The statistical structure of forecast errors and its representation in
The Met. Office Global 3-D Variational Data Assimilation Scheme

By N. BRUCE INGLEBY*

Met Office, UK

(Received 9 November 1999; revised 19 May 2000)

SUMMARY

Previous studies and different methods of estimating short-range forecast errors are summarized. Zonally
and temporally averaged statistics based on differences of one- and two-day forecasts valid at the same time are
presented and an attempt is made to explain many of the features by reference to dynamical concepts.

Vertical correlation length-scale tends to increase with horizontal correlation scale but to be very short in
the Tropics; horizontal scale is longest in the Tropics and in the stratosphere. The variations in vertical correlation
are much more pronounced for largely balanced variables such as rotational wind and temperature than they are
for divergent wind or humidity. The extratropics are dominated by an equivalent barotropic mode with the level
of the maximum wind amplitude (and the zero crossing of the temperature correlation) being determined by the
tropopause. Surface pressure is negatively correlated with low-level temperature as expected (except over the
Antarctic plateau where it is positively correlated); it is also negatively correlated with temperatures near/above
the tropopause in the extratropics.

The covariance model used in The Met. Office Global Three-Dimensional Variational (3D-Var) Data
Assimilation system represents the variation of vertical covariances with latitude reasonably well, but the longer
horizontal scales in the stratosphere are not currently reproduced. The implied covariances used operationally
have been modified so that the correlation length-scales, both horizontal and vertical, are somewhat shorter than
those direct from the forecast differences. Recent changes to the representation are briefly described, with an
indication of their impact on the forecasts. The impacts are significant relative to other changes tested, and the
covariance model has played a major role in the successful implementation and subsequent improvement of our
3D-Var system.

KEYWORDS: Error covariances Forecast errors Variational data assimilation

1. INTRODUCTION

(a) Overview

Atmospheric data assimilation systems combine information from observations and
a short-range (typically six-hour) forecast referred to as the ‘background’. Estimates are
needed of the observation- and background-error covariances. The background-error
covariance estimates are important, particularly in data-sparse regions. They determine
the spreading of information between the observation locations, including extrapolation
into relative data voids, and also multivariate balance relationships. The covariances
are often decomposed into correlations, assumed to be constant globally or over large
regions, and standard deviations which are allowed more geographical variation. Daley
(1991) provides an overview of this field.

Estimating the forecast-error covariances is not straightforward as we never have the
truth available, only different approximations to it. The remainder of section 1 summa-
rizes the different ways of estimating background-error covariances, the results of pre-
vious studies and physical relationships between horizontal and vertical length-scales.
Section 2 shows the covariances estimated directly from forecast differences and con-
centrates particularly on the vertical structure and its variation with latitude. Section 3
describes the representation of the background errors within our Three-Dimensional
Variational Data Assimilation (3D-Var) system, its strengths and weaknesses. Selected
results of recent analysis/forecast experiments are presented in section 4 and there is a
brief conclusion in section 5.

* Corresponding address: Met Office, London Road, Bracknell, Berkshire RG12 2SZ, UK. e-mail: nbingleby@meto.gov.uk
(b) Forecast errors—determining factors and evolution

Forecast errors depend on the observations available, the data assimilation system and the forecast model; they also vary with the synoptic and seasonal situation. In ‘sensitive’ areas the forecast errors may grow rapidly whereas in other areas the errors may grow slowly or even decay initially. The analysis/forecast systems of operational forecasting centres are sufficiently good (or sufficiently similar) that their forecast-error statistics have much in common.

Boer (1994) differentiates three regimes of forecast skill. The large-scale regime (global wave number $n < 10$) is dominated by stationary (largely zonal) structures that are relatively uncontaminated by error in forecasts out to ten days. The synoptic scales ($10 < n < 80$) exhibit classical predictability behaviour in which error, initially concentrated at smaller scales, penetrates up the spectrum and saturates at values roughly twice the observed variance. Thus there is some similarity between the covariances of short-range forecast error and of observed fields, but the latter have rather longer scales. Somewhat surprisingly, high wave numbers ($n > 100$) exhibit some forecast skill out to ten days—owing to local topographic forcing.

In practice we are largely concerned with errors in the synoptic scales, so that forecast-error length-scales tend to increase as the length of the forecast increases. Initially this will be most noticeable in data-dense areas which have shorter analysis-error length-scales than data-sparse areas. Thus, using differences between longer period forecasts to estimate background errors, we may overestimate the length-scales, particularly for data-dense areas.

(c) Previous studies

Until recently the main way of estimating background error was by comparison with observations. The observation errors were assumed to be uncorrelated so that any spatial correlation was due to the forecast error. In this way Hollingsworth and Lönnberg (1986) and Lönnberg and Hollingsworth (1986) investigated the structure of wind and height errors using the North American radiosonde network. They found that the background errors were comparable in magnitude to the observation errors. Forecast errors were dominated by synoptic-scale rotational-wind and height perturbations in approximate geostrophic balance. There were some problems with the separability assumption (treating correlations as the product of horizontal and vertical correlation functions); in particular horizontal length-scales were significantly longer in the stratosphere. Work elsewhere largely confirmed these features. For example, Bartello and Mitchell (1992) concentrated on the increase of horizontal scale with height.

Ghil et al. (1979) and Mitchell et al. (1990), comparing forecasts with satellite temperatures and radiosonde heights, respectively, found some evidence of shorter horizontal length-scales at higher latitudes. Dee and Gaspari (1996), using forecast differences, found longer horizontal scales for height in the Tropics. The variation of horizontal scale with latitude was examined using stochastic–dynamic models by Balgovind et al. (1983) and Jiang and Ghil (1993). Both papers conclude that horizontal scales vary in a similar way to the Rossby radius of deformation—but they take the scale-height to be independent of latitude. In data-dense areas, error variances will be smaller, length-scales generally shorter and the covariances will be more isotropic and less dependent on the dynamics (Bouttier 1994).

The Met. Office’s previous data assimilation system, the Analysis Correction (AC) scheme (Lorenc et al. 1991), had longer horizontal scales in the Tropics and southern hemisphere. Based on unpublished studies both the AC and the European Centre
for Medium-Range Weather Forecasts (ECMWF) Optimal Interpolation analysis systems used narrower vertical correlations in the Tropics. Ingleby and Bromley (1990) compared wind-error covariances for two different areas and two different horizontal resolutions of The Met. Office forecast system and found that the upstream data density appeared to make more difference than the model horizontal resolution.

Parrish and Derber (1992) introduced the use of forecast differences as a proxy for forecast errors (used here in sections 2 and 3). Rabier et al. (1998) noted sharper vertical correlations at smaller horizontal scales using this so-called ‘NMC method’.

The three main ways of estimating forecast errors each have their own strengths and weaknesses. Comparison with observations is the only independent calibration available, but the observation errors have to be taken into account and, more importantly, there are severe limitations on the information available for data-sparse areas. Stochastic–dynamic or Kalman-filter methods are powerful tools, but necessarily include an approximate model error term and may not treat nonlinear error saturation correctly. The use of forecast differences is somewhat heuristic. In practice it provides reasonable estimates of forecast error in a form suitable for use in variational analysis systems. However, the variances need rescaling (by coincidence the factor is often near 1*), and in data-sparse areas the variances may be underestimated. In data-dense areas the length-scales will tend to be overestimated as mentioned above.

(d) Horizontal and vertical length-scales

In the studies mentioned above and in our forecast difference statistics (section 2) it was found that vertical correlations of rotational wind and temperature (largely ‘balanced’ variables) vary strongly with both horizontal scale and latitude, but most other variables show much less variation of the vertical correlation. The physical explanation for this, outlined below, is largely taken from Lindzen and Fox-Rabinovitz (1989). Note that for forecast errors it will be modified by data density as already discussed.

For quasi-geostrophic (‘balanced’) flow on a beta-plane, the horizontal scale $\Delta L$ (Rossby radius of deformation) is related to the vertical scale $\Delta z$ by

$$\Delta L = (N/f_0) \Delta z,$$

(1)

where $f_0$ is the characteristic Coriolis parameter and $N$ is the Brunt–Väisälä frequency. This suggests that $\Delta z$ will increase both as $\Delta L$ increases and as higher latitudes are approached, and is consistent with larger $\Delta L$ in the more stable stratosphere. Near the equator

$$ (\Delta L)^2 = (N/2\Omega) \Delta za,$$

(2)

where $\Omega$ is the earth’s rotation rate and $a$ its radius. Consistent with these equations we observe both small $\Delta z$ and large $\Delta L$ in the Tropics. For gravity waves there are less clear theoretical relations between vertical scale and either horizontal scale or latitude, but divergent wind and ageostrophic pressure may be weakly coupled with the balanced flow and show some of its characteristics.

The analysis above is relevant for free dynamics and may be modified by any forcing applied. It also neglects $\beta$ and mean-flow effects; these are summarized in Eq. (1) of Charney (1969). In particular the transmission of large amounts of energy from the troposphere into the upper atmosphere is prevented through most of the year (except near the equinoxes) by easterly or large westerly winds above the tropopause, and the

* We currently take it to be 1.0; Rabier et al. (1998) scale their standard deviations by 0.9.
transmissivity of the stratosphere increases with wavelength (Charney and Drazin 1961). This appears to be the main reason for longer horizontal length-scales in the stratosphere and the continued increase in length-scale above 100 hPa (the larger value of $N$ in the stratosphere will also play a role but it is almost constant above 100 hPa). In an analogous way, disturbances propagating from mid-latitudes into the Tropics will tend to be confined to the high troposphere and lower stratosphere where the zonal winds are weak easterly or westerly (Charney 1969; Tomas and Webster 1994).

2. **Forecast-difference covariances**

   (a) *Horizontal covariances*

   The statistics shown here are mainly taken from differences of T+24 and T+48 operational forecasts (one- and two-day forecasts valid at the same time) for 29 days each in July 1998 and January 1999; those showing variation with latitude are for January 1999, except Fig. 1. The mean values have not been removed from the statistics, so strictly speaking we are calculating cross-products rather than covariances (for rationale see section 3(a) of Lorenc et al. 2000); in practice this makes the horizontal scales slightly longer, particularly for temperature and pressure. As already stated, such forecast differences are an imperfect proxy for forecast errors, brief validation of some aspects against observation-minus-background statistics is mentioned later, but a fuller comparison is outside the scope of the current work.
TABLE 1. **DIFFERENTIAL LENGTH-SCALES (km) FOR FORECAST DIFFERENCES, SELECTED MODEL LEVELS**

<table>
<thead>
<tr>
<th>Level</th>
<th>P</th>
<th>ψ</th>
<th>RKE</th>
<th>χ</th>
<th>DKE</th>
<th>Ap</th>
<th>RH</th>
<th>p</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>30</td>
<td>1353</td>
<td>366</td>
<td>1879</td>
<td>418</td>
<td>977</td>
<td>893</td>
<td>756</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>102</td>
<td>1057</td>
<td>277</td>
<td>1343</td>
<td>282</td>
<td>611</td>
<td>234</td>
<td>683</td>
<td>409</td>
</tr>
<tr>
<td>19</td>
<td>251</td>
<td>670</td>
<td>273</td>
<td>983</td>
<td>222</td>
<td>490</td>
<td>207</td>
<td>541</td>
<td>297</td>
</tr>
<tr>
<td>11</td>
<td>519</td>
<td>599</td>
<td>229</td>
<td>827</td>
<td>179</td>
<td>495</td>
<td>174</td>
<td>500</td>
<td>282</td>
</tr>
<tr>
<td>4</td>
<td>867</td>
<td>628</td>
<td>217</td>
<td>1045</td>
<td>195</td>
<td>536</td>
<td>171</td>
<td>499</td>
<td>251</td>
</tr>
<tr>
<td>2</td>
<td>961</td>
<td>581</td>
<td>199</td>
<td>724</td>
<td>181</td>
<td>433</td>
<td>164</td>
<td>466</td>
<td>211</td>
</tr>
</tbody>
</table>

Data for July 1998 and January 1999 combined.
The second column gives the average pressures in hPa for the levels. Due to the vertical grid staggering T and RH are at slightly lower average pressures than those indicated. See text for definition of symbols.

Until January 1998 the forecast model (described by Milton and Wilson (1996)) was run at 19 levels on a 1.25° longitude × 0.833° latitude grid; it was then upgraded to 30 levels, 0.833° longitude × 0.556° latitude grid spacing. During the development of our 3D-Var system, forecast-difference covariances have been calculated for several different periods and for both resolutions. The statistics have been fairly consistent; some exceptions are noted later. As explained in the appendix of Lorenc *et al.* (2000), we convert fields to the 'New Dynamics' grid. This involves changing from a pressure-based to a height-based coordinate, and also vertically interpolating temperature and humidity fields onto model half-levels. For storage reasons all fields have been interpolated onto the previous 1.25° × 0.833° horizontal grid—for the spectral statistics they were transformed to spectral T143 resolution. The control variables used in our 3D-Var system are: stream function, ψ; velocity potential, χ; unbalanced pressure, denoted Ap (essentially ageostrophic pressure, see appendix of Lorenc *et al.* (2000) for details); and relative humidity, RH. Other variables used below are total pressure, p; temperature, T; and horizontal wind components, (υ, υ). All the figures are presented on model levels.

Let the spectrum of the stream function be given by $D_n(ψ)$, where $n$ is the global wave number; $D_n(ψ)$ is essentially the variance at that wave number (normalized by the total variance in the case of a correlation spectrum). The rotational kinetic energy at that wave number is given by

$$RKE_n = 0.5D_n(ψ)n(n+1)/a^2,$$

and the vorticity ($ζ$) spectrum is

$$D_n(ζ) = D_n(ψ)n^2(n+1)^2/a^4$$

(Boer 1983). The differential length-scale for $ψ$ is given by

$$\left[2a^2 \sum_n D_n(ψ)/\left(\sum_n \{D_n(ψ)n(n+1)\}\right)\right]^{0.5}$$

and so is directly proportional to the square root of the ratio between stream-function variance and RKE. Similar relationships apply to velocity potential, divergent kinetic energy (DKE) and divergence. The forecast differences have significant variance of $ψ$ and $χ$ at large wavelengths (small $n$), but the associated wind is rather small. For this reason the RKE and DKE length-scales following are felt to be more meaningful than those for $ψ$ and $χ$.

The differential length-scales in Table 1 all show the well-known increase with height in the stratosphere (see sections 1(c) and (d)). There is also a secondary maximum
Figure 2. Vertical correlation with level 11 (approximately 500 hPa) as a function of horizontal wave number, July 1998 and January 1999 combined. (a) Rotational wind, (b) divergent wind and (c) temperature. Contours as Fig. 1.
at approximately 800 hPa, particularly for velocity potential, the reasons for this being unclear. One might expect slightly shorter scales very close to the surface; this is seen to a limited extent. In moving from 19 to 30 levels, horizontal scales of most variables were reduced by about 20% in mid-troposphere and by 35% or more at 100 hPa—the scales in the stratosphere were presumably constrained by the vertical resolution. East-west correlations for July and latitudes between 60°N and 60°S are shown in Fig. 1. The most notable feature is that the correlations are broadest in the Tropics, particularly for pressure and temperature. The north–south correlations (not shown) are also longer in the Tropics, but not by as much. The temperature and pressure differences in the Tropics appear to have large-scale biases, or perhaps errors in the model’s treatment of tides, because the correlations remain positive and fairly flat out to 3000 km. For most of the variables there is some change of correlation shape with latitude—the curves are not described adequately by a single length-scale.

In July the correlations are longer scale in the southern extratropics than in the northern extratropics. In January (not shown) the extratropical correlations are similar in both hemispheres and intermediate between the northern and southern July correlations. This corresponds to what one might expect: longer scales, on average, in the more data-sparse and oceanic southern mid-latitudes and in winter when the large-scale forcing is more dominant. In the Tropics the very longest scales tend to be just within the winter hemisphere. Stream function (not shown) has similar correlations to ρ, but not as broad in the Tropics. Velocity potential (not shown) has rather different characteristics to the other variables examined, with largest scales at the poles. There is some indication that, for most variables in the extratropics, correlation scales decrease slightly towards the poles, consistent with Eq. (1). However, any latitudinal and seasonal variations in the extratropics are much less than the differences from the Tropics.

Figure 2 shows the vertical correlations with level 11 (approximately 500 hPa) as a function of n, the global wave number. Note that Fig. 2(a) is valid for stream function, RKE and vorticity—for each n they have the same correlation matrix, but different variances (similarly for velocity potential, DKE and divergence in Fig. 2(b)). Apart from global scales (n < 10, where the term involving β and mean flow becomes important), stream function and temperature vertical scales decrease with horizontal scale, in qualitative agreement with Eq. (1). Stream-function vertical correlations are very broad at large scales (except n < 5) and are essentially non-negative at all scales. Pressure vertical correlations (not shown) are qualitatively similar, but show signs of
Figure 4. Standard deviations of forecast differences for January 1999: (a) $u$, contours every 1 m s$^{-1}$ with extra dashed contours at 0.5 and 1.5 m s$^{-1}$; (b) as (a) but for $u_{div}$; (c) temperature, contours every 0.4 degC with extra dashed contour at 0.6 degC; (d) relative humidity in %, contours every 5%, extra dashed contours at 1 and 3%; (e) pressure in hPa, contours every 0.5 hPa, extra dashed contours at 0.1 and 0.3 hPa; and (f) ratio of variances (unbalanced pressure/total pressure), dashed contours above 0.5.
Figure 4. Continued.
broadening again, particularly into the stratosphere, for \( n > 80 \). Temperature vertical correlations (Fig. 2(c)) are narrower and show significant negative correlations with higher levels for \( 10 < n < 50 \).

Apart from \( n < 8 \), velocity-potential vertical correlations (Fig. 2(b)) are very narrow, and narrow more with increasing \( n \), with slight negative lobes. Relative-humidity vertical correlations (not shown) also show gradual narrowing with increasing \( n \), at least for \( n > 15 \).

\((b)\) Vertical covariances

Daley (1991, section 4.7) notes that “the vertical aspects of objective analysis have never been very satisfactory”. In our 3D-Var system the vertical aspects are rather more sophisticated than the previous AC scheme, but are lacking a theoretical basis. The presentation here is a first step towards rectifying that.

Figure 3 sets the scene showing \( N \) from the T+24 forecast zonal-mean temperatures; \( N \) is slightly greater than 0.01 s\(^{-1}\) throughout much of the troposphere with minima (below 0.01 s\(^{-1}\)) in the Tropics between approximately 450 and 200 hPa and also near the surface. At high latitudes the stability is higher near the surface, particularly near the winter pole with its low-level inversion. Values in the stratosphere are typically about 0.02 s\(^{-1}\) with the tropopause height clearly varying with latitude. Apart from changes of tropopause height and polar inversions, there is little seasonal variation of \( N \), but stability tends to be slightly lower over continents in summer.

The standard deviation (SD) of the westerly wind component \( u \) (Fig. 4(a)) has maxima in mid-latitudes close to the tropopause and sloping upwards towards the subtropics. There are generally smaller SDs in the Tropics and in the stratosphere, although at high levels there are two localized maxima centred on the equator. The \( v \) component (not shown) has very similar vertical covariances, with slightly smaller magnitude in the Tropics but slightly larger mid-latitude maxima. The SD of \( u_{div} \) (the \( u \) component of the divergent wind, Fig. 4(b)) is generally much smaller, illustrating that the wind is largely rotational. Largest values are at high levels particularly in the Tropics. The divergent wind shows more variability with sample than other variables examined and the tropical maximum at about 125 hPa became more pronounced after the introduction of a parametrization of convective momentum transport.

Temperature (Fig. 4(c)) has broad maximum SDs in the mid-latitude troposphere, extending into the lower stratosphere but with slightly lower values round the tropopause itself. Near the surface the SDs are much larger in northern mid-latitudes than southern mid-latitudes (also true in July) presumably reflecting the larger temperature variations over land. Tropical magnitudes are much less.

Relative-humidity SDs (Fig. 4(d)) are quite low in the boundary layer rising to maxima in the upper troposphere; the gradient of the transition and the maximum values are rather large compared to our previous estimates from radiosonde-minus-background statistics. Pressure SDs (Fig. 4(e)) decrease with height (as might be expected given the reduction of pressure values), gradually in the troposphere then more rapidly in the stratosphere; magnitudes are much less in the Tropics. Figure 4(f) shows that the pressure is largely balanced polewards of 30° below 100 hPa with the southern extratropics being slightly more balanced than the northern extratropics. Ongoing investigations suggest that noise in the forecast model is at least partly responsible for the lack of balance in the mid-stratosphere. Stratospheric balance is also sensitive to the details of the vertical regression applied to the balanced pressure (see appendix of Lorenc et al. 2000).
The January SDs shown are reasonably symmetric about the equator—extratropical errors tend to be larger in winter than summer—but overall errors are smaller in the northern extratropics than the southern extratropics. In July (not shown) there is significant asymmetry with the largest SDs in the southern extratropics.

For comparison, Figure 5 shows root-mean-square (r.m.s.) surface marine observation-minus-background (O−B) statistics for January 1999. The errors of the ships and buoys used should be reasonably independent of latitude (gross errors have been removed by quality control), and biases are relatively small at most latitudes. The wind r.m.s. O−B magnitudes are smaller in the Tropics than mid-latitudes; when the observation error is taken into account there is reasonable agreement with the lowest level of Fig. 4(a). Temperature r.m.s. differences show large values at high latitudes and broad secondary maxima in the subtropics; to some extent this pattern can be seen in the southern hemisphere of Fig. 4(c). As already suggested, the land in the northern hemisphere will give larger overall variances there: surface land stations (not shown) have a similar pattern of temperature r.m.s. but with values about 2 degC larger. The pressure r.m.s. statistics are relatively flat north of 50°S, and the mid-latitude maxima in Fig. 4(e) are too large (partly due to the horizontal length-scales being too large?). For pressure, and to a lesser extent temperature, the significantly lower SDs in the Tropics seen in Fig. 4 are overstated, probably because error growth rates (affecting the T+24 and T+48 forecasts) are lower in the Tropics.

The $u$ component of the wind has very narrow vertical correlations in the Tropics; Fig. 6(a) shows the correlation with level 11 (approximately 500 hPa). In the extratropics there are much broader vertical scales with positive correlations extending almost to the top of the model. Looking at correlation greater than 0.6, the vertical scale tends to increase at higher latitudes and is largest in the southern hemisphere. In July (not shown) the vertical scale is even larger in high southern latitudes, but is fairly similar elsewhere; $u_{\text{div}}$ has much narrower vertical correlations (Fig. 6(b)) which are approximately independent of latitude, but show some narrowing near the equator and some broadening over Antarctica.

Temperature (Fig. 6(c)) shows a distinct narrowing in the Tropics. In the extratropics the correlations with level 11 are positive throughout the troposphere, and they change sharply at the tropopause, with correlations down to −0.4 (or very locally −0.6) in the lower stratosphere. In contrast with the generally positive vertical correlations of most
Figure 6. Vertical correlations of forecast differences with level 11 (approximately 500 hPa) for (a) $u$, (b) $u_{div}$, and (c) temperature. Contours as Fig. 1.
variables, such large negative correlations are quite notable; they were also found by Rabier et al. (1998, Fig. 12). The correlation between pressure differences at the top and bottom of the model is approximately zero, so that positive temperature correlations have to be compensated by negative correlations at a different level.

Pressure vertical correlations (not shown) are very broad but narrow slightly in the Tropics. Relative-humidity vertical correlations (not shown) are moderately narrow and approximately independent of latitude, but with a slight narrowing about 20°N—also seen in previous versions of the statistics.

(c) **Mid-latitude vertical structure**

Figure 7 shows the first vertical mode for various variables, and also the percentage of total variance explained. These are calculated as eigenvectors of the global vertical covariances, pressure weighted to take the mass of each layer into account; they are used in our representation of the background-error covariances—see section 3 of Lorenc et al. (2000). (The percentage of total variance, and to a lesser extent the shape of the leading modes, are sensitive to the weighting used.) The global covariance matrices are dominated by the larger SDs in the extratropics, so the global modes are typical of mid-latitude conditions. The leading modes in particular appear to be physically meaningful, but they have somewhat larger scale than apparent in the correlations—see discussion in Richman (1986) and Jolliffe (1987).

The vertical modes have been calculated for the different variables independently, but for the largely balanced variables the leading modes are clearly interrelated to some extent. Looking at $u$ we have an “equivalent barotropic” vertical structure with flow in the same sense at all levels and maximum amplitude near the tropopause” (Hoskins 1987, p. 59). This will be in balance with (horizontally displaced) temperature perturbations of different sign in the troposphere and stratosphere. The first stream-function mode is similar to that for $u$, but the maximum is displaced upwards slightly and there is more amplitude in the stratosphere—these effects are due to the longer scales and hence larger stream function: KE ratio in the stratosphere (see Table 1 and discussion). There is also some broadening in the vertical as stream function emphasises larger horizontal scales and hence larger vertical scales (see Fig. 2(a)). For pressure, the first mode explains almost 75% of the total variance (Fig. 7); in the troposphere the magnitude increases gradually towards the surface (it is only partially in balance with the first wind mode).

Velocity potential has a negative lobe (expected as the vertically integrated divergence should be close to zero) and slightly odd structure in the stratosphere. Relative humidity has a mode of the same sign at all levels with a maximum in mid-troposphere.
Together with temperature these have shorter vertical scales than \( u \) or pressure and also a smaller proportion of variance explained by the first mode.

\( d \)  **Tropical aspects**

Figure 8 shows leading vertical modes calculated just for the region between 15°S and 15°N; they are significantly different from the global modes in Fig. 7. The leading modes for stream function, velocity potential and \( u \) are quite similar and suggest strong flow in the upper troposphere associated with weak flow in the opposite direction at low levels (more so in \( v \) than \( u \)), similarly to Hoskins (1987). The leading pressure mode still dominates the pressure variance but has a simpler structure corresponding to a height perturbation approximately constant in the vertical. The first temperature mode has a curious shape and only 12% of the variance, and is probably not physically meaningful. The first relative-humidity mode has a broad maximum at about 250 hPa.

The vertical scales in the Tropics are generally very short, but the small proportion of the Tropics with deep convection, or particularly with tropical storms, will have much more coupling in the vertical. In the lower troposphere there are some indications of larger wind vertical scales at about 15° in the summer hemisphere (see Fig. 6(a), although it is more noticeable at lower levels), which is tentatively identified with the effect of tropical cyclones.

In our system there is effectively no mass–wind coupling in the Tropics. Daley (1996) discussed the possibility that the westerly wind and mass fields might be correlated in the equatorial zone but, in our forecast differences, collocated \( u \)–pressure correlations are generally less than 0.1 in magnitude in the troposphere within 10° of the equator.

\( e \)  **Surface pressure/temperature correlations and polar regions**

Figure 9 shows correlations between model surface pressure \( p_s \) and temperature as a function of latitude. In mid-latitudes there are negative correlations in the lower troposphere and upper troposphere/upper stratosphere, with zero or slightly positive values in mid-troposphere. The negative correlation is strongest, \(-0.3\) or locally \(-0.4\), at about 900 hPa. In the Tropics the correlations are generally slightly weaker. The strength of the negative correlation between \( p_s \) and lower-stratospheric temperatures in the southern extratropics is surprising. It is related to the dominance of the ‘equivalent barotropic’ mode: the first pressure mode in Fig. 7 shows that surface pressure increments are associated with vertical pressure gradients—and hence temperature changes—at upper levels. During tests it was noted that satellite soundings at these latitudes could have a
large effect on $p_s$ in our 3D-Var system (unlike in the AC system); in the case examined the impact came largely from high levels.

Another remarkable feature of Fig. 9 is the strong positive correlation of tropospheric temperatures with $p_s$ over Antarctica, with some compensating negative correlations in the stratosphere above. There is some sign of a similar feature in the Arctic, but tropospheric correlations are close to zero. The positive correlation over Antarctica is slightly weaker in July (not shown) when the low-level inversion and associated katabatic winds are stronger. The forecast-difference statistics for the previous 19-level model did not have such positive correlations over Antarctica, but a comparison with the observational studies reported later suggests that the positive correlation is ‘real’. In the AC scheme, ‘hydrostatic’ potential-temperature increments, of opposite sign and largest near the surface, were derived from the $p_s$ increments (Lorenc et al. 1991). The negative rather than positive correlation at low levels was apparently responsible for some rather poor analyses at the south pole in March/April 1998 shortly after the introduction of the 30-level model and before 3D-Var was operational.

The synoptic relationship between pressure and temperature in the Antarctic winter is investigated by Wendler and Kodama (1993). They find significant positive correlation between surface temperature and pressure—“at first glance a very astonishing result”. This appears to be at least partly due to occasional large-scale incursions of mid-latitude air. A very large warming of this type, associated with high pressure over east Antarctica, is documented by Enomoto et al. (1998). Perhaps the most obvious explanation among those suggested by Wendler and Kodama (1993) is that anticyclonic circulation is associated with descent and that for Antarctica the air aloft is relatively warm. The positive correlation of tropospheric temperature (errors) and surface pressure (errors) might be expected in any ‘warm’ anticyclone, including many extratropical blocking episodes.

On a seasonal time-scale, Antarctic surface temperature and pressure are positively correlated in that pressures are highest in summer. Parish and Bromwich (1997) note that seasonal temperature changes are largest in the stratosphere and near the surface and less in the middle and upper troposphere. They also note similarities with Greenland, and Rogers et al. (1997) describe an abrupt springtime temperature rise over Greenland in several years, preceded by a significant pressure rise. The large-scale nature of the phenomenon is consistent with Derber and Bouttier (1999). Their Fig. 8 shows negative
correlations in the lower troposphere at all scales except the very large ones; at large scales there are negative correlations at/above the tropopause and some positive correlations below. The height of the Antarctic plateau prevents most synoptic disturbances from penetrating the interior, so the covariances there are probably more representative of very large scales.

Temperature–$p_*$ correlations were calculated for a few radiosonde stations in Antarctica and, for comparison, in Britain. The results (not shown) are reasonably consistent with Fig. 9—if anything they suggest that the low-level negative correlation is overstated relative to the high-level negative correlation. The south pole station shows positive correlations between tropospheric temperatures and $p_*$ for both observed values and observed-minus-background values, and a number of other stations show positive correlations for observed values but not for observed-minus-background values.

3. REPRESENTING BACKGROUND-ERROR COVARIANCES

(a) Basic method

The design of the background term in our 3D-Var system is described in section 3 and the appendix of Lorenz et al. (2000) and it is partly based on that of Parrish and Derber (1992). We use the forecast-difference statistics (section 2) for the ‘control variables’ ($\psi$, $\chi$, Ap and RH). We calculate vertical modes from global vertical covariance
matrices, and allow their variances to vary with latitude and season. For each vertical mode we calculate the horizontal correlation spectrum, but we replace these with second-order autoregressive (SOAR) functions with related length-scales. We then have to rescale the stream-function and velocity-potential variances so that the implied global kinetic energy is equal to that from the forecast differences.

(b) Implied covariances

Figure 10(a) shows implied standard deviations (SDs) for $u$ using the default options and can be compared with Fig. 4(a). The main features are moderately well captured: the mid-latitude jet level maxima are there but are slightly weak, SDs are weaker in the Tropics as ‘observed’ but somewhat too large below about 500 hPa (level 11). A feature that shows up slightly here is that the slope of the maxima with latitude is less marked than the ‘observed’ slope which follows the tropopause; the dominant global mode is most representative of mid-latitudes and largely determines the location of the maximum. The implied vertical correlations with level 11 in Fig. 10(b) can be compared with those in Fig. 6(a), as for the SDs the main features are present but somewhat smoothed, and the implied vertical correlations in the Tropics are narrow but not as narrow as in the forecast difference statistics. Implied temperature SDs, Fig. 11(a), are less than those in Fig. 4(c), partly due to the use of shorter $\psi$ length-scales. The vertical temperature correlations in Fig. 11(b) have shorter vertical scales.
and correspondingly less negative correlations in the stratosphere than in Fig. 6(c), but the gross features are similar. The correlations between $p_\ast$ and temperature (not shown) capture the main features of Fig. 9, particularly the negative low-level correlations. The negative high-level correlations are represented to some extent; they are sensitive to some of the modifications that have been made to the system—so are the correlations over Antarctica.

The dominant equivalent barotropic mode is well modelled—it is probably most important in data-sparse areas and may be exaggerated in the forecast differences relative to six-hour forecast errors. Even if it is dominant over large areas of the globe, there is an argument that it may be more important to get the analysis correct in the active baroclinic areas—this is part of the motivation for experimenting with shorter vertical correlations (section 4).

Our system does represent the gross features of the relationship between vertical and horizontal scale for rotational wind (Fig. 12 compared to Fig. 2(a)), but horizontal length-scales increase very little in the stratosphere and actually decrease somewhat for unbalanced pressure—Table 2. In the troposphere the lengths are somewhat less than those in Table 1; this is due to the use of SOAR functions rather than direct use of the forecast-difference spectra.

(c) Discussion

Various approximations are made in the modelling of the covariances which result in differences between the input covariances (section 2) and the implied covariances. For a three-dimensional field the approximations made are:
(i) We project latitudinally and seasonally varying covariance matrices onto a set of modes derived from the global average covariance.

(ii) For each vertical mode we specify a horizontal correlation function.

(iii) The decomposition into SD and correlation implies that the SD should not vary on short scales.

(iv) For each vertical mode the horizontal correlation is homogeneous and isotropic because we are using a spectral representation.

(v) We have replaced the horizontal spectra from the forecast differences by SOAR spectra with related length-scales.

Through (i) we model some, but not all, of the variation with latitude. The semi-separability assumption (ii) is probably the worst approximation as currently implemented. The orthogonality condition on the vertical modes means that they have different vertical scales, and the relationship between vertical and horizontal scales is modelled reasonably well, see Fig. 12. However, some modes contain contributions from both the stratosphere and the short vertical scales in the troposphere—their horizontal spectra tend to be dominated by the tropospheric component and hence the implied horizontal scales in the stratosphere are inadequate (Table 2). Approximation (iii) does not seem too restrictive for the ‘climatological’ covariances; and (iv) works quite well for mid-latitude conditions, but there should be longer horizontal scales in the Tropics, especially for the mass field (which has correlations elongated east–west there). There are some variations in length-scale due to data density, and small variations with season. Effectively (v) reduces horizontal length-scales—this is empirically justified on the basis that our forecast-difference statistics probably overestimate the error length-scales at shorter range (section 1(b)). However, there is some evidence that the reduction is too large and a smaller reduction (also retaining some of the ‘shape’ from the forecast-difference statistics) is being tested.

There are also multivariate aspects to consider. The wind is represented using stream function and velocity potential, partly because the rotational and divergent wind do have different covariance characteristics but also because it is much easier to deal with the covariances of a scalar quantity. In the extratropics, wind and temperature are largely derived from the horizontal and vertical gradients of stream function so they are sensitive to the details of its covariances. Stream function and velocity potential are integrated quantities and their use emphasises large scales—see discussion of Fig. 7. Their non-local nature is illustrated graphically by Hendon (1986). Stream-function SD maxima tend to be smoothed and slightly displaced in latitude and level compared with wind SD maxima—also reflected in the implied wind covariances. We now use vertical modes derived from rotational (and divergent) kinetic-energy covariances which alleviate some of these problems. However, it seems better to use more localized control variables; consideration is being given to the use of potential vorticity. The assumption that our control variables are uncorrelated is reasonably accurate, but minor improvements are possible.

The ECMWF covariances described by Derber and Bouttier (1999) differ mainly in that vertical correlations are treated as a function of horizontal wave number. This automatically gives longer horizontal scales in the stratosphere and a relationship between horizontal and vertical scales. The disadvantage is that for their control variables (e.g. vorticity) the vertical correlations are homogeneous and notably do not narrow in the Tropics. Through their balance operator and the ratio of vorticity to divergence variances, their wind and temperature vertical correlations do have some latitudinal variation. They lack a representation of tropopause sloping in mid-latitudes which can cause
increments from aircraft data near the tropopause to be incorrect, particularly for temperature (F. Bouttier, personal communication). They include some coupling between the divergence and vorticity fields especially near the surface and the tropopause.

The Derber and Bouttier correlation formulation is, like ours, the sum of a number of separable functions. It has more terms and significantly more coefficients/parameters to specify than our formulation, and within the constraint of horizontal homogeneity and isotropy it is fully non-separable. Our formulation makes a modified separability assumption allowing some representation of the horizontal inhomogeneities. Unfortunately, neither of these formulations can replicate the full three-dimensional structure suggested by Eq. (1) and seen in section 2.

The lack of increased horizontal length-scales in the stratosphere is possibly the biggest flaw in our implied covariances. We have performed some experiments using 'rotated' vertical modes (see e.g. Richman 1986) in order to try to obtain more localized (and perhaps more physically meaningful) vertical modes. This did give longer horizontal scales, and better forecasts, in the stratosphere but the tropospheric impact was more mixed. A modified form (with tropospheric modes re-orthogonalized) is more successful, but the details are outside the scope of this paper. Work is underway on the humidity covariances and on a vertically extended model. A practical disadvantage of the NMC method is that significant effort is needed if forecast-model levels change; a more theoretically based model of vertical structure would be useful as well as interesting.

Other possible changes to our representation of background errors are mainly related to modelling 'synoptic dependent errors', via a 'geostrophic transform' and/or 'bred modes'. With our covariance formulation it would be relatively easy to vary the vertical correlations (within limits) as a function of longitude as well as latitude, it might be useful to specify shorter vertical scales in baroclinically active areas, and also some account could be taken of data density. Allowing the first stream-function mode to vary slowly as a function of latitude (or tropopause height) could also be considered.

4. IMPACT ON FORECASTS

We briefly present four experiments that show the development of our covariance modelling over the first year of operations. Selected verification results are shown in Table 3. Experiments 1 to 3 use the previous sample of forecast differences based on 20 days each in July 1997 and January/February 1998. Experiment 1 (operational from late March to July 1999) used stream-function and velocity-potential vertical modes. Experiment 2 used rotational/divergent kinetic-energy vertical modes giving slightly smaller vertical scales and a slightly better position of the wind variance maxima. Experiment 3 (operational between late July and October 1999) differs from experiment 2 in the scaling of rotational wind magnitudes, which gives slightly smaller vertical scales. In general, experiment 3 appears better than experiment 1 (tested on June 1999 data the impact was neutral) and experiment 2 (although the impact is slightly negative measured against observations).

Experiment 4 is essentially the same as experiment 3 but with more recent forecast-difference samples (presented in section 2, they are slightly smoother and there are fairly subtle changes to the SDs and vertical correlations). This gives a moderate improvement relative to experiment 3, smaller when the experiment is extended for another week, with generally similar behaviour tested on June data. These statistics became operational in October 1999. The magnitude of the impact is comparable to major changes in observation usage.
TABLE 3. Root-mean-square verification against observations and analyses for the period 5–19 March 1999

<table>
<thead>
<tr>
<th>Reference Experiment:</th>
<th>Observations</th>
<th></th>
<th></th>
<th></th>
<th>Analyses</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>NH PMSL T+24</td>
<td>2.03</td>
<td>2.02</td>
<td>2.03</td>
<td>2.04</td>
<td>2.06</td>
<td>2.06</td>
<td>2.06</td>
<td>2.04</td>
</tr>
<tr>
<td>NH PMSL T+48</td>
<td>3.07</td>
<td>3.05</td>
<td>3.09</td>
<td>3.09</td>
<td>3.23</td>
<td>3.26</td>
<td>3.23</td>
<td>3.19</td>
</tr>
<tr>
<td>NH PMSL T+72</td>
<td>4.72</td>
<td>4.61</td>
<td>4.69</td>
<td>4.61</td>
<td>4.48</td>
<td>4.50</td>
<td>4.49</td>
<td>4.43</td>
</tr>
<tr>
<td>NH PMSL T+96</td>
<td>5.88</td>
<td>5.70</td>
<td>5.75</td>
<td>5.64</td>
<td>5.79</td>
<td>5.70</td>
<td>5.79</td>
<td>5.64</td>
</tr>
<tr>
<td>NH PMSL T+120</td>
<td>6.95</td>
<td>6.80</td>
<td>6.77</td>
<td>6.51</td>
<td>6.80</td>
<td>6.72</td>
<td>6.82</td>
<td>6.58</td>
</tr>
<tr>
<td>NH H500 T+24</td>
<td>1.88</td>
<td>1.88</td>
<td>1.88</td>
<td>1.85</td>
<td>1.58</td>
<td>1.59</td>
<td>1.55</td>
<td>1.53</td>
</tr>
<tr>
<td>NH H500 T+48</td>
<td>2.81</td>
<td>2.79</td>
<td>2.80</td>
<td>2.78</td>
<td>2.69</td>
<td>2.71</td>
<td>2.64</td>
<td>2.61</td>
</tr>
<tr>
<td>NH H500 T+72</td>
<td>4.26</td>
<td>4.27</td>
<td>4.22</td>
<td>4.17</td>
<td>4.03</td>
<td>4.05</td>
<td>4.00</td>
<td>3.91</td>
</tr>
<tr>
<td>NH W250 T+24</td>
<td>7.76</td>
<td>7.75</td>
<td>7.77</td>
<td>7.72</td>
<td>5.26</td>
<td>5.37</td>
<td>5.22</td>
<td>5.19</td>
</tr>
<tr>
<td>TR W850 T+24</td>
<td>4.01</td>
<td>4.01</td>
<td>4.00</td>
<td>4.00</td>
<td>2.20</td>
<td>2.19</td>
<td>2.15</td>
<td>2.18</td>
</tr>
<tr>
<td>TR W850 T+48</td>
<td>4.41</td>
<td>4.43</td>
<td>4.41</td>
<td>4.40</td>
<td>2.80</td>
<td>2.81</td>
<td>2.77</td>
<td>2.79</td>
</tr>
<tr>
<td>TR W850 T+72</td>
<td>4.61</td>
<td>4.63</td>
<td>4.64</td>
<td>4.59</td>
<td>3.23</td>
<td>3.25</td>
<td>3.22</td>
<td>3.23</td>
</tr>
<tr>
<td>TR W250 T+24</td>
<td>5.80</td>
<td>5.80</td>
<td>5.80</td>
<td>5.76</td>
<td>3.73</td>
<td>3.95</td>
<td>3.78</td>
<td>3.84</td>
</tr>
<tr>
<td>SH PMSL T+24</td>
<td>1.82</td>
<td>1.79</td>
<td>1.79</td>
<td>1.80</td>
<td>1.92</td>
<td>1.87</td>
<td>1.84</td>
<td>1.86</td>
</tr>
<tr>
<td>SH PMSL T+48</td>
<td>2.71</td>
<td>2.61</td>
<td>2.65</td>
<td>2.63</td>
<td>3.24</td>
<td>3.16</td>
<td>3.10</td>
<td>3.12</td>
</tr>
<tr>
<td>SH PMSL T+72</td>
<td>3.33</td>
<td>3.21</td>
<td>3.21</td>
<td>3.19</td>
<td>4.36</td>
<td>4.22</td>
<td>4.21</td>
<td>4.22</td>
</tr>
<tr>
<td>SH PMSL T+96</td>
<td>3.61</td>
<td>3.54</td>
<td>3.50</td>
<td>3.44</td>
<td>5.27</td>
<td>5.07</td>
<td>5.08</td>
<td>5.06</td>
</tr>
<tr>
<td>SH PMSL T+120</td>
<td>4.65</td>
<td>4.47</td>
<td>4.61</td>
<td>4.46</td>
<td>6.61</td>
<td>6.02</td>
<td>6.09</td>
<td>5.94</td>
</tr>
<tr>
<td>SH H500 T+24</td>
<td>2.09</td>
<td>2.15</td>
<td>2.11</td>
<td>2.02</td>
<td>1.77</td>
<td>1.76</td>
<td>1.66</td>
<td>1.68</td>
</tr>
<tr>
<td>SH H500 T+48</td>
<td>2.74</td>
<td>2.72</td>
<td>2.71</td>
<td>2.62</td>
<td>3.14</td>
<td>3.05</td>
<td>2.97</td>
<td>3.00</td>
</tr>
<tr>
<td>SH H500 T+72</td>
<td>3.65</td>
<td>3.65</td>
<td>3.57</td>
<td>3.48</td>
<td>4.33</td>
<td>4.17</td>
<td>4.19</td>
<td>4.13</td>
</tr>
<tr>
<td>SH W250 T+24</td>
<td>7.48</td>
<td>7.39</td>
<td>7.45</td>
<td>7.35</td>
<td>5.32</td>
<td>5.43</td>
<td>5.07</td>
<td>5.09</td>
</tr>
<tr>
<td>NH Skill</td>
<td>-0.12</td>
<td>0.09</td>
<td>46.77</td>
<td>0.25</td>
<td>-0.03</td>
<td>-0.02</td>
<td>47.93</td>
<td>0.32</td>
</tr>
<tr>
<td>TR Skill</td>
<td>-0.06</td>
<td>-0.11</td>
<td>4.11</td>
<td>0.12</td>
<td>0.09</td>
<td>-0.18</td>
<td>9.82</td>
<td>-0.07</td>
</tr>
<tr>
<td>SH Skill</td>
<td>-0.20</td>
<td>0.08</td>
<td>22.05</td>
<td>0.18</td>
<td>-0.32</td>
<td>-0.09</td>
<td>23.54</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Root-mean-square verification against observations (radiosonde and surface land stations) and analyses. Variables are: pressure at mean sea level (PMSL) (hPa), height at 500 hPa (H500) (dm), and vector wind at 850 and 250 hPa (W850 and W250) (m s⁻¹). Weighted average skill scores relative to persistence are given for experiment 3, and for the differences from experiment 3. Maximum possible skill scores are 56 for NH (20°N–90°N), 16 for TR (20°S–20°N) and 28 for SH (20°S–90°S).

5. CONCLUSIONS

A good specification of background-error covariances is an essential part of any state-of-the-art data assimilation system. A major part of this paper is concerned with trying to understand the structure of the covariances, and also the extent to which forecast-difference statistics can be used to represent them. The answer seems to be that they work well if used with caution—we have taken steps to narrow the horizontal and vertical correlations somewhat. Further comparison with observation-based statistics is desirable.

The covariances of our 3D-Var system have a better, more consistent, representation of balance than our previous AC scheme. A corollary of this is that a change to one aspect cannot be made in isolation, e.g. a change to horizontal stream-function length-scales has effects on wind and mass variances and also, to some extent, on vertical correlations. The covariances have most impact in data-sparse areas and it seems that, in some sense, the correlations are more important than the variances. (Although if background-error variances are significantly underestimated the assimilation can lose track of reality—from this point of view it is safer to slightly overestimate them.) Certainly forecast quality is sensitive to fairly small changes in the correlations.
ACKNOWLEDGEMENTS

This work has depended in various ways on others in The Met. Office VAR team including A. Lorenc, P. Andrews, S. Ballard, S. Bell, D. Barker, J. Bray, A. Clayton and M. Thurlow. In understanding the covariances, conversations with, or information from, F. Bouttier, P. Courtier, M. Cullen, G. Desroziers, L. Fillion, M. Fisher, D. Parrish, F. Rabier and R. Swinbank have been useful. Thanks also to A. White and two anonymous referees for constructive comments on the manuscript. Many people have supplied pieces of the jigsaw.

REFERENCES


