THE STATISTICAL STRUCTURE OF SHORT-RANGE FORECAST ERRORS AS DETERMINED FROM RADIOSONDE DATA PART II: THE COVARIANCE OF HEIGHT AND WIND ERRORS

P.Lönnberg and A.Hollingsworth
European Centre for Medium Range Weather Forecasts
Reading, U.K.

ABSTRACT

Part I of this study analysed the statistical structure of the mid-latitude errors of the short-range wind forecasts used in the global data assimilation system at ECMWF, by comparing the forecasts with verifying radiosonde data over North America. After an analysis of the corresponding statistics for the errors of the height forecasts, this paper studies the covariance of the height and wind forecast errors.

The methods of Part I, based on the theory of homogeneous turbulence, are used to provide a spectral description of the height autocovariance function and of the cross-covariances of height with stream function and velocity potential. Particular attention is paid to the question of the degree of geostrophy of the non-divergent forecast errors. As a by-product, the calculations provide estimates of the vertical covariance matrices for prediction error and radiosonde observational error in the height field, where the term observational error includes both instrumental error and errors of representativeness.

The forecast errors for height are comparable in magnitude with the observation errors, and there are good grounds for increasing the resolution of the analysis system, both in the horizontal and the vertical. The height errors have a substantial large-scale component whose vertical structure has a very broad scale; the geostrophic wind errors are dominated by synoptic

scales. There is a high directional correlation (0.89) between the geostrophic wind and the stream function wind. The magnitudes of the geostrophic and non-divergent wind errors agree to within 15% in the troposphere. In the stratosphere the geostrophic wind errors are somewhat smaller than the non-divergent wind errors, indicating a possible aliasing from the large scales to synoptic scales in our calculations there. The correlation of height and velocity potential is such as to imply convergence in lows in the troposphere, but divergence in lows in the stratosphere.

The methods developed here and in Part I offer a powerful set of diagnostic tools with which to improve both analysis and short-range forecast performance.

1. INTRODUCTION

The first part of this study (Hollingsworth and Lönnberg 1985, hereafter called Part I) analysed the statistical structure of the mid-latitude errors of the short-range wind forecasts used in the global data assimilation system at ECMWF, by comparing the forecasts with verifying radiosonde data over North America. This paper presents the corresponding results for the height field errors, and for the cross-covariances of the height and wind field errors.

The analysis of the height field errors follows the general approach of Rutherford (1972) and Hollett (1975), but we present a more complete analysis of the vertical correlation structure of both horizontally correlated prediction error and horizontally uncorrelated observational error. The fact that the calculated vertical correlation structure for radiosonde thickness error agrees with expectations gives confidence in the results. Throughout this work the term observational error includes both instrumental error and sampling error or errors of representativeness.

The analysis of the cross-correlations of the height and wind errors uses the formalism introduced in Part I to examine the height-stream function correlation and the height-velocity potential correlation separately. The results are presented in a spectral form, which has interesting potential applications for high resolution work.

As discussed in Part I, observational data gives point information on forecast errors. An accurate specification of the statistical structure of the forecast error is essential to make a good interpolation of the observational information to a three-dimensional grid, so as to produce the analysis from which the next forecast can start. Knowledge of the structure of forecast errors can justify the imposition of constraints on the analytic forms of the

structure functions. As discussed by Daley (1983), such constraints have a profound effect on the ability of the analysis to extract useful information from noisy data.

Despite its importance for statistical analysis, the determination of the statistical structure of forecast errors has received little attention in the literature. The unpublished thesis by Hollett (1975) on the forecast errors of the Canadian 5-level model in the fall of 1974 has provided the main statistical basis for the specification of mid-latitude forecast errors in the ECMWF assimilation system (Lorenc 1981).

The main features of the results from our study are that the forecast errors for height are comparable in magnitude with the observation errors, and that there are good grounds for increasing the resolution of the height analysis, both in the horizontal and the vertical. The geostrophic wind errors are dominated by synoptic scales; the height errors have a substantial large-scale component whose vertical structure has a very broad scale. There is a high directional correlation between the geostrophic wind and the stream function wind. The geostrophic wind and non-divergent wind magnitudes agree to within 15% in the troposphere. The correlation of height and velocity potential is such as to imply convergence in lows in the troposphere, but divergence in lows in the stratosphere.

Section 2 discusses the assimilation system, and the data used for the study. The autocorrelation structures of the height field prediction and observational errors are discussed in Section 3. The mathematical methods for establishing the cross-correlations of height with the stream function and velocity potential are described in Sect.4. Sect.5 describes the results of the height-stream function calculations, with particular emphasis on the

degree of geostrophy of the non-divergent wind errors. The cross-correlation of the height and divergent wind field errors is discussed in Sect. 6. We conclude in Sect. 7 with a discussion of the results.

2. THE ASSIMILATION SYSTEM AND THE STATISTICAL DATA

The assimilation system is an intermittent insertion system consisting of three main steps - the analysis step, the initialisation step, and the forecast step. The analysis system is described by Lorenc (1981) and is an application of the optimal interpolation technique discussed by Gandin (1963), Rutherford (1972) Schlatter (1975) and Bergman (1979). Similar methods are used in several operational centres (Gustavsson 1981). The notable feature of the ECMWF implementation is that the analysis is performed for a large number of grid-points and variables simultaneously, which requires the selection of a large quantity of data for each analysis volume. The demands on computer power are correspondingly large; a typical analysis requires the inversion of several thousand matrices with orders between 100 and 200.

The initialisation scheme is an application of the non-linear normal mode scheme proposed by Machenhauer (1977), and described by Temperton and Williamson (1981), Williamson and Temperton (1981); it has been modified to include diabatic effects by Wergen (1982, pers.comm.).

The data studied here are the differences between the observations and the 6-12 hour grid-point model forecasts for the period 1 January-30 March, 1983. Only radiosonde data for 1200GMT is used; a particular station is used only if a minimum of 60 acceptable reports were available from the station. Acceptable means that the data was accepted as probably correct by all the stages of the operational quality control procedure. Attention is concentrated on the North American region between 30°N and 60°N. Later

papers will consider the results for other regions. Further information on the assimilation system and the data may be found in Part I.

The mean difference between the station reports and the forecast is removed separately for each station. The average variance of the resulting ensemble of station time series is defined as the ensemble mean of the station variances. We use the notation <a,b> for the correlation of a with b, cov<a,b> for the covariance and E_a, E_b for the standard deviations so that

$$\text{cov}\langle a,b\rangle = E_a E_b \langle a,b\rangle$$

Spherical geometry is used to define distances and angles.

3. THE HEIGHT FIELD

3.1 Mathematical formulation

The mathematical formalism used in Part I for the stream function and velocity potential fields will be adopted here. We seek a representation of the auto-correlation data in terms of cylindrical harmonic expansions. Positive definiteness requires that the expansions be truncated before the first negative term in the isotropic component, and also requires that the phase functions of the anisotropic components be continuous. In fact we shall only consider the isotropic component in this paper. Then the isotropic component of the prediction error auto-covariance may be written in the form.

$$\text{cov}\langle\phi,\phi\rangle = E_\phi^2 \sum_{n=o}^{N} \phi_n^2 J_o (k_n r/D),$$

where E_ϕ is the rms prediction error of height for the level in question, and the ϕ_n, by definition, satisfy

$$\sum_{n=o}^{N} \phi_n^2 = 1.$$

The normalised spectrum of the error variance is then given by $\phi_n^2 J_o^2(k_n)$ (Hildebrand 1962, p230) where

$$\phi_n^2 = \int_o^1 \frac{r}{D} \langle\phi,\phi\rangle \, J_o(k_n \frac{r}{D}) \, \frac{dr}{D} / J_o^2(k_n)$$

The implied wavenumbers in the Bessel function expansion for the radial dependence are determined by the requirement of vanishing radial derivative at some large distance D (3000 km in this paper), where the slope of the auto-correlation is close to zero. The gravest term in the expansion is a constant term corresponding to forecast errors which are perfectly correlated over the domain. We shall speak of this as the large-scale component, and we shall speak of the remaining terms as the synoptic scale components. The equivalent total spherical wavenumber of each of the terms in the series is given in Table 1 of Part I.

The expansion coefficients are determined by a least squares fit of the series to the empirical correlation data. The extrapolation of the fitting function to zero separation partitions the total perceived forecast error into a horizontally correlated part which is described as prediction error, and a horizontally uncorrelated part which is described as observational error (Rutherford 1972). The calculation is performed for the height errors at each level. A similar calculation is made for the thickness errors for all pairs of levels in order to establish the vertical correlation matrices for observational error and for each term separately in the prediction error

expansion (see I for details). In this way we can evaluate the validity of an assumption of separability.

(a) Partitioning of observation and prediction errors

Fig. 1 shows the data for the 500 mb $\langle\phi,\phi\rangle$ correlation, together with an 8-term fitting function. The data have been grouped in 100 km bins using the Fisher (1921) averaging algorithm. The criterion for truncating the expansion was that the first omitted term had a negative sign, unless the negative term was less than 10^{-3} of the variance. In such a case the term was reset to a very small positive number and the expansion was continued. This happened at one level (500 mb) and allowed the inclusion of the 7th and 8th terms (cf Fig.7). The area mean of the correlation in Fig.1 is clearly non-zero, so that the constant term will be non-zero. The intercept of the fitting curve with the ordinate gives the partition of the perceived forecast error into prediction error and observation error.

Fig. 2a shows the vertical profiles of perceived error, prediction error and observation error. The observation error and prediction error are of similar magnitude throughout the atmosphere. This creates new possibilities for monitoring the behaviour of the analysis system and the observing system (Hollingsworth et al. 1985).

Fig. 2b shows the profile of prediction error together with the profiles of the contributions from the constant term and the non-constant terms; we shall speak of these as the large-scale and the synoptic-scale components.

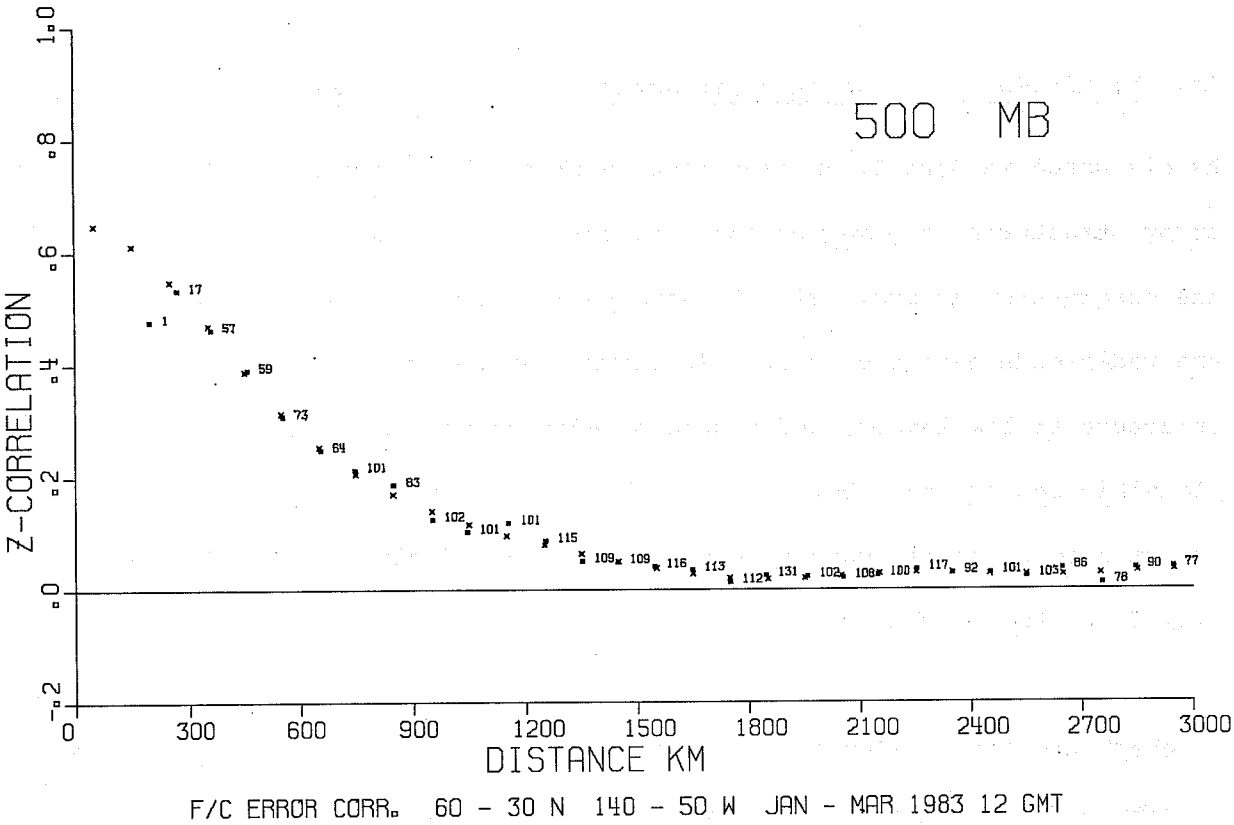


Fig. 1 The correlation of 500mb height forecast errors as a function of station separation. The empirical data, averaged over 'bins' of 100km, is shown by the squares. The figures indicate the number of station pairs in each bin. The smooth curve (x) is obtained by a least squares fit to the empirical data; the truncation is 8 terms in the synoptic-scale component

The peak in the prediction errors at 300 and 250 mb is mainly due to the synoptic-scale part of the prediction error. The ratio of the synoptic-scale to the large-scale part of the ϕ forecast errors is larger than 1.0 in the upper troposphere and less than 1.0 in the stratosphere, suggesting that the forecast errors arise for different reasons in the two regions.

(b) Magnitude and correlation structure of observational thickness errors

As discussed in Part I, an important check on the estimates of the prediction error structures is given by the structure of the corresponding results for the observation errors. Fig. 3 shows the vertical correlation structure of the radiosonde height errors. It shows a narrow vertical correlation structure at low levels, and a much broader structure at high levels. This is plausible because the observational thickness errors for standard layers are expected to be nearly uncorrelated, and so should give the structure shown in Fig. 3 for the height error.

To check the latter argument, Fig. 4 shows the observational error correlation structure for thickness, as derived from the height observation error covariance matrix. The layers are defined as those between adjacent standard levels. For most layers the correlation of observational error with adjacent layers is rather small. The inter-layer observational error correlation for thickness exceeds 0.4 in magnitude only for the 500/700 and 700/850 mb layers.

The estimated observation errors for thickness show hardly any variation of the rms thickness error in the upper troposphere, and give a value of between 6 and 9 metres for most layers. Fig. 5 shows the estimated prediction error and observation error for standard layer virtual temperatures, derived from

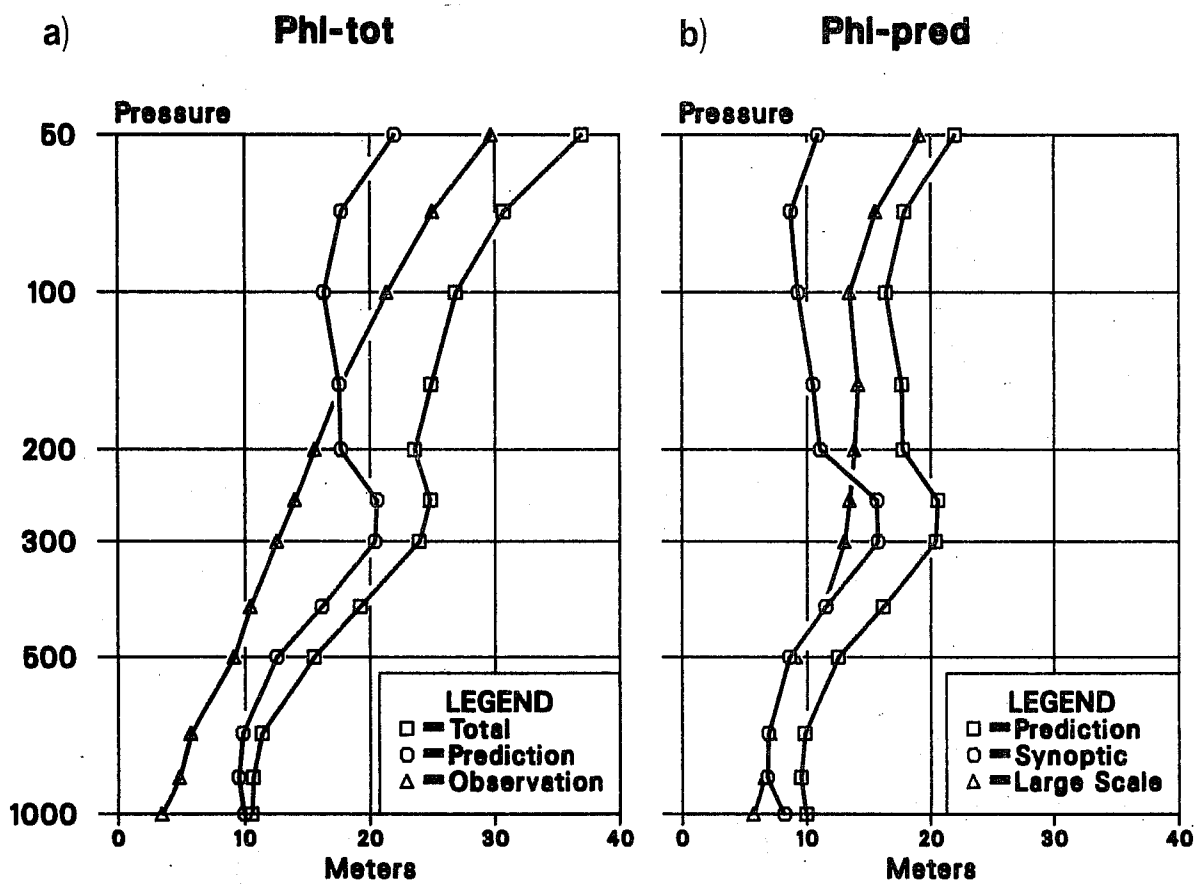


Fig. 2a: Vertical profiles of the total, or perceived, forecast error of height, together with the contributions to this error from the the prediction error, and the observation error. The unit is metre.

2b: Vertical profiles of the prediction error (copied from 2a) and of the contributions of the synoptic-scale and large-scale components to the prediction error. The sum of the squares of the components gives the square of the prediction error.

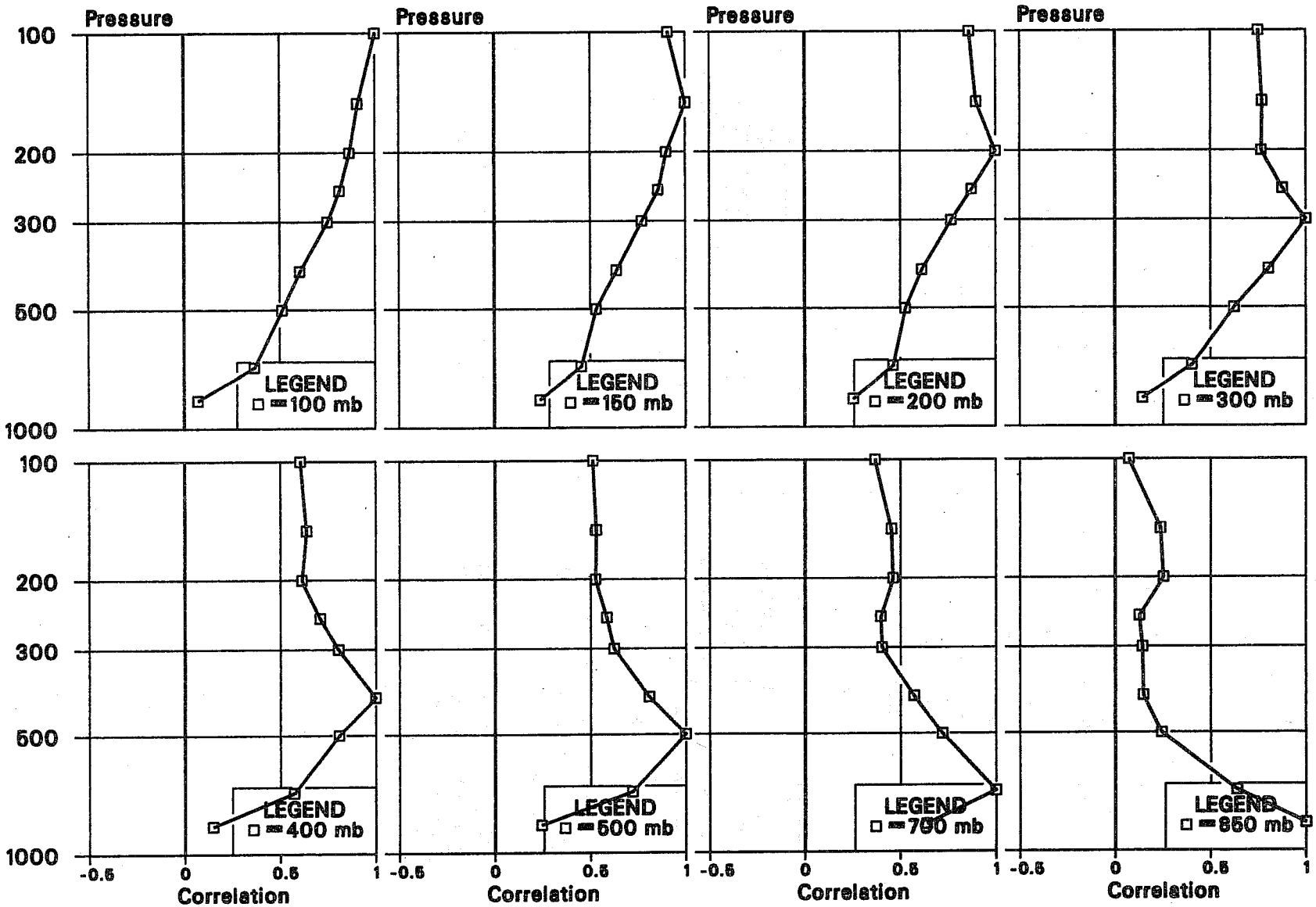


Fig. 3 Vertical profiles of the vertical correlation of radiosonde observational error for height, for the set of selected standard levels indicated in the legend of each frame of the plot.

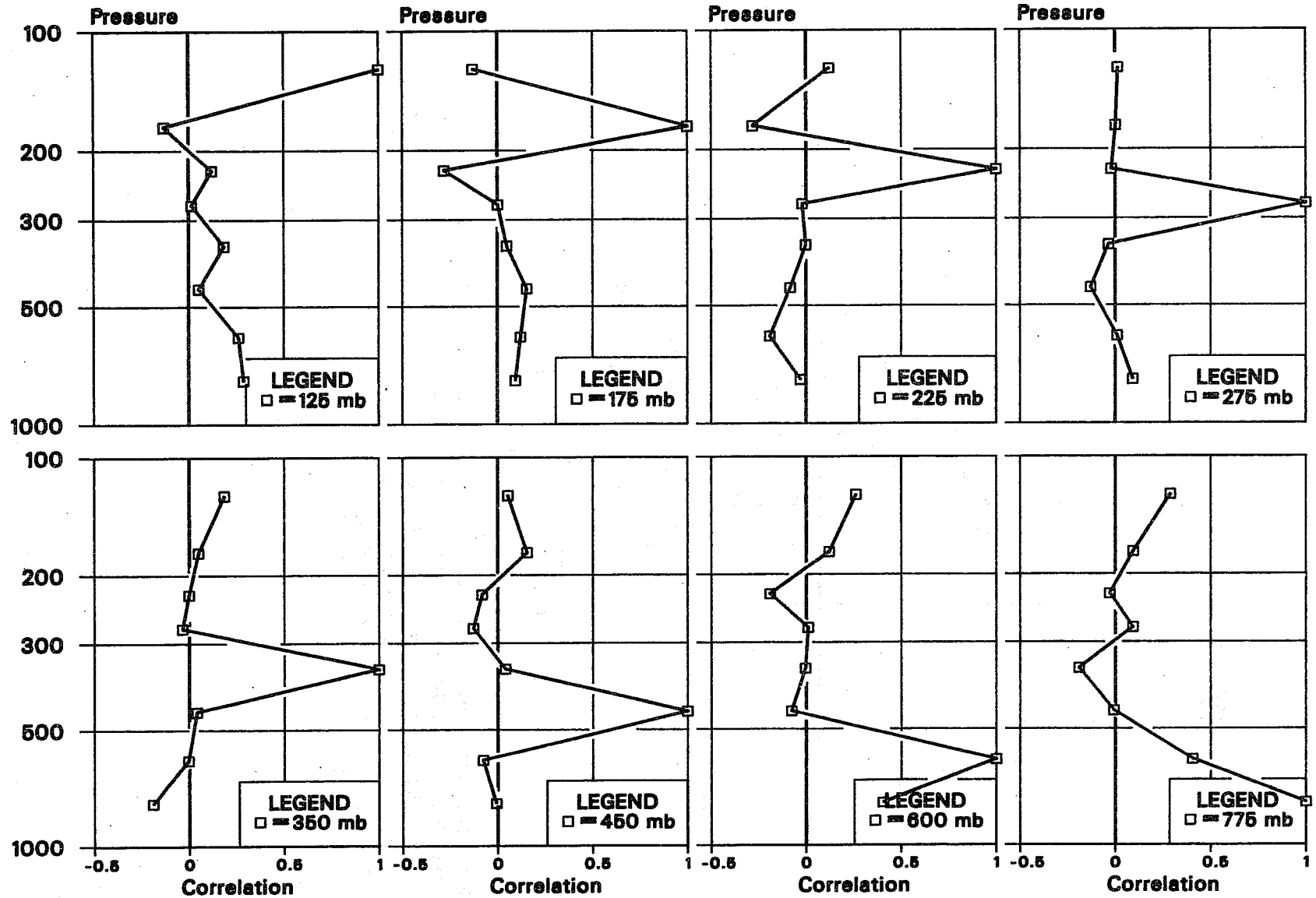


Fig. 4 Vertical profiles of the vertical correlation of radiosonde observational error for the thickness of layers between adjacent standard levels, for the set of selected layers indicated by the legend on each frame of the plot. Thus 125mb corresponds to the 100-150mb thickness and so on down to the 700-850mb layer, labelled by 775mb.

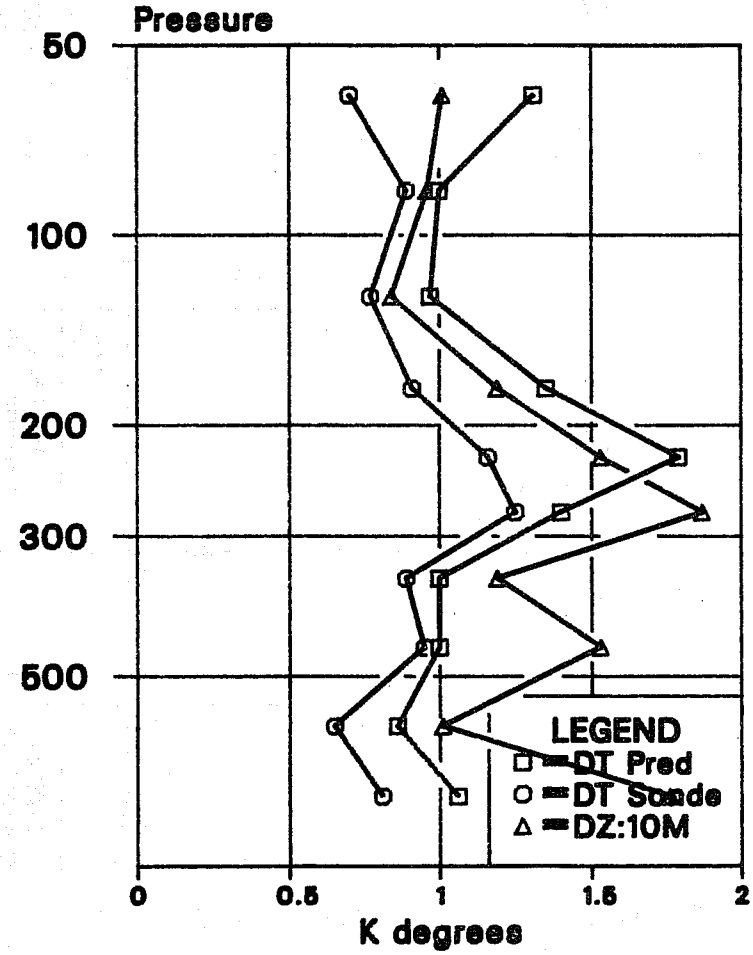


Fig. 5 Vertical profiles of the prediction error for the layer-mean virtual temperature for layers between adjacent standard levels (DT pred) and the corresponding observation error (DT Sonde). Also shown is the layer mean virtual temperature error corresponding to a thickness error of 10m (DZ:10m). This is the minimum detectable thickness forecast error above 500mb, because of reporting practices. The unit is K

the thickness errors. The prediction and observation errors have quite comparable magnitudes, of order 1K, for most layers except near the tropopause. There the prediction errors are particularly large. It is surprising that the observational error is as large as 1K for layers in excess of 1km thick, when the error for spot temperatures is quoted at about 1K (WMO 1983). Heights are rounded to the nearest decametre above 500 mb in the internationally agreed coding practice. The rounding error in height is therefore uniformly distributed between ±5m. The rounding error in thickness then has a triangular distribution with an rms value of ~4m. This rounding error is not negligible compared to the rms thickness error. If this error were eliminated, then the rms observational error would be reduced by a substantial amount.

It is hard to quarrel with the reasonableness of either the magnitude or correlation structure of the thickness errors and height errors. We conclude therefore that the prediction errors are also reasonably defined.

(c) The minimum detectable forecast error in thickness

Fig. 5 includes a plot of the temperature corresponding to a thickness of 10m, which is the minimum detectable forecast error in thickness above 500 mb. The relationship between this curve and the curves for both forecast error and observational error is clear. It is difficult to make a good thickness analysis near the tropopause since the minimum detectable difference between the forecast and the observation is of the same order as the rms forecast error, i.e. ~1.5 K. Attempts by an analysis centre to recalculate the thickness from the reported temperature are bound to be inadequate. It ought to be possible to remedy a situation where the minimum signal that can be

reported for thickness temperature is about twice as large as the rms instrumental error for temperature.

3.3 Vertical variation of the length scale and spectrum of the height prediction errors

Fig. 6 shows the vertical variation of the length scale (defined in Part I) of the $\langle\phi,\phi\rangle$ correlation for truncations of 6, 8 and 10 terms in the synoptic scale components. The results are very similar in their general shape to the corresponding results for the stream function shown in Part I (Fig.13). In the calculations for Fig. 6 the effect of the large-scale component of the height errors was excluded. The length scales are increased by 10 to 20 km in the troposphere if the constant term is included, and by 30 to 100 km in the stratosphere. The effect of retaining more terms in the expansion is noticed mainly in the troposphere (at 500 mb the $\langle\phi,\phi\rangle$ expansion terminated at 8 terms), as was also the case for the stream function.

The scale of the ϕ errors is nearly constant in the troposphere with a value of ~300 km, and shows an increase to about 450 km in the stratosphere. This result implies that the assumption of separability will be more reasonable if the troposphere and stratosphere are analysed separately, with appropriate correlation functions, than if they are analysed together, with the same correlation function for ϕ.

Fig. 7 shows a plot of the normalised spectra of the $\langle\phi,\phi\rangle$ correlation, as a function of height. It is evident from the normalised spectra that there is good level to level consistency between the results. There is a marked tendency for larger scales to predominate at higher levels in the atmosphere.

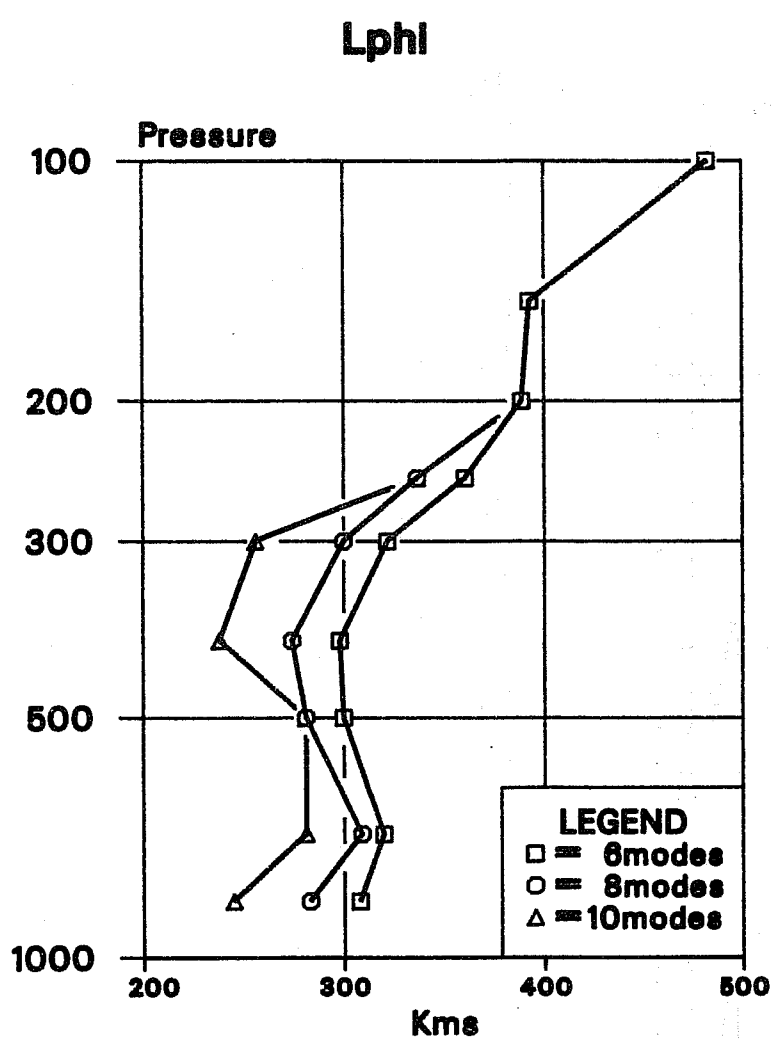


Fig. 6 Vertical profiles of the length scale of the height auto-correlation for truncations of 6, 8, and 10 modes in the least squares procedure. At 500mb the expansion was truncated at 8 terms. The unit is Km.

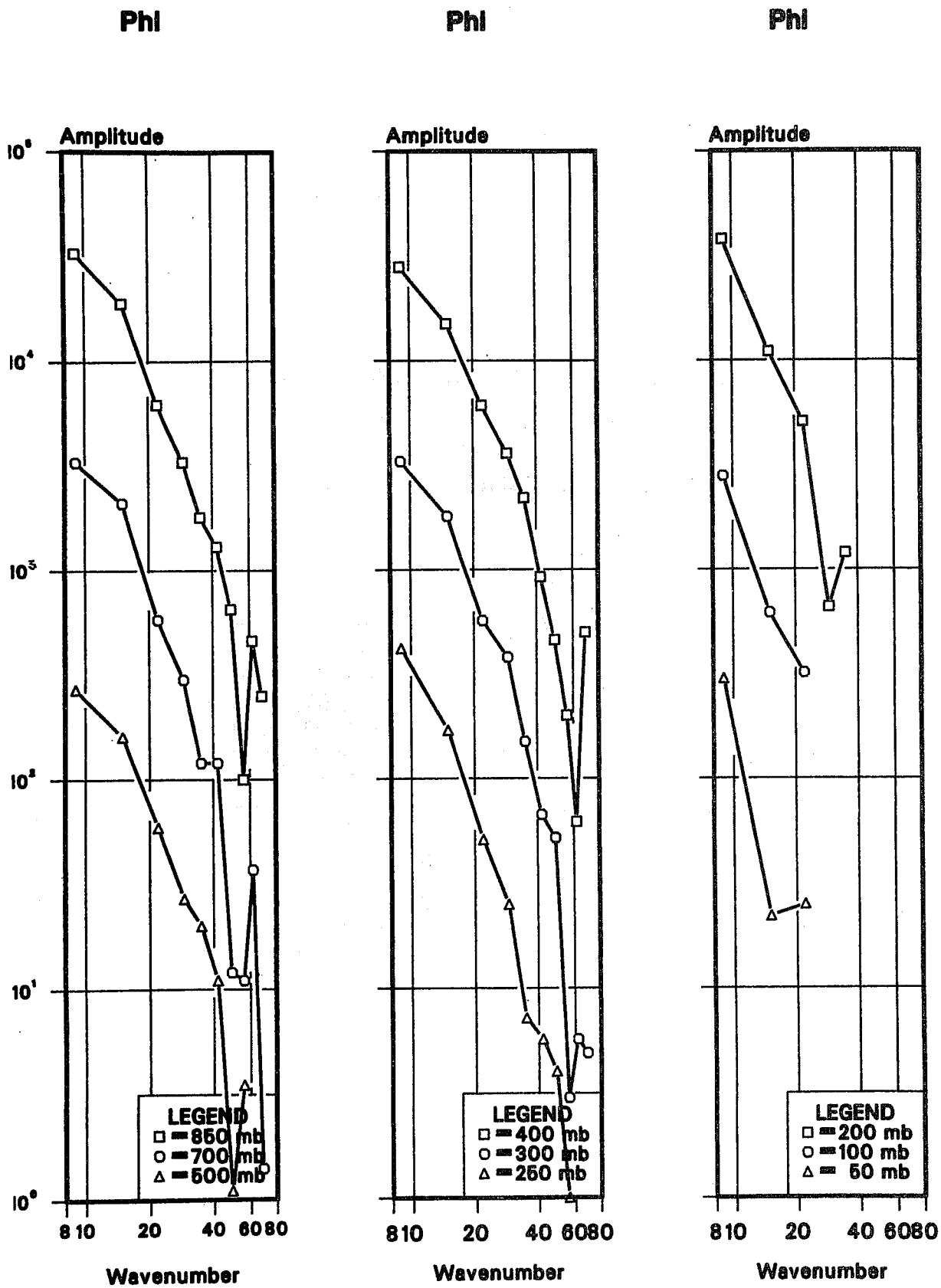


Fig. 7 Normalised spectra of the height auto-correlation at selected levels. The curves have been separated for clarity, so the origin of the ordinate is arbitrary; horizontal lines mark decades. The abscissa is equivalent planetary wave-number.

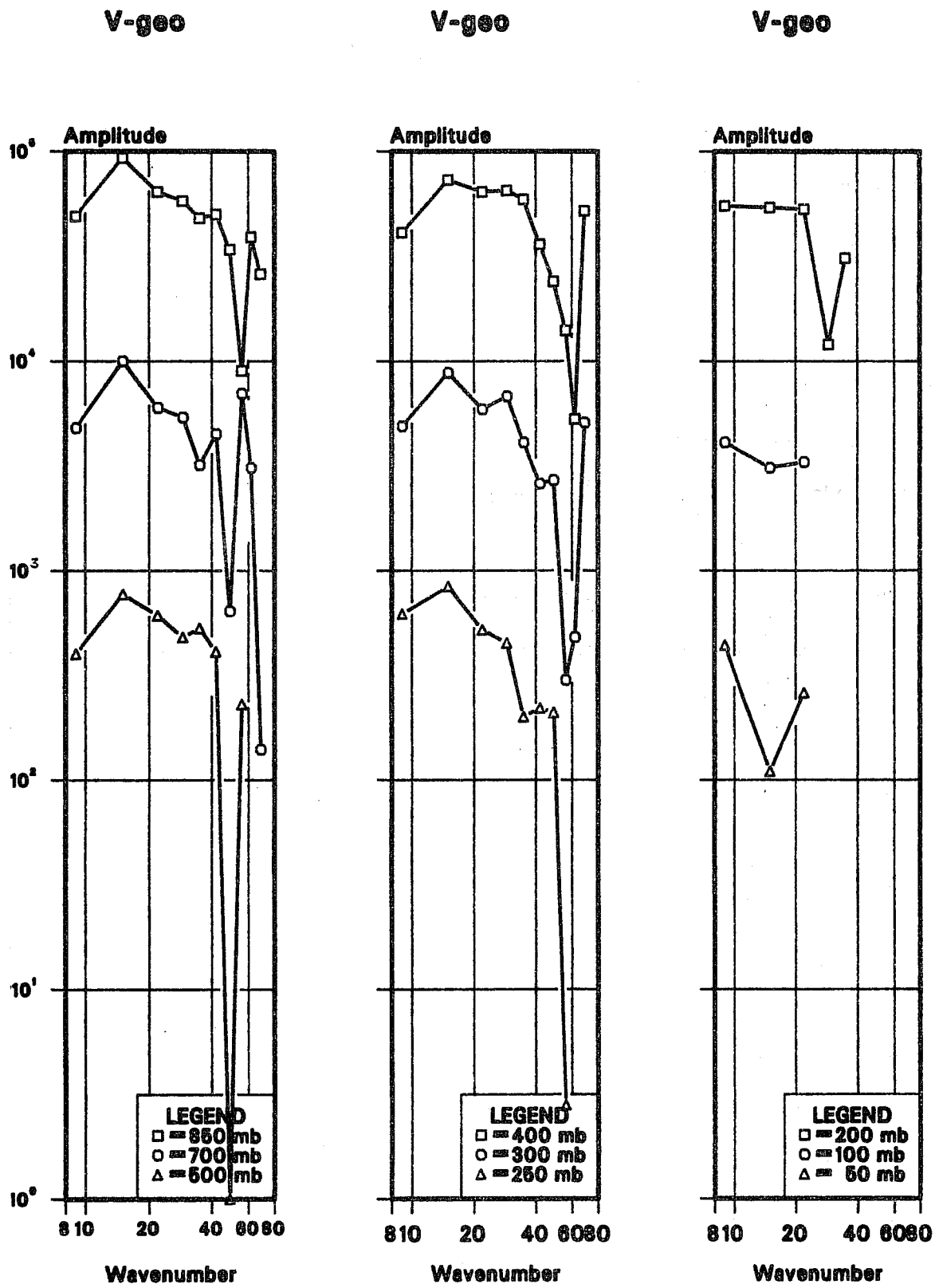


Fig. 8. Normalised spectra of the Laplacian of the height auto-correlation. This gives the normalised energy spectrum of the geostrophic wind auto-correlation implied by the height auto-correlation, if the Coriolis parameter is treated as constant. See Fig. 7 for remaining details.

The spectra in Fig. 7 fall rather steeply with wavenumber. To study this more closely we present the corresponding spectra for $\nabla^2(-\langle\phi,\phi\rangle)$, Fig. 8. These spectra correspond to the geostrophic vector wind error spectra. The weighting by the square of the wavenumber amplifies the inevitable noise in the determinations of the high wavenumber (and small amplitude) components. Nevertheless there is remarkable level-to-level consistency in the geostrophic wind spectra. At most tropospheric levels, the largest component of the geostrophic wind spectrum is mode 2 (equivalent to planetary wavenumber 15). In general these levels show a power law dependence of the spectrum for the higher modes in the troposphere, corresponding to slopes between $k_n^{-\frac{1}{2}}$ and k_n^{-1}.

At 200, 100 and 50 mb, only a few terms of the synoptic scale expansion could be determined. At these three levels, the first three components of the geostrophic wind have similar magnitude.

3.4 Vertical correlation of height prediction error

To consider the validity of the separability assumption we examine the vertical correlations for each mode with the 250 mb level in Fig. 9a. Each curve corresponds to the appropriate column of the vertical correlation matrix for that mode. In the stratosphere the higher horizontal modes do not occur with sufficiently large amplitudes to allow a reliable determination of the amplitudes; this restricts the determination of the corresponding correlations. The correlations of the 250 mb level with other tropospheric levels indicate that the modes with larger horizontal scales tend to have broader correlation structure in the vertical. This is also found for other levels, and suggests that the assumption of separability is not as well justified for the height field as it is for the non-divergent wind field.

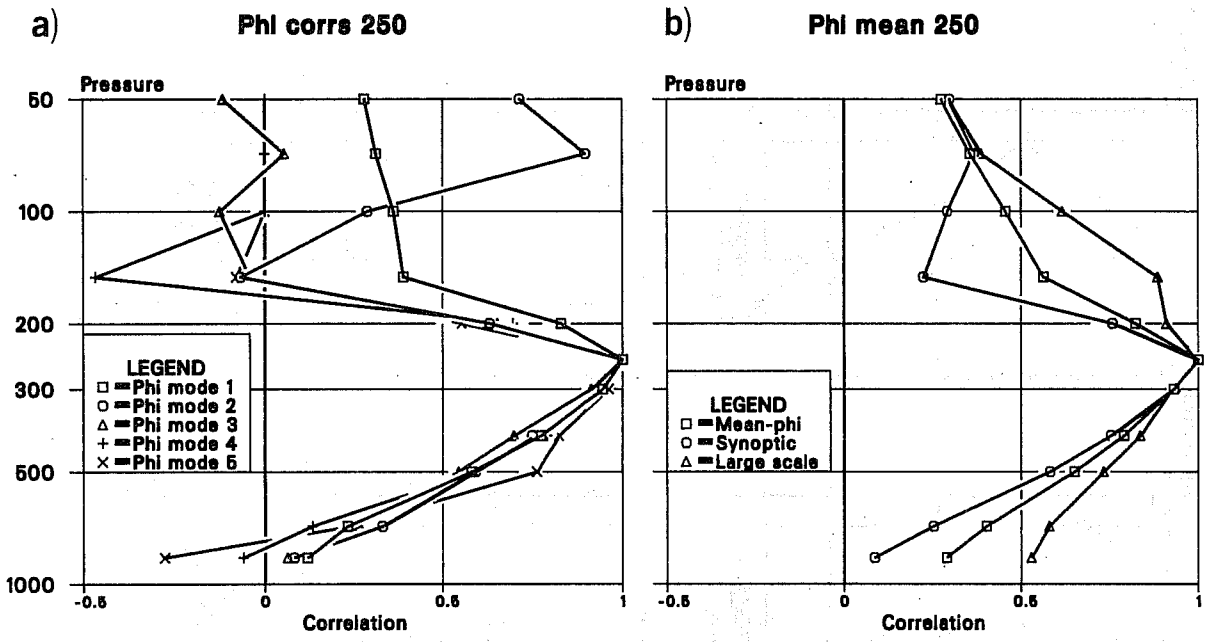


Fig. 9a: Vertical correlations with 250 mb of the first five terms of the expansion of the horizontal geopotential correlation. Terms 4 and 5 are undefined above 100mb and are set to zero.

9b: Vertical correlations with 250mb of the overall mean geopotential forecast errors (Mean-phi) together with the correlations for the synoptic-scale and large-scale components.

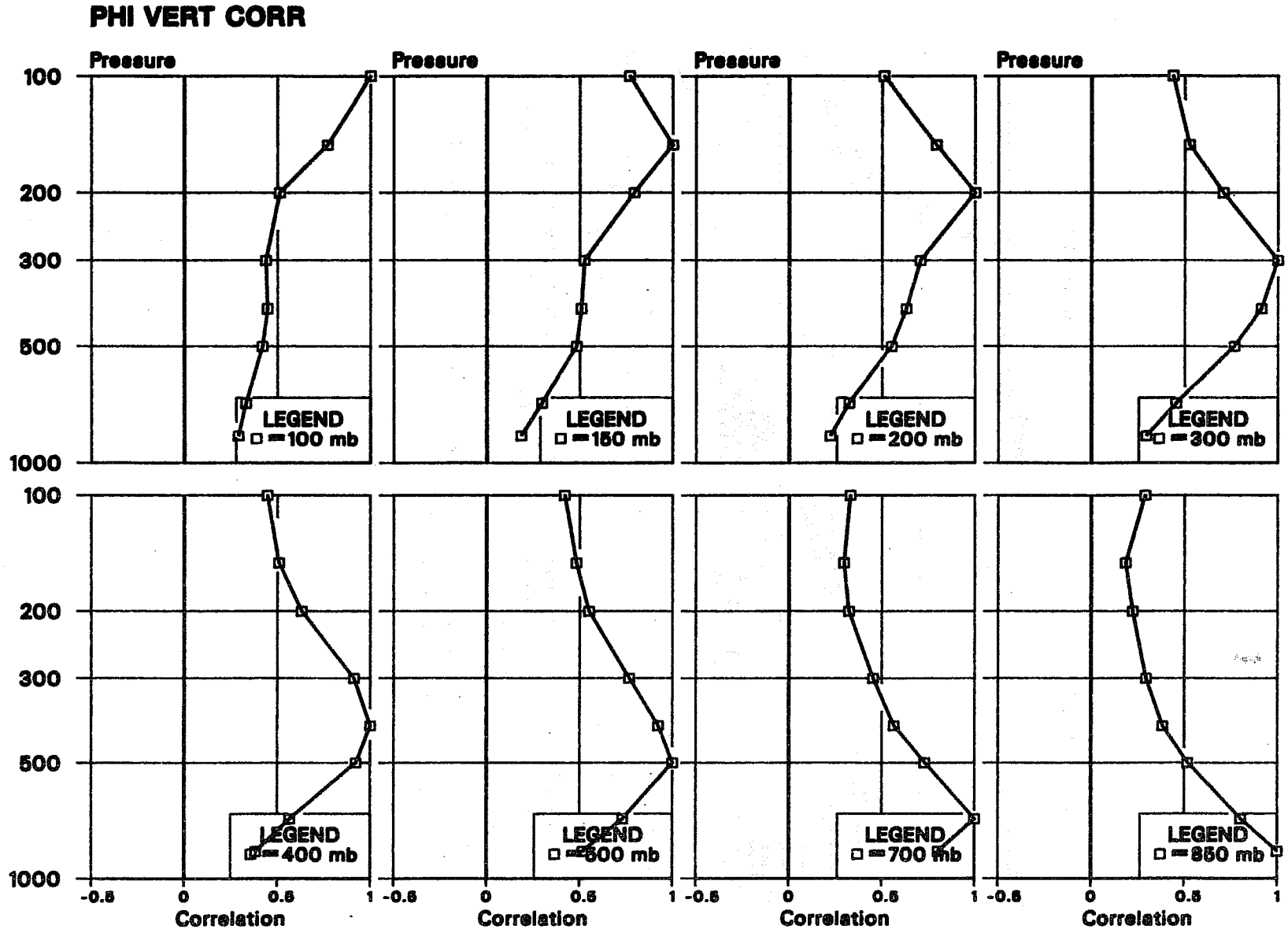


Fig. 10 Height prediction error vertical correlations for a selected set of standard levels, indicated in the legend of each frame of the plot.

a)

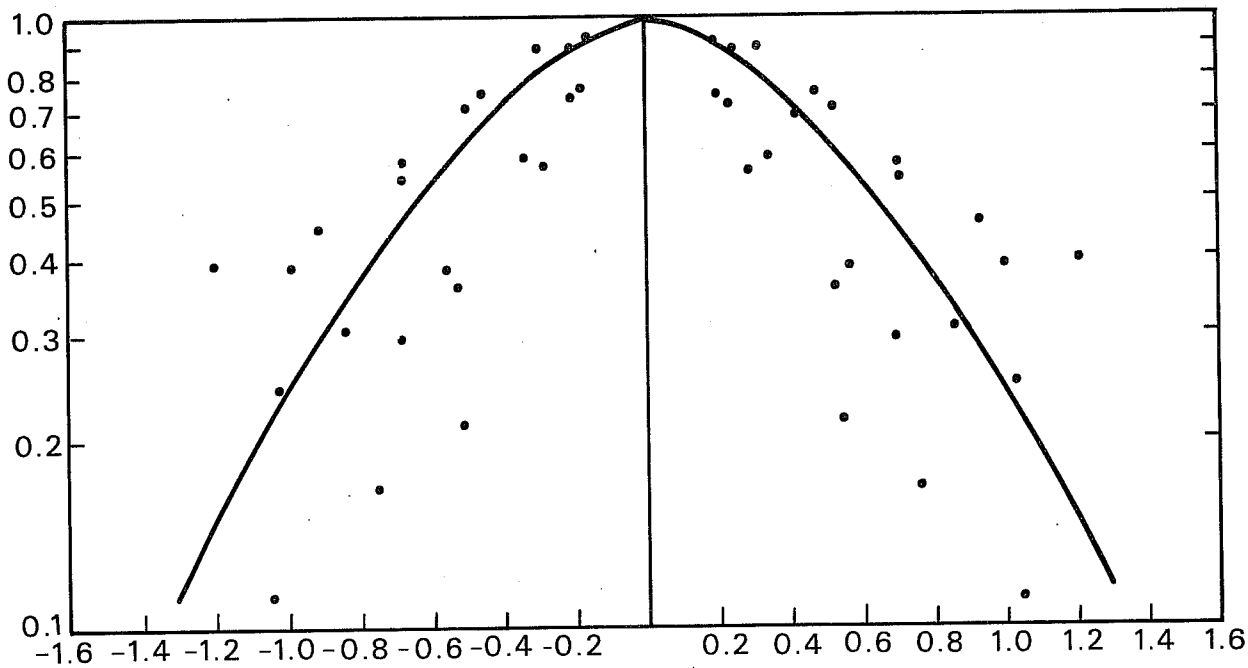
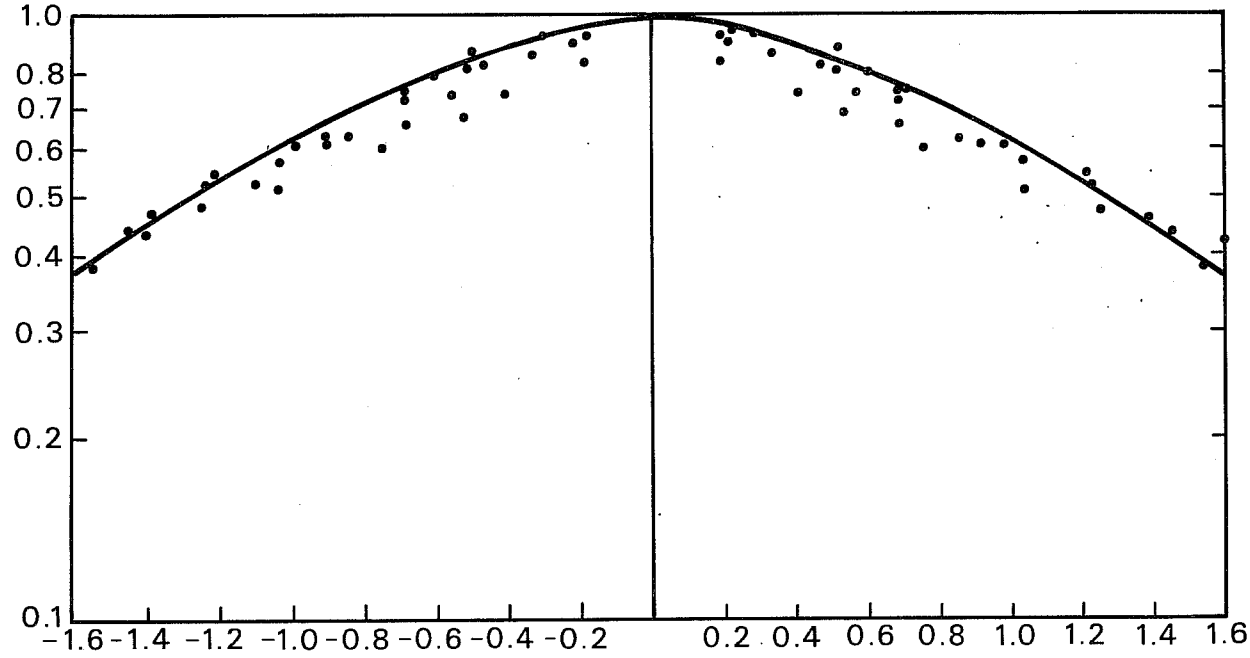


Fig. 11 a: Plot of the logarithm of the synoptic-scale height prediction error vertical correlation against the difference in the logarithms of the pressures of the levels involved in the correlation. The selection of the pairs of levels is described in the text, as is the curve defining the parameterisation of the correlations.

b)



11b: As (a) for the large-scale component of the height prediction errors.

Fig.9b plots the average correlation with 250 mb for the synoptic components, the correlation for the large scale mode, and the overall vertical correlation for the height field errors. The large scale component has a much broader vertical correlation than the synoptic scale component. This is quite different from the behaviour of the large scale component in the wind field, which showed a reversal of sign between the troposphere and lower stratosphere. Because of the magnitude of the large-scale component, the overall $\langle\phi,\phi\rangle$ vertical correlation is very nearly a mean of the large scale and synoptic scale components.

Fig.10 shows the overall vertical $\langle\phi,\phi\rangle$ correlation for a set of selected levels. The correlations are all positive and are generally much broader than the corresponding non-divergent wind correlations. The correlations tend to fall off fastest near the surface, indicating a degree of non-homogeneity in the vertical.

(a) Parameterisation of the vertical correlations

A summary of the mean vertical $\langle\phi,\phi\rangle$ correlations with tropospheric levels is given in Fig. 11(a,b). This plots the logarithm of the correlation with tropospheric levels against $\pm\left|\log(p_1/p_2)\right|$, for values of the correlation larger than 0.1. We excluded data from levels where the correlations were not monotonically decreasing with distance, as these only occurred for large separations. In the plots we have separated the large-scale, Fig. 11b, from the mean synoptic-scale, Fig. 11a, components.

Despite the fact that there is no complete vertical homogeneity for the covariances we see that there is approximate homogeneity for the correlations, provided the level separation is not too large. It is clear

that the large-scale component has a much broader correlation structure than the synoptic-scale component. There is very little scatter in the large-scale results, and rather more scatter in the synoptic-scale results.

The shapes of the curves suggest power law relationships of the form

$$C = \exp - \left(\frac{|\Delta \ell np|}{a_\phi}\right)^{b_\phi}$$

For theoretical reasons (existence of derivatives at zero separation) it would be desirable to seek a representation of this form with $b_\phi \geq 2.0$. We found that such a representation led to values of the correlations which were too large at short separations. Since this implied an underestimate in the representation of well-defined errors in the thickness field, we ignored the restriction on b_ϕ. A similar problem arose in the representation of the thermal wind errors as discussed in Part I. In order to have comparability of the representations of the thickness and thermal wind errors, we made the following empirical determination of the constants in the above representation of the vertical correlation of ϕ:

b_ϕ = 1.6 for both the large-scale and synoptic-scale correlations;

a_ϕ = 1.6 for the large-scale correlation;

a_ϕ = 0.8 for the synoptic-scale correlation.

Note that this representation is a homogeneous representation in the coordinate ln(p), which is closely related to geometric height. The conclusion is that the average correlations are sufficiently homogeneous in ln(p), provided the separation of levels is not too large.

In practical applications one works with differences of geopotential over finite layers and so the lack of derivatives at the origin is not a serious weakness of the representation. The practical virtue of these formulations is that they are positive definite.

3.5 Prediction errors of thickness

(a) Thickness errors in the forecasts

The vertical distribution of the thickness errors, converted to virtual temperature, has been seen already in Fig. 5. The results show errors of about 1K in mid-troposphere and the stratosphere, with a peak of 1.8K at 225 mb. The rms thickness prediction errors near the tropopause have been discussed earlier where it was noted that they are similar in magnitude to the smallest detectable forecast error.

(b) The implicit reference level

An important practical question in the use of thickness data concerns the definition of a reference level, when no reference level data is available. In the simplest case the question is: Given a single thickness observation in a column, which geopotential level in the forecast should be regarded as the most accurate for use as a reference level? The height-thickness covariance matrix provides an implicit definition of the reference level. In general ϕ is positively correlated with $\Delta\phi$ at lower levels, and negatively correlated with $\Delta\phi$ at higher levels, so that the $\langle\phi,\Delta\phi\rangle$ vertical matrix has an anti-symmetric structure. Given an observation of thickness between two adjacent standard levels, the implicit reference level is defined by the level at which the appropriate $\langle\phi,\Delta\phi\rangle$ correlation changes sign in the vertical. Fig. 12 is a plot of the reference level, thus defined. In the

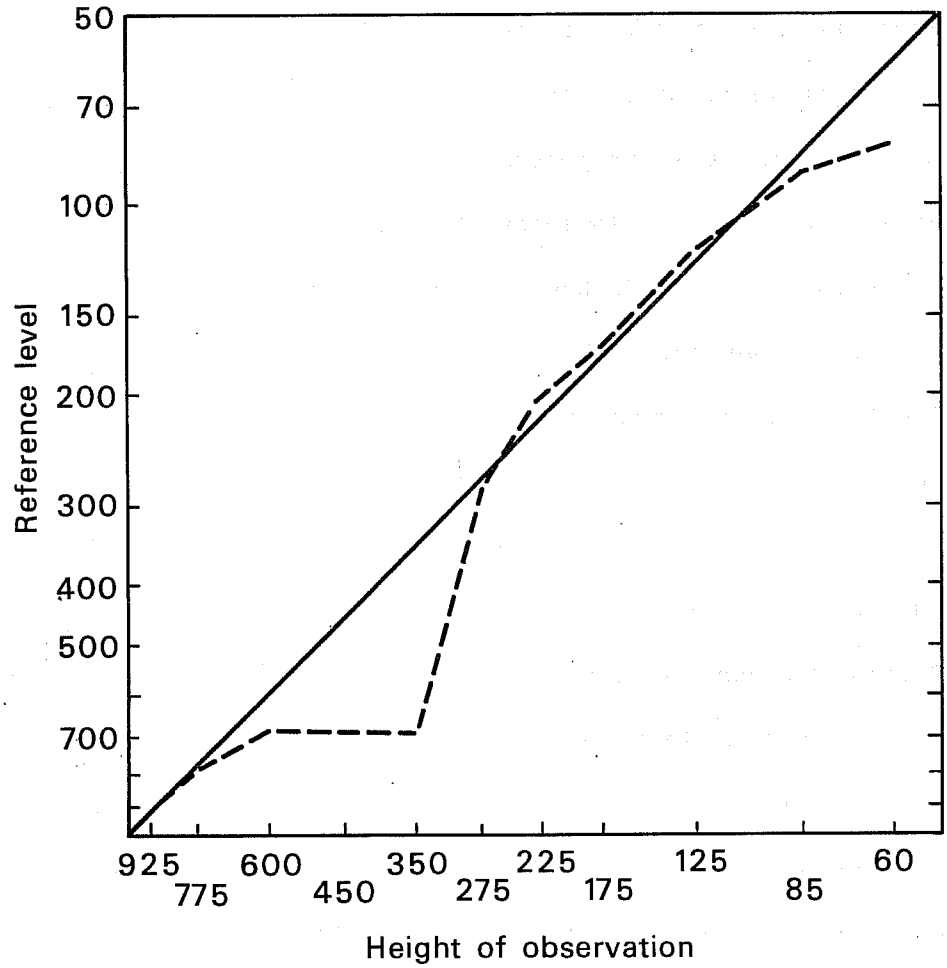


Fig. 12. The reference level implied by the height-thickness covariance matrix for prediction error: the abscissa is the pressure of an isolated thickness observation; the ordinate is the level at which the corresponding column in the matrix changes sign, thereby defining the level of the model to be used as a reference level in the absence of reference level data. The dashed curve is derived from the calculations. The solid curve with unit slope defines an idealised situation where the reference level and observation level for thickness coincide.

lower troposphere, and above 250 mb, the reference level is close to the observation level. For observations in the middle and upper troposphere the reference level is close to 700 mb.

Some insight into this result is offered by Fig. 13 which shows the $\langle \Delta\phi, \Delta\phi \rangle$ vertical correlation for a selection of levels. For most tropospheric layers the correlation with other tropospheric layers is positive, but it is negative with stratospheric layers. Moreover the stratospheric correlations decrease from 1.0 more rapidly than do the tropospheric correlations. These are the underlying reasons why the implicit reference level lies close to the observation level in the stratosphere, but lies in the lower troposphere for most tropospheric levels.

3.6 Discussion

In considering the results of the calculations, the most basic question is whether or not the partition between observational and prediction error is reasonable. We find that inter-layer correlations of the radiosonde observational errors of thickness are small, as would be expected. This result, together with the results on the magnitudes of the observational errors for height and thickness (which are also very reasonable), suggest that the separation between observational error and prediction error is soundly based.

The height prediction errors have a horizontal component which is essentially constant on the largest resolvable scale. This component also has a broad correlation structure in the vertical. Physically, this component corresponds to errors in the mean height over the data search volume. The large-scale component predominates in the stratosphere, but the synoptic-scale components

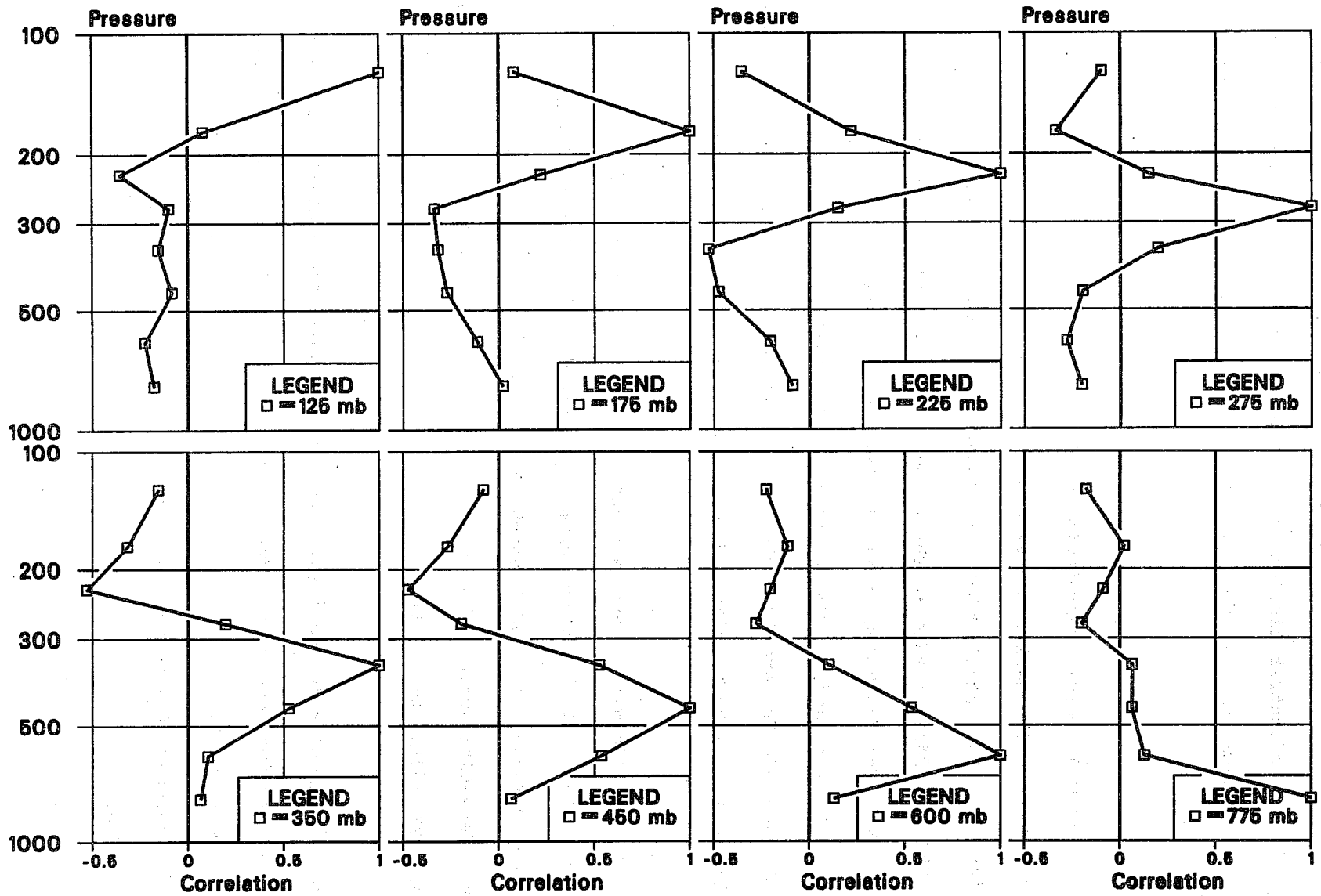


Fig. 13 With thicknesses defined between adjacent standard pressure levels, the plot shows the vertical correlation of thickness prediction error for a selected set of standard layers, as indicated in the legend. Thus 125mb corresponds to the 100-150mb thickness and so on down to the 700-850mb layer, labelled by 775mb.

predominate in the upper troposphere. The normalised spectrum of the synoptic-scale component shows a marked variation with height, with small scales making a contribution to the ϕ error variance in the lower troposphere, but not at higher levels. This result is not consonant with the assumption of separability.

A study of the vertical correlation structures of the horizontal modes shows some variation with horizontal wavenumber in the width of the vertical correlation. This result would not support an assumption of separability. These results further suggest that if one wishes to use the assumption of separability for the height analysis, one should analyse the troposphere and stratosphere separately. The behaviour of the average vertical correlations suggests that for separations of order a scale height or less, one may model the logarithm of the correlations by a power law in the layer separation.

The forecast thickness errors are largest at the tropopause, with magnitudes of order 1.75K. Elsewhere the errors are of order 1K. The variation of thickness forecast error with height is very similar to the variation of the minimum detectable thickness signal with height. The latter is determined by current WMO coding practices.

In the absence of reference level data, the Optimum Interpolation algorithm implicitly defines the model level to be used as a reference level. In the stratosphere the implicit reference level lies close to the observation level, but for tropospheric observations the reference level is in the lower troposphere.

4. THE MATHEMATICAL FORMULATION OF THE HEIGHT-WIND CORRELATIONS

4.1 Homogeneity

As discussed earlier in Part I, it is common in studies of this kind (Panchev 1971, Rutherford 1972, Julian and Thiebaux 1975) to work with correlation functions, rather than covariance functions, because the correlation functions tend to be more homogeneous than the covariances. All of the work presented above has used this approach. However in considering the correlations between heights and winds, one encounters the problem of non-homogeneity in a new form.

Our first calculations of the $<\phi,\psi>$ correlation, from the height and wind data, implied surprisingly low values (~0.65) for a typical geostrophic coupling coefficient in mid-troposphere. Detailed study showed that the low values arose in part because the wind error variances were definitely inhomogeneous, while the height error variances were quite homogeneous.

Fig. 14 shows a plot of the estimated 300 mb prediction error for ϕ and vector wind, as determined by all available radiosondes between the equator and North Pole, grouped in 10 degree latitude bands, for the three month period. In plotting these curves, reasonable estimates of the observation errors (5.25 m/s for the vector wind, 12.2 m for the height) have been removed from the total perceived forecast errors. The vector wind error varies strongly with latitude, while the height error is nearly independent of latitude. A similar result is found for the North American data. Fig. 14 also includes a plot of the vector wind error scaled by $\tilde{f} = f(\lambda)/f(\lambda_o)$ with $\lambda_o = 25$. Poleward of latitude 25, the scaled vector wind error is almost independent of latitude, and is therefore much more nearly homogeneous. The introduction of the scaling was motivated by the fact that one is seeking correlations between $f\,\underline{k}\wedge\underline{u}$ and $-\nabla\phi$.

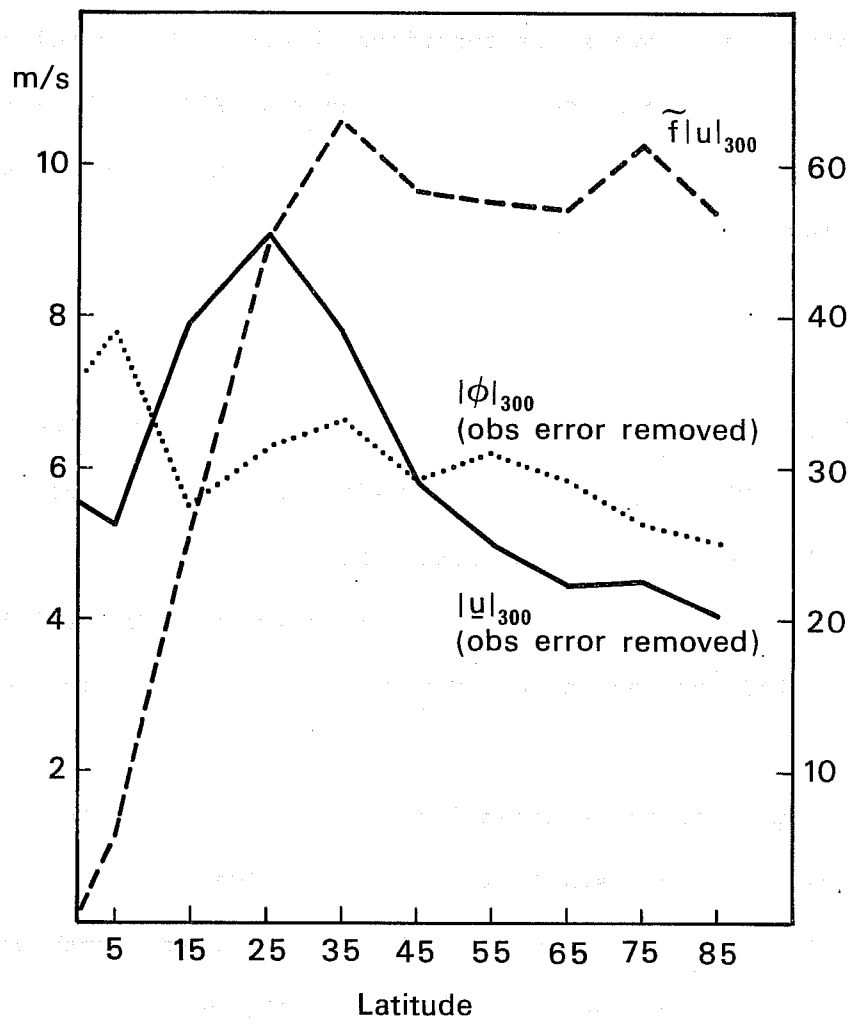


Fig. 14 The solid and dotted lines show the northern hemisphere meridional profiles of the estimated prediction error for 300mb vector wind (solid) and 300mb height (dotted). The estimates were obtained by verifying the assimilating forecasts against all accepted radiosonde data, and subtracting reasonable estimates of observational error. The dashed line shows the effect of multiplying the vector wind error by $\mathrm{Sin}(\theta)/\mathrm{Sin}(25)$. The units are m/s for wind (left) and m for height (right)

The introduction of a stream function Ψ with the property that

$$\tilde{f}u = -\frac{\partial \Psi}{\partial y} \text{ and } \tilde{f}v = \frac{\partial \Psi}{\partial x}$$

is likely to increase the height-stream function correlations relative to the usual definition of the stream function ψ as

$$u = -\frac{\partial \psi}{\partial y} \text{ and } v = \frac{\partial \psi}{\partial x}$$

It should be noted that the use of this scaling throughout Part I only changes the definition of the total perceived forecast error, which is a scaling parameter for all the other results. Since one is working with correlations, the earlier results for ψ, χ (e.g. normalised spectra, vertical correlations, length scales, etc.) apply equally well to Ψ,X.

4.2 Mathematical formulation

Given the assumption that the statistics are homogeneous, one may proceed to the determination of the cross-correlations of the height with the stream function and velocity potential.

Let $U=\tilde{f}u$, and $V=\tilde{f}v$, where $\tilde{f}=f(\lambda)/f(\lambda_o)$, and λ_o is a reference latitude chosen to be the mean latitude for the radiosondes. Define the correlation functions I, J as

$$I(r,\theta) = \gamma\langle\phi,\Psi\rangle \text{ and } J(r,\theta) = \delta\langle\phi,X\rangle$$

where Ψ, X are the stream function and velocity potential for U,V, while $\gamma = E_\psi/E_\ell$, $\delta = E_\chi/E_\ell$ and E_ψ, E_χ, E_ℓ are the rms amplitudes of ψ, χ, and the synoptic scale wind component. We use r and θ as great circle distance and angle from local east.

Then using the turbulence theory approach of Part I we have

$$\text{cov}\langle\phi,U\rangle = -\text{cov}\langle\phi,\Psi_y\rangle + \text{cov}\langle\phi,X_x\rangle$$

which in polar coordinates become, after division by E_ℓ,

$$\langle\phi,U\rangle = \sin\theta\left(I_r + \frac{1}{r} J_\theta\right) + \cos\theta\ \left(J_r - \frac{1}{r} I_\theta\right)$$

Similarly,

$$\text{cov}\langle\phi,V\rangle = \text{cov}\langle\phi,\Psi_x\rangle + \text{cov}\langle\phi,X_y\rangle$$

becomes

$$\langle\phi,V\rangle \quad = -\cos\theta\ \left(I_r + \frac{1}{r} J_\theta\right) + \sin\theta\left(J_r - \frac{1}{r} I_\theta\right)$$

The determinant of the right hand side is 1 and so one may solve the set of linear equations to obtain

$$\begin{aligned} I_r + \frac{1}{r} J_\theta &= +\sin\theta\langle\phi,U\rangle - \cos\theta\langle\phi,V\rangle = -\langle\phi,\hat{t}\rangle \\ J_r - \frac{1}{r} I_\theta &= -\cos\theta\langle\phi,U\rangle - \sin\theta\langle\phi,V\rangle = -\langle\phi,\hat{\ell}\rangle \end{aligned} \tag{4.1}$$

where $\hat{\ell}$, $\hat{t}$ are the longitudinal and transverse velocity components corresponding to U, V.

Expanding the right hand sides in Fourier Bessel series, and posing corresponding expansions for I and J, one can find least squares expressions for the expansion coefficients of I, J. The problem is particularly simple for the isotropic part of the problem, as the equations reduce to

$$\begin{aligned} \frac{d}{dr} I(r) &= -\langle\phi,\hat{t}\rangle \\ \frac{d}{dr} J(r) &= -\langle\phi,\hat{\ell}\rangle \end{aligned} \tag{4.2}$$

The physical interpretation of this result for the isotropic component is simple and appealing. The height-stream function correlation only involves the swirl or transverse component of the wind about a height observation

point, while the height-velocity potential correlation only involves the radial component.

If the basic correlations I,J are anisotropic, then the set (4.1) must be used. As in the more complex case of the winds, one solves term by term for the Fourier expansion of the azimuthal dependence. The cosine terms for I are obtained with the sine terms for J, and vice-versa.

In this paper we consider only the isotropic components of the height-wind correlation. To complete the discussion of this case we need to consider the definition of the expansions for I and J. The height covariance is defined as

$$\mathrm{cov}\langle\phi,\phi\rangle = E_\phi^2 \sum_0^N \phi_n^2 \; J_o\,(k_n \tfrac{r}{D}).$$

where E_ϕ is the rms height error,

$$\sum_0^N \phi_n^2 = 1.$$

and $$\phi_n^2 = \int_0^1 \tfrac{r}{D} \langle\phi,\phi\rangle \; J_o\,(k_n \tfrac{r}{D})\; d\,\tfrac{r}{D} / J_o^2(k_n)$$

Similarly the velocity covariances are defined so that

$$\mathrm{cov}\langle\hat{\ell},\hat{\ell}\rangle + \mathrm{cov}\langle\hat{t},\hat{t}\rangle = E_u^2 \sum_0^N k_n^2(\Psi_n^2 + \chi_n^2)\; J_o(k_n \tfrac{r}{D})$$

where E_u is the rms vector wind error and

$$\sum_0^N k_n^2(\Psi_n^2 + \chi_n^2) = 1.$$

With E_t, E_ℓ as the rms wind component errors ($E_t^2 = E_\ell^2 = \tfrac{1}{2}E_u^2$) it is natural to pose expressions for the mass-wind covariances in the form

$$\mathrm{cov}\langle\phi,\hat{t}\rangle = E_\phi E_t \langle\phi,\hat{t}\rangle,\quad \mathrm{cov}\langle\phi,\hat{\ell}\rangle = E_\phi E_\ell \langle\phi,\hat{\ell}\rangle \text{ and}$$

$$\frac{dI}{dr} = -\langle\phi,\hat{t}\rangle = -\sum_1^N I_n \frac{d}{dr} J_o(k_n \tfrac{r}{D}).$$

(4.3)

$$\frac{dJ}{dr} = -\langle\phi,\hat{\ell}\rangle = -\sum_1^N J_n \frac{d}{dr} J_o(k_n \tfrac{r}{D}).$$

One may then determine I_n, J_n by a least squares procedure.

Fig. 15 shows the variation with radial distance of the 850 mb isotropic component of I, together with the empirical data. The function has the expected form of a linear growth near the origin, a peak at some distance from the origin, and a gradual decay at large distance. Fig. 16 shows the corresponding data for the 850 mb isotropic component of J. In general the correlation associated with I is much larger than that associated with J.

4.3 Discussion

The problem of determining the height-wind correlations has been formulated in a simple fashion, using the turbulence theory approach in Part I. The magnitude of the wind errors in mid-latitudes varies almost inversely with the sine of the latitude. A scaling of the mid-latitude wind errors is introduced which takes account of the latitudinal variation of the Coriolis parameter. This rescaling leads to a simplification of the correlations of height and wind, as will be shown presently. The scaling suggests that it may be advantageous to analyse $\tilde{f}\underline{u}$ rather than $\underline{u}$ in mid-latitudes. Such a modification of the usual technique would have a simpler wind law, and might also provide the large-scale divergence field which is implicit in geostrophic motion on a sphere (Phillips 1963).

5. THE HEIGHT-STREAM FUNCTION CORRELATION

In this section we discuss the coupling of the height and stream function errors with a view to determining the degree of geostrophy of the forecast

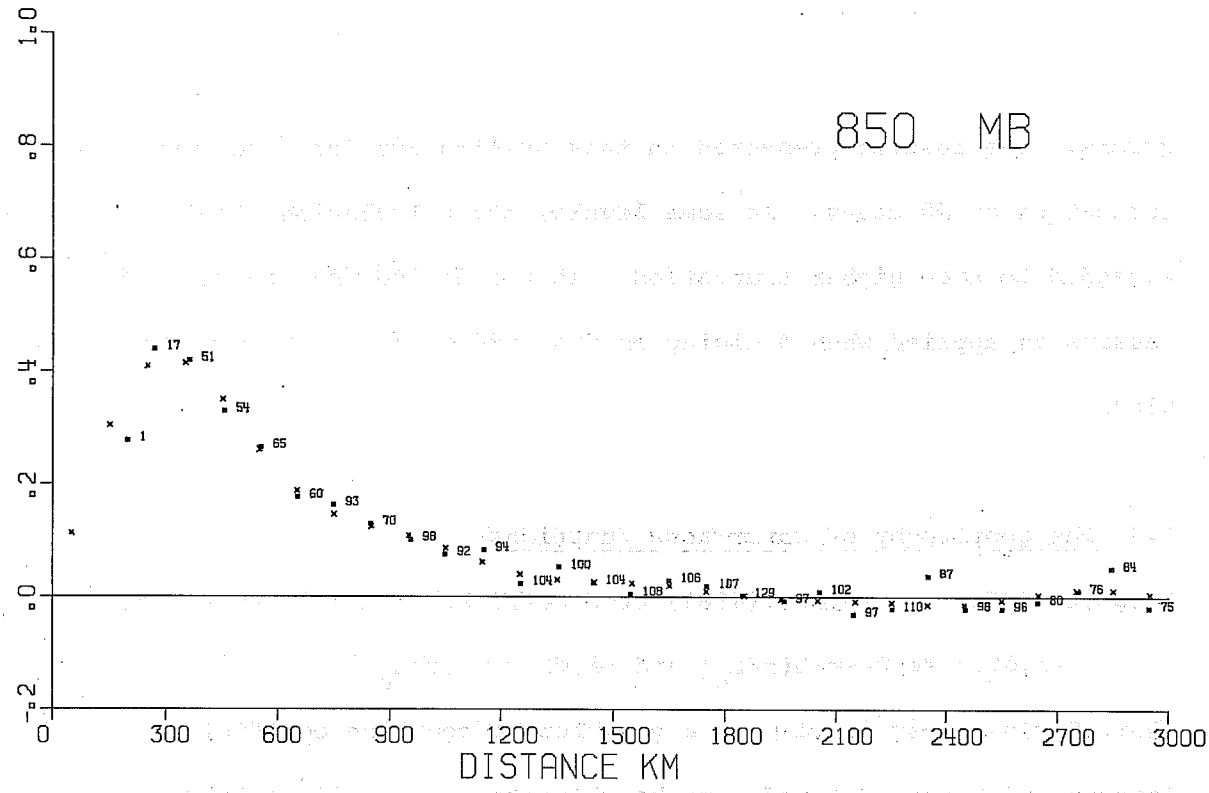


Fig. 15 The variation of the height/transverse-velocity correlation with station separation at 850mb. The squares show the empirically determined average values for each 100km 'bin', while the figures show the number of station pairs in that bin. The smooth curve (x) shows the result of the least squares calculation, using 10 terms.

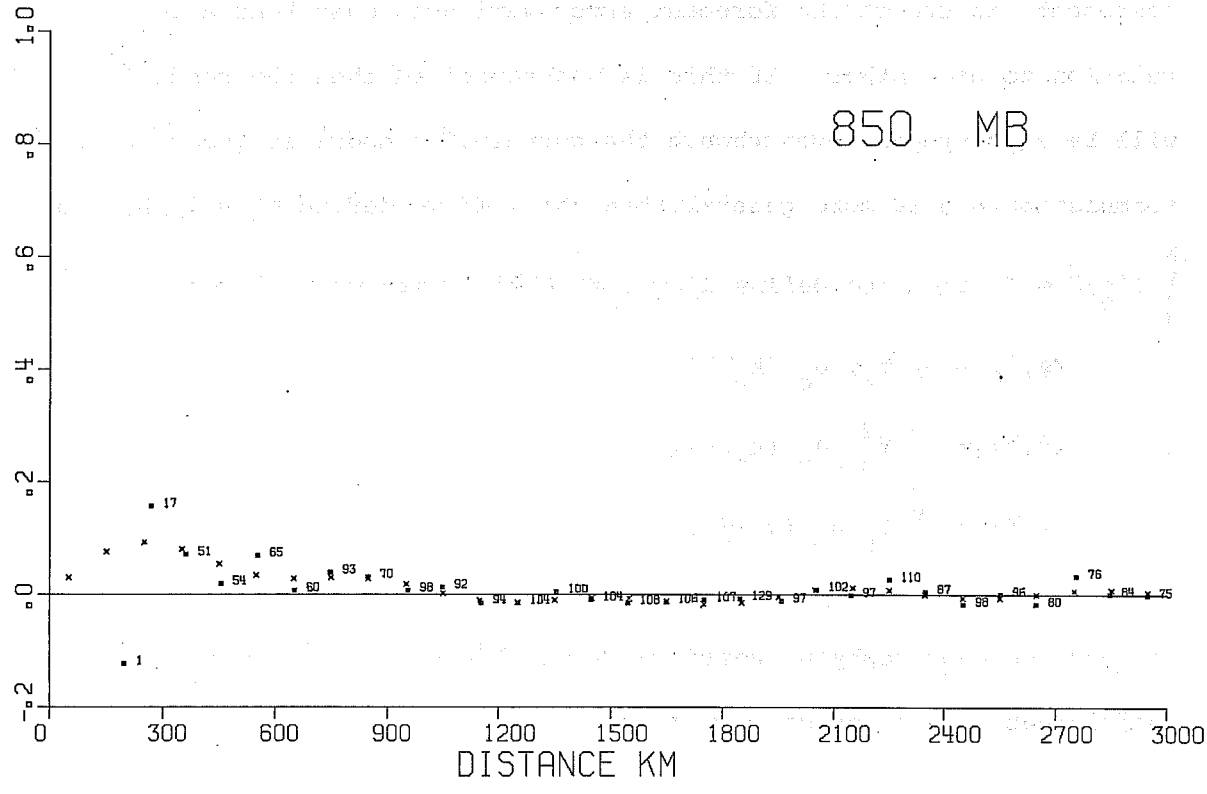


Fig. 16 As Fig. 15 for the height/radial-velocity correlation.

errors. The results presented in this section are for a maximum permitted truncation of 10 modes. At some levels, the calculations could have been extended to even higher truncation. It was decided not to do so for the same reasons as applied when deciding on the truncation of the expansions for the wind.

5.1 The geostrophy of covariance functions

Bergman (1979) and Lorenc (1981) used correlation models of the form

$$\langle\phi,\phi\rangle = \langle\psi,\psi\rangle = G(r/L_c) \text{ and } \langle\phi,\psi\rangle = \mu G(r/L_c)$$

where $G(r)=\exp(-\frac{1}{2}r^2)$ and μ is a geostrophic coupling coefficient. Both Bergman and Lorenc defined μ to be close to 1.0 in mid-latitudes. Lorenc pointed out that the covariances will be geostrophic (for $\mu=1$) provided

$$\frac{E_\phi}{f_o E_u} = \frac{L_c}{\sqrt{2}} \quad \text{or} \quad \frac{E_\phi}{f_o E_t} = L_c$$

where E_ϕ, E_u, E_t are the rms forecast errors for height, vector wind and wind component; in effect the forecast error variances must bear a geostrophic relation to each other. If this is not satisfied then the resulting analysis will be ageostrophic even though the correlation model is geostrophic. The formulation 4.3 is more general than this. If we define $\psi'_n = \psi_n/L_\psi$, so that $\sum_1^N (\psi'_n)^2 = 1$, and also define $I'_n = I_n/\gamma$ then we may write, dropping primes:

$$\langle\phi,\phi\rangle = \sum \phi_n^2 \; J_o\left(k_n \frac{r}{D}\right).$$

$$\langle\Psi,\Psi\rangle = \sum \Psi_n^2 \; J_o\left(k_n \frac{r}{D}\right).$$

$$\langle\phi,\Psi\rangle = \sum I_n \; J_o\left(k_n \frac{r}{D}\right).$$

For perfect geostrophy we would require all three sets of expansion coefficients to be the same, viz:

$$\phi_n = \Psi_n \text{ and } \phi_n \Psi_n = I_n .$$

We would also require a geostrophic relation between the height and velocity magnitudes:

$$\frac{E_\phi}{f_o E_t . L_c} = 1,$$

where L_c would be the common component length scale of the three correlations.

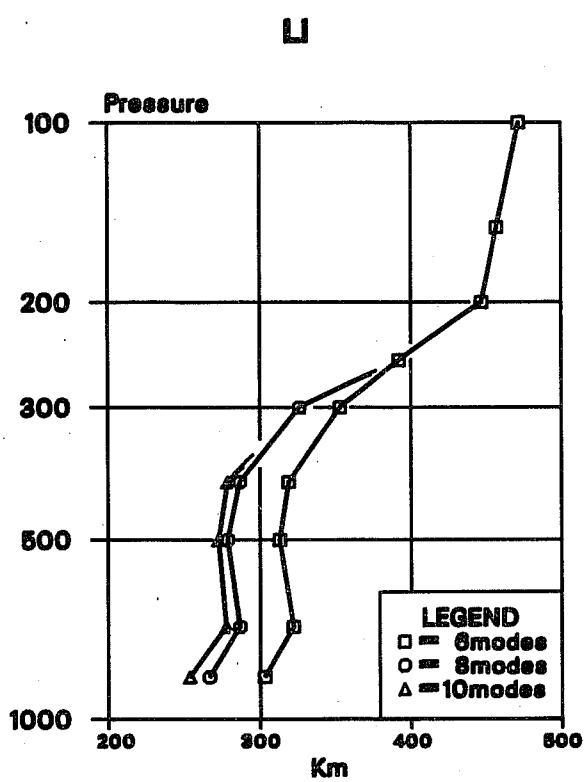
None of these requirements is strictly satisfied by our data.

To discuss the extent to which they are approximately satisfied we compare the spectra and length scales of all three functions, we examine the scale dependent geostrophic coupling coefficient μ_n defined by $I_n = \mu_n \phi_n \Psi_n$, and we examine the extent to which $\frac{E_\phi}{f_o E_t L_\phi^c}$ approximates 1, where L_ϕ^c is based on the height data.

5.2 Spectra and length scales

Fig. 17a presents the length scale of the $\langle \phi, \Psi \rangle$ function corresponding to truncations at 6, 8 and 10 modes. These results may be compared with the corresponding results for $\langle \phi, \phi \rangle$ (Fig. 6) and for $\langle \psi, \psi \rangle$ (Fig.13, Part I). To facilitate comparison, Fig. 17b plots all three curves corresponding to a truncation of 6 terms. The effect of the barotropic component in the height field is neglected, since it plays no role in the coupling of height and wind. The three curves in Fig. 17b exhibit the same major features - a modest decrease of scale from 700 mb to 500 mb, and a more rapid increase from 400 mb to 250 mb. The same general features are also found at higher truncations.

The most interesting aspect of these results, from a practical point of view, is that the length scales in the troposphere are substantially shorter than



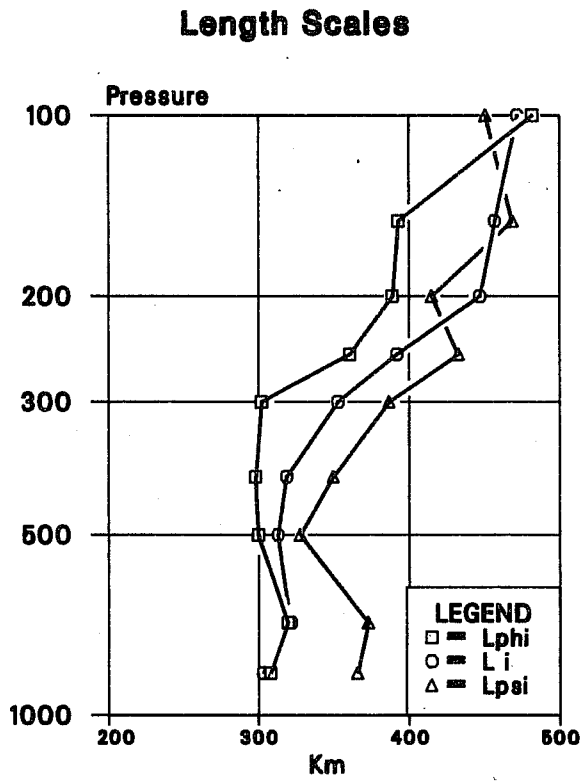


Fig. 17a: Vertical profile of the horizontal length scale of the height-streamfunction correlation for truncations of 6, 8, and 10 terms in the least squares procedure.

17b: A comparison of the length scales of the height-height (Lphi), height-streamfunction (Li), and streamfunction-streamfunction (Lpsi) length scales, when the least squares procedure uses 6 terms in the synoptic scale components.

the length scales used to produce the analyses (L_c=600 km). This implies that the resolution of the analysis system can be increased significantly, even if one uses only 6 terms in the correlation functions. The inclusion of even higher terms has negligible effect in the stratosphere, but it has significant effects in shortening the scales of the correlations in the troposphere.

The Ψ length scale is somewhat longer than the ϕ scale in the troposphere at all truncations. In the lower troposphere the $\langle\phi,\Psi\rangle$ correlation has a scale which is as short as the scale of the ϕ correlation. However in the upper troposphere the scale of the cross-correlation is intermediate between the scales of the other two correlations. At and above 100 mb the ϕ and ψ correlations have very similar length scales, with the scale of the cross-correlation being somewhat shorter.

The spectrum of the Laplacian of the $\langle\phi,\Psi\rangle$ correlation is shown, as a function of height, in Fig. 18, and may be compared with Fig. 8 and Fig. 14 of Part I for the Laplacians of the $\langle\phi,\phi\rangle$ and $\langle\psi,\psi\rangle$ correlations. These plots give the spectra of an equivalent wind field. We have seen already that the spectrum of the Laplacian of the height correlation is flatter than that of the stream function. Fig. 18 indicates that the spectrum of the Laplacian of the $\langle\phi,\Psi\rangle$ correlation is intermediate in steepness between the other two, particularly in the upper troposphere. This indicates, as we have seen, that the scale of the cross-correlation is intermediate between the scales of the $\langle\phi,\phi\rangle$ correlation and $\langle\psi,\psi\rangle$ correlation.

5.3 The geostrophic coupling coefficient

For each term in the expansion (4.3) one may define a geostrophic coupling

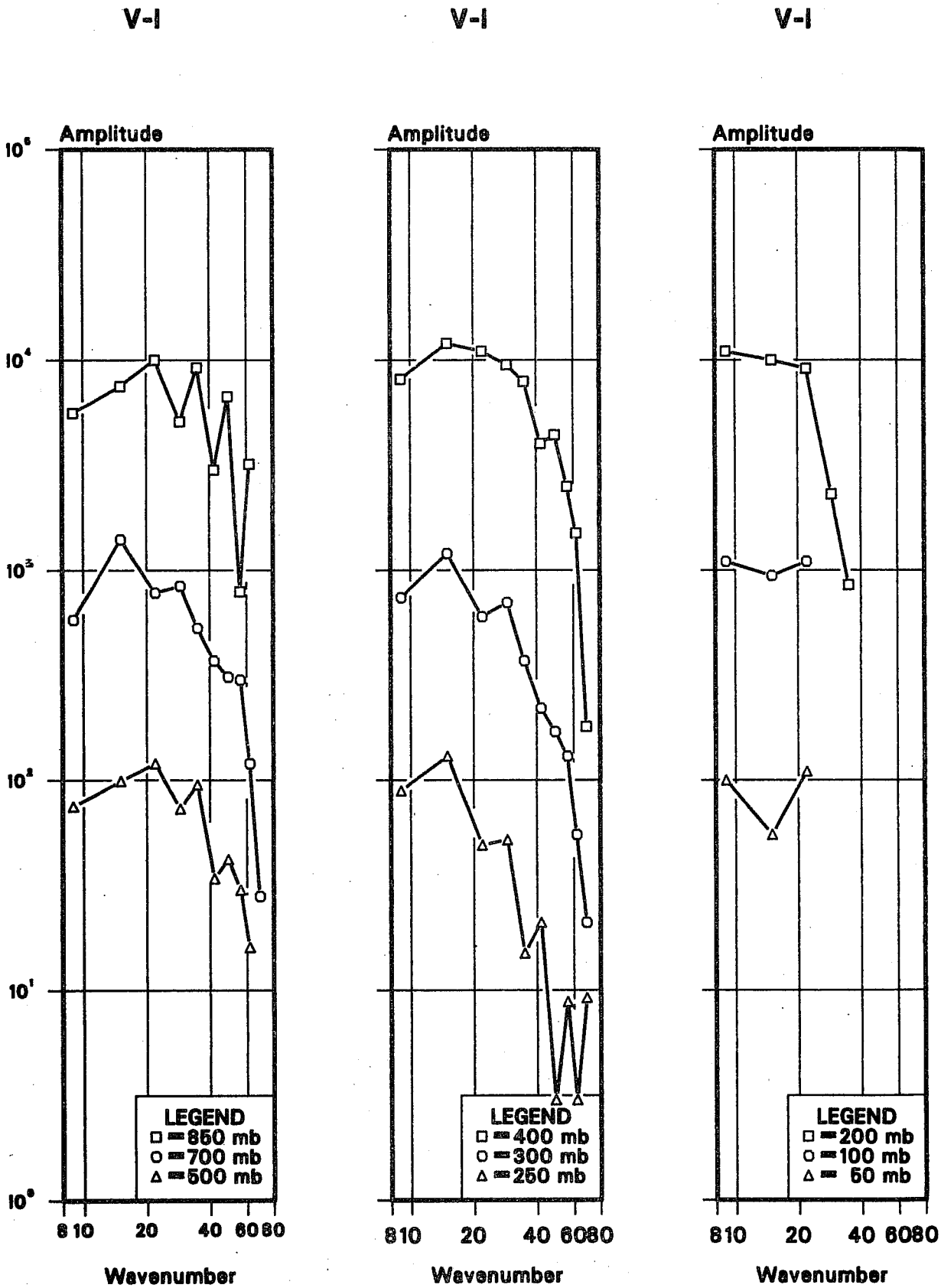


Fig. 18 As Fig. 8 for the Laplacian of the height-streamfunction correlation. The gives the normalised spectrum of the wind field implied by the height/streamfunction correlation, and is based on both height and wind data. It may be interpreted as a hybrid of the geostrophic wind spectrum in Fig 8 based purely on height data, and the non-divergent wind spectrum in Fig 14 of Part I based purely on wind data.

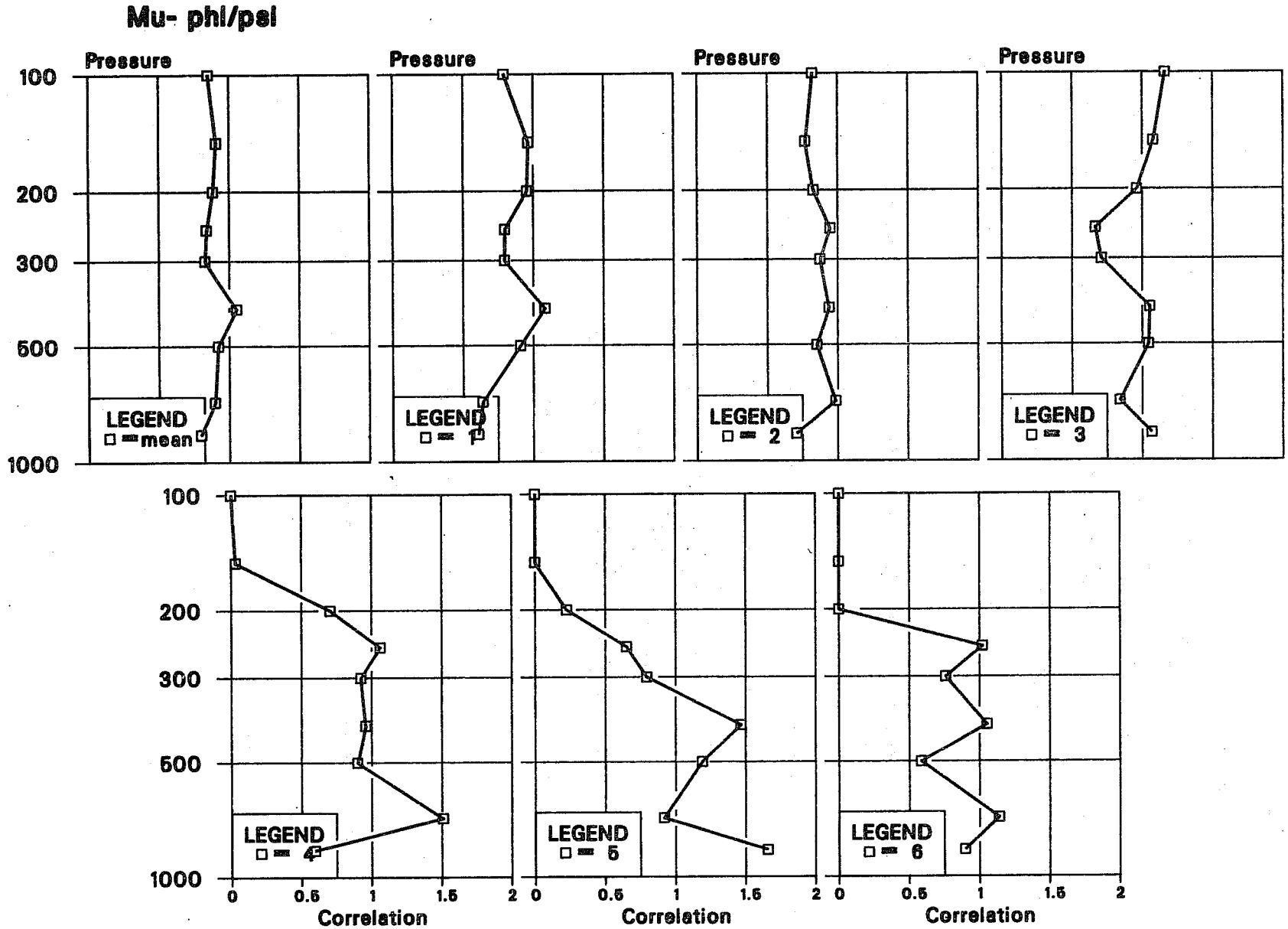


Fig. 19 The height-streamfunction coupling coefficient. The upper left frame shows the vertical profile for the overall value, while the remaining frames show the vertical profiles of the coupling coefficient for each of the first 6 terms in the expansion of the height-streamfunction correlation.

coefficient μ_n by

$$\mu_n = \frac{I_n}{\phi_n \psi_n}$$

Fig. 19 shows a plot of μ_n as a function of height and mode number, together with the average μ. The average value of μ increases from 0.8 at 850 mb to 1.05 at 400 mb. It then decreases to 0.83 at 300 mb, and varies between that value and 0.91 from 250 mb to 100 mb. The vertical average value is 0.89. For most modes μ reaches its maximum value in the mid to upper troposphere.

The values of μ_n become noiser with increasing n, probably because of noise in the height field calculations. Recall that the Laplacian of the stream function was determined directly from the wind data, and so the high wavenumber amplitudes of the ψ_n are more reliable than the corresponding amplitudes of the ϕ_n. However only 6 of the 63 values plotted in Fig.19 exceed 1.1. Since it is natural to interpret μ_n as the correlation between the directions of the geostrophic wind and the non-divergent wind, the results are rather reasonable, given the difficulties of the calculation.

It is hard to discern a clear pattern in the variation of μ_n with wavenumber, even if the values are averaged over several levels. If the values exceeding 1 are treated as being precisely 1, then there is some suggestion that modes 4 and 5 (planetary wave numbers 29 and 35) are the most geostrophic, with mode 1 (planetary wave number 9) being the least geostrophic. The low values of μ for this mode near the lower boundary are striking (.61 and .64 at 850 and 700 mb). If tangent plane geometry is used instead of spherical geometry to calculate angles, then these values are as low at 0.4.

5.4 Height and velocity variances

The second important facet of the geostrophy of the covariances concerns the relative magnitudes of the geostrophic and non-divergent wind errors.

The ratio, α, of the geostrophic wind to the stream function wind is given by

$$\alpha = \frac{E_\phi}{E_t \, f_o L_\phi^c}$$

The results of two calculations of α are shown in Fig. 20; one where the large-scale contributions to L_ϕ^c, E_ϕ and E_t are completely excluded and one where they are included. If the synoptic scale components only are included, then α lies between .70 and 1.15 for the levels between 700 and 200 mb. If the effects of the large-scale modes are included, then α lies between 0.94 and 1.35 for the same levels. The reason for presenting both calculations is to examine the possibility of aliasing, in that large scale components of the height field might be associated with synoptic scale components of the wind field, because of our truncation procedures (Julian and Thiebaux 1975).

From 700 to 300 mb α is 1 ± 0.15. To within this accuracy the synoptic scale non-divergent wind errors are geostrophic. If we believe that the non-divergent wind forecast errors above 200 mb are also geostrophic then Fig. 20 suggests that there may indeed be an aliasing problem arising from the fact that we have chosen to work with a radius of 3000 km. Even if all the large-scale height errors represent an alias from fields having associated geostrophic winds, there is still a 15% imbalance at 100 mb. One may speculate that part of this imbalance occurs because the analysis system used the same horizontal scale for its Gaussian structure function in both troposphere and stratosphere. There would then be a tendency to introduce too

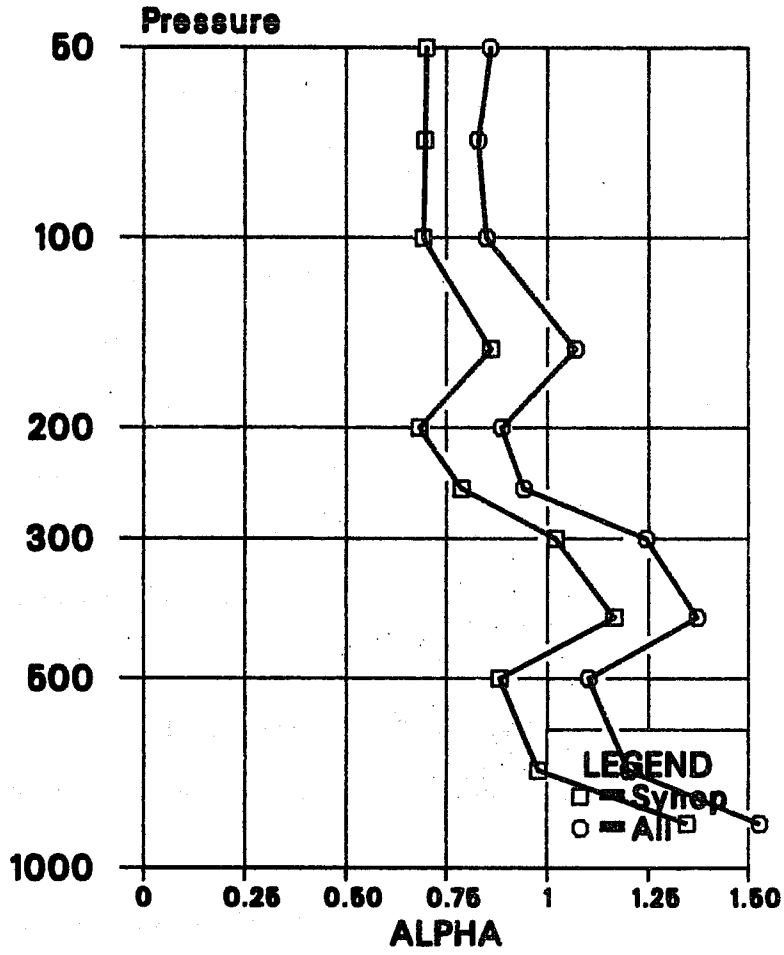


Fig. 20 Vertical profiles of the ratio of the geostrophic wind amplitude (derived purely from the height-height covariance) to the non-divergent wind amplitude (derived purely from the wind data). The cases shown are i) when only the synoptic-scale components are included in the calculation of the geostrophic wind amplitude (Synop) and ii) when both large-scale and synoptic-scale components are included in the calculation of height variance and length scale which together define the geostrophic wind amplitude (All).

much wind variance on shorter scales in the analysis of the stratosphere, which could lead to the sort of imbalances seen there.

5.5 The vertical correlation structure of $\langle\phi,\Psi\rangle$

One may calculate the vertical correlations of the $\langle\phi,\Psi\rangle$ function in the same way as the vertical correlations of the $\langle\phi,\phi\rangle$ and $\langle\psi,\psi\rangle$ functions. One might expect that the vertical correlation of the $\langle\phi,\Psi\rangle$ correlation should lie between the rather broad curve of the $\langle\phi,\phi\rangle$ vertical correlation, and the much sharper curve of the $\langle\psi,\psi\rangle$ vertical correlation. This proves not to be the case because the effective μ for thermal winds over thick layers is smaller than that for thin layers. This has the effect of broadening the vertical correlations of $\langle\phi,\Psi\rangle$. A detailed discussion would add little to our main results and so is omitted.

5.6 Discussion

The results on the geostrophic coupling of height and wind show that there is a high directional correlation between the stream function wind errors and the geostrophic wind errors in the troposphere. The magnitudes of the geostrophic winds and the non-divergent winds agree to within ±15% in the troposphere.

In the stratosphere the directional correlation of the geostrophic and non-divergent winds is also good, but there is an imbalance in the magnitudes of the two wind fields which may be partly explained as an aliasing problem. The use of inappropriate structure functions in the assimilation system may also contribute to the problem.

6. THE HEIGHT-VELOCITY POTENTIAL CORRELATION

The substantial divergent component of the forecast wind error (see Part 1) is mainly a large scale phenomenon, dominated by horizontal mode 1 (planetary wave 9). The same is true of the height-velocity potential correlation, and so we confine discussion of the $\langle\phi,\chi\rangle$ correlation to this mode. Fig. 21 shows the vertical profile of the height-velocity potential correlation ν_1 for horizontal mode 1. Below 250 mb ν_1 has peak values of order -0.25 while above 150 mb it is positive and of order 0.25. In discussing this correlation it is helpful to recall the sign convention that if $\langle\phi,\chi\rangle$ is negative, then $\langle\phi,\nabla^2\chi\rangle$ is positive, so that high geopotential tends to occur with divergent winds and vice-versa. We recall that the $\langle\psi,\chi\rangle$ correlation (Fig.20, Part I) for this mode is small and of uncertain sign.

The available evidence suggests that the divergent component of the forecast errors above the tropopause occurs because of the lack of a diurnal cycle in the forecast model. The magnitude of the error (~2m/s), its large horizontal scale, its vertical auto-correlation structure and the cross-correlation with ϕ (Fig. 21) are all consistent with the view that the stratospheric divergent wind error is a large scale forced motion.

There is less confidence in speculation about the origin of the divergent wind errors in the troposphere. Fig. 16 shows a clear and consistent signal of convergence in lows and divergence in highs in the 850 mb error fields. Fig. 21 shows that similar horizontal correlations are found at most tropospheric levels. Fig. 19 of Part I showed that the χ field auto-correlations show a phase reversal in the vertical within the troposphere. All of these results suggest that the forecast errors in the χ field in the troposphere are

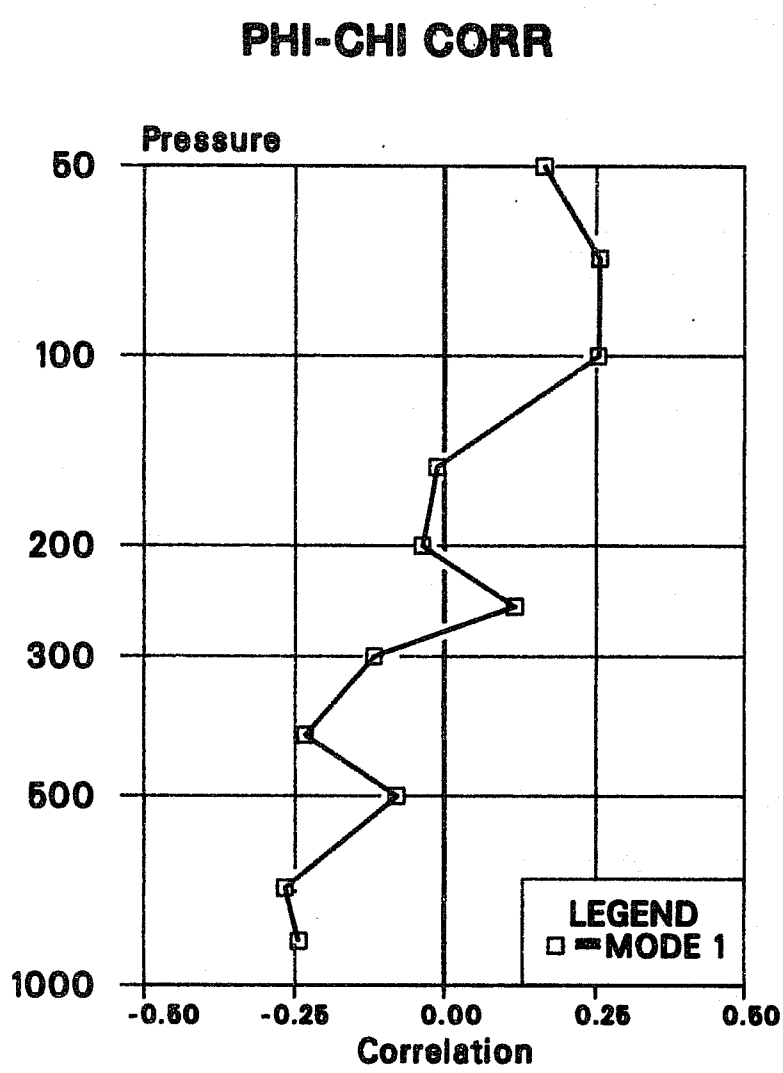


Fig. 21 The vertical profile of the height/velocity-potential correlation for the gravest (first synoptic-scale) term in the expansion. A negative sign implies that low geopotential is associated with convergence.

probably associated with the synoptic scale baroclinic disturbances. The use of $\tilde{f}\underline{u}$ rather than $\underline{u}$ as the analysis variable might well have an effect on the amplitude of this divergence field.

7. SUMMARY AND CONCLUSIONS

The methods developed in this paper and in Part I provide a comprehensive description of the covariance structure of the forecast and observation errors, for both height and wind. The determination of the structure of the height field errors follows earlier work. The decomposition of the wind field errors into divergent and non-divergent components, the determination of the auto-correlations of these components and of each with the height field, is new.

The main features of the results are that the forecast errors are comparable in magnitude with the observation errors, and that there are good grounds for increasing the resolution of our analysis system, both horizontally and vertically, in areas where data is available.

The tropospheric forecast errors in the height field are dominated by the baroclinically unstable wavelengths, as was found already for the wind field in Part I. The forecast errors in geopotential also contain an effectively barotropic component of large horizontal scale, which is uncorrelated with the wind field. The synoptic scale forecast errors may be resolved into a spectrum of horizontal modes, each of which has a somewhat different vertical correlation structure. Separability is therefore not as accurate an approximation for the height forecast errors as for the non-divergent wind. The horizontal scale of the height field errors is significantly shorter than that in operational use at the time the forecasts were produced.

The $\langle\phi,\Psi\rangle$ correlation showed a good agreement between the directions of the geostrophic and non-divergent wind errors. In the troposphere a representative value for the overall correlation of the ϕ and ψ fields is 0.89 with variations in the value depending on which mode or which level is examined. The gravest horizontal mode showed a markedly lower correlation than higher modes near the lower boundary.

Comparison of the magnitudes of the geostrophic and non-divergent wind errors showed agreement to within 15% at most levels in the troposphere. In the stratosphere the geostrophic wind errors were 15% less than the non-divergent wind errors even if allowance is made for aliasing of large scale height fields. A window of 3000 km is apparently inadequate to resolve all the scales of forecast error in the stratospheric height field which have associated geostrophic wind-fields.

The $\langle\phi,\chi\rangle$ correlation is weak ($\sim$-.25 in the troposphere), and shows the expected relationship between low pressure, cyclonic vorticity and convergence, in the troposphere. The correlation reverses sign above 150 mb, where it is suggestive of forced motion. The suggestion is that the absence of a diurnal cycle in the forecast model is readily detectable in the verifying observations.

The results in Part I and the present paper suggest many areas for improvement in our analysis system - better specification of observational error, better horizontal and vertical resolution, changes in the data selection area, modification of the analysed wind variable in mid-latitudes, scale dependent constraints on non-divergence and geostrophy, separation of the analysis of the troposphere and stratosphere etc. The methods also provide powerful

diagnostic tools to examine dynamical balance in the assimilation. Many other aspects of the new statistical formulation have only been adverted to in passing, such as the questions of anisotropy and non-separability. We have not dwelt at length on these matters as it is likely that they are sensitive to the parameters currently used in the analysis. The upgrading of an O/I system ought to be an iterative learning process, where at each stage one makes the most profitable improvements and then re-derives the statistics.

The methods presented here and in Part I can be extended in a variety of directions. Current studies are concerned with the forecast errors in other regions including in the tropics. The methods can also be adapted to provide spectral estimates of analysis accuracy. This problem is more difficult than the estimation of forecast accuracy because the analysis errors are correlated with the observation errors, unlike the forecast errors.

The indications that in data rich areas the accuracy of the 6-12 hour forecast is frequently comparable with that of the observations offers considerable potential for monitoring the performance of both the observational system and the analysis system (Hollingsworth et al. 1985). The fact that the coding practices for the exchange of data lead to rounding errors which are not negligible compared to the forecast errors must cause some concern; improvements in this area would be welcome. Systematic exploitation of these methods is likely to lead to all-round improvements in performance.

ACKNOWLEDGEMENTS

We are grateful to our colleagues at ECMWF for many discussions and helpful comments, particularly D.B.Shaw, G.J.Cats, D.M.Burridge and A.J.Simmons. R.Daley and N.A.Phillips provided much stimulation through early access to their work. H.Tennekes and A.C.Wiin-Nielsen provided valuable guidance to the literature. We are particularly grateful to Iris Rhodes for her skill and patience in iterating the manuscript.

REFERENCES

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

Daley, R., 1983: Spectral characteristics of the ECMWF objective analysis system. ECMWF Tech.Rep.No.40., 119 pp.

Fisher, R.A., 1921: On the probable error of a correlation deduced from a small sample. Metron., 1, Part 4, 3-32.

Gandin, L.S. 1963 Objective analysis of meteorological fields. Translated from Russian by the Israeli Program for Scientific Translations (1965).

Gustavsson N. 1981 'A review of methods for objective analysis' in Dynamic meteorology- Data assimilation methods ed L Bengtsson, M Ghil E Kallen Pub Springer-Verlag

Hildebrand F.B. 1962 Advanced calculus for applications. Prentice Hall pp646

Hollett, S.R. 1975 3 dimensional spatial correlations of PE forecast errors. M.Sc. thesis. Department of meteorology, McGill University, Montreal, Canada. 73 pp.

Hollingsworth A., D.B.Shaw, P.Lönnberg, L.Illari, K.Arpe, A.J.Simmons 1985: Monitoring of Observation quality by a data assimilation system. To appear

Hollingsworth, A., and P.Lönnberg, 1985: The statistical structure of short range forecast errors as determined from radiosonde data. Part I: The Wind Field. To appear.

Julian, P.R. and H.J. Thiebaux, 1975: On some properties of correlation functions used in optimum interpolation schemes. Mon.Wea.Rev., 103, 605-616.

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

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model, with application to normal mode initialisation. Contrib.Atmos.Phys., 50, 253-271.

Panchev, S., 1971: Random functions and turbulence. Pergamon Press, 444 pp.

Phillips, N.A., 1963: Geostrophic Motion. Rev.Geophys. 1, 123-176.

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

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

WMO 1983: Guide to meteorological instruments and methods of observation, 5th edn, WMO pubn. No.8, Geneva.