Estimation of Three-Dimensional Error Covariances. Part III: Height–Wind Forecast Error Correlation and Related Geostrophy

QIN XU

NOAA/National Severe Storms Laboratory, Norman, Oklahoma

LI WEI

Cooperative Institute for Mesoscale Meteorological Studies and University of Oklahoma, Norman, Oklahoma

(Manuscript received 5 July 2001, in final form 24 September 2001)

ABSTRACT

The method of statistical analysis of wind innovation (observation minus forecast) vectors is extended and applied to the innovation data collected over North America for a 3-month period from the Navy Operational Global Atmospheric Prediction System to estimate the height–wind forecast error correlation and to evaluate the related geostrophy. Both single-level and multilevel analyses are performed. The single-level analysis shows that the geostrophy is well satisfied in the middle troposphere but is not well satisfied in the boundary layer and around the tropopause. The multilevel analysis indicates that the cross correlation between height and tangential wind forecast errors at different vertical levels is not small and thus should not be neglected.

1. Introduction

The statistical analysis of innovation (observation minus forecast) vectors has been widely used for estimating observation and forecast error covariances in large-scale data assimilation (Gandin 1965; Rutherford 1972; Hollingsworth and Lönnberg 1986; Lönnberg and Hollingsworth 1986; Thiebaux et al. 1986; Bartello and Mitchell 1992). The method was recently refined by Xu et al. (2001, henceforth referred to as Part I) for the analysis of height innovation vectors and by Xu and Wei (2001, henceforth referred to as Part II) for the analysis of wind innovation vectors. As a follow-up study, this paper applies the method to estimate the cross covariance and cross correlation between the height and wind forecast errors and to evaluate the related geostrophy (that can be measured statistically by a single parameter introduced in the next section). As in Part I and Part II, the height and wind innovation data were collected from the Navy Operational Global Atmospheric Prediction System (NOGAPS; Hogan and Rosmond 1991) over North America (between 25°–65°N and 60°–130°W) during the period from 1 March to 31 May 1999. The data quality control and assimilation system were briefly described in section 2 of Part I. The basic assumptions and formulations required by the wind–height correlation analysis are described in the next section. The method is illustrated for the single-level analysis in section 3 and then extended to the multilevel analysis in section 4. Conclusions follow in section 5.

Corresponding author address: Dr. Qin Xu, National Severe Storms Laboratory, 1313 Halley Circle, Norman, OK 73069.
E-mail: qin.xu@nssl.noaa.gov

2. Basic assumptions and forecast error geostrophy

Following Lönnberg and Hollingsworth (1986, henceforth referred to as LH), it is convenient to introduce the pseudostreamfunction, ψ, and pseudovelocity potential, χ, defined in pressure coordinates (x, y, p) by

$$u'f/f_o = -\partial_y\psi + \partial_x\chi \quad \text{and}$$

$$v'f/f_o = \partial_x\psi + \partial_y\chi, \tag{2.1}$$

where $f_o = 0.91 \times 10^{-4}\ \text{s}^{-1}$ is the Coriolis parameter at the central latitude of the domain, (u', v') is the vector wind forecast error as in Part II, and $(u', v')f/f_o$ is the pseudovelocity. Clearly, ψ and χ are not exactly the same as those defined in (2.1) of Part II. To measure how well the geostrophy is satisfied statistically by the forecast error fields, it is also convenient to introduce the following nondimensional parameter:

$$a \equiv \langle(\psi - \phi)^2\rangle(\sigma_\psi\sigma_\phi)^{-1}, \tag{2.2}$$

where $\langle(\cdot)\rangle$ denotes the ensemble mean of $(\cdot)$ and will be computed as the time mean under the ergodicity assumption, $\phi = gz'/f_o$ is the geopotential forecast error scaled by f_o, z' is the height forecast error, $\sigma_\psi^2 = \langle\psi^2\rangle$ is the variance of ψ, and $\sigma_\phi^2 = \langle\phi^2\rangle$ is the variance of

ϕ. Clearly, if $a = 0$, then the forecast error fields satisfy the geostrophy exactly. The smaller the value of a, the better the geostrophy is satisfied statistically by the forecast error fields.

The random fields of ϕ, ψ, and χ are assumed to be jointly homogeneous and isotropic in the horizontal and so should be the pseudovelocity. This assumption is consistent with that made for the height forecast error in Part I, but they are slightly different from that made for the velocity (u', v') itself in Part II. With the above assumption, the cross covariances between ϕ and (ψ, χ) are functions of only three independent variables (r, p_m, p_n), where $r = [(x_j - x_i)^2 + (y_j - y_i)^2]^{1/2}$ is the horizontal distance between the two points, (x_i, y_i) and (x_j, y_j), along the great circle of the earth's surface (see appendix A of Daley and Barker 2000). Their general forms are

$$C_{\phi\psi} = C_{\phi\psi}(r, p_m, p_n) = \langle \phi(x_i, y_i, p_m)\psi(x_j, y_j, p_n)\rangle, \quad (2.3a)$$

$$C_{\phi\chi} = C_{\phi\chi}(r, p_m, p_n) = \langle \phi(x_i, y_i, p_m)\chi(x_j, y_j, p_n)\rangle. \quad (2.3b)$$

For simplicity the covariances can sometimes be written as functions of r only as long as their implicit dependence on (p_m, p_n) is well understood.

The cross-covariance functions in (2.3a) and (2.3b) cannot be directly estimated from the data, but they can be related to the following cross covariances between z' and the tangential and radial components of the pseudovelocity as follows [(5.3.7) of Daley (1991)]:

$$dC_{\phi\psi}/dr = (g/f_o)C_{zt}, \quad (2.4a)$$

$$dC_{\phi\chi}/dr = (g/f_o)C_{zl}, \quad (2.4b)$$

where

$$C_{zt}(r) = C_{zt}(r, p_m, p_n) = \langle z'(x_i, y_i, p_m)t'(x_j, y_j, p_n)\rangle, \quad (2.5a)$$

$$C_{zl}(r) = C_{zl}(r, p_m, p_n) = \langle z'(x_i, y_i, p_m)l'(x_j, y_j, p_n)\rangle, \quad (2.5b)$$

with l' denoting the tangential component and t' the radial component of the pseudovelocity with respect to the r direction from point (x_i, y_i) to point (x_j, y_j). It is easy to verify that $C_{\phi\chi} = C_{\chi\phi}$, $C_{\phi\psi} = C_{\psi\phi}$, $C_{zt} = -C_{tz}$, and $C_{zl} = -C_{lz}$.

The Hankel transformations of $C_{\phi\psi}$ and $C_{\phi\chi}$ give the following cross spectra:

$$S_{\phi\psi}(k) = \int C_{\phi\psi}(r)J_0(kr)r\, dr, \quad (2.6a)$$

$$S_{\phi\chi}(k) = \int C_{\phi\chi}(r)J_0(kr)r\, dr. \quad (2.6b)$$

The inverse transformations are

$$C_{\phi\psi}(r) = \int S_{\phi\psi}(k)J_0(kr)k\, dk, \quad (2.7a)$$

$$C_{\phi\chi}(r) = \int S_{\phi\chi}(k)J_0(kr)k\, dk. \quad (2.7b)$$

Substituting (2.6a) and (2.6b) into (2.4a) and (2.4b) gives

$$kS_{\phi\psi}(k) = (g/f_o)S_{zt}(k), \quad (2.8a)$$

$$kS_{\phi\chi}(k) = (g/f_o)S_{zl}(k), \quad (2.8b)$$

where

$$S_{zt}(k) = \int C_{zt}(r)J_1(kr)r\, dr, \quad (2.9a)$$

$$S_{zl}(k) = \int C_{zl}(r)J_1(kr)r\, dr, \quad (2.9b)$$

and $J_1(\cdot) = J_0'(\cdot) \equiv \partial J_0(\cdot)/\partial(\cdot)$ is used. The inverses of (2.9a) and (2.9b) are given by

$$C_{zt}(r) = \int S_{zt}(k)J_1(kr)k\, dk, \quad (2.10a)$$

$$C_{zl}(r) = \int S_{zl}(k)J_1(kr)k\, dk. \quad (2.10b)$$

Truncated discrete forms of (2.10a) and (2.10b) will be used to estimate $C_{zt}(r)$ and $C_{zl}(r)$ over a finite radial range in the later sections.

In addition to the above basic assumption for the forecast errors, the (radiosonde) observation errors are assumed to be independent of the forecast errors and to be uncorrelated between different stations and between different variables. With these assumptions, the forecast error cross covariances in (2.5a) and (2.5b) can be estimated directly by innovation cross covariances. The detailed analysis and results are shown in the following sections.

3. Single-level analysis

Over the finite radial range of $r \le D$ (=3000 km), (2.10a) and (2.10b) reduce to the following truncated spectral expansions:

$$C_{zt}(r) = \sum\nolimits_1 S_{zt}(k_i)J_1(k_i r), \quad (3.1a)$$

$$C_{zl}(r) = \sum\nolimits_1 S_{zl}(k_i)J_1(k_i r), \quad (3.1b)$$

where the summation Σ_1 is from $i = 1$ to $i = M$. The wavenumbers $k_i (i = 1, 2, \ldots, M)$ are zeros of $J_1(kD)$ or, equivalently, zeros of $J_0'(kD)$. Since $J_1(k_0 r) = 0$ for $k_0 = 0$, there is no zero-order spectral component ($i = 0$) although the summation can be formally extended to Σ_0 (from $i = 0$ to $i = M$) as in (4.1b) of Part II. Since the observation errors are assumed to be uncorrelated

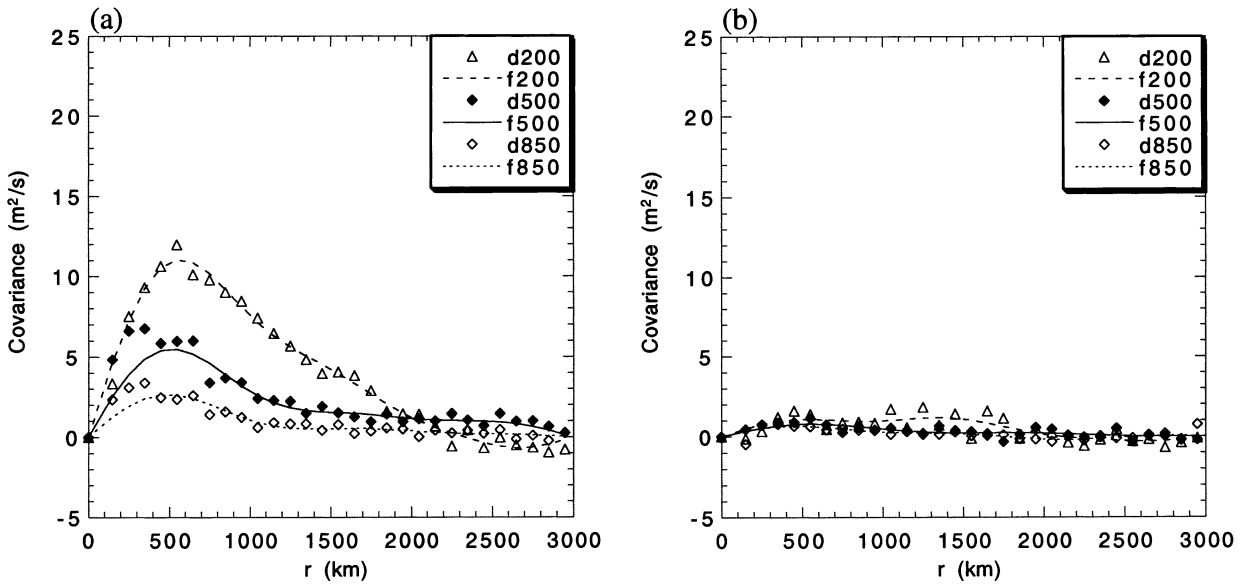

FIG. 1. Binned innovation cross covariances and fitted forecast error cross-covariance functions: (a) $C_{zt}(r)$ and (b) $C_{zl}(r)$, at $p_m = p_n =$ 850 (dotted), 500 (solid), and 200 mb (dashed).

between different variables and also independent of the forecast errors, the forecast error cross covariances $C_{zt}(r)$ and $C_{zl}(r)$ should be the only nonzero part of their associated innovation cross covariances. Thus, the spectral coefficients $S_{zt}(k_i)$ and $S_{zl}(k_i)$ can be estimated by least squares fitting of (3.1) directly to the binned innovation cross covariances.

Note that (p_m, p_n) are implicit for the function forms in (3.1a) and (3.1b). For the single-level analysis considered in this section, we have $p_m = p_n$. In this case, the spectral coefficients $S_{zt}(k_i)$ and $S_{zl}(k_i)$ are constrained by the following conditions:

$$|S_{zt}(k_i)| \le [S_{zz}(k_i)S_{tt}(k_i)]^{1/2}, \tag{3.2a}$$

$$|S_{zl}(k_i)| \le [S_{zz}(k_i)S_{ll}(k_i)]^{1/2}. \tag{3.2b}$$

Here, the spectral coefficients $S_{zz}(k_i)$ are given by least squares fitting of (5.1) in Part I. The spectral coefficients $S_{ll}(k_i)$ and $S_{tt}(k_i)$ can be obtained by least squares fitting of the same spectral expansions as in (4.1a) and (4.1b)

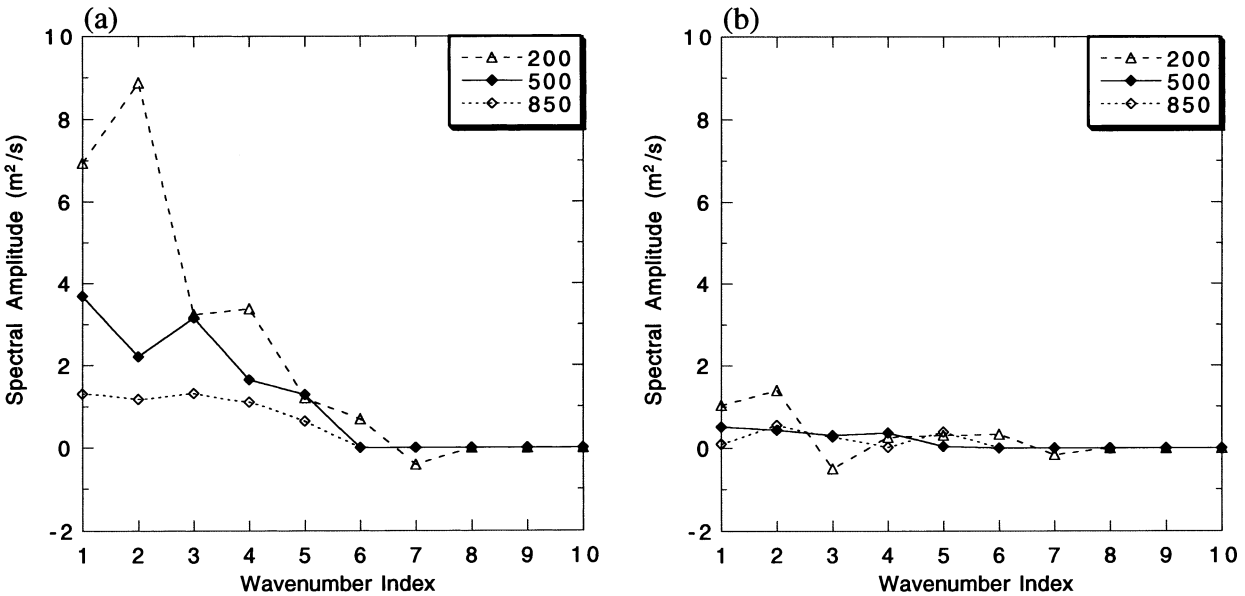

FIG. 2. Cross spectra: (a) $S_{zt}(k_i)$ and (b) $S_{zl}(k_i)$ at $p_m = p_n =$ 850 (dotted), 500 (solid), and 200 mb (dashed).

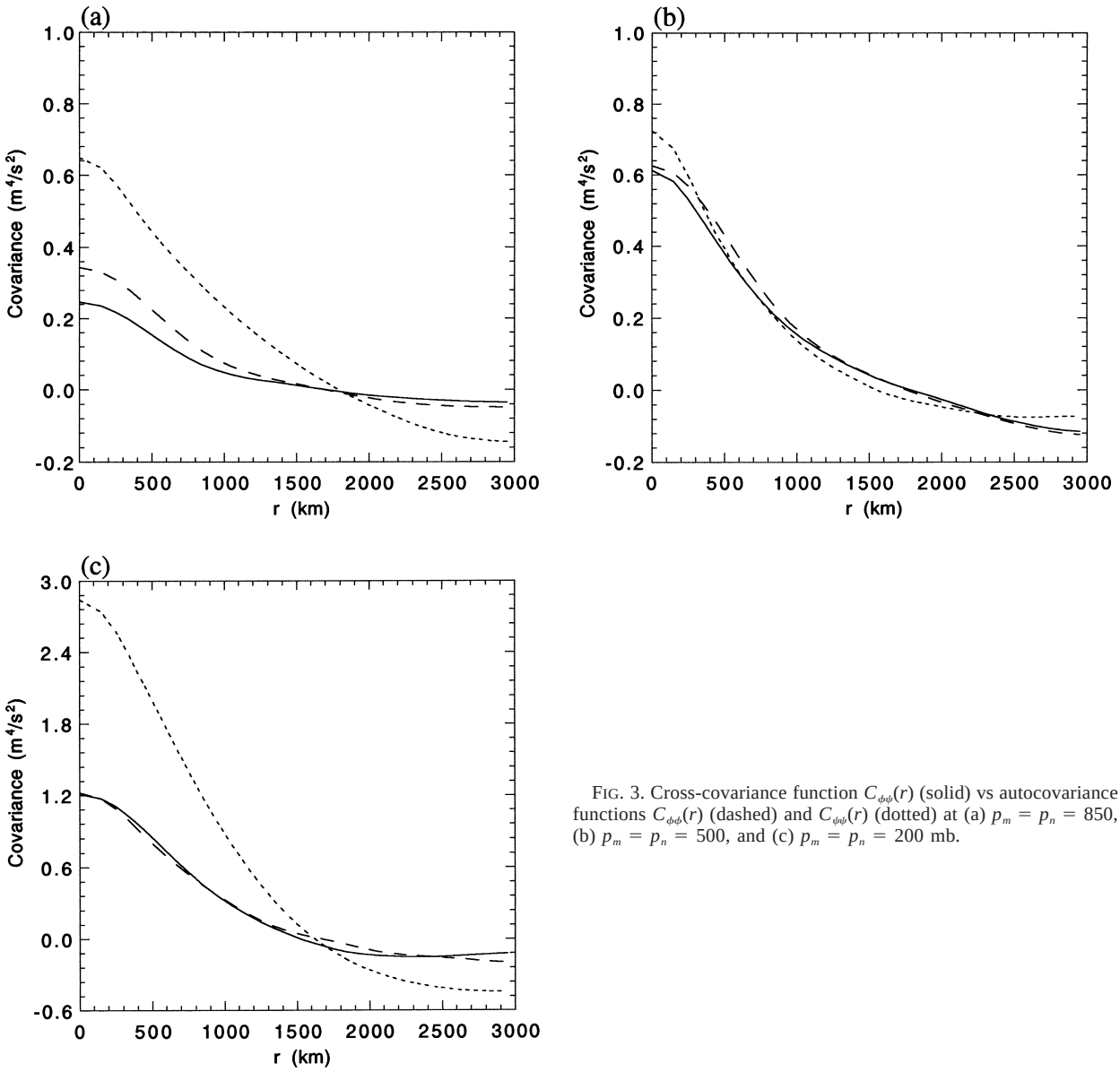

FIG. 3. Cross-covariance function $C_{\phi\psi}(r)$ (solid) vs autocovariance functions $C_{\phi\phi}(r)$ (dashed) and $C_{\psi\psi}(r)$ (dotted) at (a) $p_m = p_n = 850$, (b) $p_m = p_n = 500$, and (c) $p_m = p_n = 200$ mb.

of Part II but applied to the pseudovelocity (instead of the original velocity). The condition in (3.2a) is derived from the semidefinite positiveness of the autocovariances: $\langle(\zeta_i + \tau_i)\,(\zeta_j + \tau_j)\rangle$ and $\langle(\zeta_i - \tau_i)\,(\zeta_j - \tau_j)\rangle$, where $\zeta = \beta z'$ and $\tau = \beta^{-1}t'$ for any positive number β. The condition in (3.2b) is derived from the semidefinite positiveness of the autocovariances: $\langle(\zeta_i + \lambda_i)\,(\zeta_j + \lambda_j)\rangle$ and $\langle(\zeta_i - \lambda_i)\,(\zeta_j - \lambda_j)\rangle$ where $\zeta = \beta z'$ and $\lambda_i = \beta^{-1}l'$ for any positive number β. The detailed derivations follow the same steps as in (6.1) and (6.2) of Part I.

The estimated cross-covariance functions are plotted for $p_m = p_n = 850$, 500, and 200 mb in Fig. 1, where the binned innovation covariances are shown by symbols. For $p_m = p_n = 850$ mb, the structure of $C_{zt}(r)$ in Fig. 1 is similar to that in Fig. 16 of LH, and the structure of $C_{zl}(r)$ in Fig. 2 is flatter than that in Fig. 17 of LH. As shown by the cross spectra plotted in Figs. 2a and 2b, $S_{zl}(k_i)$ is also very flat and negligibly small in comparison with $S_{zt}(k_i)$. Thus, according to (2.8b), the cross-covariance function $S_{\phi\chi}(k_i)$ and thus $C_{\phi\chi}(r) = \Sigma_1 S_{\phi\chi}(k_i)J_0(k_i r)$ are very small. Because $C_{\phi\chi}(r)$ is small and irrelevant to the geostrophy, we will only examine the cross covariance between ϕ and ψ; that is, $C_{\phi\psi}(r) = \Sigma_1 S_{\phi\psi}(k_i)J_0(k_i r)$, where $S_{\phi\psi}(k_i)$ is obtained from $S_{zt}(k_i)$ using (2.8a).

Since the large-scale component ($i = 0$) is not contained in the above spectral expansions, $C_{\phi\psi}(r)$ is the synoptic-scale part of the cross covariance between ϕ and ψ. Similarly, we denote by $C_{\phi\phi} = (g/f_o)^2 C^S_{zz}$ the synoptic-scale part of the height forecast error covariance [see (6.7a) of Part I], and by $S_{\psi\psi} = S_{tt}(k_i)/k_i^2$ the

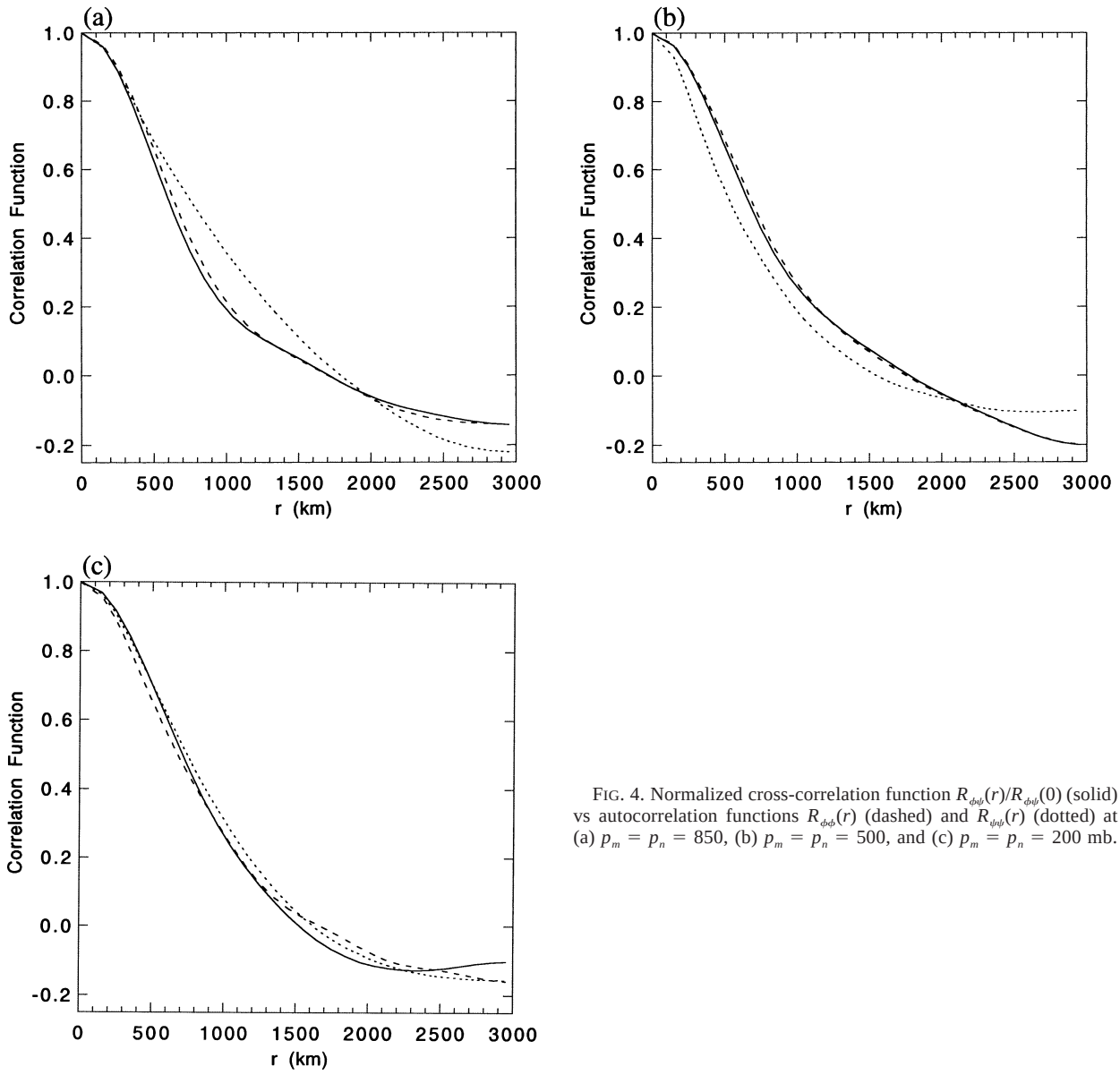

FIG. 4. Normalized cross-correlation function $R_{\phi\psi}(r)/R_{\phi\psi}(0)$ (solid) vs autocorrelation functions $R_{\phi\phi}(r)$ (dashed) and $R_{\psi\psi}(r)$ (dotted) at (a) $p_m = p_n = 850$, (b) $p_m = p_n = 500$, and (c) $p_m = p_n = 200$ mb.

synoptic-scale part of the pseudostreamfunction forecast error covariance [see (2.11a) and (4.5) of Part II]. The perfect geostrophy corresponds to $a = 0$ in (2.2). One can show that this condition ($a = 0$) is equivalent to $C_{\phi\psi}(r) = C_{\phi\phi}(r) = C_{\psi\psi}(r)$ (see the appendix). Thus, by comparing $C_{\phi\psi}(r)$ with $C_{\phi\phi}(r)$ and $C_{\psi\psi}(r)$, we can see how closely the geostrophy is satisfied by the forecast error fields. The three covariance functions are compared in Figs. 3a–c at the vertical levels of $p_m = p_n =$ 850, 500, and 200 mb. As shown in Fig. 3a, $C_{\phi\psi}$ (solid) is flatter than $C_{\phi\phi}$ (dashed) and $C_{\phi\phi}$ is flatter than $C_{\psi\psi}$ (dotted), indicating that the geostrophy is not well satisfied and the variance of the error field ψ is larger than that of ϕ at 850 mb. In Fig. 3b, $C_{\phi\psi}$ nearly coincides with $C_{\phi\phi}$ and $C_{\psi\psi}$, so the geostrophy is well satisfied at 500 mb. In Fig. 3c, $C_{\phi\psi}$ is very close to $C_{\phi\phi}$ but much flatter than $C_{\psi\psi}$, so the geostrophy is not well satisfied and the variance of the error field ψ is much larger than that of ϕ at 200 mb.

Associated with the above three covariance functions, three correlation functions can be defined by $R_{\phi\psi}(r) \equiv C_{\phi\psi}(r)(\sigma_\phi\sigma_\psi)^{-1}$, $R_{\phi\phi}(r) \equiv C_{\phi\phi}(r)\sigma_\phi^{-2}$, and $R_{\psi\psi}(r) \equiv C_{\psi\psi}(r)\sigma_\psi^{-2}$, where σ_ϕ and σ_ψ are as defined in (2.2). The normalized cross-correlation function, defined by $R_{\phi\psi}(r)/R_{\phi\psi}(0)$, is compared with the two autocorrelation functions $R_{\phi\phi}(r)$ and $R_{\psi\psi}(r)$ in Figs. 4a–c at the vertical levels of $p_m = p_n =$ 850, 500, and 200 mb. As shown, the three functions are quite similar especially for $p_m = p_n = 200$ mb (Fig. 4c). For $p_m = p_n =$ 850 and 500 mb (Figs. 4a,b), $R_{\phi\psi}(r)/R_{\phi\psi}(0)$ is not very close to $R_{\phi\phi}(r)$

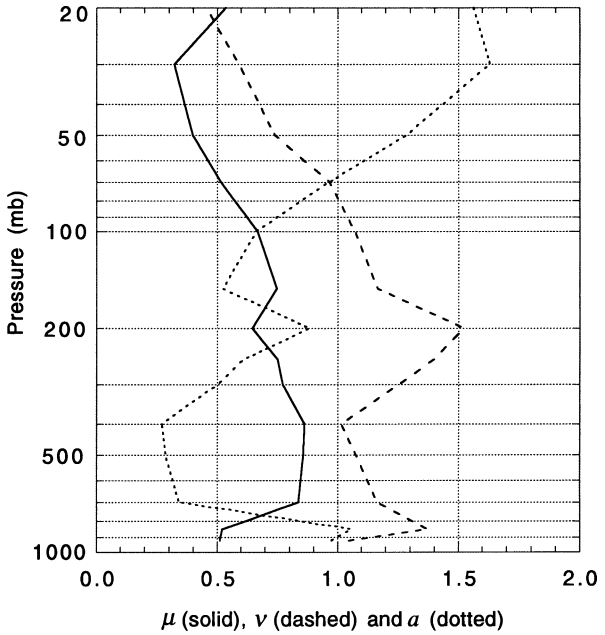

FIG. 5. Vertical profiles of μ (solid), ν (dashed), and a (dotted).

but nearly coincides with $R_{\psi\psi}(r)$. Since these correlation functions have similar shapes, how well the geostrophy is satisfied by the forecast error fields can be largely described by the relative amplitudes of the three covariance functions or, equivalently, by the following two parameters:

$$\mu = R_{\phi\psi}(0) = \sum_1 S_{\phi\psi}(k_i)[\sigma_\phi\sigma_\psi]^{-1}, \qquad (3.3a)$$

$$\nu = \sigma_\psi/\sigma_\phi, \qquad (3.3b)$$

where the dependence on p_m $(=p_n)$ is implicit. Here, μ is the correlation coefficient as in LH and ν may be called the relative amplitude between the error fields ψ and ϕ. It is easy to see that

$$a = \nu + \nu^{-1} - 2\mu, \qquad (3.4)$$

where a is defined in (2.2). Note that $\nu + \nu^{-1} \geq 2$ and $\mu \leq 1$, so $a = 0$ if and only if $\mu = 1$ and $\nu = 1$. This means that $a = 0$ is equivalent to $\mu = \nu = 1$, which is consistent with the result in the appendix. If $\mu = 1$ but $\nu \neq 1$, then ψ and ϕ are fully correlated but do not satisfy perfect geostrophy.

The computed parameter values of μ, ν, and a are plotted as functions of p_m $(=p_n)$ in Fig. 5. As shown, both μ and ν are close to unity and, consequently, a is small in the middle troposphere. Clearly, the geostrophy is well satisfied by the forecast error fields in the middle troposphere and this is consistent with the weak divergence of the wind forecast error in the middle troposphere (see Fig. 6 of Part II). The geostrophy is not well satisfied in the boundary layer as indicated by the jump of a from 0.34 at 700 mb to 1.06 at 850 mb in Fig. 5. The geostrophy is also not well satisfied around the tropopause as shown by the local maximum of a $(=0.88)$ at 200 mb (where $\mu = 0.65$ and $\nu = 1.52$). Around the tropopause, air masses come from two different origins (troposphere and stratosphere) and the divergent part of the wind forecast error reaches its maximal intensity (see Fig. 6 of Part II). This may explain why the geostrophy is not well satisfied around the tropopause in the pressure coordinates.

Above 100 mb, the spectral expansion of $C_{\phi\phi}(r)$ is severely truncated under the semidefinite positive constraint [see (3.6) of Part I] and consequently the spectral expansion of $C_{\phi\psi}(r)$ is also severely truncated under the constraint (3.2). The truncated high-wavenumber components are negative for $C_{\phi\phi}(r)$ but are not so for $C_{\phi\psi}(r)$. Thus, σ_ϕ could be overestimated and μ could be underestimated above 100 mb. The spectral expansion of $C_{\psi\psi}(r)$, however, is not severely truncated under the semidefinite positive constraint [see (2.11) and (4.2) of Part II], so σ_ψ is unlikely to be overestimated, at least, not as much as σ_ϕ. This implies that $\nu = \sigma_\psi/\sigma_\phi$ could be underestimated above 100 mb. Thus, the decreases of μ and ν and associated rapid increase of a above 100 mb in Fig. 5 could be spurious features.

The gross structures of the correlation functions—$R_{\phi\psi}(r)$, $R_{\phi\phi}(r)$, and $R_{\psi\psi}(r)$—can be measured by their associated horizontal length scales. As in Hollingsworth and Lönnberg (1986) [or see (6.8) of Part I], these length scales are defined by

$$L_{\psi\phi} \equiv \left[\sum\nolimits_1 S_{\phi\psi}(k_i)\right]^{1/2}\left[\sum\nolimits_1 k_i^2 S_{\phi\psi}(k_i)\right]^{-1/2} = \left[\sum\nolimits_1 S_{zt}(k_i)/k_i\right]^{1/2}\left[\sum\nolimits_1 k_i S_{zt}(k_i)\right]^{-1/2}, \qquad (3.5a)$$

$$L_{\phi\phi} \equiv \left[\sum\nolimits_1 S_{\phi\phi}(k_i)\right]^{1/2}\left[\sum\nolimits_1 k_i^2 S_{\phi\phi}(k_i)\right]^{-1/2} = \left[\sum\nolimits_1 S_{zz}(k_i)\right]^{1/2}\left[\sum\nolimits_1 k_i^2 S_{zz}(k_i)\right]^{-1/2}, \qquad (3.5b)$$

$$L_{\psi\psi} \equiv \left[\sum\nolimits_1 S_{\psi\psi}(k_i)\right]^{1/2}\left[\sum\nolimits_1 k_i^2 S_{\psi\psi}(k_i)\right]^{-1/2} = \left[\sum\nolimits_1 S_{tt}(k_i)/k_i^2\right]^{1/2}\left[\sum\nolimits_1 S_{tt}(k_i)\right]^{-1/2}. \qquad (3.5c)$$

where the summation Σ_1 is from $i = 1$ to $i = M$ for the synoptic-scale forecast error components resolved by the current analysis. Here, (2.11a) of Part II is used in the derivation of (3.5c). Clearly, $L_{\phi\phi}$ is essentially the same as L_z defined in (6.8) of Part I. The computed length scales are plotted as functions of p_m $(=p_n)$ in Fig. 6. Above 50 mb, the spectral expansion of $C_{\phi\phi}(r)$ is truncated to only one component under the semidefinite positive constraint [see (3.6) of Part I] and so does

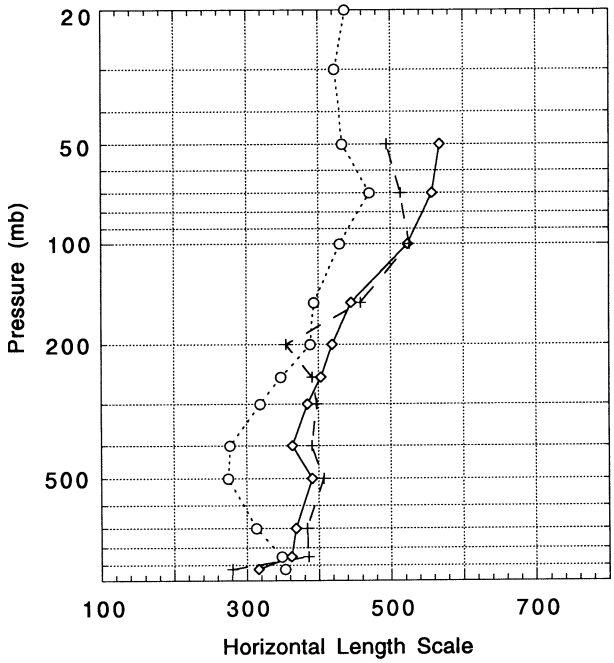

FIG. 6. Vertical profiles of horizontal length scales: $L_{\psi\phi}$ (solid), $L_{\phi\phi}$ (dashed), and $L_{\psi\psi}$ (dotted).

the spectral expansion of $C_{\phi\psi}(r)$ under the constraint (3.2). Because of this, $L_{\psi\phi}$ and $L_{\phi\phi}$ cannot be reliably estimated and thus are not plotted above 50 mb. As shown in Fig. 6, the ranges of these length scales and their general increases with height are similar to those (from 925 to 100 mb) in Fig. 18b of LH. The detailed structures and values, however, are different. For example, in the middle troposphere, we have $L_{\phi\phi} > L_{\psi\phi} > L_{\psi\psi}$ in Fig. 6, while LH obtained $L_{\psi\psi} > L_{\psi\phi} > L_{\phi\phi}$ in their Fig. 18b, although the three length scales are close to each other in both figures.

4. Multilevel analysis

When $p_m \neq p_n$, the analysis can be done similarly to the above single-level analysis. However, since $p_m \neq p_n$, the spectral coefficients $S_{zt}(k_i) = S_{zt}(k_i, p_m, p_n)$ and $S_{zl}(k_i) = S_{zl}(k_i, p_m, p_n)$ in (3.1a) and (3.1b) are not constrained by (3.2a) and (3.2b). Instead, they are constrained by

$$|S_{zt}(k_i, p_m, p_n)| \leq [S_{zz}(k_i, p_m, p_m)S_{tt}(k_i, p_n, p_n)]^{1/2}, \quad (4.1a)$$

$$|S_{zl}(k_i, p_m, p_n)| \leq [S_{zz}(k_i, p_m, p_m)S_{ll}(k_i, p_n, p_n)]^{1/2}. \quad (4.1b)$$

These conditions can be derived similarly to that in (3.2a) and (3.2b) by using the identities $C_{zt}(r, p_m, p_n) = C_{tz}(r, p_n, p_m)$ and $C_{zl}(r, p_m, p_n) = C_{lz}(r, p_n, p_m)$. The details are omitted.

The computed correlation function $R_{\psi\phi}(r, p_m, p_n)$ is plotted in Fig. 7a for fixed $p_m = 850$ mb (vertical level of ψ) and different p_n (vertical levels of ϕ). The plotted correlation function has a local maximum of 0.52 at $r = 0$ and $p_n = p_m$ (=850 mb). The ψ field at 850 mb is not only maximally correlated to the ϕ field at 850 mb but also positively correlated to the ϕ fields between 950 and 300 mb, although the correlation decreases gradually as the vertical level p_n is shifted away from

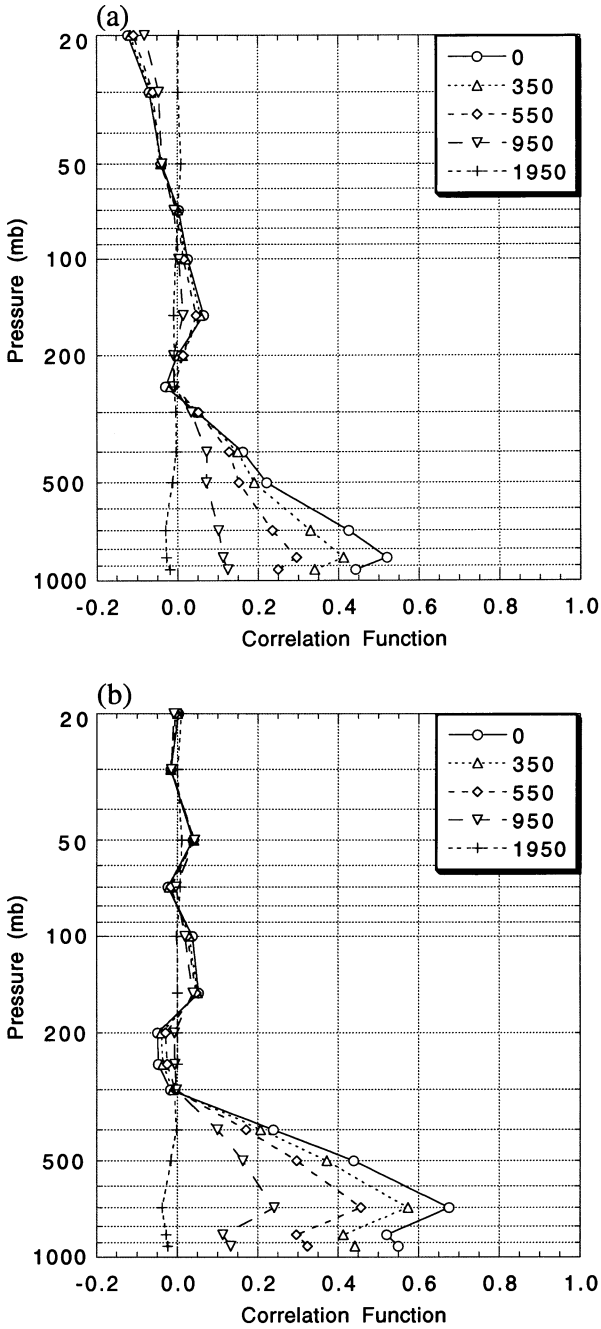

FIG. 7. Cross-correlation functions (a) $R_{\psi\phi}(r, p_m, p_n)$ and (b) $R_{\phi\psi}(r, p_m, p_n)$ plotted as functions of p_n for fixed $p_m = 850$ mb and $r = 0$, 350, 550, 950, and 1950 km.

850 mb. Note that $R_{\phi\psi}(r, p_n, p_m) = R_{\psi\phi}(r, p_m, p_n) \neq R_{\psi\phi}(r, p_n, p_m) = R_{\phi\psi}(r, p_m, p_n)$. The latter is plotted in Fig. 7b for fixed $p_m = 850$ mb (vertical level of ϕ) and different p_n (vertical levels of ψ). As shown, the correlation function in Fig. 7b has a local maximum of 0.68 at $r = 0$ and $p_n = 700$ mb, indicating that the ϕ field at 850 mb is correlated more closely to the ψ field at 700 mb than the ψ field of the same (850 mb) level. This is different from that in Fig. 7a. As shown in Fig. 5, the single-level geostrophic coupling, measured by $\mu = R_{\phi\psi}(0, p_m, p_n) = R_{\psi\phi}(r, p_n, p_m)$ with $p_n = p_m$, drops sharply from 0.84 to 0.52 as the vertical level ($p_n = p_m$) decreases toward the boundary layer from 700 to 850 mb. In comparison with this, the multilevel geostrophic coupling, measured by $R_{\phi\psi}(0, p_m, p_n) = R_{\psi\phi}(0, p_n, p_m)$, drops more sharply from 0.84 (see Fig. 5) to 0.43 (see Fig. 7a) as the vertical level of ψ decreases from 700 to 850 mb while the vertical level of ϕ is fixed at 700 mb, but it drops less sharply from 0.84 (see Fig. 5) to 0.68 (see Fig. 7b) as the vertical level of ϕ decreases from 700 to 850 mb while the vertical level of ψ is fixed at 700 mb.

Plotted in Figs. 8a and 8b are the correlation functions $R_{\psi\phi}(r, p_m, p_n)$ and $R_{\phi\psi}(r, p_m, p_n)$ for fixed $p_m = 500$ mb and different p_n. As shown, for $0 \leq r \leq 950$ km, $R_{\psi\phi}(r, p_m, p_n)$ in Fig. 8a is positive for all the vertical levels of p_n, but $R_{\phi\psi}(r, p_m, p_n)$ in Fig. 8b is positive mainly for the vertical levels of p_n below 200 mb. As the vertical level p_n is shifted away from 500 mb, the vertical profiles of $R_{\phi\psi}(r, p_m, p_n)$ in Fig. 8b drop more rapidly than those in Fig. 8a, and their values decrease to virtually zero as the vertical level of p_n (for the ϕ field) moves up to 200 mb. The major difference between Figs. 8a and 8b is that the ψ field at 500 mb is positively correlated with the ϕ fields at all the vertical levels but the ϕ field at 500 mb is positively correlated with the ψ fields mainly at the vertical levels within the troposphere (below 200 mb).

Figures 9a and 9b are the correlation functions $R_{\psi\phi}(r, p_m, p_n)$ and $R_{\phi\psi}(r, p_m, p_n)$ for fixed $p_n = 200$ mb. As shown, for $0 \leq r \leq 950$ km, $R_{\psi\phi}(r, p_m, p_n)$ in Fig. 9a is positive only when p_n (the vertical level of ϕ) is between 500 and 50 mb, while $R_{\phi\psi}(r, p_m, p_n)$ in Fig. 9b is positive as long as p_n (the vertical level of ψ) is above 850 mb. This is the major difference between Figs. 9a and 9b. From the symmetry of $R_{\psi\phi}(r, p_m, p_n) = R_{\phi\psi}(r, p_n, p_m)$, it is easy to see that the difference between Figs. 9a and 9b is consistent with the difference between Figs. 8a and 8b.

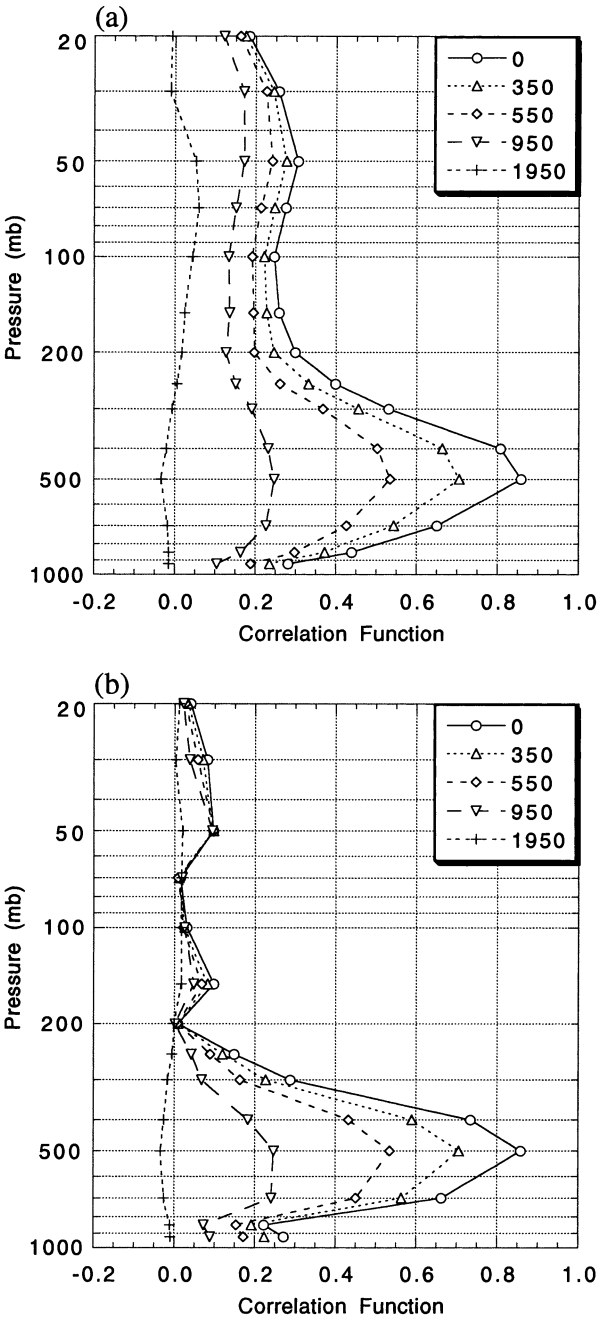

FIG. 8. As in Fig. 7 but for fixed $p_m = 500$ mb.

5. Conclusions

The method of statistical analysis of wind innovation (observation minus forecast) vectors is applied to the height and wind innovation data collected over North America for a 3-month period from NOGAPS to estimate the height–wind forecast error correlation and to evaluate the related geostrophy. The method and related analysis are extended upon the work of LH in three aspects: (i) new constraints are derived for the spectral representations of height–wind forecast error covariance functions [see (3.2) and (4.1)], (ii) the single-level analysis of cross covariance is extended to multilevel analysis [see section 4], and (iii) a new parameter a [see (2.2)] is introduced to measure the geostrophy. Zero value of this parameter ($a = 0$) defines the perfect geostrophy and this condition is shown to be the same as the equivalence condition between the cross covariance

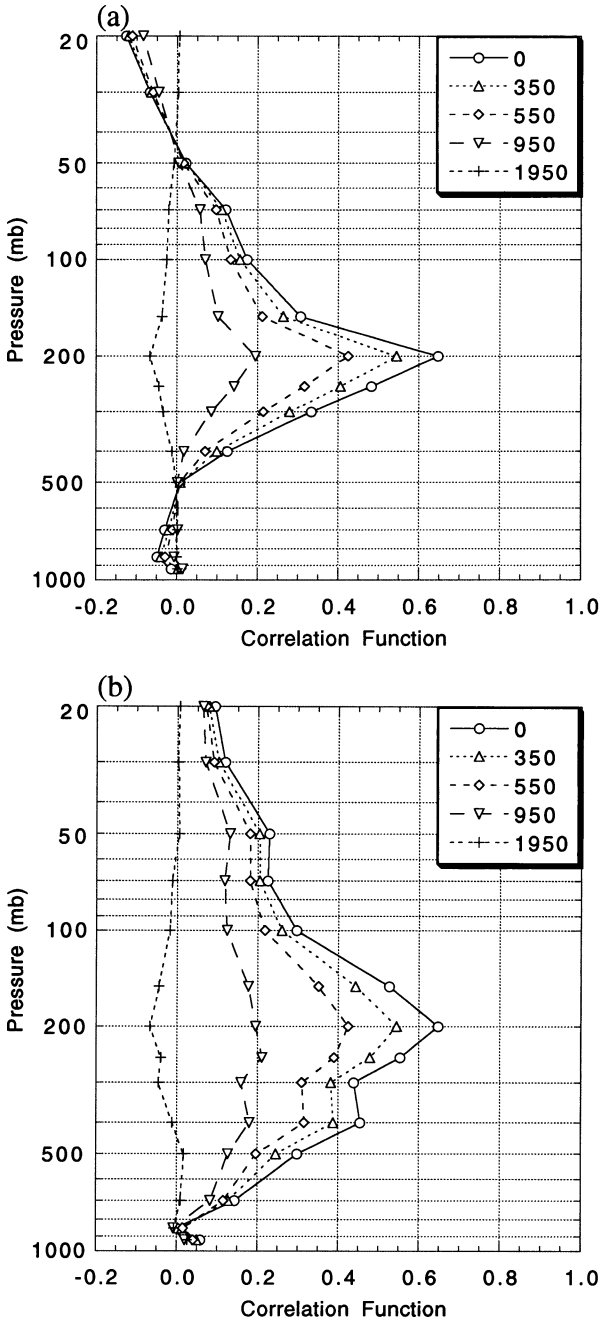

FIG. 9. As in Fig. 7 but for fixed p_m = 200 mb.

and two autocovariances of the pseudostreamfunction and geopotential forecast error [see the appendix]. It is also equivalent to $\mu = \nu = 1$, where μ is the correlation coefficient and ν is the standard deviation ratio between the pseudostreamfunction and geopotential forecast error fields [see (3.3) and (3.4)]. The smallness of a measures the closeness to perfect geostrophy and requires both μ and ν be close to unity. Conventionally, however, only μ is used for the geostrophic coupling in the multivariate optimal interpolation (Bergman 1979; Lorenc 1981; Daley 1991).

The single-level analysis performed in this paper shows that the height forecast error is correlated to the tangential wind forecast error but not (at least, not obviously) to the radial wind forecast error. This implies that the geopotential forecast error is correlated to the pseudostreamfunction but not to the pseudovelocity potential [defined in (2.1)]. This result is consistent with LH. It indicates that the geostrophy is satisfied, to a certain degree, by the forecast error fields. The degree to which the geostrophy is satisfied is examined quantitatively and the results can be summarized as follows (also see Figs. 5 and 6).

1) The geostrophy is well satisfied by the forecast error fields in the middle troposphere, and this is consistent with the weak divergence of the wind forecast error in the middle troposphere.
2) The geostrophy is not well satisfied in the boundary layer where the eddy diffusivity is strong.
3) The geostrophy is not well satisfied around the tropopause where air masses come from troposphere and stratosphere, and this is consistent with the maximal divergence of the wind forecast error at the tropopause (see Fig. 6 of Part II).
4) Above 100 mb, the geostrophy is not well satisfied by the synoptic-scale forecast error components resolved by the current analysis, but it may be well satisfied by the large-scale forecast error components not included in the current analysis (due to the limited data coverage).
5) The horizontal length scale of the cross correlation between the pseudostreamfunction and geopotential forecast error is found to be close to and mostly between the length scales of the two autocorrelation functions. The ranges of these length scales and their general increases with height are similar to those in LH, but the detailed structures are quite different.

When the National Meteorological Center (NMC) method (Parrish and Derber 1992) was extended and used to estimate the European Centre for Medium-Range Weather Forecasts (ECMWF) global model forecast error correlation structures, Derber and Bouttier (1999) found that a part of the divergent wind error can be explained by the rotational wind error in the middle latitudes near the surface and near the jet streams (at small scales) around the tropopause. A similar correlation between the vorticity and divergence forecast errors was also detected by Berre (2000) when the NMC method was further extended and applied to a limited-area model. Thus, the correlation between the pseudovelocity potential and geopotential forecast error may not be negligibly small. This was also suggested by Polavarapu (1995), based on divergent wind analyses in the marine boundary in the vicinity of cyclones. Nevertheless, the correlation between the pseudovelocity potential and geopotential forecast error appears to be very

subtle and cannot be detected by the analysis performed in this paper due to the limited resolution of the radiosonde data. Apparently, the radiosonde and associated innovation data used in this paper are too coarse to detect the aforementioned small-scale correlation between the vorticity and divergence forecast errors near the jet streams around the tropopause and the radiosonde standard vertical levels are too coarse to detect the aforementioned correlation near the surface.

The multilevel analysis shows that the pseudostreamfunction and geopotential forecast error fields are correlated between different vertical levels, although the correlation decreases gradually as the two vertical levels are shifted away from each other. It is found that the pseudostreamfunction at the middle troposphere (500 mb) is positively correlated with the geopotential forecast error fields at all the vertical levels, but the geopotential forecast error field at the middle troposphere is positively correlated with the pseudostreamfunctions only at the vertical levels below 200 mb. The pseudostreamfunction around the tropopause (200 mb) is positively correlated with the geopotential forecast error fields between 500 and 50 mb, while the geopotential forecast error field around the tropopause is positively correlated with the pseudostreamfunctions above 850 mb. These features are consistent with the cross-correlation symmetry [i.e., $R_{\psi\phi}(r, p_m, p_n) = R_{\phi\psi}(r, p_n, p_m)$]. The results obtained from the multilevel analysis suggest that the geostrophic coupling between different vertical levels are very significant and should not be neglected. Multilevel geostrophic coupling has complex vertical structures. When these vertical structures are described by the vertical normal modes, the geostrophic coupling between different vertical modes may remain to be significant. Thus, considering multilevel geostrophic coupling or multimode geostrophic coupling may improve the conventional multivariate optimal interpolation and three-dimensional variational data assimilation.

Acknowledgments. This study was benefited from discussions with Roger Daley, Andrew Van Tuyl, Edward Barker, James Goerss, and Keith Sashegyi at NRL Monterey. The innovation data were collected by Andrew Van Tuyl. Suggestions from Robert Davies-Jones at NSSL and the anonymous reviewers improved the presentation of the results. The research work was partially supported by NRL Grant N00173-98-1-G903, by NSF Grant ATM-9983077, and by FD-SDSU Contract N66001-97-D-5028 to the University of Oklahoma.

APPENDIX

Proof of $C_{\phi\psi}(r) = C_{\phi\phi}(r) = C_{\psi\psi}(r)$ for Perfect Geostrophy Defined by $a = 0$

Denote by $C_{\varepsilon\varepsilon}(r)$ the autocovariance function of $\varepsilon = \psi - \phi$. Substituting this into (2.2) gives

$$a\sigma_\psi\sigma_\phi = C_{\varepsilon\varepsilon}(0) = \int S_{\varepsilon\varepsilon}(k)k\, dk, \qquad \text{(A.1)}$$

where $S_{\varepsilon\varepsilon}(k) = \int C_{\varepsilon\varepsilon}(r)J_0(kr)r\, dr$ is the power spectrum of $\varepsilon = \psi - \phi$ and the dependence on $p_m = p_n$ is implicit. As a power spectrum, $S_{\varepsilon\varepsilon}(k)$ must be a non-negative function of k. Thus, if $a = 0$ in (A.1), then $C_{\varepsilon\varepsilon}(0) = \int S_{\varepsilon\varepsilon}(k)k\, dk = 0$, which implies that $S_{\varepsilon\varepsilon}(k) = 0$ for any k. Consequently, $C_{\varepsilon\varepsilon}(r) = \int S_{\varepsilon\varepsilon}(k)J_0(kr)k\, dk = 0$ or, equivalently, $C_{\varepsilon\varepsilon}(r) = C_{\phi\phi}(r) + C_{\psi\psi}(r) - 2C_{\phi\psi}(r) = 0$ for any r. The Hankel transformation of this result gives $S_{\varepsilon\varepsilon}(k) = S_{\phi\phi}(k) + S_{\psi\psi}(k) - 2S_{\phi\psi}(k) = 0$ or, equivalently,

$$S_{\phi\psi}(k) = [S_{\phi\phi}(k) + S_{\psi\psi}(k)]/2 \quad (\geq 0) \quad \text{for perfect geostrophy.} \qquad \text{(A.2)}$$

On the other hand, by using the method that derives (3.2), one can show that

$$|S_{\phi\psi}(k)| \leq [S_{\phi\phi}(k)S_{\psi\psi}(k)]^{1/2}. \qquad \text{(A.3)}$$

Note that the geometric mean cannot exceed the algebraic mean, that is, $[S_{\phi\phi}(k)S_{\psi\psi}(k)]^{1/2} \leq [S_{\phi\phi}(k) + S_{\psi\psi}(k)]/2$ and the equality stands only for $S_{\phi\phi}(k) = S_{\psi\psi}(k)$. Thus, (A.2) and (A.3) imply that $S_{\phi\psi}(k) = S_{\phi\phi}(k) = S_{\psi\psi}(k)$ or, equivalently, $C_{\phi\psi}(r) = C_{\phi\phi}(r) = C_{\psi\psi}(r)$ for perfect geostrophy.

REFERENCES

Bartello, P., and H. L. Mitchell, 1992: A continuous three-dimensional model of short-range forecast error covariances. *Tellus,* **44A,** 217–235.

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

Berre, L., 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Mon. Wea. Rev.,* **128,** 644–667.

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

——, and E. Barker, 2000: *NRL Atmospheric Variational Data Assimilation System (NAVDAS) Source Book 2000.* NRL Publ. NRL/PU/7530-00-418, 153 pp.

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

Gandin, L. S., 1965: *Objective Analysis of Meteorological Fields* (in Russian). Israel Program for Scientific Translation, 242 pp.

Hogan, T. F., and T. E. Rosmond, 1991: The description of the Navy Operational Global Atmospheric Prediction System's spectral forecast model. *Mon. Wea. Rev.,* **119,** 1786–1815.

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus,* **38A,** 111–136.

Lönnberg, P., and A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus,* **38A,** 137–161.

Lorenc, A., 1981: A global three-dimensional multivariate statistical analysis system. *Mon. Wea. Rev.,* **109,** 701–721.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.,* **120,** 1747–1763.

Polavarapu, S. M., 1995: Divergent wind analysis in the oceanic boundary layer. *Tellus,* **47A,** 221–239.

Rutherford, I. D., 1972: Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sci.*, **29,** 809–815.

Thiebaux, H. J., H. L. Mitchell, and D. W. Shantz, 1986: Horizontal structure of hemispheric forecast error correlations for geopotential and temperature. *Mon. Wea. Rev.*, **114,** 1048–1066.

Xu, Q., and L. Wei, 2001: Estimation of three-dimensional error covariances. Part II: Analysis of wind innovation vectors. *Mon. Wea. Rev.*, **129,** 2939–2954.

——, ——, A. Van Tuyl, and E. H. Barker, 2001: Estimation of three-dimensional error covariances. Part I: Analysis of height innovation vectors. *Mon. Wea. Rev.*, **129,** 2126–2135.