Q. J. R. Meteorol. Soc. (2005), **131**, pp. 987–1011 doi: 10.1256/qj.04.54

Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors

By NEDJELJKA ŽAGAR[1]*, ERIK ANDERSSON[2] and MICHAEL FISHER[2]

[1]*Stockholm University, Sweden*

[2]*European Centre for Medium-Range Weather Forecasts, Reading, UK*

(Received 2 April 2004; revised 25 November 2004)

SUMMARY

This paper seeks to represent the tropical short-range forecast error covariances of the European Centre for Medium-Range Weather Forecasts (ECMWF) model in terms of equatorial waves. The motivation for undertaking this investigation is increasing observational evidence indicating that a substantial fraction of the tropical large-scale variability can be explained by equatorially trapped wave solutions known from shallow-water theory. Short-range forecast differences from a data-assimilation ensemble were taken to serve as a proxy for background errors.

It was found that the equatorial waves coupled to convection can explain on average 60–70% of the error variance in the tropical free atmosphere. The largest part of this explained variance is represented by the equatorial Rossby (ER) modes, and a significant percentage pertains to the equatorial inertio-gravity (EIG) modes. Eastward-propagating EIG modes have maximum variance in the stratosphere, where the short-wave variance in westward-moving waves is particularly small. This feature is most likely related to the phase of the quasi-biennial oscillation during the study period, suggesting that significant temporal variations could be present in longer-term time series of such statistics.

The vertical correlations for ER modes display characteristics similar to those of their extratropical counterparts: correlations narrow towards shorter scales and in the stratosphere. However, the present statistics do not display the significant increase with altitude of the horizontal correlation scale for the height field which is typical for global, quasi-geostrophic statistics commonly used in current data-assimilation schemes. Furthermore, tropospheric ER correlations are vertically asymmetric and deeper for the $n = 1$ mode than for higher modes. Most likely, deep convection, acting as a generator of equatorial wave motion, is the dominant mechanism underlying these results.

In spite of its relatively small contribution to the tropospheric variance, the Kelvin-wave coupling plays a decisive role for determining the characteristics of the horizontal correlation near the equator. EIG modes also play an important role for the tropical mass–wind coupling; these waves have a major impact by reducing the meridional correlation scales and the magnitudes of the balanced height-field increments.

KEYWORDS: Covariance modelling Ensemble methods Mass–wind coupling Tropics Variational data assimilation

1. INTRODUCTION

Numerical weather-prediction (NWP) models are never going to be error-free. A reliable estimate of their forecast errors poses a real challenge, since the knowledge of the true state of the atmosphere is beyond our grasp. In data assimilation for NWP, short-range forecast errors are commonly referred to as background errors; they are frequently represented by surrogate quantities with statistical and dynamical properties assumed similar to those of the unknown forecast errors (e.g. Parrish and Derber 1992). Derived dependencies are built into the background-error covariance matrix for data assimilation; the purpose of these relationships is to spread observed information from a point to nearby grid points and levels. Moreover, the observed information is also distributed to other variables. In this way observations of the temperature field carry information about the wind field, and vice versa. This is the fundamental reason why the balance relationships between the mass- and the wind-field variables are of such great importance for data assimilation, especially in regions where observations are sparse and in a Global Observing System (GOS) dominated by mass-field information.

* Corresponding author: Department of Meteorology, Stockholm University, SE-106 91 Stockholm, Sweden.
e-mail: nedjeljka@misu.su.se

In the midlatitudes, the basic balance relationship is geostrophy, which has been extensively used in analysis procedures (e.g. Courtier *et al.* 1998; Gustafsson *et al.* 2001). The most important elements of atmospheric motion are Rossby waves; their accurate analysis is, therefore, the primary concern in data assimilation for NWP. To ensure that the observational information is assimilated mainly in terms of Rossby modes, initialization procedures and methods for generating geostrophically balanced increments have been developed. As a result, the excessive generation of inertio-gravity (IG) waves is suppressed.

In the tropics, on the other hand, a dominant relationship similar to geostrophy is lacking; the analysis here has thus traditionally been undertaken in the univariate fashion. Consequently, large-scale divergence fields, such as the Hadley and Brewer–Dobson circulations, are analysed nearly univariately. Since GOS in the tropics relies on mass-field information, uncertainties in the analysed wind field are significant (e.g. Kistler *et al.* 2001). Furthermore, large-scale motion in the tropics cannot be considered without taking into account IG waves (e.g. Browning *et al.* 2000). In addition, the change of sign of the Coriolis parameter, f, at the equator gives rise to important types of large-scale non-rotational motion, which are absent in the midlatitude atmosphere: the Kelvin and mixed Rossby–gravity (MRG) modes (Matsuno 1966).

Indeed, equatorially trapped Kelvin, MRG and equatorial IG (EIG) waves have regularly been detected in observations since the 1960s. Recent observational studies (Wheeler and Kiladis 1999; Yang *et al.* 2003) identified equatorially trapped wave structures in long-term satellite observations of outgoing long-wave radiation, a proxy for deep tropical cloudiness. The waves have been denoted the 'convectively-coupled equatorial waves', as their presence in areas of moist convection implies an interaction between convection and the dynamics. This causes the wave characteristics to change: their frequency is lowered, i.e. the equivalent depth is becoming smaller. Available equivalent-depth estimates according to linear shallow-water theory are in the range 12–50 m (Wheeler and Kiladis 1999, and references therein). These observational studies thus corroborate an earlier numerical model study by Ko *et al.* (1989), which revealed that 'the gravity waves associated with the shallow vertical modes and long zonal waves play an important role in the balanced gravitational energy'. Normal-mode-based 3D-Var and initialization techniques (Cats and Wergen 1983; Wergen 1988; Heckley *et al.* 1992; Andersson *et al.* 1998) have therefore been unsuccessful in the tropics.

In the present study we consider Kelvin, MRG and EIG waves, in addition to equatorial Rossby (ER) modes, as balanced tropical motion, and investigate the extent to which these equatorially trapped wave solutions are present in the ECMWF (European Centre for Medium-Range Weather Forecasts) short-range forecast errors. An ultimate goal is to construct a background-error covariance matrix comprising the coupling between the mass and the wind field in the tropics.

In the background-error modelling, an important, commonly made, assumption is that the forecast errors are dominated by the balances of the model's slow modes (Phillips 1986); this implies that background-error covariances are dominated by structures similar to that of growing quasi-geostrophic perturbations in a model atmosphere. In accordance with these assumptions, we employ equatorial-wave theory to replace the quasi-geostrophic relationships that have proven to be successful in extratropical latitudes. The idea has previously been presented in Žagar *et al.* (2004, henceforth ŽGK), summarized here in section 2. This section also describes a data-assimilation ensemble, used to extract short-range forecast differences to serve as a proxy for background errors; the justification is given in the appendix. We investigate how accurately a spectral, mode-based covariance model can represent the dominant characteristics of

the ECMWF short-range tropical forecast errors. The resulting covariance statistics are presented in section 3. In section 4, the covariance model is implemented in the variational data-assimilation scheme developed in ŽGK and utilized for an examination of horizontal structures in the troposphere and stratosphere through idealized single-observation experiments. Conclusions are presented in section 5.

2. Background-error covariance modelling for the tropics

(*a*) *An ensemble-based dataset*

The current background-error covariance model in operational use within the 4D-Var system at ECMWF (Fisher 2003) is based on statistics of concurrent short-range forecast differences between members of a data-assimilation ensemble. The justification for this approach is given in the appendix. The ensemble consists of ten independent data assimilations that each use different sets of randomly perturbed observations over a period of 31 days in October 2000. For the purposes of this study, forecast differences were extracted from the ensemble, for the tropical belt 20°S−20°N, which is commonly used in observational studies. A one-degree resolution is used in both horizontal directions. The two wind components and geopotential height were available at 60 model levels, for different forecast lengths: 3-, 12- and 24-hour forecasts have been used. The results presented here are with respect to 12-hour forecast errors, but the outputs for other ranges are very similar.

The ensemble standard deviation of error for the zonal wind is shown in Fig. 1 at three model levels: in the lower and upper troposphere and near the tropopause. A striking feature in the lower troposphere is the frequent occurrence of errors along the intertropical convergence zone (ITCZ) and over Indonesia (Fig. 1(a)). Local maxima are located west of the tropical continents. The absolute error maximum is found in the upper troposphere, near the 200 hPa level, just west of South America (Fig. 1(b)). Higher up, the Indonesian region displays larger errors than the rest of the domain (Fig. 1(c)), whereas above the tropopause the errors become largely homogeneous.

The errors in the meridional wind field display a similar structure in the lower troposphere (not shown), but higher up the error maxima near the South American coast and over Indonesia are weaker (up to 50 %) and the errors become more homogeneous at lower altitude as compared to the zonal wind errors. Errors in geopotential (not shown) are larger over the continents than over the oceans in the lower troposphere, with peaks associated with orographic features. There is an error maximum over Indonesia in the upper troposphere, but above the tropopause the spatial error variations become small.

Figure 1 shows standard deviations of 12-hour forecast errors; however, patterns and magnitudes are similar for the other ranges (3 and 24 hours). The wind-field errors grow in time, at least in the troposphere. The growth is larger during the later 12-hour period (12–24 hours), and it is largest for levels between 100 and 300 hPa. Areas of largest error growth coincide with the areas of significant error in Fig. 1, i.e. the largest errors are located just west of South America between levels 28 (154 hPa) and 35 (353 hPa). The geopotential-error growth is almost negligible. The stratospheric forecast errors do not grow during 24 hours: on the contrary, they are on average 2% smaller than the 3-hour errors for all fields.

(*b*) *Covariance-model formulation*

In the present study, forecast errors are represented in terms of tropical eigenmodes based on a parabolic cylinder function expansion. The expansion for the meridional

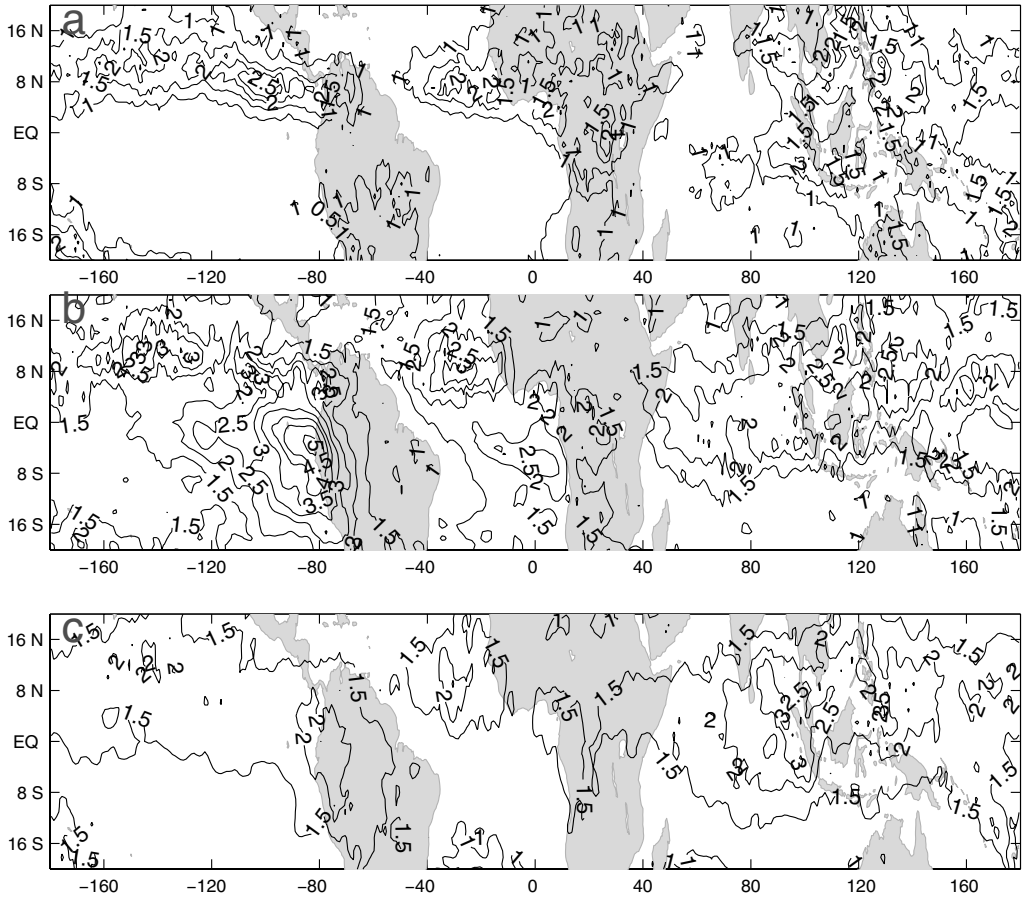

Figure 1. Standard deviations of 12-hour forecast errors in the zonal wind field (in m s^{-1}) of the ECMWF model in the tropics, as estimated with a data-assimilation ensemble. Shown are model levels (a) 48 (approximately 828 hPa), (b) 30 (approximately 202 hPa) and (c) 25 (approximately 96 hPa). Isolines every 0.5 m s^{-1}, starting from 0.5 m s^{-1}.

wind-field error (v) at vertical level z assumes the following form:

$$v(x,\, y) = \sum_{k=-N_k}^{N_k} \sum_{n=0}^{N_n} \sum_{m=1}^{3} \Psi_{knm} v_{knm}(k,\, y) e^{\mathrm{i}kx}, \tag{1}$$

where k is a zonal wave-number index, n is the number of nodes in the meridional direction of v (thus being denoted the meridional mode number), and m is an index for wave type. For each single equatorial mode, the three modal indices define a modal index $\nu(k, n, m)$, which will be used in what follows. A meridional structure function $v_\nu(k,\, y)$ is given in terms of a parabolic cylinder function of degree n; k-dependence enters through the equatorial trapping scale a_e (see ŽGK for details). The corresponding ranges are defined as follows: N_k is defined by the minimal wavelength resolved (3Δx) and the elliptic truncation criterion (Eq. (1) in ŽGK); N_n is defined by the equivalent modal truncation criterion (Eq. (10) in ŽGK with $c_1 = c_2$). The three wave types are ER waves, westward- (WEIG) and eastward-propagating (EEIG) EIG modes for $n \geq 1$, and

Kelvin modes as well as westward- (WMRG) and eastward-propagating (EMRG) MRG waves for $n = 0$.

Since a common expansion coefficient, Ψ, is used for v, u (the zonal wind field) and h (geopotential height), statistics apply to combinations of these quantities, each representing a single equatorial mode. The inverse projection consists of the inverse Fourier transform in the zonal direction and the projection inverse for the parabolic cylinder functions, details of which are described in ŽGK. Utilizing also the energy norm (Eq. (6) in ŽGK), Ψ_ν is obtained as

$$\Psi_\nu = \sum_x \sum_y \{v(x, y)v_\nu^*(k, y) + u(x, y)u_\nu^*(k, y) + gh_0^{-1}h(x, y)h_\nu^*(k, y)\}e^{ikx}. \quad (2)$$

Here, the asterisk (*) denotes the complex conjugate, and the equivalent depth h_0 represents the depth of the shallow layer of fluid required to give the correct horizontal and time-varying structure of each mode. The mean zonal mode, $k = 0$, is not included in the expansion (1). Furthermore, for reasons of symmetry, only half of the modes need to be computed (for $k \geq 0$).

The independence of different modes, expressed as

$$\overline{\Psi_\nu \Psi_{\nu'}^*} = \delta_k^{k'} \delta_n^{n'} \delta_m^{m'} \overline{\Psi_\nu \Psi_{\nu'}^*}, \quad (3)$$

is an essential assumption leading to a diagonal covariance matrix in spectral space. Applying the energy norm and the homogeneity assumption in the zonal direction, the covariance for mode ν between model levels z and z' is obtained as the following sum:

$$\sum_x \sum_y \left\{ \overline{v(x, y, z)v(x, y, z')} + \overline{u(x, y, z)u(x, y, z')} + gh_0{}^{-1}\overline{h(x, y, z)h(x, y, z')} \right\} = \sum_k \sum_n \sum_m \overline{\Psi_\nu^{r,z}\Psi_\nu^{r,z'} + \Psi_\nu^{i,z}\Psi_\nu^{i,z'}}. \quad (4)$$

The superscript i denotes the imaginary and r the real part of the spectral expansion coefficient Ψ. The overbar represents the ensemble average and the mean is removed from the Ψ series. The variance is obtained for $z = z'$. The modal variance is denoted γ_ν^2 to retain consistency with Eq. (11) in ŽGK.

(c) *The ability of the covariance model to represent ECMWF short-range tropical forecast errors*

Before using the derived covariance model it is essential to investigate the completeness of (1), i.e. how representative the assumed structures are for ECMWF model forecast errors. We define the unexplained variance ratio (Derber and Bouttier 1999) at level z as:

$$\epsilon(x, y, z) = \frac{\sum_{i=1}^{N_{\text{ens}}} \left\{ \overline{v_p(x, y, z)} - v_{p,i}(x, y, z) \right\}^2}{\sum_{i=1}^{N_{\text{ens}}} \left\{ \overline{v(x, y, z)} - v_i(x, y, z) \right\}^2}.$$

Here v is a proxy for forecast errors in the meridional wind, obtained as a difference between two forecasts of the same length. The number of samples extracted from the data-assimilation ensemble, N_{ens}, is 225, and $\overline{v}$ is the mean error at the model grid point (x, y, z). The part of the error field not explained by equatorial wave theory, $v_p(x, y, z)$, is obtained as a difference between the total error (evaluated at the physical grid points)

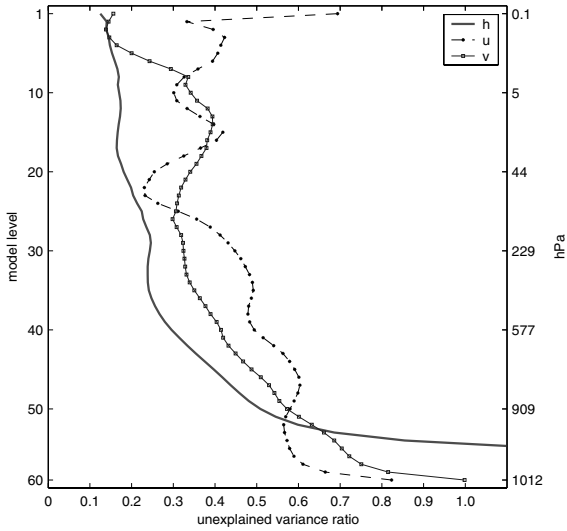

Figure 2. Vertical distribution of the unexplained variance ratio (area mean), representing an error of the equatorial wave approach to modelling the height and wind field errors at model levels. See text for definition. The right y-axis shows average pressure of model levels marked on the left y-axis.

and the inverse of the projection (2), analogous to (1):

$$v_p(x, y, z) = v(x, y, z) - \sum_{\nu=1}^{N_{\text{mode}}} \Psi_\nu v_\nu(k, y) e^{ikx} .$$

N_{mode} is the number of eigenmodes. The unexplained variance ratio, ϵ, measures the ratio between the part of the error variance which is not successfully represented by the selected equatorial waves and the total variance. This ratio is similar to the time-mean analysis error used by Yang *et al.* (2003).

We choose 15 m s^{-1} as a typical value of the phase speed c; then the equivalent depth, related to c by $c = (gh_0)^{1/2}$, is $h_0 = 23$ m, and the trapping scale, i.e. the equatorial radius of deformation, is $a_e = 5°$. Under these conditions, the lowest four meridional modes are completely trapped within the analysed area, i.e. the orthogonality condition is exactly fulfilled for those modes.

Equatorial eigenmodes are either symmetric or anti-symmetric with respect to the equator. The error fields, on the other hand, tend to be localized within the ITCZ (Fig. 1), i.e. in the centre of the northern part of the analysis domain. To resolve most of the important part of the error field in the lower troposphere, use of a linear combination of a number of anti-symmetric meridional modes is thus crucial. Furthermore, one needs a sufficiently large N_n to resolve smaller meridional structures. In the case of the full fields, rather than their differences, a sufficiently large N_n is needed for resolving structures offset from the equator (Yang *et al.* 2003).

The degree of completeness of (1) is evident from the values of ϵ for varying choices of N_n and when different types of waves are included. In Fig. 2 we show ϵ for $N_n = 10$, including all modes consonant with the equivalent meridional truncation criterion (Eq. (10) in ŽGK, with $c_1 = c_2 = 1$). Although this criterion is not important for the value of ϵ, it is nevertheless applied for consistency with the covariance model used in the idealized assimilation tests to be discussed later on. With a zonal truncation $N_k = 119$,

a total of $N_{\text{mode}} = 2937$ eigenmodes are included in the expansion. The values shown in the diagram are area means at each model level.

It can be seen that above the 500 hPa level (level 39), equatorial eigenmodes successfully represent 75–80% of the height- and 50–75% of the wind-field variances. The methodology is most successful at the levels where the errors, as measured by forecast differences, are largest. Below 500 hPa the amount of unexplained variance increases and becomes as large as 60% near 900 hPa (level 50). The unexplained variance is generally largest for the zonal wind component, except near the tropopause and below 900 hPa. At these lowest levels, the methodology fails for all fields, especially that representing the height. Consequently, we shall not consider these levels in the subsequent discussion.

A closer inspection of the spatial patterns of unexplained errors (the v_p, u_p and h_p fields) reveals that most of the tropospheric unexplained variance is located in the ITCZ (not shown). This is the region where the assumed forecast errors do not display the expected symmetry with respect to the equator. Obviously, the number of anti-symmetric modes included in (1) do not suffice to reproduce the error pattern in the ITCZ. We did not, however, try to further increase the value of N_n because of the requirement that eigenmodes remain equatorially trapped.

Figure 2 would assume a different character for other choices of eigenmodes. For example, in the case of $N_n = 3$ the unexplained variance ratio is much larger since three meridional nodes only poorly resolve the structure of the error fields within the ITCZ. The zonal truncation, on the other hand, is of little importance since most of the spectral variance is found below zonal wave number 20. An important part of the explained variance is due to EIG modes; including these has a major effect on the percentage of the explained wind-field variance, an effect larger than can be achieved by increasing N_n. For example, in the case of $N_n = 10$, but without the EIG waves included in (1), the explained variance ratio for wind components is less than 40% in the troposphere, and above the tropopause it sharply decreases, invalidating the approach. On the other hand, the representation of height errors is only slightly affected by a change of N_n, and almost wholly impervious to the choice of modes. This is due to the geopotential gradients in the tropical atmosphere being weak and much more symmetric with respect to the equator than the wind field. Changes in c (as derived from the observed range of h_0) are unimportant for ϵ as compared to the two factors discussed above.

Of the three fields, the height field is best represented by the covariance model, except close to the surface, while the wind-field representation is most successful in the upper troposphere and near the tropopause. Except for the lowest ten model levels, the mean values of ϵ range between 0.3 and 0.4, implying that in what follows we will be concerned with the statistical structure of about 60−70% of the tropical variance.

(*d*) *Meridional variation of errors*

A spectral representation of errors retains full information about the scale variations of the covariances but, as a consequence of the homogeneity assumption, provides no information regarding their spatial variations. In the present case the homogeneity assumption is applied only in the zonal direction, which thus retains the meridional structure in the statistics. The implied error variances in grid-point space can be investigated by applying a randomization method (Fisher 1996) to the spectral mode-based covariance model, followed by a transformation to grid-point space. This approach is based on the fact that the usual 3D-Var formulation in terms of normalized non-dimensional variables (e.g. Ψ_ν/γ_ν) has a background-error covariance matrix equal to the identity matrix. Thus, one can take an ensemble of random vectors for Ψ_ν/γ_ν, drawn from a Gaussian

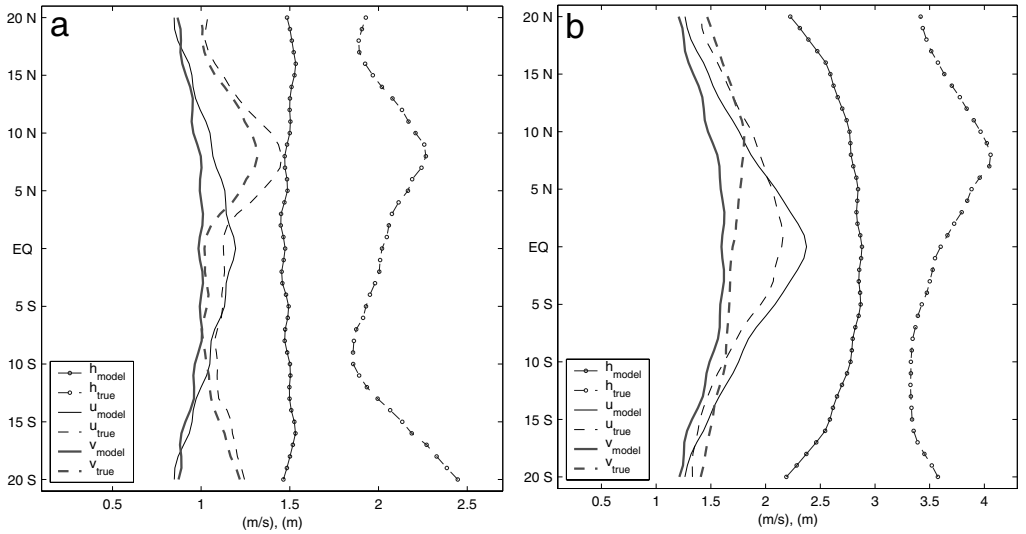

Figure 3. Meridional profiles of 12-hour forecast errors in the zonal wind (thin black lines) and the meridional wind fields (thick grey lines) and the geopotential height (line with ○ symbols). Dashed lines apply to the 'truth' whereas full lines apply to the model (see text for further details). (a) Model level 43 (approximately 654 hPa), (b) model level 27 (approximately 133 hPa).

distribution of zero mean and variance one, and obtain randomization estimates of the error variances first in spectral space, and after further transformation, also in grid-point space. An application of this methodology is provided by Andersson *et al.* (2000).

The resulting randomization variances yield meridional profiles which can be compared with the zonally averaged profiles of the original error fields (generated by the data-assimilation ensemble). The result, presented in Fig. 3 for two model levels, highlights the main shortcomings of our methodology. In the lower troposphere (Fig. 3(a)), the error profiles of the ECMWF model have maxima at 8°N, as illustrated in Fig. 1(a); in addition, errors increase south of 15°S, especially those of the geopotential height. In contrast, the error profiles modelled by equatorial waves display maxima at the equator, in particular for the wind. Error minima are located close to the meridional boundaries. The covariance model yields the best results in the upper troposphere (Fig. 3(b)), where the ECMWF forecast errors are not concentrated along the ITCZ (Fig. 1(b)). In the stratosphere our model retains the error maxima at the equator (not shown) while the 'true' errors of all the fields to a large extent become homogeneous. Equatorially-centred error profiles are caused by Kelvin-wave correlations, to be illustrated in section 4.

Another shortcoming noticeable in Fig. 3 arises from the wind and mass fields being analysed together. Their error coupling is responsible for the insufficiently large amplitudes of the modelled height errors in the troposphere. This interaction also yields unrealistically large amplitudes of the wind-field errors in the stratosphere, where the ECMWF geopotential errors increase.

The spectral mode-based covariance model thus captures a significant part of the error variance that is symmetric around the equator, but far less of the asymmetric part. Note, however, that in a practical implementation of the derived covariance model in a variational data-assimilation scheme, the unexplained part of the tropical error variance would be analysed in a univariate fashion.

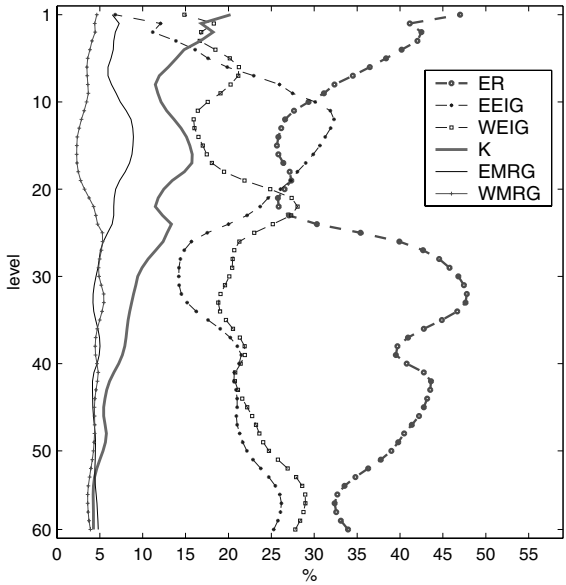

Figure 4. Vertical distribution of the variance among various equatorial eigenmodes: equatorial Rossby (ER) modes, eastward-propagating (EEIG) and westward-propagating (WEIG) equatorial inertio-gravity modes, Kelvin (K) modes, eastward- (EMRG) and westward-propagating (WMRG) mixed Rossby–gravity waves. The x-axis notation applies to the percentage of the total explained variance.

3. VARIANCES AND VERTICAL CORRELATIONS

(*a*) *Variances*

In Fig. 4 we illustrate the tropical covariance model in terms of its variance distribution between the different types of equatorial waves. Among individual modes, the largest part of the variance pertains to Rossby-type motion. Its profile displays a sharp decrease of the variance in the vicinity of the tropopause, with a broad minimum in the range from 5 to 50 hPa (levels 11 to 22). Such a profile shape is in agreement with strong easterlies in the stratosphere during the study period of October 2000; in these circumstances westward-propagating waves do not satisfy the criterion for vertical wave propagation (Charney and Drazin 1961). Within the troposphere, the variance assigned to ER waves is between 40% and 50%, and is largest at levels between 150 and 300 hPa (levels 28 and 34, respectively). Contributions from various meridional modes decrease as n increases. The variance profile for the lowest two modes adheres to the shape of the total profile. For higher modes, stratospheric variance is small. Modes five and higher thus make no contribution to the variance in the stratosphere, whereas they contribute a few percent of the variance in the troposphere (not shown).

Profiles of ER and EIG modes have the opposite variance phase in the troposphere, while EEIG and WEIG modes are 180° out of phase in the stratosphere. Otherwise, the percentage of variance contributed by EEIG and WEIG modes, respectively, is similar in the troposphere, except at levels 26–34 corresponding to the range 100–300 hPa, where WEIG modes play a more important role. In the stratosphere, EEIG waves have an absolute maximum of variance around the 6 hPa level (model level 12). Together, EEIG and WEIG modes constitute 40–50% of the variance in the troposphere. In contrast to ER waves, EIG modes demonstrate the same vertical profile of variance for all n (not shown).

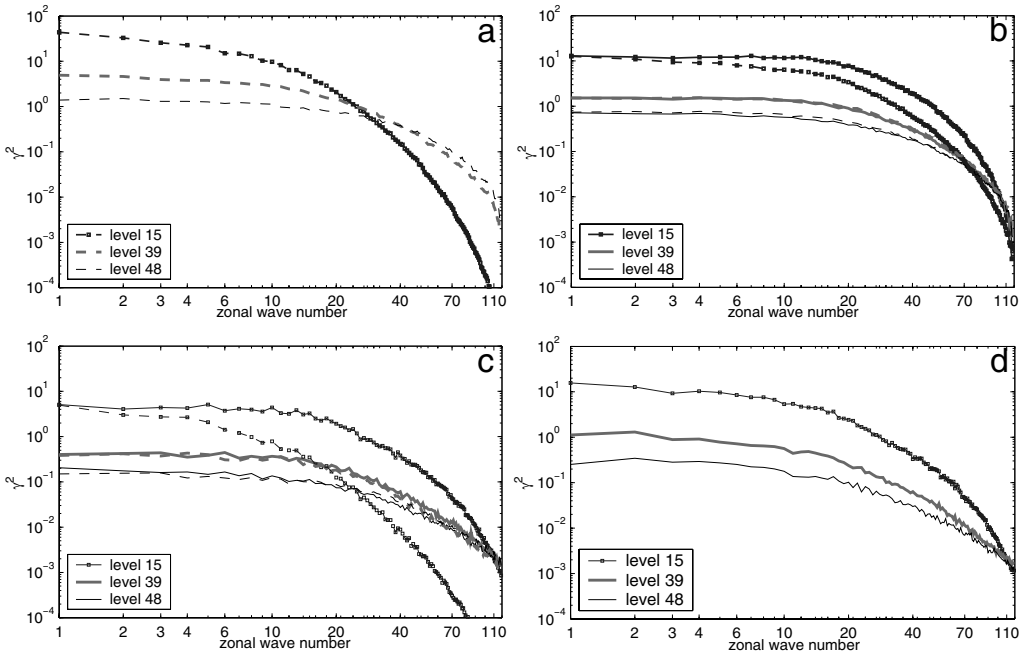

Figure 5. Spectral variance density in (a) ER, (b) EIG, (c) MRG and (d) Kelvin modes at various model levels, as a function of zonal wave number. Full lines denote the eastward-propagating waves, dashed lines the westward-propagating waves.

Other eastward-propagating modes (Kelvin and EMRG waves) have the maximum variance in the stratosphere, similar to EEIG waves. On the other hand, all westward-propagating modes (ER, WMRG and WEIG waves) have minima at these levels. Figure 4 suggests that MRG and Kelvin modes only provide a minor contribution to the variance in the tropical troposphere of the ECMWF model, around 10% and 5–10%, respectively. This, however, does not suggest that they are unimportant for the balanced structure functions, as will be shown in what follows. The variance of Kelvin waves increases steadily with altitude and constitutes 15–20% of the explained stratospheric variance, suggesting that at these altitudes Kelvin waves are more important for the dynamics.

In Fig. 5, the zonal-scale dependence of the spectral variance density is shown at three selected levels representing the stratosphere, the middle troposphere and the lower troposphere, 12 hPa (level 15), 500 hPa (level 39) and 828 hPa (level 48), respectively. In agreement with Fig. 4, the most interesting feature in Fig. 5 is the difference between the eastward- and westward-propagating waves in the stratosphere and the troposphere. At long horizontal scales, the variance is much larger in the stratosphere than in the troposphere, as previously found in global model studies (e.g. Rabier *et al.* 1998). A distinct feature of the present statistics is a decrease of the stratospheric variance for westward-propagating waves at zonal wave numbers larger than 20 ($\sim$2000 km); the variance reduction is less sharp for WEIG modes than for the ER and WMRG modes, although it is still noticeable. This pronounced decrease of the spectral density characterizing the westward-propagating waves in the stratosphere is most likely a feature of this specific dataset and is related to the zonal-wind shear region in the easterly phase of the quasi-biennial oscillation (QBO) during the study period. Since we are

using a sample representative of only one phase of the QBO, the present results may not be representative of the long-term statistics for the tropics. It is likely that the statistics for the dataset from the opposite phase of the QBO would look different, suggesting that significant temporal variations could be present in longer-term time series of such statistics. A more thorough investigation of this issue is, however, beyond the scope of this study.

Among the six different wave types, ER modes dominate the spectra at all scales except in the stratosphere. Here, all eastward-propagating modes have a larger variance than the ER waves for $k > 10$, with Kelvin waves becoming relatively more important at long scales; the latter feature is in agreement with observational and theoretical studies of the role of Kelvin waves in the equatorial stratosphere. The greatest part of the spectral variance in the stratosphere is, however, associated with EIG waves. Their dominance over the Kelvin and MRG modes is in agreement with recent observational and modelling studies (e.g. Dunkerton 1997; Giorgetta *et al.* 2002), suggesting that the gravity waves carry approximately half of the momentum flux required to drive the QBO.

Based on a comparison between 3-, 12- and 24-hour forecasts (figure not shown), the partitioning of variance between modes and scales does not change much with forecast time. The main difference between the 3- and 24-hour forecasts is an increase of the Kelvin- and ER-wave variance in the stratosphere (2–3% and ∼5%, respectively), at the expense of the EIG-wave variance. This quantity reduces by a few percent also in the troposphere, with a corresponding increase of the ER-wave variance. MRG-wave variance remains virtually unchanged. Based on a comparison between the results for the standard case of $h_0 = 23$ m and those obtained when h_0 is taken equal to 50 and 250 m, the sensitivity with respect to c is not significant.

It might be argued that a significant fraction of the variance associated with EIG modes could arise from model deficiencies, especially in the early stages of the forecast. Possibly the model has systematic errors, but those are ignored in strong-constraint 4D-Var. The background-error term should describe the *de facto* statistics of background errors. If these contain more EIG-wave motion than deemed desirable, such model deficiencies could be addressed through a model-error term (i.e. weak-constraint 4D-Var). In this paper, we only attempt to diagnose the equatorial background errors in the current ECMWF analysis, and do not attempt to formulate a background-error term for use in an analysis system. This latter task would first require dealing with the 30–40% of the variance not described by the projection onto equatorial modes.

(*b*) *Vertical correlations*

Vertical correlations are calculated for all model levels, but we disregard levels 50–60 (900–1012 hPa) due to the small percentage of variance here which can be explained using the equatorial wave approach. The shape of correlation depends on the wave type as well as the altitude (Figs. 6–8). Moving upwards and towards smaller scales generally gives rise to more narrow correlations. The narrowing of vertical correlations for smaller scales is seen mainly for the ER and Kelvin modes, to a lesser extent for the MRG modes and almost not at all for the EIG waves.

The broadening at large scales of the ER-wave correlations (Fig. 6) begins approximately at zonal wave number 20, and is larger at $n = 1$ (Fig. 6(a)) than for the higher modes (Fig. 6(b)). However, comparing ER correlations for the same n it may be noted that the narrowing of the correlations with altitude primarily applies to the model-level space. If the vertical coordinate is height, the shapes of the correlations for small zonal wave numbers would be more-or-less similar at various altitudes. Such narrowing

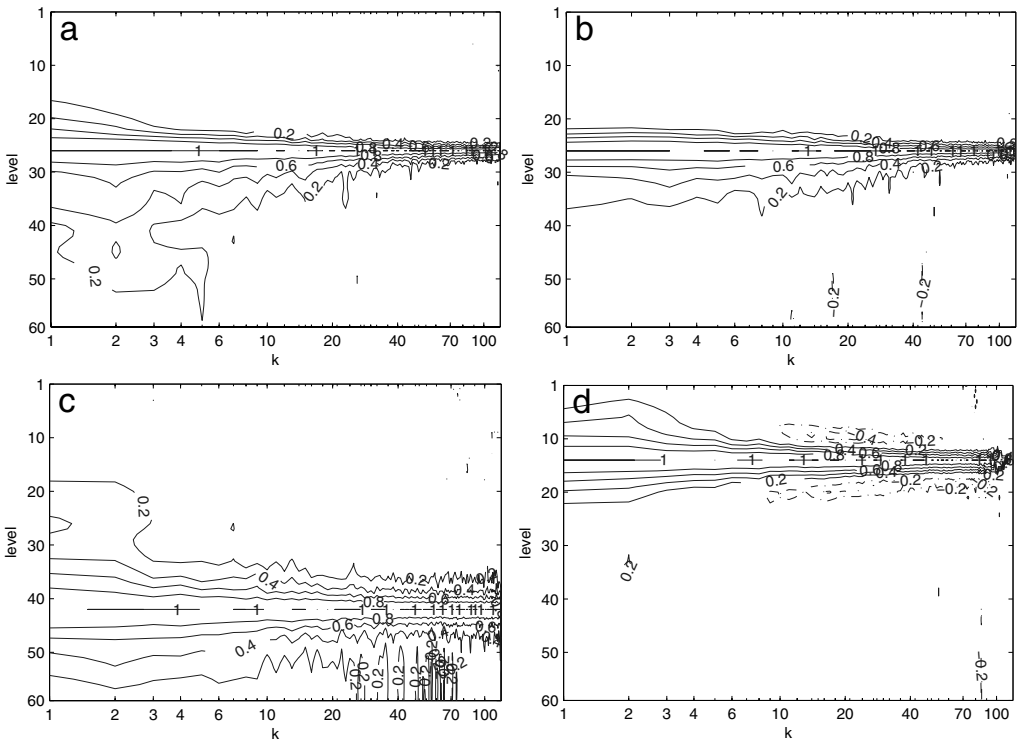

Figure 6. Vertical correlations for ER modes, as a function of zonal wave number k. (a) $n = 1$ and (b) $n = 2$ at model level 27 (approximately 133 hPa); (c) $n = 1$ at model level 43 (approximately 654 hPa); (d) $n = 1$ at model level 15 (approximately 12 hPa). Isolines every 0.2 with zero isoline omitted.

in model space could indicate an insufficient vertical resolution. Thus, Lindzen and Fox-Rabinovitz (1989) demonstrated that the consistency requirement for the horizontal and vertical resolution demands a higher vertical resolution in the tropics than at mid-latitudes.

Vertical correlations in the troposphere are not symmetric; throughout this region correlations with levels below are much stronger than those with levels above (Figs. 6(a)–(c)). The underlying physical mechanism is, most likely, convection. This hypothesis is given added weight by the fact that the stratospheric correlations do not display such an asymmetry.

A prominent feature of stratospheric correlations is the appearance of 'wings' of negative correlations for wave numbers above $k = 10$ (Fig. 6(d)). This feature characterizes all modes in the stratosphere, although it is strongest for the EIG modes. In the case of EEIG modes, 'negative wings' are found for all horizontal scales (Fig. 7(a)). This may be a consequence of the highly stable stratification typical of the stratosphere. Vertical perturbations tend to be of limited extent, and generally integrate to zero over a relatively shallow layer. Lower in the troposphere, EIG-wave correlations show negligible change with scale and have a symmetric shape (Fig. 7(b)).

MRG- and Kelvin-mode correlations display intermediate characteristics, i.e. the correlations are more narrow towards larger wave numbers and are somewhat asymmetric in the troposphere (Fig. 8). Kelvin waves display characteristics more similar to those of the ER modes than do the MRG waves.

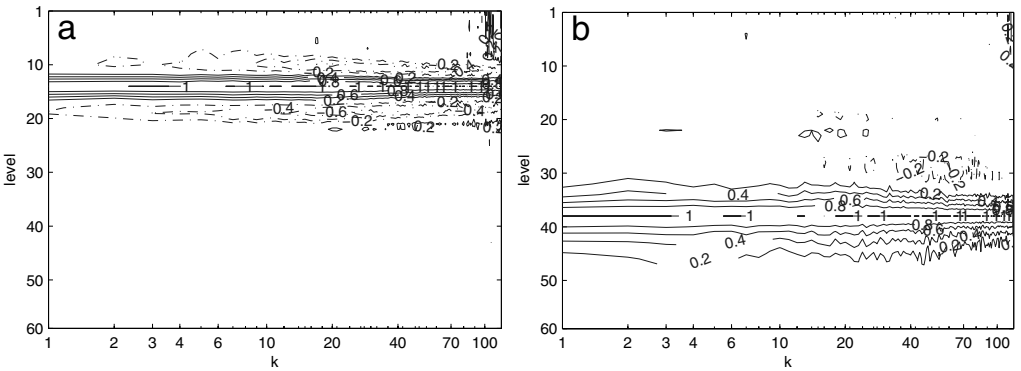

Figure 7. Vertical correlations, as a function of zonal wave number k, with (a) model level 15 for $n = 1$ EEIG modes and (b) with model level 39 (approximately 500 hPa) for $n = 1$ WEIG modes.

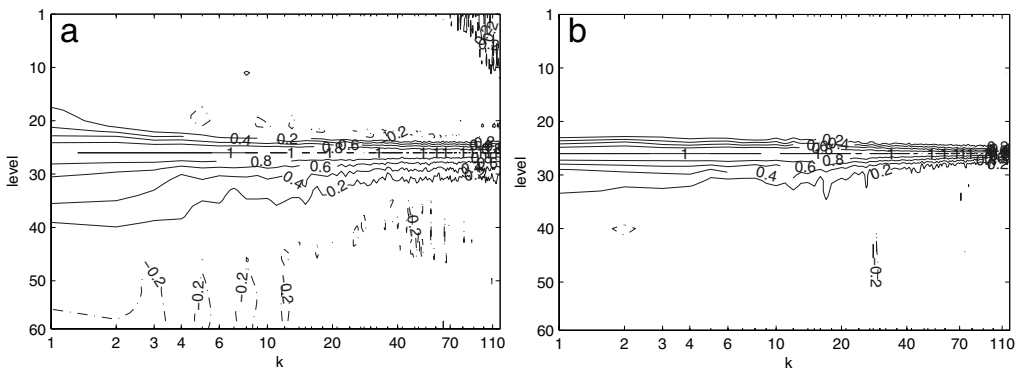

Figure 8. Vertical correlations, as a function of zonal wave number k, with model level 27 for (a) Kelvin modes and (b) WMRG modes.

4. HORIZONTAL STRUCTURES

(*a*) *Single-observation experiments*

In this section we use the assimilation system developed in ŽGK, including the mode-based error-covariance model described above, to carry out single-observation experiments at various levels. First, we compare the results with those obtained using the analytical covariance model in ŽGK. In particular, the earlier study suggested that the presence of Kelvin and MRG modes is critical for the mass–wind coupling near the equator. We ask whether this is also a prevailing feature for a state-of-the-art NWP system. Secondly, we study the variation of horizontal scales with height. Analysing height and wind fields together precludes a direct evaluation of their respective horizontal correlations; carrying out single-observation experiments serves as an alternative method of comparing their scale variations. In Figs. 9–11, examples of single-observation analyses are shown for three vertical levels representing the lower troposphere, the middle/upper troposphere and the stratosphere.

Figure 9 shows analysis increments due to a single height observation. A striking feature is that the horizontal scale does not increase significantly with altitude. This is an intriguing result, in apparent disagreement with earlier theoretical studies (Phillips 1986;

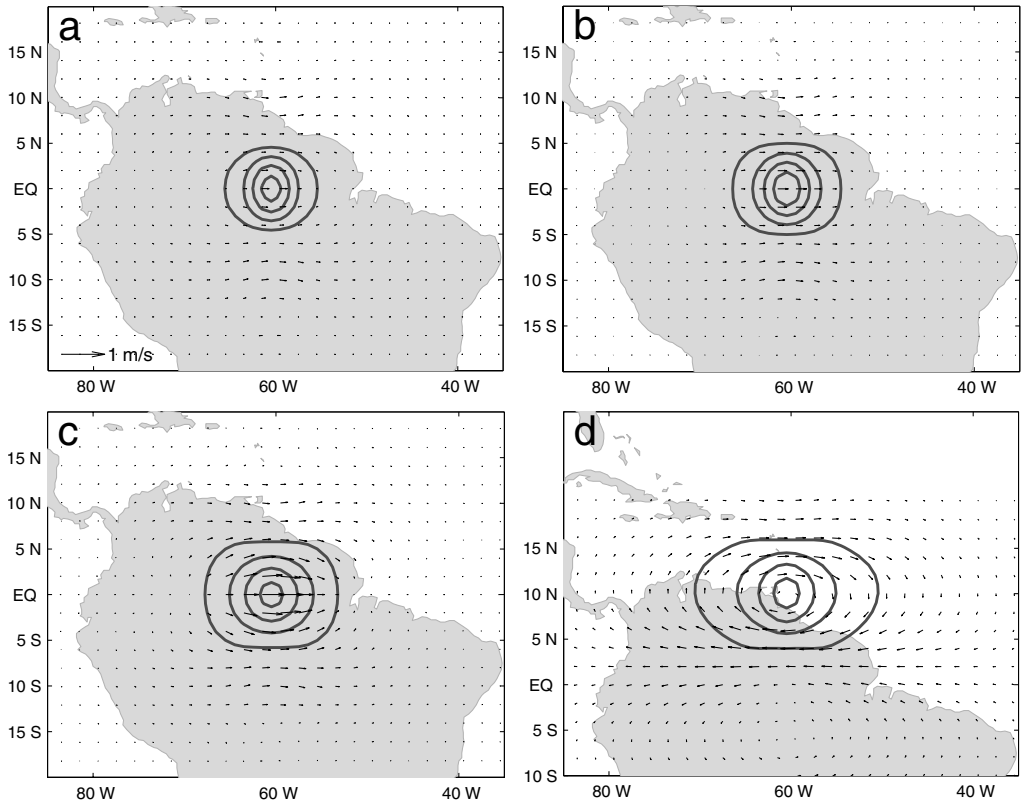

Figure 9. Horizontal structure of the analysis increments due to a single height observation located (a–c) at the equator and (d) at 10°N. The background error variance corresponds to the ECMWF model levels (a) 43 (approximately 654 hPa), (b) 39 (approximately 500 hPa), (c) 15 (approximately 12 hPa), and (d) 31 (approximately 229 hPa). The height observation is 4 m higher than the background, corresponding to a motionless fluid of depth h_0. In all cases the observation error is taken equal to the background field error at the same point, as obtained from the randomization experiment. Isoline spacing is ± 0.5 m and the zero contour is omitted. Heavy lines correspond to positive values, while thin lines are used for negative values. Wind vectors are shown in every second grid point in both horizontal directions.

Bartello and Mitchell 1992) and studies of global statistics (e.g. Rabier *et al.* 1998; Ingleby 2001). These studies instead showed an increase of horizontal scales with altitude, particularly for the height field, and the broadest correlations were found in the tropics. What makes these studies less suitable for comparative purposes is their common assumption that forecast errors are represented by the ensemble of quasi-geostrophic normal modes (Phillips 1986). Moreover, NWP studies utilize the NMC (US National Meteorological Center) methodology (Parrish and Derber 1992), the disadvantages of which are most pronounced in the tropics. The NMC method assumes that the forecast-error covariances can be represented by the differences between forecasts of different range valid at the same instant of time. Due to the length of forecasts, the method tends to generate covariances which are broader in both horizontal and vertical direction than those of the background errors. Furthermore, in data-sparse region such as the tropics, two forecasts may be very similar, resulting in an underestimate of the variances. Finally, the NMC technique is prone to breaking down in the tropics, due to the small temporal variability of the tropical large-scale flow (Rabier *et al.* 1998). Based on a study using

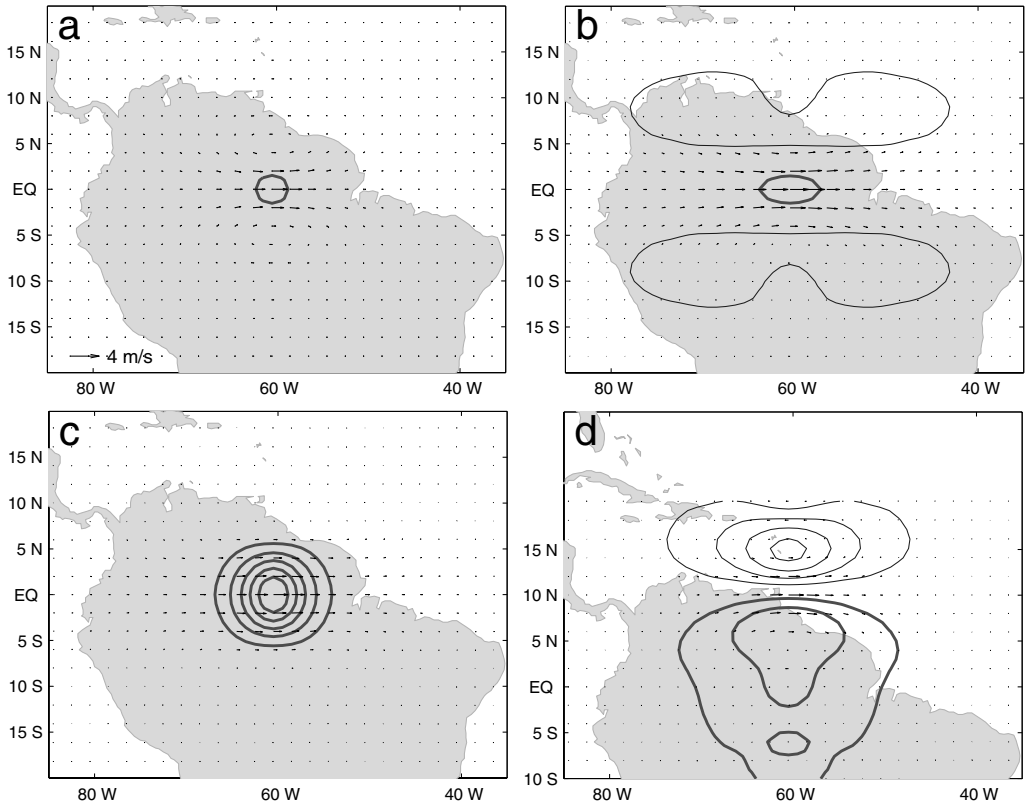

Figure 10. As Fig. 9, but for a single westerly wind observation at model levels (a) 43, (b) 31, (c) 15, and (d) 31. The wind observations are 4 m s^{-1}. Isoline spacing for the height field increments is ± 0.2 m.

the NMC method it has also been noted that 'any latitudinal and seasonal variations in the extra-tropics are much smaller than the differences from the tropics' (Ingleby 2001).

Global error statistics are dominated by the midlatitude errors in the Rossby modes at synoptic scales. Modelling tropical forecast errors using global statistics or in terms of quasi-geostrophic modes is not wholly appropriate. Convective processes, not described by mass–wind relationships such as the linear-balance equation, could lead to localized tropical correlations such as those shown in Fig. 9. Convection is a generator of equatorial wave disturbances, which serve as the basis of our methodology. Therefore, our results for the horizontal structure of the height-field increments most likely reflect that part of the tropical height-field variability which is coupled to convection. We can see that the tropical horizontal correlations are strongly influenced by the Kelvin and MRG waves (which do not contribute in the midlatitudes) and by EIG waves (with a significantly larger relative contribution than that due to gravity waves in the mid-latitudes).

Another feature seen in Fig. 9 is that the mass–wind coupling, although weak, increases with altitude. Wind-field increments in the lower troposphere are, for practical purposes, negligible. In the stratosphere, the magnitude of the wind increment is increased by a factor three at the observation point (Fig. 9(c)). Even so, the kinetic energy of the increment field is only around 17% of the total energy (being around 6% at a lower-tropospheric level in Fig. 9(a)).

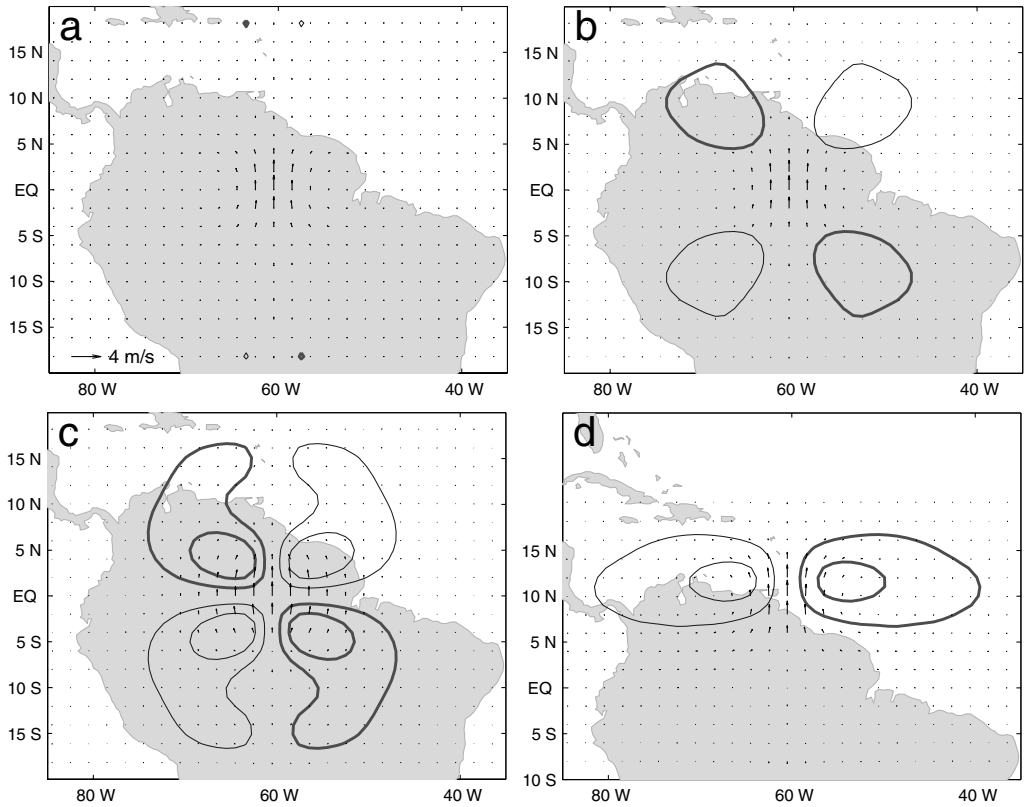

Figure 11. As Fig. 9, but for a single northerly wind observation. The wind observations are 4 m s^{-1}. Isoline spacing is ± 0.05 m in (a) and (b), and ± 0.2 m in (c) and (d).

The mass–wind balance is suggestive of a Kelvin wave, although the fraction of the total error variance associated with Kelvin waves is comparatively small. The reason why the Kelvin-wave coupling is so effective near the equator is that this wave type is the only mode with a significant amplitude in height at the equator coinciding with the strongest zonal wind. Away from the equator the importance of Kelvin waves diminishes, at least in the troposphere, and the ER type of mass–wind balance becomes more pronounced (Fig. 9(d)). It can also be noticed in Fig. 9(d) that the subtropical increments are zonally elongated for the height as well as the wind fields. Stratospheric increments show a stronger influence of the Kelvin wave, as its contribution to the variance is three times larger here than in the troposphere.

Contrary to what holds true for the height data, assimilation of wind observations results in increments with horizontal scales increasing with altitude (Figs. 10–11). In addition, the equatorial wave balance gives rise to a coupling with the height field, especially in the stratosphere. A zonal wind observation at the equator is associated with a positive height-field increment through most of the troposphere; the coupling is weakest in the upper troposphere around model level 31 (229 hPa) (Fig. 10(b)). In the stratosphere, both height and wind increments have larger scales. The latter are zonally elongated, in contrast to the balanced height-field increments (Fig. 10(c)). The maximal magnitude of the balanced height-field increment at model level 15 (12 hPa)

is approximately 3.6 times larger than at model level 43 (654 hPa); the corresponding potential-energy contribution is increased from 2% at level 43 to 7% at level 15. The balance at the equator is dominated by Kelvin waves, but at a distance from the equator, height-field increments become geostrophic and zonally elongated (Fig. 10(d)), in full agreement with quasi-geostrophic theory. By comparing Fig. 10(d) with Fig. 10(b) it can also be seen that equator-centred modes considerably reduce the magnitude of the balanced height field.

The increase of the amplitude of the balanced height increments with altitude is largest for a meridional wind observation (Fig. 11). When this is centred on the equator (Figs. 11(a)–(c)), the shape of the increments resembles an MRG wave. With the observation located at 10°N, the increments appear nearly geostrophic (Fig. 11(d)) and very similar to those found in the midlatitude case (e.g. Courtier *et al.* 1998), except for a broader zonal scale. A significant enlargement of the horizontal scale of a height-field increment with altitude is visible (note that the isoline spacing in Figs. 11(c)–(d) is four times larger than in Figs. 11(a)–(b)). The potential-energy content increases from zero at level 43 (654 hPa) to around 3% of the total energy at level 15 (12 hPa).

(*b*) *Sensitivity experiments*

In ŽGK it was shown that Kelvin waves largely determine the character of the height–zonal-wind coupling at the equator, while MRG waves play a similar role for the height–meridional-wind coupling. This behaviour is now confirmed from the ECMWF variance spectra. In the experiments illustrated in Fig. 12, the Kelvin waves are removed from the spectrum in such a way that their variance for each zonal wave number is distributed equally between other modes. The resulting increments at model level 39 (500 hPa) (Fig. 12(a), to be compared with Fig. 9(b)) are still centred on the equator, but the balanced winds have changed sign. The new configuration comprises 'wings' of negative correlations south and north of the equator and the whole structure is most reminiscent of an $n = 1$ WEIG wave. The EEIG waves, to a similar degree present in the variance spectrum, would be associated with the opposite height–zonal-wind balance at the equator. However, correlations due to ER modes have the same effects as those due to WEIG modes. At other levels, the structures appear quite similar.

Kelvin and MRG modes represent trapped motion centred on the equator; that they have an impact on the mass–wind coupling has previously been suggested also by other authors (Parrish 1988; Daley 1993, 1996). On the other hand, it has been more difficult to estimate the impact of EIG waves. Their meridional structure is more complex and their role in the large-scale tropical dynamics has been less evident, although some observational evidence has been provided (e.g. Wheeler and Kiladis 1999; Wada *et al.* 1999; Clayson *et al.* 2002) and theoretical scaling arguments furthermore confirm their physical relevance (Browning *et al.* 2000).

Figure 12(b) indicates that EIG waves exert an important influence on the tropical background-error correlations. In this experiment, EIG variance for each k and n has been removed and assigned to the corresponding ER mode. In the case of a height-field observation, the meridional scale of the increments is increased by a factor of 2–3 at all altitudes. Balanced zonal winds are stronger and show maxima displaced from the equator. The prevailing structure is still a Kelvin wave. When both Kelvin and EIG waves are removed from the spectrum, the analysis increments become rotational (Fig. 12(c)). The amplitude of the height-field increments is reduced by a factor 3 and the increments are displaced from the equator. Along the equator, strong (see Fig. 9(b)) easterly winds are produced.

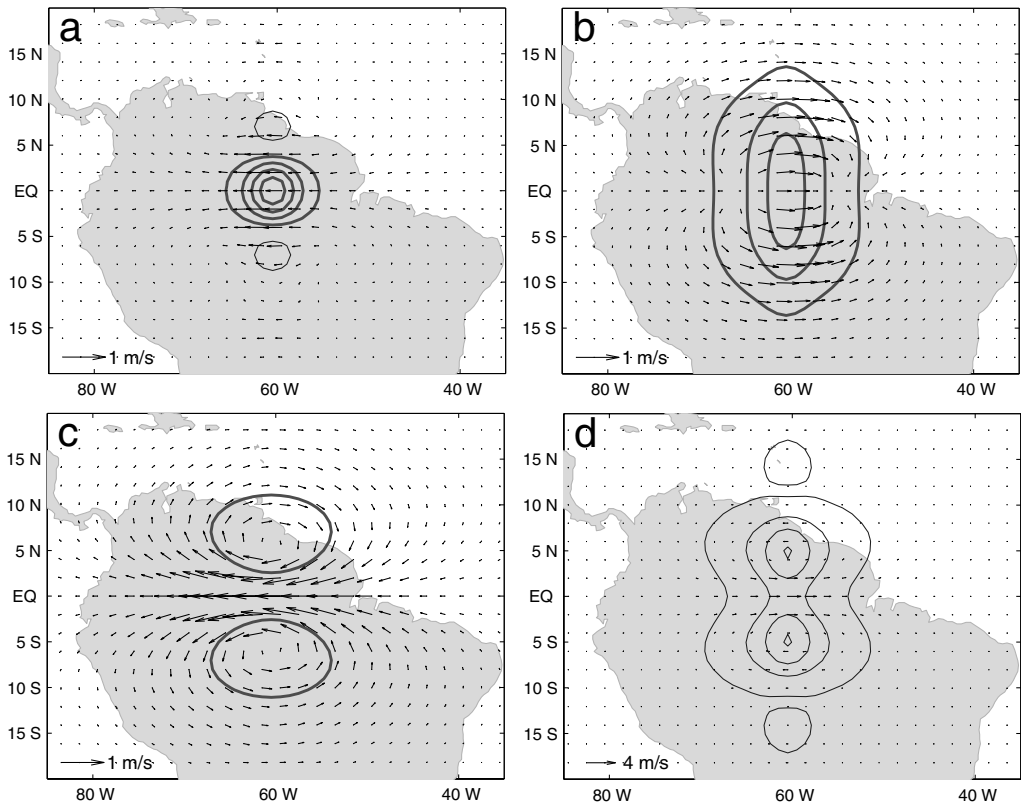

Figure 12. As Fig. 9, but for (a–c) a height observation and (d) a zonal wind observation at model level 39. (a, d) without Kelvin wave variance, (b) without EIG wave variance, (c) without Kelvin and EIG wave variance. Isoline spacing is ±0.5 m in (a–c), and ±0.2 m in (d).

Without the Kelvin-wave balance in the background-error term, increments due to a zonal wind observation (Fig. 12(d)) have a structure of an n=1 ER wave throughout the atmosphere. The horizontal scale of balanced height-field increments increases upwards more than it does in Fig. 10. To summarize, the impact of the Kelvin waves is to change the structure and to reduce the horizontal correlation scales for balanced analysis increments near the equator. The effect of the EIG modes is similar to that of the Kelvin modes, but weaker.

Removal of the MRG-wave variance from the spectrum does not produce dramatic changes similar to those for the Kelvin-mode variance for zonal winds. The balanced height increments have longer correlation scales, especially in the stratosphere. The wind-field structure is less suggestive of an MRG wave due to a counterflow generated at a distance from the equator (Fig. 13(a), to be compared with Fig. 11(b)). However, EIG modes have a major impact by reducing the meridional correlation scales and the magnitudes of the balanced height-field increments (Fig. 13(b)).

In view of these pronounced sensitivities, it is not surprising that progress with balanced tropical data assimilation has been slow, and that the most adequate treatment of this issue in global NWP systems so far has been to regard it univariately!

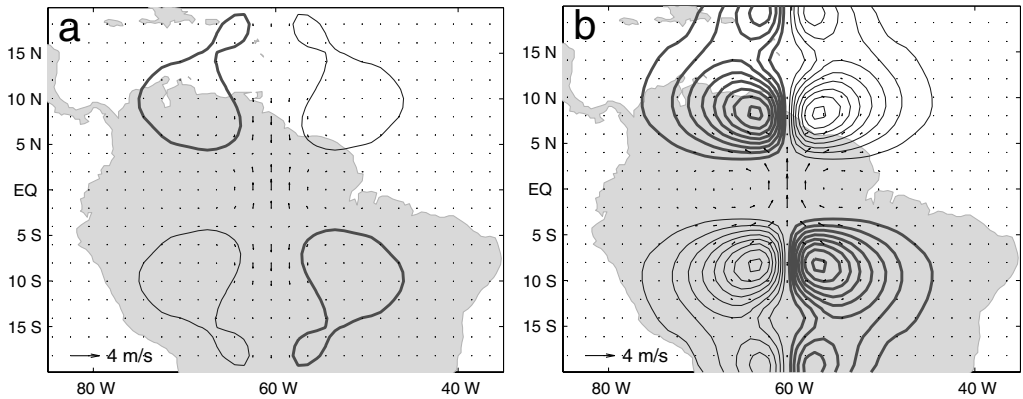

Figure 13. As Fig. 11, but for model level 39. (a) Without MRG wave variance, (b) without EIG wave variance. Isoline spacing is ±0.05 m.

5. CONCLUSIONS

The theory for linear equatorial wave motion coupled to deep convection has been successfully applied in a study of ECMWF short-range tropical forecast errors simulated using an ensemble of data assimilations. It was found that the equatorial waves (Equatorial Rossby, Equatorial Inertio-Gravity, Kelvin and Mixed Rossby–Gravity waves) explain a substantial fraction (60–70%) of the tropical large-scale variability, which has consequences for tropical data assimilation. Using a mode-based background-error covariance model, idealized single-observation analysis experiments were undertaken, illustrating the main features of balanced analysis increments at various horizontal levels in the tropics.

Among the various tropical modes, the highest percentage of variance pertains to ER waves (30–40%) and to EIG waves (20–30% to each westward- and eastward-propagating mode). The vertical correlations for ER waves display characteristics similar to those of their extratropical counterparts: correlations narrow in the stratosphere as well as towards shorter scales. On the contrary, EIG-mode correlations do not display vertical correlation broadening with increased zonal scales. Kelvin and MRG modes explain a relatively small fraction of the variance in the troposphere, but the Kelvin-wave variance increases significantly with altitude and constitutes 15–20% of the stratospheric variance.

In spite of its relatively small variance contribution in the troposphere, the Kelvin-wave coupling plays a decisive role for determining the characteristics of the horizontal correlations near the equator. (This was earlier suggested on the basis of idealized experiments in ŽGK and has been confirmed in the present study using statistics from an operational NWP system.) Another important feature is that EIG modes affect the mass–wind coupling in a similar way as do the Kelvin modes. In particular, EIG modes have a major impact by reducing the meridional correlation scales and the magnitudes of the balanced height-field increments.

It was found that the westward-propagating waves are much less pronounced in the stratosphere, which is most likely related to the phase of the quasi-biennial oscillation (QBO) during the study period. As the QBO is the dominant dynamical feature of the equatorial lower and middle stratosphere, its impact on stratospheric correlations, as well as the fact that the largest part of tropospheric errors is associated

with the intertropical convergence zone (ITCZ), suggest that the background-error statistics should vary with time. The high concentration of errors along the ITCZ, i.e. an asymmetric distribution with respect to the equator, has been found to be the main impediment for a methodology based on equatorial wave theory.

Application of quasi-geostrophic theory in the tropics results in zonally elongated horizontal structures, since according to this theory the horizontal scales vary in accordance with the Rossby deformation radius. However, quasi-geostrophic considerations do not take into account the possible influence exerted by non-rotational tropical modes on the horizontal correlations near the equator. The covariance model derived in the present study retains this influence, as illustrated by the single-observations assimilation results. In particular, it was found that the increase with altitude of the horizontal-correlation scale for height is less pronounced than that which would result from the quasi-geostrophic approach or from globally-averaged error statistics. It was also found that the correlations are vertically asymmetric. Both features are linked to deep tropical convection acting to generate equatorial wave motion.

The background-error covariance model in the ECMWF 4D-Var scheme is effectively univariate near the equator, similar to many other global data-assimilation schemes. An appropriate projection of tropical analysis increments onto equatorial modes is therefore not ensured. Excessive large-scale divergence in the increments may be a contributing factor for the overestimated Hadley circulation observed in ECMWF operational analysis and in the ECMWF 40-year re-analysis (ERA-40) (Uppala 2001). Too large analysed amounts of gravity waves in the lower stratosphere/upper troposphere may also contribute to the overestimate of the Brewer–Dobson circulation also observed in the operational analyses as well as in ERA-40. An approach that combines tropical and midlatitude dynamics and produces balanced increments globally is thus required for further progress.

To summarize, a diagnostic study of the tropical background errors in the current ECMWF analysis has revealed a number of distinct features compared to global error statistics. This work can be regarded as one of several steps towards improving the background-error term used in a global analysis system. Remaining tasks include dealing with the 30–40% of the variance not described by the method, analysis of the vertical eigenvectors, the study of longer-term variations and not least the mathematical challenges involved in formulating a global three-dimensional covariance model which includes these equatorially trapped modes.

ACKNOWLEDGEMENTS

The authors would like to thank Erland Källén and Nils Gustafsson (MISU) for advice and stimulating discussions during the course of this work, and to them, Adrian Simmons (ECMWF) and Peter Lundberg (MISU) for reading the paper and their relevant comments. We also thank Agathe Untch (ECMWF) for discussions of the results, and her and Joan Alexander (CORA) for commenting on the role of the QBO.

APPENDIX

A data-assimilation ensemble method for generation of background error samples

Let us consider the analysis system as a 'black box', which delivers an analysis $\mathbf{x}^{\mathrm{a}}$ given a background $\mathbf{x}^{\mathrm{b}}$ and a set of observations represented by the vector $\mathbf{y}$:

$$\mathbf{x}^{\mathrm{a}} = \mathbf{f}(\mathbf{x}^{\mathrm{b}}, \mathbf{y}). \tag{A.1}$$

Note that, with the exception of the background, $\mathbf{y}$ includes all the inputs to the analysis system. In particular, it may include quantities such as sea surface temperature.

We will also consider the forecast for time T with initial conditions provided by the analysis:

$$\mathbf{x}^{\mathrm{f}}(T) = \mathcal{M}_T(\mathbf{x}^{\mathrm{a}}). \tag{A.2}$$

The analysis, background and observation errors are

$$\epsilon^{\mathrm{a}} = \mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{t}} \tag{A.3}$$

$$\epsilon^{\mathrm{b}} = \mathbf{x}^{\mathrm{b}} - \mathbf{x}^{\mathrm{t}} \tag{A.4}$$

$$\epsilon^{\mathrm{o}} = \mathbf{y} - \mathbf{y}^{\mathrm{t}} \tag{A.5}$$

where $\mathbf{x}^{\mathrm{t}}$ is the discretization of the true state, and where $\mathbf{y}^{\mathrm{t}}$ is the vector of true values of the observed quantities. The forecast error is

$$\epsilon^{\mathrm{f}}(T) = \mathbf{x}^{\mathrm{f}}(T) - \mathbf{x}^{\mathrm{t}}(T), \tag{A.6}$$

where $\mathbf{x}^{\mathrm{t}}(T)$ is the true state at time T.

We will denote by ϵ^{s} the error of an analysis made with perfect background and observations, and by $\epsilon^{\mathrm{m}}(T)$ the error of a forecast made with perfect initial conditions:

$$\epsilon^{\mathrm{s}} = \mathbf{f}(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}}) - \mathbf{x}^{\mathrm{t}} \tag{A.7}$$

$$\epsilon^{\mathrm{m}}(T) = \mathcal{M}_T(\mathbf{x}^{\mathrm{t}}) - \mathbf{x}^{\mathrm{t}}(T). \tag{A.8}$$

Assuming that $\mathbf{f}$ is differentiable at $(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})$, the analysis error is given by a Taylor expansion of $\mathbf{f}$ about $(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})$:

$$\epsilon^{\mathrm{a}} = \frac{\partial \mathbf{f}(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})}{\partial \mathbf{x}^{\mathrm{b}}}\epsilon^{\mathrm{b}} + \frac{\partial \mathbf{f}(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})}{\partial \mathbf{y}}\epsilon^{\mathrm{o}} + \epsilon^{\mathrm{s}} + O(\epsilon^2), \tag{A.9}$$

where $\partial\mathbf{f}/\partial\mathbf{x}^{\mathrm{b}}$ and $\partial\mathbf{f}/\partial\mathbf{y}$ denote the Jacobian matrices of partial derivatives of $\mathbf{f}$ with respect to the elements of $\mathbf{x}^{\mathrm{b}}$ and $\mathbf{y}$. Similarly, if $\mathcal{M}_T$ is differentiable at $\mathbf{x}^{\mathrm{t}}$, then the forecast error is

$$\epsilon^{\mathrm{f}}(T) = \frac{\partial \mathcal{M}_T(\mathbf{x}^{\mathrm{t}})}{\partial \mathbf{x}}\epsilon^{\mathrm{a}} + \epsilon^{\mathrm{m}}(T) + O(\epsilon^2). \tag{A.10}$$

Now consider an analysis made by adding perturbations ζ and η to $\mathbf{x}^{\mathrm{b}}$ and $\mathbf{y}$, with the result further perturbed by the addition of ω:

$$\widehat{\mathbf{x}^{\mathrm{a}}} = \mathbf{f}(\mathbf{x}^{\mathrm{b}} + \zeta, \mathbf{y} + \eta) + \omega. \tag{A.11}$$

The error in $\widehat{\mathbf{x}^{\mathrm{a}}}$ can be expressed as a Taylor expansion about $(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})$:

$$\widehat{\epsilon^{\mathrm{a}}} = \frac{\partial \mathbf{f}(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})}{\partial \mathbf{x}^{\mathrm{b}}}(\epsilon^{\mathrm{b}} + \zeta) + \frac{\partial \mathbf{f}(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})}{\partial \mathbf{y}}(\epsilon^{\mathrm{o}} + \eta) + (\epsilon^{\mathrm{s}} + \omega) + O(\epsilon^2). \tag{A.12}$$

If we now consider the difference $\delta\widehat{\mathbf{x}^{\mathrm{a}}}$ between two perturbed analyses, made with different perturbations, we have:

$$\delta\widehat{\mathbf{x}^{\mathrm{a}}} = \frac{\partial \mathbf{f}(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})}{\partial \mathbf{x}^{\mathrm{b}}}\delta\zeta + \frac{\partial \mathbf{f}(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})}{\partial \mathbf{y}}\delta\eta + \delta\omega + O(\epsilon^2), \tag{A.13}$$

where $\delta\zeta$, $\delta\eta$ and $\delta\omega$ are the differences between the corresponding perturbations for the two analyses.

Suppose now that the perturbations are chosen so that the covariance matrix for the vector $(\delta\zeta, \delta\eta, \delta\omega)^{\mathrm{T}}$ is twice the corresponding covariance matrix for $(\epsilon^{\mathrm{b}}, \epsilon^{\mathrm{o}}, \epsilon^{\mathrm{s}})^{\mathrm{T}}$. This may be achieved by perturbing each analysis with different random errors drawn from the distribution of $(\epsilon^{\mathrm{b}}, \epsilon^{\mathrm{o}}, \epsilon^{\mathrm{s}})^{\mathrm{T}}$. Comparing (A.9) and (A.13), we see that, to first order in ϵ, the covariance matrix for $\delta\widehat{\mathbf{x}^{\mathrm{a}}}$ is twice the covariance matrix for the analysis error ϵ^{a} of the unperturbed analysis.

Next consider a forecast with initial condition $\widehat{\mathbf{x}^{\mathrm{a}}}$, and with an additional perturbation, ξ:

$$\widehat{\mathbf{x}^{\mathrm{f}}} = \mathcal{M}_T(\widehat{\mathbf{x}^{\mathrm{a}}}) + \xi. \tag{A.14}$$

The error of this forecast is:

$$\widehat{\epsilon^{\mathrm{f}}}(T) = \frac{\partial \mathcal{M}_T(\mathbf{x}^{\mathrm{t}})}{\partial \mathbf{x}} \widehat{\epsilon^{\mathrm{a}}} + \xi + \epsilon^{\mathrm{m}}(T) + O(\epsilon^2). \tag{A.15}$$

By an argument similar to that given above, the difference between two forecasts made from differently perturbed analyses is:

$$\delta\widehat{\mathbf{x}^{\mathrm{f}}}(T) = \frac{\partial \mathcal{M}_T(\mathbf{x}^{\mathrm{t}})}{\partial \mathbf{x}} \delta\widehat{\mathbf{x}^{\mathrm{a}}} + \delta\xi + O(\epsilon^2). \tag{A.16}$$

Comparing (A.10) and (A.16), we see that if the vector $(\delta\widehat{\mathbf{x}^{\mathrm{a}}}, \delta\xi)^{\mathrm{T}}$ has a covariance matrix equal to twice that of $(\epsilon^{\mathrm{a}}, \epsilon^{\mathrm{m}}(T))^{\mathrm{T}}$, then to first order, $\delta\widehat{\mathbf{x}^{\mathrm{f}}}(T)$ will have a covariance matrix equal to twice the covariance matrix of an unperturbed forecast with initial conditions provided by an unperturbed analysis.

Of particular interest for the purposes of this paper are forecasts which provide the background for the next cycle of analysis. In this case, differences between forecasts have precisely the covariance matrix required for $\delta\zeta$ at the next analysis cycle. So, given an *initial* pair of background perturbations for which $\delta\zeta$ has the required covariance matrix, together with a sequence of perturbations $(\delta\eta, \delta\omega, \delta\xi)$, a sequence of perturbed analyses and forecasts may be generated such that the difference between any pair of contemporaneous analyses or forecasts will have a covariance matrix equal to twice the covariance matrix of errors in the corresponding unperturbed analysis or forecast.

The derivation presented above is based on truncated Taylor expansions, and consequently assumes that the evolution of errors in the analysis–forecast system is weakly nonlinear. However, it is worth noting that Evensen (1997) has suggested that the closely related Ensemble Kalman filter method is suitable for generating samples of backgrounds and analyses for a system whose dynamics are strongly nonlinear.

Note that, in addition to perturbations representative of observation and background error, the method also requires perturbations which are representative of model error and of the residual analysis error which would arise for an analysis given a perfect background and perfect observations. These perturbations should have covariance matrices equal to those of the corresponding actual error. In practice, these matrices are very poorly known. For the ensemble of analyses presented in this paper, the perturbations $\delta\omega$ and $\delta\zeta$ were set to zero.

In the ECMWF analysis system, the covariance matrix of observation error is approximated by a diagonal matrix. The perturbations $\delta\eta$ were drawn from a Gaussian distribution with this approximate observation error covariance matrix, whereas in principle they should be drawn from the true distribution of observation error.

The sea surface temperature used in the ECMWF analysis system is taken from an independent analysis (Thiebaux *et al.* 2001). The sea surface temperatures used in

the ensemble were perturbed using an estimate of the random error in the sea-surface-temperature analysis (Vialard *et al.* 2005).

In principle, perturbations to the backgrounds are required for the first cycle of analysis. In practice, the analysis on a given day is effectively independent of the background fields used several days earlier. The statistics presented in this paper were generated by discarding the first six days of an ensemble whose initial background fields were not perturbed.

The ensemble consisted of ten independent runs of the ECMWF 4D-Var analysis/forecast system for the period 1–31 October 2000. In most respects, the analysis/forecast system resembled the system that became operational at ECMWF in June 2001. However, the resolution was T319, which is lower than that used operationally. Furthermore, a finite element vertical formulation was used (Untch and Hortal 2003), additional ozone data was assimilated, and background-error variances for stratospheric humidity were increased to realistic values.

References

Andersson, E., Haseler, J., Undén, P., Courtier, P., Kelly, G., Vasiljević, D., Branković, Č., Cardinali, C., Gaffard, C., Hollingsworth, A., Jakob, C., Janssen, P., Klinker, E., Lanzinger, A., Miller, M., Rabier, F., Simmons, A., Strauss, B., Thépaut, J.-N. and Viterbo, P. 1998 The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part III: Experimental results. *Q. J. R. Meteorol. Soc.*, **124,** 1831–1860

Andersson, E., Fisher, M., Munro. R. and McNally, A. 2000 Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Q. J. R. Meteorol. Soc.*, **126,** 1455–1472

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

Browning, G. L., Kreiss, H.-O. and Schubert, W. H. 2000 The role of gravity waves in slowly varying in time tropospheric motions near the equator. *J. Atmos. Sci.*, **57,** 4008–4019

Cats, G. and Wergen, W. 1983 'Analysis of large scale normal modes by the ECMWF analysis scheme'. Pp. 343–372 in Proceedings of the ECMWF Workshop on Current problems in data assimilation, 8–10 November 1982, Reading, UK

Charney, J. G. and Drazin, P. G. 1961 Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *J. Geophys. Res.*, **66,** 83–109

Clayson, C. A., Strahl, B. and Schrage, J. 2002 2–3-day convective variability in the tropical western Pacific. *Mon. Weather Rev.*, **130,** 529–548

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljević, D., Hamrud, M., Hollingsworth, A., Rabier, F. and Fisher, M. 1998 The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Q. J. R. Meteorol. Soc.*, **124,** 1783–1807

Daley, R. 1993 Atmospheric data analysis on the equatorial beta plane. *Atmos.–Ocean*, **31,** 421–450

1996 Generation of global multivariate error covariances by singular-value decomposition of the linear balance equation. *Mon. Weather Rev.*, **124,** 2574–2587

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

Dunkerton, T. J. 1997 The role of gravity waves in the quasi-biennial oscillation. *J. Geophys. Res.*, **102,** 26053–26076

Evensen, G. 1997 Advanced data assimilation for strongly nonlinear dynamics. *Mon. Weather Rev.*, **125,** 1342–1354

Fisher, M. 1996 'The specification of background error variances in the ECMWF variational analysis system'. Pp. 645–652 in Proceedings of the ECMWF Workshop on Non-linear aspects of data assimilation, 9–11 September 1996, Reading, UK

2003 'Background error covariance modelling'. Pp. 45–64 in Proceedings of the ECMWF Workshop on Recent developments in data assimilation for atmosphere and ocean, 8–12 September 2003, Reading, UK

Giorgetta, M. A., Manzini, E. and Roeckner, E. 2002 Forcing of the quasi-biennial oscillation from a broad spectrum of atmospheric waves. *Geophys. Res. Lett.,* **29,** 861–864

Gustafsson, N., Berre, L., Hörnquist, S., Huang, X.-Y., Lindskog, M., Navascuès, B., Mogensen, K. S. and Thorsteinsson, S. 2001 Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus*, **53A,** 425–446

Heckley, W., Courtier, P., Pailleux, J. and Andersson, E. 1992 'The ECMWF variational analysis: General formulation and use of background information'. Pp. 49–93 in Proceedings of the ECMWF Workshop on Variational assimilation, with special emphasis on three dimensional aspects, 9–12 November 1992, Reading, UK

Ingleby, N. B. 2001 The statistical structure of forecast errors and its representation in The Met. Office Global 3-D Variational Data Assimilation Scheme. *Q. J. R. Meteorol. Soc.*, **127,** 209–231

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R. and Fiorino, M. 2001 The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bull. Am. Meteorol. Soc.*, **82,** 247–267

Ko, S. D., Tribbia, J. J. and Boyd, J. P. 1989 Energetics analysis of a multilevel global spectral model. Part I: Balanced energy and transient energy. *Mon. Weather Rev.*, **117,** 1941–1953

Lindzen, R. S. and Fox-Rabinovitz, M. 1989 Consistent vertical and horizontal resolution. *Mon. Weather Rev.*, **117,** 2575–2583

Matsuno, T. 1966 Quasi-geostrophic motions in the equatorial area. *J. Meteorol. Soc. Jpn*, **44,** 25–43

Parrish, D. 1988 'The introduction of Hough functions into optimum interpolation'. Pp. 191–196 in Proceedings of the Eighth Conference on NWP, American Meteorological Society, Boston, USA

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

Phillips, N. 1986 The spatial statistics of random geostrophic modes and first-guess errors. *Tellus*, **38A,** 314–322

Rabier, F., McNally, A., Andersson, E., Courtier. P., Undén, P., Eyre, J. R., Hollingsworth , A. and Bouttier, F. 1998 The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Q. J. R. Meteorol. Soc.*, **124,** 1809–1829

Thiebaux, J., Katz, B. and Wang, W. 2001 'New sea-surface temperature analysis implemented at NCEP'. Pp. J159–J163 in Preprints of 18th Conf. on Weather Analysis and Forecasting, American Meteorological Society, Ft. Lauderdale, FL, USA

Untch, A. and Hortal, M. 2004 A finite-element scheme for the vertical discretization of the semi-Lagrangian version of the ECMWF forecast model. *Q. J. R. Meteorol. Soc.*, **130,** 1505–1530

Uppala, S. 2001 ECMWF Re-analysis, 1957–2001, ERA-40. *ERA-40 Project Series*, **3,** 1–10

Vialard, J., Vitart, F., Balmaseda, M. A., Stockdale, T. N. and Anderson, D. L. T. 2005 An ensemble generation method for seasonal forecasting with an ocean-atmosphere coupled model. Submitted to *Mon. Weather Rev.* (in press)

Wada, K., Nitta, T. and Sato, K. 1999 Equatorial inertia-gravity waves in the lower stratosphere revealed by TOGA-COARE IOP data. *J. Meteorol. Soc. Jpn*, **77,** 721–736

Wergen, W. 1988 The diabatic ECMWF normal mode initialization scheme. *Beitr. Phys. Atmos.*, **61,** 274–302

Wheeler, M. and Kiladis, G. N. 1999 Convectively coupled equatorial waves: analysis of clouds and temperature in the wavenumber-frequency domain. *J. Atmos. Sci.*, **56,** 374–399

Yang, G.-Y., Hoskins, B. J. and Slingo, J. 2003 Convectively coupled equatorial waves: a new methodology for identifying wave structures in observational data. *J. Atmos. Sci.*, **60,** 1637–1654

Žagar, N., Gustafsson, N. and Källén, E. 2004 Variational data assimilation in the tropics: The impact of a background-error constraint. *Q. J. R. Meteorol. Soc.*, **130,** 103–125