p_m, p_n) \beta^2 + 2 \mathcal{S}_{zz}(k_i, p_m, p_n) \geq 0,
\]

which leads to

\[
\left| \mathcal{S}_{zz}(k_i, p_m, p_n) \right| \leq \left[ \mathcal{S}_{zz}(k_i, p_m, p_n) \mathcal{S}_{zz}(k_i, p_m, p_n) \right]^{1/2},
\]

for any \( \beta > 0 \).
It is easy to verify that the right-hand side of (6.2) reaches a minimum and gives the right-hand side of (6.1) when \( B^2 = [S_{m}(k, p_m, p_n)]^{-1/2}[S_{m}(k, p_m, p_n)]^{1/2} \).

Under the condition (6.1), the upper bound \( M \) of the spectral truncation for \( S_{m}(k, p_m, p_n) \) should not exceed the upper bound \( M \) of the power spectral truncation; that is, \( M = 8 \) as obtained in the previous section. With \( M = 8 \), the obtained spectra are found to satisfy the condition (6.1) for most of the wavenumbers \((0 \leq i \leq 5)\) in the troposphere (at and below 200 mb). When a spectral component of \( S_{m}(k, p_m, p_n) \) does not satisfy the condition (6.1), a minimal adjustment is applied to this component to meet one of the two marginal conditions (equalities) of (6.1). The involved adjustments are found to be small and cause no significant change but to smooth the structure of \( C_{zz}(r, p_m, p_n) \).

Examples of the above multilevel fittings are shown in Fig. 5 for two cases: (i) \( p_m = 500 \text{ mb} \) and \( p_n = 850 \text{ mb} \) and (ii) \( p_m = 500 \text{ mb} \) and \( p_n = 200 \text{ mb} \). The associated spectra \( S_{m}(k, p_m, p_n) \) are plotted in Fig. 6. The amplitudes of these spectra are smaller than the geometric mean of their associated power spectra in Fig. 3a, so the condition (6.1) is satisfied. As shown, the zero-distance limit of \( C_{zz}(r, p_m, p_n) \) gives the vertical covariance between forecast errors \( C_{zz}(0, p_m, p_n) \). Subtracting this from the total vertical covariance between innovations gives the vertical covariance between observation errors, that is, \( C_{zz}(p_m, p_n) \). The associated vertical correlation functions are defined and computed by

\[
R_{zz}(p_m, p_n) = C_{zz}(p_m, p_n)\left[\sigma_{p}(p_m)\sigma_{p}(p_n)\right]^{-1},
\]

\[
R_{zz}^{e}(p_m, p_n) = C_{zz}^{e}(p_m, p_n)\left[\sigma_{p}^{e}(p_m)\sigma_{p}^{e}(p_n)\right]^{-1}. \tag{6.3b}
\]
nonnegative and so are their right-hand sides. The non-
negative conditions for the right-hand sides of (6.4a) and (6.5a) can be written into the following combined
form:

\[ 2 \left| C_z(0, p_m, p_n) \right| \leq \left| \sigma_z(p_m) \right|^2 + \left| \sigma_z(p_n) \right|^2, \quad (6.6a) \]

or, equivalently, the following spectral form:

\[ 2 \sum S_z(k, p_m, p_n) \]

\[ \leq \sum [S_z(k, p_m, p_n) + S_z(k, p_n, p_m)]. \quad (6.6b) \]

Since \( \left| \sigma_z(p_m) \right|^2 + \left| \sigma_z(p_n) \right|^2 \geq 2\sigma_z(p_m)\sigma_z(p_n), \) the combined condition (6.6a) is weaker than the general condition of \( \left| \sum S_z(k, p_m, p_n) \right| \leq \sigma_z(p_m)\sigma_z(p_n) \) or, equivalently, \( \left| R_z(0, p_m, p_n) \right| \leq 1 \). It is also easy to see that the combined condition (6.6b) is weaker than (6.2) with \( \beta = 1 \) while the latter is weaker than (6.1). The method of LH86 is constrained only by the positive part of the combined condition (6.6b), which gives an upper bound of \( \sum S_z(k, p_m, p_n) \), while the above counterpart method is constrained only by the negative part of (6.6b), which gives a lower bound of \( \sum S_z(k, p_m, p_n) \). Thus, the mul-
tilevel analysis here is more rigorously constrained than the method of LH86 and the counterpart method.

The multilevel analysis in this paper can be consid-
ered as a direct method. A similar direct method was

\[ R_z(r, p_m, p_n) \]

used by Bartello and Mitchell (1992) except that their
analysis was not constrained [such as by (6.1)] and the resultant vertical correlation in the space of \( (p_m, p_n) \) was projected into the space of vertical normal modes by least squares fitting. As in Bartello and Mitchell (1992), our method is also used to estimate the covariance and correlation structures in the three-dimensional space. The estimated correlation \( R_z(r, p_m, p_n) \) is plotted as functions of \( p_n \) for different \( r \) and fixed \( p_m = 500 \text{ mb} \) in Fig. 8. As shown, the vertical correlation profile be-
comes increasingly flat as \( r \) increases from 0 to 950 km. As \( r \) increases to 1950 km, the peak at \( p_n = 500 \text{ mb} \)
diminishes completely. Clearly, the horizontal and ver-
tical structures of \( R_z(r, p_m, p_n) \) are not quite separable even locally for fixed \( p_m \) or \( p_n \).

The spectrum of \( R_z(r, p_m, p_n) \) is given by \( \sum [S_z(k, p_m, p_n) + S_z(k, p_n, p_m)][\sigma_z(p_m)\sigma_z(p_n)]^{-1} \). The spectral profiles are plotted as functions of \( p_n \) for different wavenumbers and fixed \( p_m \).
FIG. 9. As in Fig. 8 but for the associated spectra. The spectral profiles are plotted as functions of $p_n$ for different wavenumbers and fixed $p_m = 500$ mb.

$= 500$ mb in Fig. 9. The spectral profile for the zeroth-order component ($i = 0$) estimates the vertical correlation for the large-scale components not included within the analysis range ($D = 3000$ km). This profile is very close to the correlation profile for $r = 1950$ km in Fig. 8. When the large-scale spectral component is removed from the covariance function, the remaining synoptic-scale part is given by

$$C_S^2(r, p_m, p_n) = \sum_i S_i(k, p_m, p_n)J_i(k, r), \quad (6.7a)$$

where the summation $\Sigma_i$ is from $i = 1$ to $M$. The associated correlation function can be defined by

$$R_S^2(r, p_m, p_n) = C_S^2(r, p_m, p_n)[\sigma^2_S(p_m)\sigma^2_S(p_n)]^{-1/2}, \quad (6.7b)$$

where $[\sigma^2_S(p_m)]^2 = C_S^2(0, p_m, p_m)$ is the synoptic-scale height forecast error variance. Using (6.7a,b), $R_S^2(r, p_m, p_n)$ is computed and plotted in Fig. 10. The similarities between the vertical profiles for different $r$ are improved in Fig. 10 compared with those in Fig. 8. Thus, by removing the large-scale spectral component, the horizontal and vertical structures of the remaining synoptic-scale part of the correlation function becomes more separable locally (for fixed $p_m$ or $p_n$).

As in HL86 and LH86, the horizontal length scale for the synoptic-scale height forecast error correlation at a single vertical level ($p_m = p_n$) can be defined by

$$L = \left[R_S^2(0)\right]^{-1/2} \left[\Delta S_i R_S^2(r)\right]^{-1/2} = \left[\sum_i S_i(k)\right]^{-1/2} \left[\sum_i k^2 S_i(k)\right]^{-1/2}, \quad (6.8)$$

where $\Delta_i$ denotes the horizontal Laplacian and the dependence of the correlation function and its spectral coefficients on $p_m$ ($= p_n$) is implicit. The computed $L$ values are listed in Table 1 for the same range of vertical levels as in Fig. 18b of LH86. [The $L$ value (465 km) computed for 1000 mb is not reliable and is not listed due to the very limited station pairs (see Fig. 1)]. The $L$ values in Table 1 are in the same synoptic-scale range as those (plotted by Lphi-curve) in Fig. 18b of LH86 and the general increase of $L$ with height (from 925 to 100 mb) is also similar to that in Fig. 18b of LH86, although the $L$ values listed for the vertical levels between 850 and 300 mb are larger than those of LH86 by about 25%. The increase of $L$ with height is a well-known feature and is consistent with the increase with height of the horizontal scale of the dominant kinetic-energy-carrying atmospheric motions. The increase of $L$ with height indicates that the horizontal and vertical structures of $R_S^2(r, p_m, p_n)$ are not separable globally, although their local structures (with $p_m$ or $p_n$ fixed) become more separable owing to the removal of the large-scale spectral component. This global type of nonse-
parability, however, can be accommodated by relatively simple modifications of separable correlation functions as implemented in many operational data assimilation systems (see e.g., Parrish and Derber 1992; Rabier et al. 1998; Daley and Barker 2000).

Franke (1999) used an empirical second-order autoregressive (SOAR) function to fit the covariances between height innovations at different levels \( (p_m \neq p_n) \). The results were found to be unsatisfactory as the fits tended to be ill-behaved for several height pairs. The SOAR function is ensured to be semidefinite positive, but it is not guaranteed to satisfy the condition \( (6.1) \) derived in this paper. Thus, although the SOAR function is suitable for a single-level analysis, it may not be suitable for multilevel analyses. When this condition is met for multilevel least squares fittings \( (p_m \neq p_n) \), the fits are expected to become ill-behaved for some height pairs. Furthermore, based on the results obtained in this paper, we see that the horizontal and vertical structures of the height forecast error correlation are not quite separable. The SOAR function, however, contains only two or three freely adjustable parameters. This makes the SOAR function even more difficult for multilevel analyses. In this paper, a truncated spectral expansion is used for multilevel least squares fittings under the condition \( (6.1) \), so the problem reported in Franke (1999) is avoided.

7. Conclusions

The technique of statistical analysis of innovation vectors is further developed and applied to NOGAPS data over North America. The major products include (i) observation error variances and vertical correlation functions, and (ii) forecast error covariances as functions of height and horizontal distance and their spectra as functions of height and horizontal wavenumber. The results are compared with HL86 and LH86, showing about a 20% reduction of forecast error and slight reduction of observation error in the NOGAPS data compared with the ECMWF global model data 16 yr ago.

The multilevel analysis used in this paper for the vertical covariance estimation is a direct method, which is similar to the single-level analysis except that the innovation data are from two different vertical levels (instead of a single level). The truncated spectral expansion used for the multilevel least squares fitting is not restricted by the nonnegative condition. Instead, the absolute value of each spectral component is bounded by the geometric mean of the two associated power spectral components [see \( (6.1) \)]. Under this condition the multilevel analysis is more rigorously constrained than the method of LH86 and the associated counterpart method [see \( (6.4)-(6.6) \)]. A similar direct method of multilevel analysis was used by Bartello and Mitchell (1992) except that their analysis was not constrained [such as by \( (6.1) \)] but projected into the space of vertical normal modes.

The direct method of multilevel analysis in this paper is simpler than that of HL86 and can be conveniently used to estimate the forecast error covariance and correlation in the full space of \( (r, p_m, p_n) \) or \( (k, p_m, p_n) \). The horizontal and vertical structures of the estimated correlation are found to be not quite separable (Fig. 8) and the nonseparability can be reduced by removing the large-scale spectral component (Fig. 10). The zeroth spectral component \( (i = 0) \) accounts for the large-scale errors whose global wavenumbers are lower than \( K_s = 8 \) [see \( (5.2) \) and Table 1 of HL86]. Since these global waves are not resolved by the current analysis, their contributions to the forecast error covariance is estimated only in a total sum. The estimated vertical correlation for the zeroth spectral component is nearly constant between any two levels above the boundary layer (see Fig. 9), indicating that the large-scale forecast errors are correlated vertically through almost the entire model atmosphere. This vertical structure may be considered as a locally smeared view of the true large-scale error correlation due to the assumed homogeneity and isotropy in the horizontal and the limited range \( (D = 3000 \text{ km}) \) of the analysis in this paper. The true large-scale errors may be highly nonhomogeneous and nonisotropic on the global scale. Estimates of their covariance and correlation structures are beyond the current method.

The technique of statistical analysis of innovation vectors will also be applied to the wind innovation vectors. The results will be reported in a follow-up paper. The studies presented in this and follow-up papers are not intended merely for academic exercises. The method and computer algorithm developed will be used to estimate error statistics in the near future for the new atmospheric variational data assimilation system developed at the Naval Research Laboratory (Daley and Barker 2000) when innovation data are collected by operational runs with this new system.

Acknowledgments. This study was motivated and benefited from many discussions with Roger Daley, James Goerss, and Keith Sashegyi at the Naval Research Laboratory (NRL), Monterey. Comments and suggestions from the anonymous reviewers improved the presentation of the results. The research work was supported by the Office of Naval Research NRL Program Element 0602435N and by NRL Grant N00173-98-1-G903 and FD-SDSU Contract N66001-97-D-5028 to the University of Oklahoma.

REFERENCES


