TECHNICAL MEMORANDUM

437

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

Nedjeljka Žagar[1], Erik Andersson and Michael Fisher

Research Department

[1]Dept. of Meteorology, Stockholm University

Submitted to Q. J. R. Meteorol. Soc.

April 2004

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, England

Abstract

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

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 the explained variance is represented by the equatorial Rossby (ER) modes, and a significant percentage pertains to the equatorial inertio-gravity (EIG) modes, especially in the lower troposphere. Eastward-propagating EIG modes have maximum variance in the stratosphere, where the short-wave variance in westward-moving waves is particularly small. This finding is consistent with the quasi-biennial oscillation (QBO) being in an easterly phase in the study period.

The vertical correlations for ER modes display characteristics similar to those of their extra-tropical counterparts: correlations narrow towards shorter scales and in the stratosphere. In addition, tropospheric ER correlations are vertically asymmetric and deeper for the n=1 mode than for higher modes. Furthermore, the present statistics do not display a significant increase with altitude of the horizontal correlation scale for the height field, typical for global, quasi-geostrophic statistics commonly used in current data-assimilation schemes. It is most likely that deep convection, acting as a generator of equatorial wave motion, is the dominant mechanism behind 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. An important role for the tropical mass-wind coupling is also played by EIG modes; these waves have a major impact by reducing the meridional correlation scales and the magnitudes of the balanced height-field increments.

Contents

1 Introduction

Numerical Weather Prediction (NWP) models are never going to be error-free, and a reliable estimate of their forecast errors poses a real challenge since the whole truth will never be within 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.

In the mid-latitudes, the basic balance relationship is geostrophy, which has been extensively used in the 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 primarily 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 missing; the analysis here has thus traditionally been undertaken in the univariate fashion. Consequently, large-scale divergence fields, such as the Hadley and Brewer-Dobson circulation, are analyzed nearly univariately. Since GOS in the tropics relies on mass-field information, uncertainties in the analyzed 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 mid-latitude atmosphere: the Kelvin and mixed Rossby-Gravity (MRG) modes (Matsuno 1966).

In this paper we consider Kelvin, MRG and EIG waves, in addition to ER modes, as balanced tropical motion and attempt to determine their appropriate relative roles based on statistics from the ECMWF forecasting system. We design a spectral, mode-based covariance model which describes the dominant characteristics seen in the ECMWF short-range tropical forecast errors. We implement the covariance model in a variational data-assimilation scheme and examine the response of the analysis to single observations in idealized assimilation experiments. The devised tropical analysis scheme maintains the appropriate balance between mass and wind, consistent with the specified statistical distribution of equatorial modes.

The failure of initialization techniques based on the cut-off frequency (Wergen 1988) and vertical normal mode-based 3D-Var assimilation in the tropics (Heckely *et al.* 1992; Andersson *et al.* 1998) suggests an important role for non-rotational equatorial waves in maintaining the divergence of the wind in the tropics. Indeed, equatorially trapped Kelvin, MRG, equatorial IG (EIG) and equatorial Rossby (ER) motion have regularly been detected in observational data (Wheeler and Kiladis 1999, and references therein). Excited by latent heat release in deep convection, these waves transmit energy in the vertical and longitudinal directions and, therefore, represent a means whereby energetic disturbances are communicated along the equator. Corresponding equivalent depths according to the linear shallow-water theory are in the range 12-50 m (Wheeler and Kiladis 1999). These observational results confirm an earlier numerical simulation 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”.

In the current paper we investigate the extent to which the various equatorial waves are present in the ECMWF short-range forecast errors in order to construct a background-error covariance matrix involving the coupling between the mass and the wind field in the tropics. In the background-error modelling, an important common 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 have employed equatorial wave theory to replace the quasi-geostrophic balance relationships that are successful in extra-tropical latitudes. The idea has been presented in Žagar *et al.* (2004a, ŽGK), summarized here in Section 2. In Section 2 we also describe a data-assimilation ensemble, from which we extract short-range forecast differences to serve as a proxy for background errors; the justification is given in Appendix A. We investigate how accurately the ensemble-generated data set can be represented in terms of equatorial wave solutions and discuss the applicability of the approach. The resulting covariance statistics are presented in Section 3. In Section 4, we utilize the variational assimilation scheme developed in ŽGK for an examination of horizontal structures in the troposphere and stratosphere through idealized single-observation experiments. Summary and conclusions are presented in Section 5.

2 Background error covariance modelling for the tropics

2.1 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 Appendix A. The ensemble consists of ten independent data assimilations that each use different sets of randomly perturbed observations, for a period of 31 days in October 2000. For the purpose of this study, forecast differences were extracted from the ensemble, for the tropical belt $20^{\circ}\mathrm{S}-20^{\circ}\mathrm{N}$ at a one-degree resolution 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 outputs for other ranges are very similar.

The ensemble standard deviation of error for the zonal wind at three model levels is shown in Fig. 1: in the lower and upper troposphere and near the tropopause. A striking feature in the lower troposphere is a concentration of errors along the intertropical convergence zone (ITCZ) and over Indonesia (Fig. 1a). Local maxima are located east 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. 1b). Higher up, the Indonesian region displays larger errors than the rest of the domain (Fig. 1c), 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 South American coast and over the 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 variations become small.

Figure 1 shows standard deviations of 12-hour forecast errors; however, patterns and magnitudes are similar for other ranges (3 and 24 hour). 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

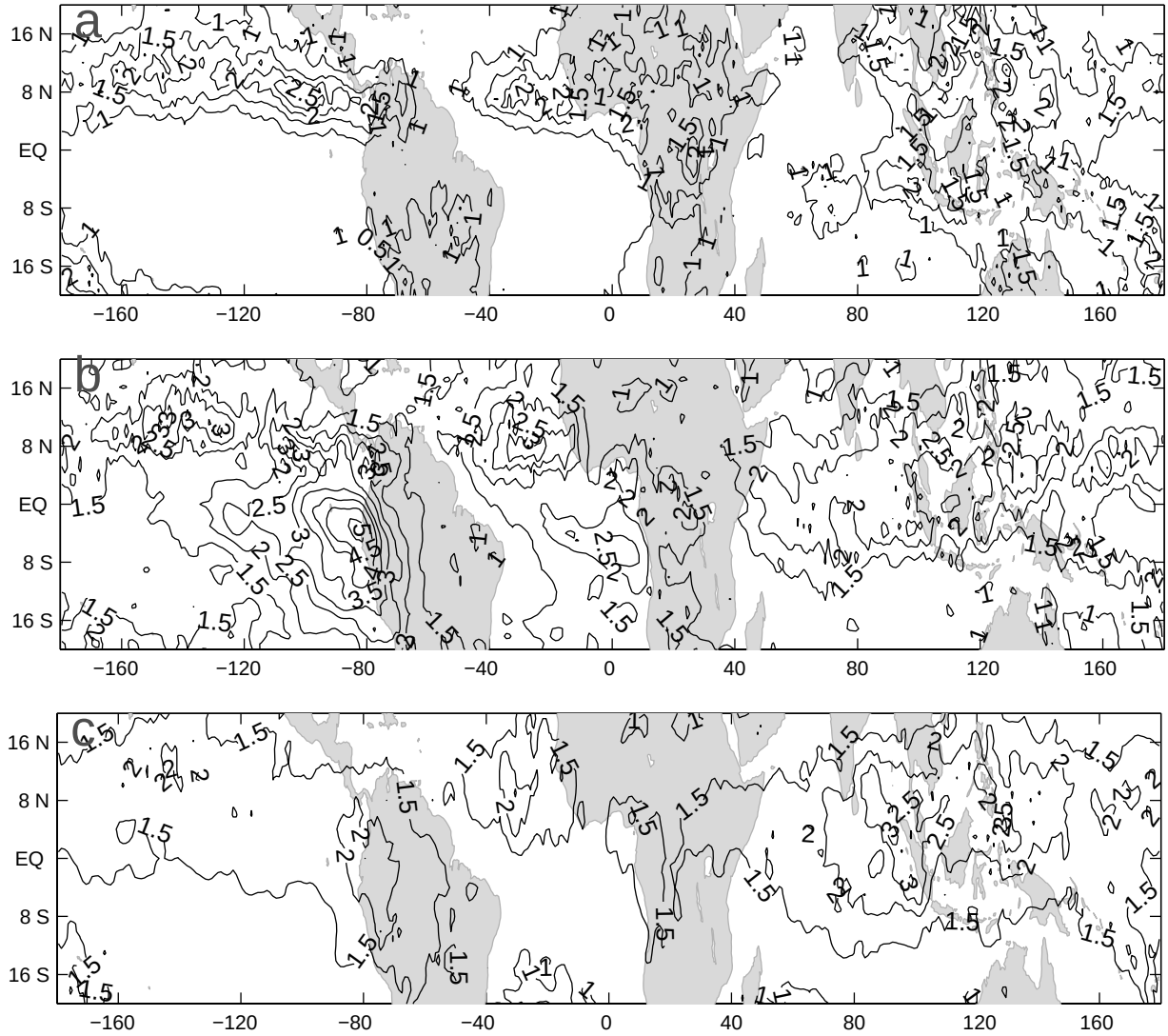

Figure 1: Standard deviations of 12-hour forecasts errors in the zonal wind field (in m s $^{-1}$) of the ECMWF model in the tropics, as estimated with a data-assimilation ensemble. Shown are model levels 48 (approximately 828 hPa) (a), 30 (approximately 202 hPa) (b) and 25 (approximately 96 hPa) (c).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

just west of South America between levels 28 and 35. The geopotential-error growth is almost negligible. The stratospheric forecast errors do not grow during 24 hours; on the contrary, they are on average up to 2% smaller than the 3-hour errors for all fields.

2.2 Covariance-model formulation

In this study forecast errors are represented in terms of tropical eigenmodes based on a parabolic cylinder function expansion. The expansion for the meridional wind field error (v) at vertical level z takes the following form:

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

where k is a zonal wave number index, n is the number of nodes in the meridional direction of v, thus it is called the meridional mode number, and m is an index for wave type. The three modal indices define a modal index $\nu(k,n,m)$ for each single equatorial mode, 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 a degree n; k dependence enters through the equatorial trapping scale a_e (see ŽGK for details). The corresponding ranges are defined in the following way: N_k is defined by the minimal wave length resolved, $3\Delta x$, 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). Three wave types are ER, westward- (WEIG) and eastward-propagating (EEIG) EIG for $n \geq 1$, and Kelvin, 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 their combinations, 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

$$\Psi_\nu = \sum_x \sum_y \left[v(x,y) v_\nu^*(k,y) + u(x,y) u_\nu^*(k,y) + g h_0^{-1} h(x,y) h_\nu^*(k,y) \right] e^{ikx}. \tag{2}$$

Here, the asterisk (*) indicates 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).

The independence of different modes is an essential assumption, and can be expressed as

$$\overline{\Psi_\nu \Psi_{\nu'}^*} = \delta_k^{k'} \delta_n^{n'} \delta_m^{m'} \overline{\Psi_\nu \Psi_{\nu'}^*}. \tag{3}$$

Applying the homogeneity assumption in the zonal direction, and the energy norm, the covariance for the mode ν between model levels z and z' is obtained as the following sum:

$$\sum_x \sum_y \left[\overline{v(x,y,z) v(x,y,z')} + \overline{u(x,y,z) u(x,y,z')} + g h_0^{-1} \overline{h(x,y,z) h(x,y,z')} \right] = \sum_k \sum_n \sum_m \overline{\Psi_\nu^{r,z} \Psi_\nu^{r,z'} + \Psi_\nu^{i,z} \Psi_\nu^{i,z'}}. \tag{4}$$

The superscript i stands for the imaginary and r for 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 by γ_ν^2 in order to keep the consistency with Eq. (11) in ŽGK.

2.3 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 NWP model forecast errors. We define the unexplained variance ratio (Derber and Bouttier 1999) at level z as:

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

Here, v is a proxy for forecast errors in the meridional wind, obtained as a difference between two forecasts of the same length. The part of the error field unexplained by the equatorial wave theory, $v_p(x,y,z)$, is obtained as a difference between the total error (evaluated in the physical grid points) and the inverse of the projection (2), analogous to (1):

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

The number of samples extracted from the data-assimilation ensemble, N_{ens}, is 225, N_{mode} is the number of eigenmodes, and $\overline{v}$ is the mean error at the model grid point (x,y,z). The unexplained variance ratio, ε, measures the ratio between the part of the error variance which is successfully represented by the selected equatorial waves and the total variance. The ratio is also similar to the time-mean analysis error used by Yang *et al.* (2003).

We choose a typical value of the phase speed c=15 ms^{-1}; 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°. With this choice made, the lowest four meridional modes are completely trapped within the analyzed area, i.e. the orthogonality condition is fulfilled exactly for those modes. Equatorial eigenmodes are either symmetric or anti-symmetric with respect to the equator. The error fields, as we have seen, are not symmetric about the equator but rather localized along the ITCZ (Fig. 1), that is, in the centre of the northern part of the analysis domain. It follows that, in order to resolve most of the important part of the error field in the lower troposphere, a linear combination of a number of anti-symmetric meridional modes is crucial. Furthermore, one needs a sufficiently large N_n in order to resolve smaller meridional structures. In the case of full fields, rather than their differences, a sufficiently large N_n is needed for resolving structures away 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 for varying types of waves included. In Fig. 2 we show ε for N_n=10, including all modes allowed by the equivalent meridional truncation criterion (Eq. (10) in ŽGK, with c_1=c_2=1). Although application of the truncation criterion is not important for the value of ε, it is nevertheless applied in order to be consistent with the covariance model used in idealized assimilation tests later on. With a zonal truncation N_k=119, a total of N_{mode}=2937 eigenmodes are included in the expansion. Values shown are area means at each model level.

It can be seen that above the 500 hPa level (∼model level 39), equatorial eigenmodes successfuly represent 75-80% of the height and 50-75% of the wind field variance. Between 500 hPa and 900 hPa, the amount of unexplained variance increases and becomes as large as 60% near model level 50. The unexplained variance generally is largest for the zonal wind component, except below 900 hPa (∼model level 50) and near the tropopause. Below ∼900 hPa, the methodology become useless for all fields, especially for the height field. For this reason we shall not consider these levels in the subsequent discussion.

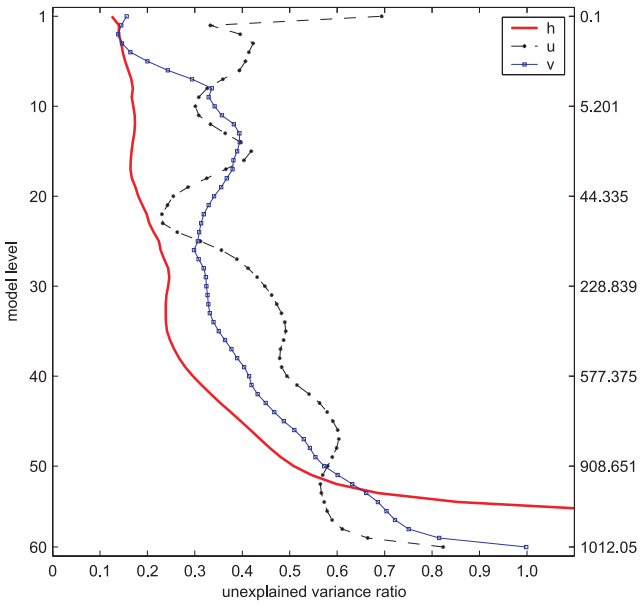

Figure 2: Vertical distribution of the unexplained variance ratio (area mean), representing an error of the equatorial wave approach to modelling the height and wind field errors at model levels. See text for definition. The right y-axis shows average pressure (hPa) of model levels marked on the left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

Looking closer at the spatial patterns of unexplained errors (v_p, u_p, and h_p fields, figures not shown), it becomes evident that most of the tropospheric unexplained variance is located in the ITCZ, a region where the assumed error structures do not display north-south symmetry with respect to the equator. One expects that including a large number of anti-symmetric modes should be sufficient to reproduce the error pattern in the ITCZ, but this appears difficult to accomplish. A compromise value for N_n is required, in order for eigenmodes to remain equatorially trapped.

Figure 2 would look different for other choices of eigenmodes. For example, in the case of N_n=3 the unexplained variance ratio is much larger since three modal points can only poorly resolve the meridional structure of the error fields. The zonal truncation, on the other hand, matters little 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 them 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 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, making the approach meaningless. On the other hand, representation of height errors is little affected by a change of N_n and almost unaffected by the mode selection. This is due to geopotential gradients in the tropical atmosphere being weak, which is a result of geostrophic adjustment (e.g. Žagar *et al.* 2004b). Changes in c are unimportant for ε as compared to the other two factors 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, mean ε is between 0.3 and 0.4, which means that in the rest of the paper we are concerned with the statistical structure of about 60−70% of the tropical variance.

2.4 Meridional variation of errors

A spectral representation of errors retains full information about the scale variations of covariances but gives no information on their spatial variation, as a consequence of the homogeneity assumption. In the present case the homogeneity assumption is applied in the zonal direction only, 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. The idea utilizes the fact that the usual 3D-Var formulation in terms of normalized and non-dimensional variables (e.g. Ψ_v/γ_v) has the background-error covariance matrix equal to the identity matrix. Thus, one can take an ensemble of random vectors for Ψ_v/γ_v, drawn from a Gaussian distribution of zero mean and variance one, and obtain randomisation estimates of the error variances first in spectral space, and after further transformation, also in grid point space. An example of the use of this methodology is provided in Andersson *et al.* (2000).

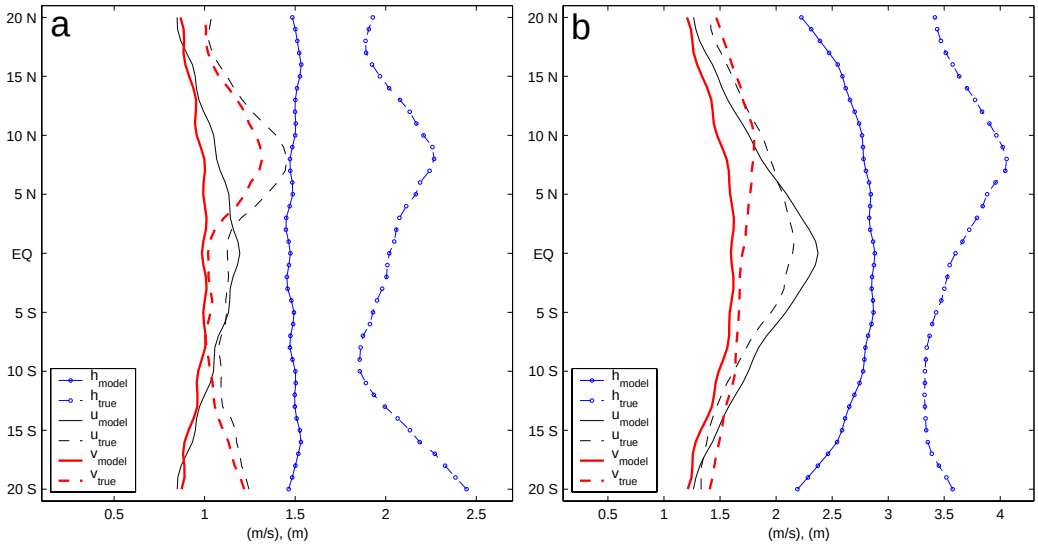

Figure 3: Meridional profiles of 12-hour forecasts errors in the zonal (black) and meridional wind fields (red) and the geopotential height (blue lines). Dashed lines apply to 'truth' whereas full lines apply to the covariance model (see text for further details). (a) model level 43 (654 hPa), (b) model level 27 (133 hPa).

The resulting randomisation variances provide meridional profiles which can be compared with the zonally averaged profiles of the error fields we started off from (i.e. those 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. 3a), the error profiles of the ECMWF model have maxima at 8°N, as illustrated in Fig. 1a; in addition, errors are increased southward of 15°S, especially the geopotential height errors. 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 performs best in the upper troposphere (Fig. 3b), where the ECMWF forecast errors are not concentrated along the ITCZ (Fig. 1b).

The discrepancy between the "true" and modelled error magnitudes is largest for the height field, especially in the stratosphere. This may seem contradictory to Fig. 2 which indicates that height errors are explained the best; the most likely reason for this is the small variability of geopotential in the tropics. Furthermore, "true" errors of all fields become largely homogeneous in the stratosphere, while our model produces the error maxima at the equator (not shown). Equatorially centred error profiles are due to Kelvin-wave correlations, to be illustrated in section 4.

Another shortcoming noticable in Fig. 3 is due to the fact that wind and mass are analyzed together. Their error coupling is responsible for too-small amplitudes of modelled height errors in the troposphere. On the other hand, it acts to unrealistically increase the amplitude of the wind field errors in the stratosphere, where the ECMWF geopotential errors increase.

In an actual implementation of the derived covariance model in a variational data-assimilation scheme, the unexplained part of the tropical error variance would be analysed univariately. The main point here is that the spectral mode-based covariance model captures a significant part of the error variance that is symmetric about the equator, but far less of the asymmetric part.

3 Variances and vertical correlations

3.1 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 the Rossby type of 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. Within the troposphere, the variance assigned to ER waves is between 40% and 50%, and it is largest at levels between 150 and 300 hPa. Contributions from various meridional modes decrease as n increases. The variance profile for the lowest two modes follows the shape of the total profile. For higher modes, stratospheric variance is small so that modes five and higher make no contribution to the variance in the stratosphere and a few percent of the variance in the troposphere (not shown).

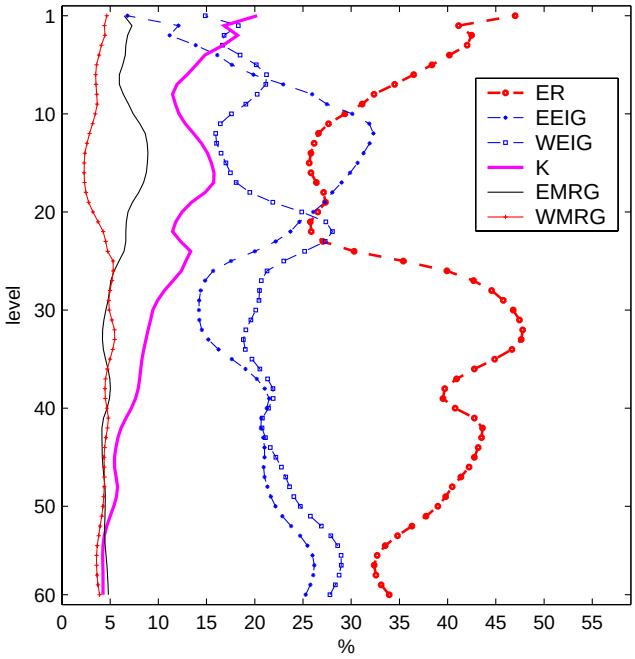

Figure 4: Vertical distribution of the variance among various equatorial eigenmodes: equtorial Rossby (ER), eastward-propagating (EEIG) and westward-propagating (WEIG) equatorial inertio-gravity modes, Kelvin (K), eastward- (EMRG) and westward-propagating (WMRG) mixed Rossby-gravity waves. The x-axis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

ER and EIG mode profiles have opposite phase of a variance in the troposphere, and EEIG and WEIG are 180° out of phase in the stratosphere. Otherwise, the percentage of variance contributed by EEIG and WEIG waves is similar in the troposphere, except at levels corresponding to the range 100−300 hPa, where WEIG have a more important role. Together, EEIG and WEIG make 40-50% of the variance in the troposphere. In the stratosphere, EEIG have an absolute maximum of variance around level 6 hPa. Opposite to ER waves, EIG modes follow the same vertical variance profile for all n (not shown).

Other eastward-propagating modes (Kelvin and EMRG) have the maximum variance in the stratosphere, similar to EEIG. On the other hand, all westward-propagating modes (ER, WMRG and WEIG) have minima at those levels. Figure 4 suggests that MRG and Kelvin modes make only a minor contribution to the variance in the tropical troposphere of the ECMWF model, around 10% and 5-10%, respectively. This does not suggest, however, that they are unimportant for the balanced structure functions, as shown later on. 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 become more important.

The zonal-scale dependence is shown in Fig. 5, at three model levels representing the stratosphere (12 hPa), middle troposphere (500 hPa) and lower troposphere (828 hPa). As can be seen, the main difference is between

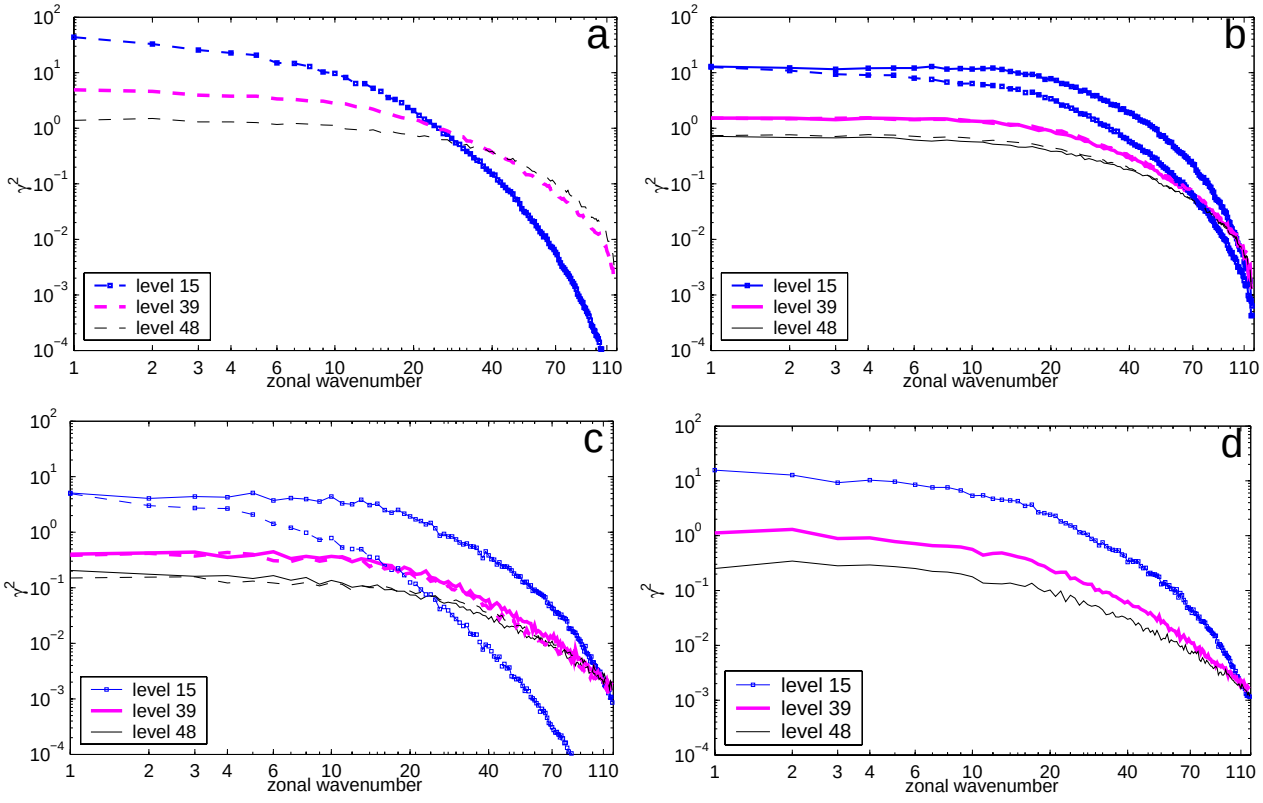

Figure 5: Spectral variance density in ER (a), EIG (b), MRG (c) and Kelvin modes (d) at various model levels, as a function of a zonal wave number. Full lines apply to the eastward-propagating whereas dashed lines denote the westward-propagating waves.

the eastward- and westward-propagating waves and between the stratosphere and the troposphere. In the stratosphere, westward-propagating waves contain little variance at zonal wave numbers larger than 20 ($\sim$ 2000 km); the variance decrease is less sharp for WEIG than for the ER and WMRG modes, but it is still notable. This marked reduction of the spectral energy in the westward-propagating waves in the stratosphere is most likely related to the strong eastward phase of the quasi-biennial oscillation (QBO) in the study period (in October

2000 there were easterly zonal winds between 30m/s and 40 m/s at model level 15). With such a mean flow, westward-propagating waves do not satisfy the criterion for vertical wave propagation (Lindzen and Holton 1968). Instead, EEIG waves dominate the spectrum (except at lowest zonal wave numbers). Our stratospheric spectra are entirely in agreement with recent observational and modelling studies (e.g. Dunkerton 1997), which suggest that the gravity waves may carry the largest part of the energy flux required to drive the QBO.

Among the six various wave types, ER modes dominate the spectra at all scales except in the stratosphere. Here, all eastward modes have larger variance than the ER waves for $k > 10$, with Kelvin waves becoming relatively more important at long scales. Below 500 hPa, the total variance in EIG modes exceeds ER at all zonal scales. Smallest variance is found in the MRG waves and it is especially small in the stratosphere for the WMRG waves.

The partition of variance among modes and scales does not change much with forecast time, based on a comparison between of 3-, 12- and 24-hour forecasts (not shown). The main difference between 3- and 24-hour forecasts is an increase in the Kelvin (2-3%) and ER ($\sim$ 5%) wave variance in the stratosphere at the expense of an EIG wave variance reduction. The EIG wave variance reduces also in the troposphere by several percent, with a corresponding increase in the ER wave variance. MRG wave variance remains virtually unchanged. Sensitivity with respect to the choice of c is not large, based on a comparison between h_0=23 m and h_0 equal to 50 and 250 m.

One may ask whether the significant fraction of the variance associated with EIG modes arises from the model deficiencies, especially in the early stages of the forecast. It may be that the model has systematic errors but those are ignored in a strong-constraint 4D-Var. The background-error term should describe the actual statistics of background errors. If these contain more EIG wave motion than it is 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 build a background-error term for use in an analysis system. The latter would first require addressing the missing 30-40% of the variance not described by the projection onto equatorial modes.

3.2 Vertical correlations

Vertical correlations are calculated for all model levels, but we leave out levels 50-60 due to the small percentage of variance explained by the equatorial wave approach. The shape of correlations depends on the wave type and altitude (Figs. 6-8). In general, correlations become more narrow moving upward and towards smaller scales. The vertical-correlation narrowing for smaller scales is seen mainly for ER and Kelvin modes, to lesser extent for MRG and almost not at all for EIG waves.

The broadening at large scales of ER wave correlations (Fig. 6) begins approximately with zonal wave number 20, and it is larger for the n=1 (Fig. 6a) than for higher modes (Fig. 6b). However, comparing ER correlations for the same n it can be noticed that correlation narrowing with altitude primarily applies to the model-level space. If the vertical coordinate is height, the shape of correlations at small zonal wave numbers would not differ that much at various altitudes. Narrowing in model space could be an indication of an insufficient vertical resolution. Lindzen and Fox-Rabinovitz (1989) demonstrated that the requirement for a consistent horizontal and vertical resolution demands a higher vertical resolution in the tropics than at the mid-latitudes.

Tropospheric vertical correlations are not symmetric; correlations with the levels below are much stronger than above, throughout the troposphere (Fig. 6a-c). The most likely physical mechanism behind this is convection. This hypothesis is supported by the fact that stratospheric correlations do not display such asymmetry.

A prominent feature of stratospheric correlations is the appearance of wings of negative correlations at scales

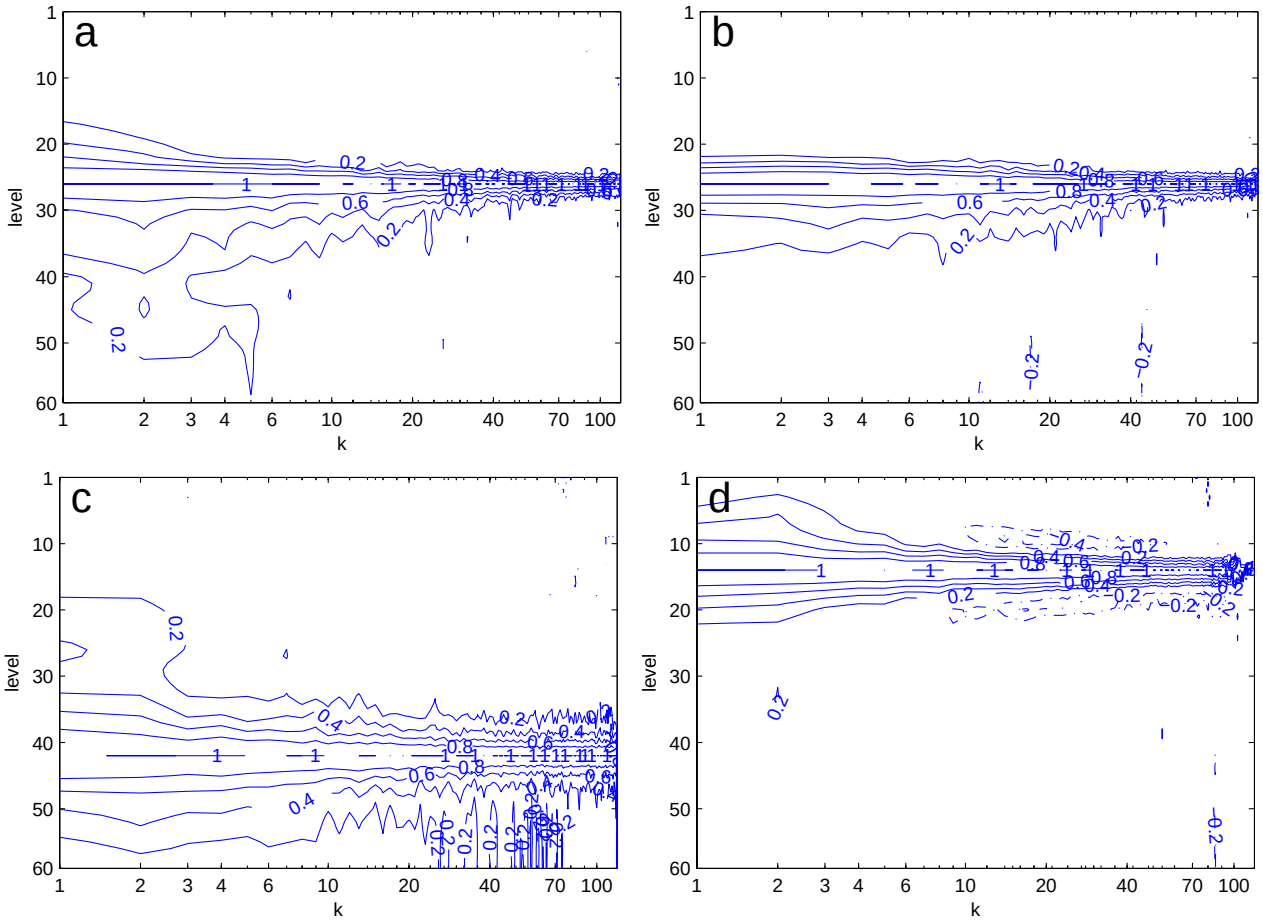

Figure 6: Vertical correlations for ER modes, as a function of zonal wave number. (a) n=1 and (b) n=2 at model level 27 (approximately 133 hPa); (c) n=1 at model level 43 (approximately 654 hPa); (d) n=1 at model level 15 (approximately 12 hPa). Isolines every 0.2 with zero isoline omitted.

smaller than k=10 (Fig. 6d). This feature is present also for all other modes in the stratosphere, although it is strongest for EIG modes. In the case of EEIG modes, negative wings extend over all horizontal scales (Figs. 7a). This may be a consequence of the strongly stable stratification in the stratosphere. Vertical perturbation-features tend to have a limited vertical extent, and tend to integrate to zero over a relatively narrow layer. Lower in the troposphere, EIG wave correlations show negligible change with scale and have a symmetric shape (Figs. 7b).

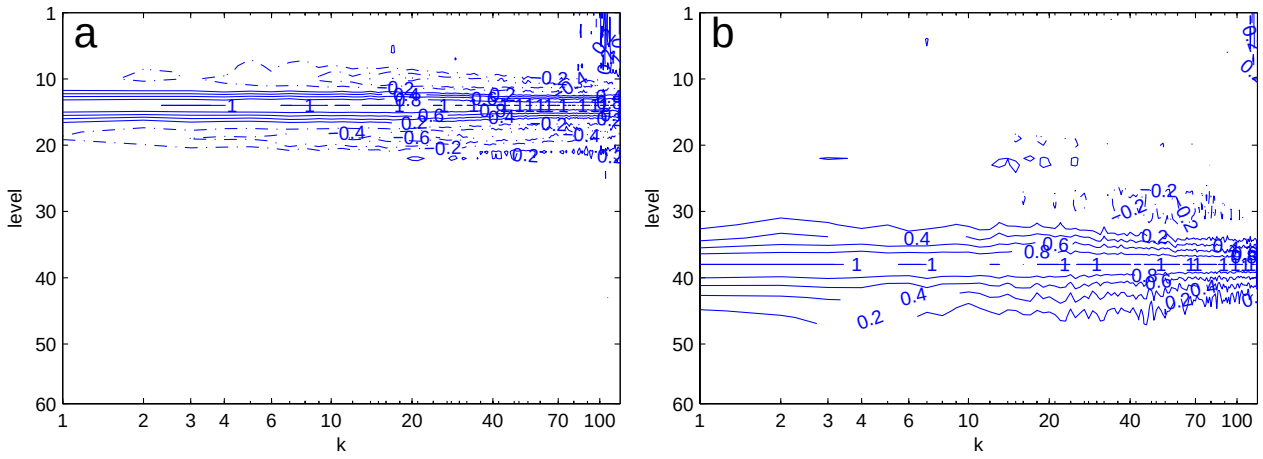

Figure 7: Vertical correlations with model level 15 for n=1 EEIG modes (a) and with model level 39 (approximately 500 hPa) for n=1 WEIG modes (b).

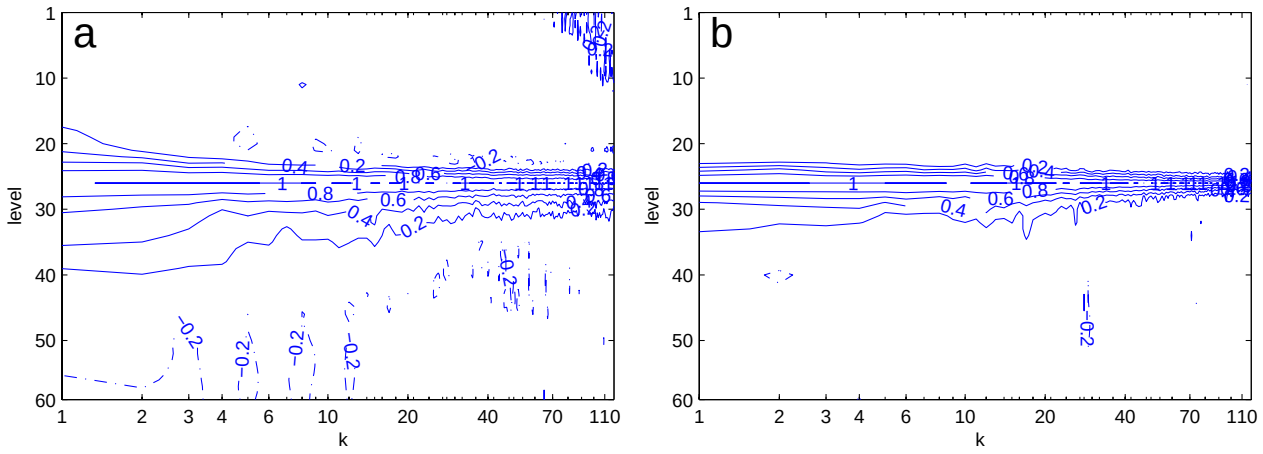

Figure 8: Vertical correlations with model level 27 for Kelvin modes (a) and WMRG (b) modes.

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

4 Horizontal structures

4.1 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 with 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? Second, we study the variation of horizontal scales with height. Analyzing height and wind fields together prevents us from directly evaluating their respective horizontal correlations; carrying out single-observation experiments is an alternative to compare their scale variations. Examples of single-observation analyses are shown in Figs. 9-11, at three vertical levels representative of the lower troposphere, the middle/upper troposphere and the stratosphere.

Fig. 9 shows analysis increments due to a single height observation. A striking feature is that the horizontal scale does not increase to a significant extent with altitude. This is an intriguing result, in apparent disagreement with earlier theoretical studies (Phillips 1986; Bartello and Mitchell 1992) and studies of global statistics (e.g. Rabier *et al.* 1998; Ingleby 2001). Namely, these studies indicate an increase of horizontal scales with altitude, particularly for the height field, and broadest correlations in the tropics. What makes the previous studies less suitable for the comparison 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 methodology (Parrish and Derber 1992), disadvantages of which are largest in the tropics. The NMC method assumes that the forecast error covariances can be represented by the differences between forecasts of different ranges valid at the same time. It tends to produce covariances which are broader in both the horizontal and the vertical direction than those of the background errors due to the length of forecasts. Furthermore, in data-sparse region such as the tropics, two forecasts may be very similar, resulting in an underestimation of variances. Finally, the NMC technique is likely to break down in the tropics, the reason being the small temporal variability of the tropical large-scale flow (Rabier *et al.* 1998). It has also been noted, based on a study using the NMC method, 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 mid-latitude errors in Rossby modes at synoptic scales. Modelling tropical forecast errors in terms of quasi-geostrophic modes or based on global statistics is inappropriate. Convective processes, not described by mass-wind balance 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 are 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 that 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 mid-latitudes), and by EIG waves (with a significantly larger relative contribution than gravity waves have in the mid-latitudes).

Another feature seen in Fig. 9 is that the mass-wind coupling, although weak, is increasing with altitude. Wind-field increments in the lower troposphere are practically negligible. In the stratosphere, the magnitude of the wind increment is increased by a factor three at the observation point (Fig. 9c). Even so, the kinetic energy of the increment field is only around 17% of the total energy (it is about 6% at level 43).

The mass-wind balance is suggestive of a Kelvin wave although the fraction of the total error variance associated with Kelvin waves is relatively small. The reason why the Kelvin-wave coupling is so effective near the equator is that it is the only mode that has significant amplitude in height at the equator, coupled with the strongest zonal

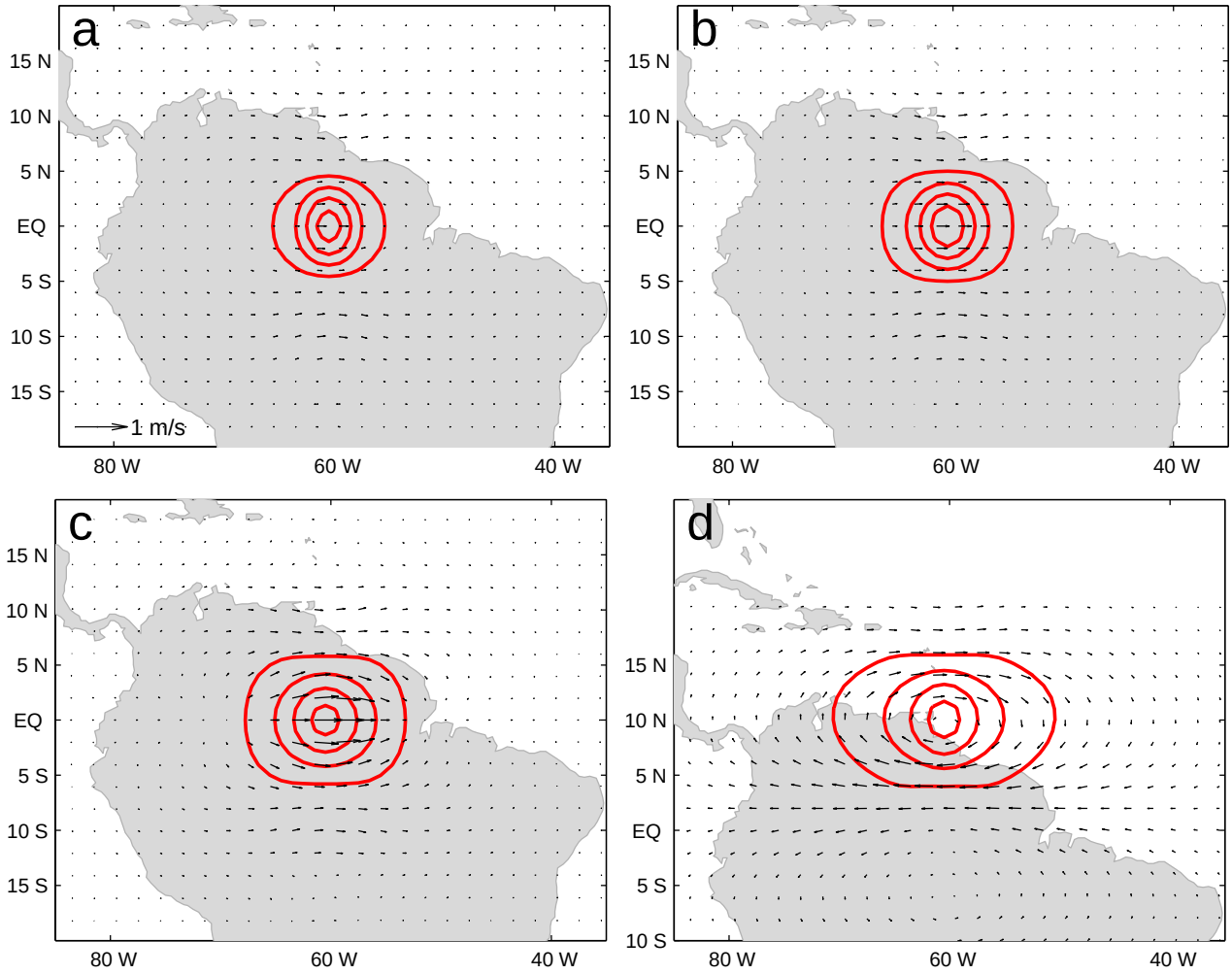

Figure 9: Horizontal structure of the analysis increments due to a single height observation located at the equator (a-c) and at 10°N (d). The background error variance corresponds to the ECMWF model levels 43 (approximately 654 hPa) (a), 39 (approximately 500 hPa) (b), 15 (approximately 12 hPa) (c), and 31 (approximately 229 hPa) (d). The height observation is 4 m higher than the background and the background state is 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. Thick grey lines correspond to positive values, while thin black lines are used for negative values. Wind vectors are shown in every second grid point in both horizontal directions.

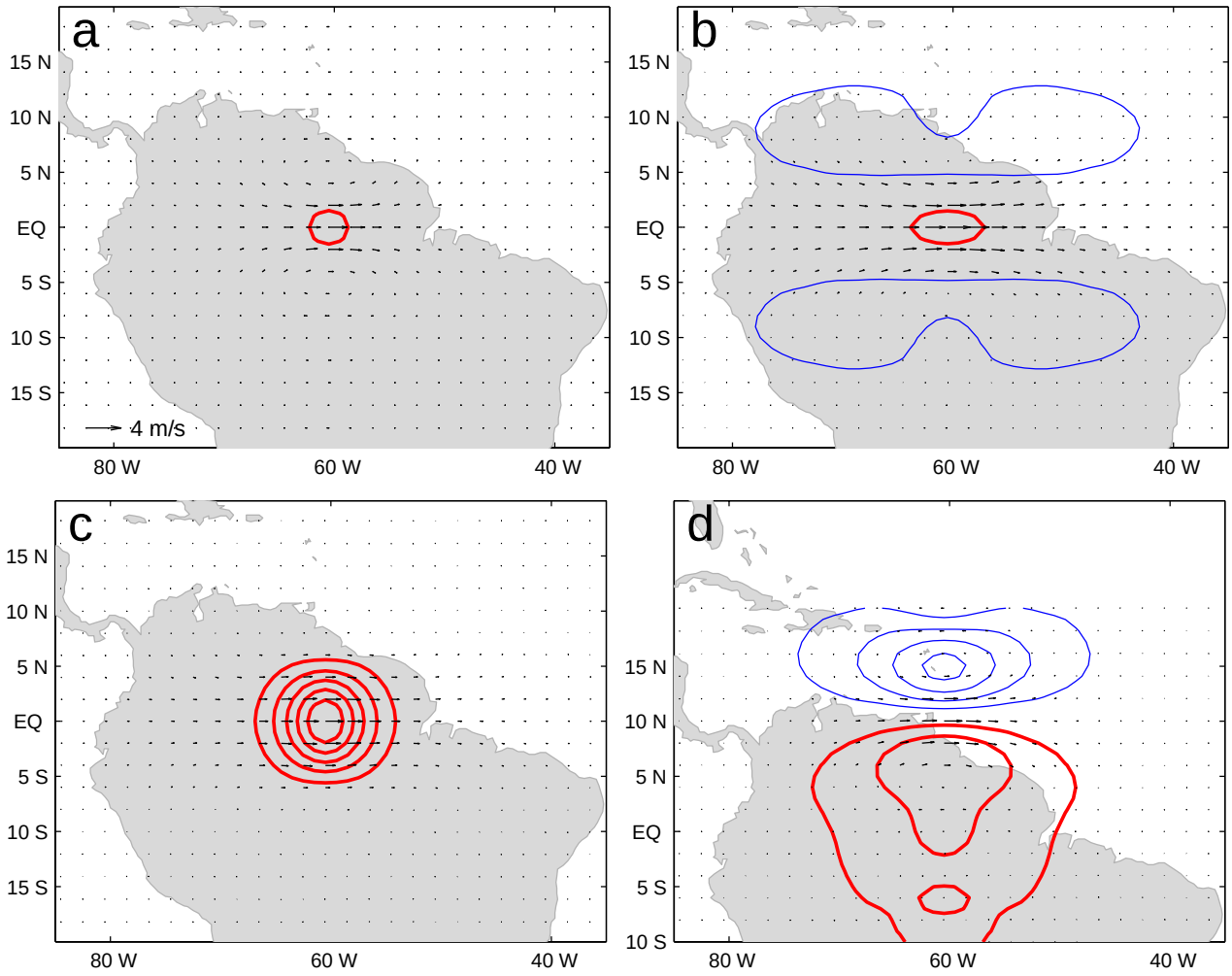

Figure 10: As Fig. 9, but for a single westerly wind observation at model levels 43 (a), 31 (b), 15 (c), and 31 (d). The wind observations is 4 ms^{-1}. Isoline spacing for the height field increments is ±0.2 m.

wind. Away from the equator the importance of Kelvin waves diminish, at least in the troposphere, and the ER type of mass-wind balance becomes more important (Fig. 9d). It can also be noticed in Fig. 9d that the subtropical increments are zonally elongated for both height and wind fields. Stratospheric increments show a stronger influence of the Kelvin wave, as its contribution to the variance is three times larger in the stratosphere than in the troposphere.

Contrary to the height data, assimilation of wind observations results in increments with horizontal scales increasing with altitude (Figs. 10-11). In addition, equatorial wave balance produces coupling with the height field, especially in the stratosphere. A zonal wind observation at the equator is coupled with a positive height field increment through most of the troposphere; the coupling is weakest in the upper troposphere around model level 31 (Fig. 10b). In the stratosphere, both height and wind increments have larger scales. Wind increments are zonally elongated, but the balanced height field increments are not (Fig. 10c). The maximal magnitude of the balanced height field increment at model level 15 is about 3.6 times larger than at model level 43; corresponding potential energy contribution is increased from 2% at level 43 up to 7% at level 15. The balance at the equator is dominated by the Kelvin waves, but further off the equator height field increments become geostrophic and zonally elongated (Fig. 10d), in agreement with quasi-geostrophic theory. By comparing Fig. 10d with Fig. 10b it can also be seen that equator-centred modes reduce the magnitude of the balanced height field to a large extent.

An increase of the amplitude of the balanced height increments with altitude is largest for a meridional wind observations (Fig. 11). When the observation is centred at the equator (Fig. 11a-c), the shape of the increments resembles an MRG wave. With observation located at 10°N, the increments appear nearly geostrophic (Fig. 11d) and very similar to the mid-latitude case (e.g. Courtier *et al.* 1998), except for a broader zonal scale. An enlargement of the horizontal scale of a height field increment with altitude is significant (notice that the isoline spacing in Fig. 11c-d is four times larger than in Fig. 11a-b). There is a factor of eleven in the magnitude of the increment between model levels 43 and 15. The potential energy content increases from zero at level 43 up to around 3% of the total energy at level 15.

4.2 Sensitivity experiments

In ŽGK it was shown that Kelvin waves largely determine the characteristics 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 with the ECMWF variance spectra. In the experiments shown in Fig. 12, the Kelvin waves are removed from the spectrum in such a way that their variance at each zonal wave number is equally distributed among other modes. Resulting increments at 500 hPa (Fig. 12a, to be compared with Fig. 9b) are still centred at the equator, but the balanced winds have changed sign. A new shape includes wings of negative correlations south and north of the equator and the whole structure is closest to an n=1 WEIG. The EEIG, present to a similar degree in the variance spectrum, would produce the opposite height-zonal wind balance at the equator. However, correlations due to ER modes act in the same direction as WEIG modes. At other levels, the structures appear quite similar.

Kelvin and MRG modes are trapped motion centred at the equator; their impact on the mass-wind coupling has previously been suggested also by other authors (Parrish 1988; Daley 1993; Daley 1996). On the other hand, it has been more difficult to estimate the impact of EIG waves. Their meridional structure is more complex while their role in the large-scale tropical dynamics is less clear, although some observational evidence has been provided (e.g. Wheeler and Kiladis 1999; Clayson *et al.* 2001) and theoretical scaling arguments (Browning *et al.* 2000) confirm their physical relevance.

Fig. 12b indicates that EIG waves have an important influence on the tropical background-error correlations. In

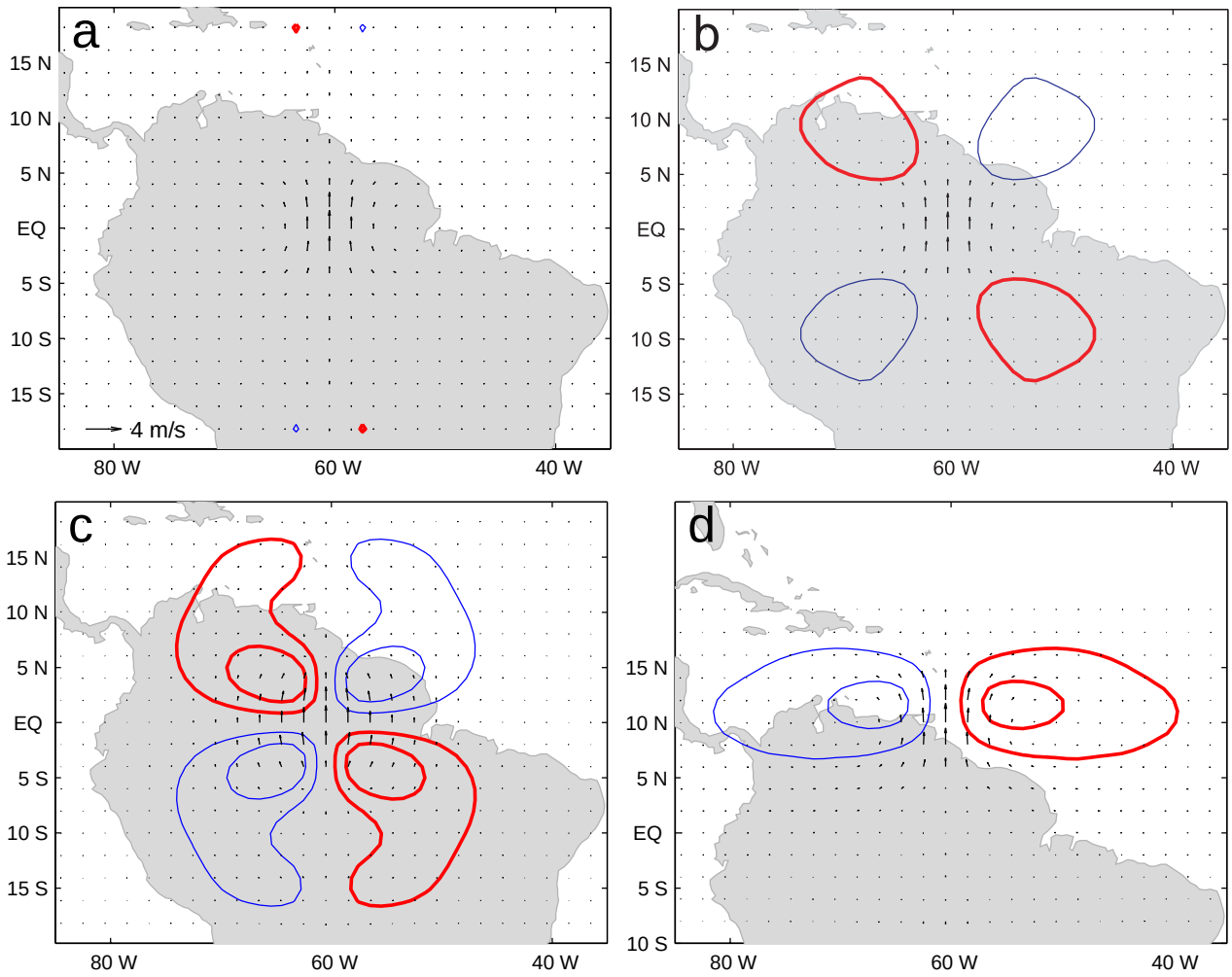

Figure 11: As Fig. 9, but for a single northerly wind observation. The wind observations is 4 ms^{-1}. Isoline spacing is ±0.05 m in (a) and (b), and ±0.2 m in (c) and (d).

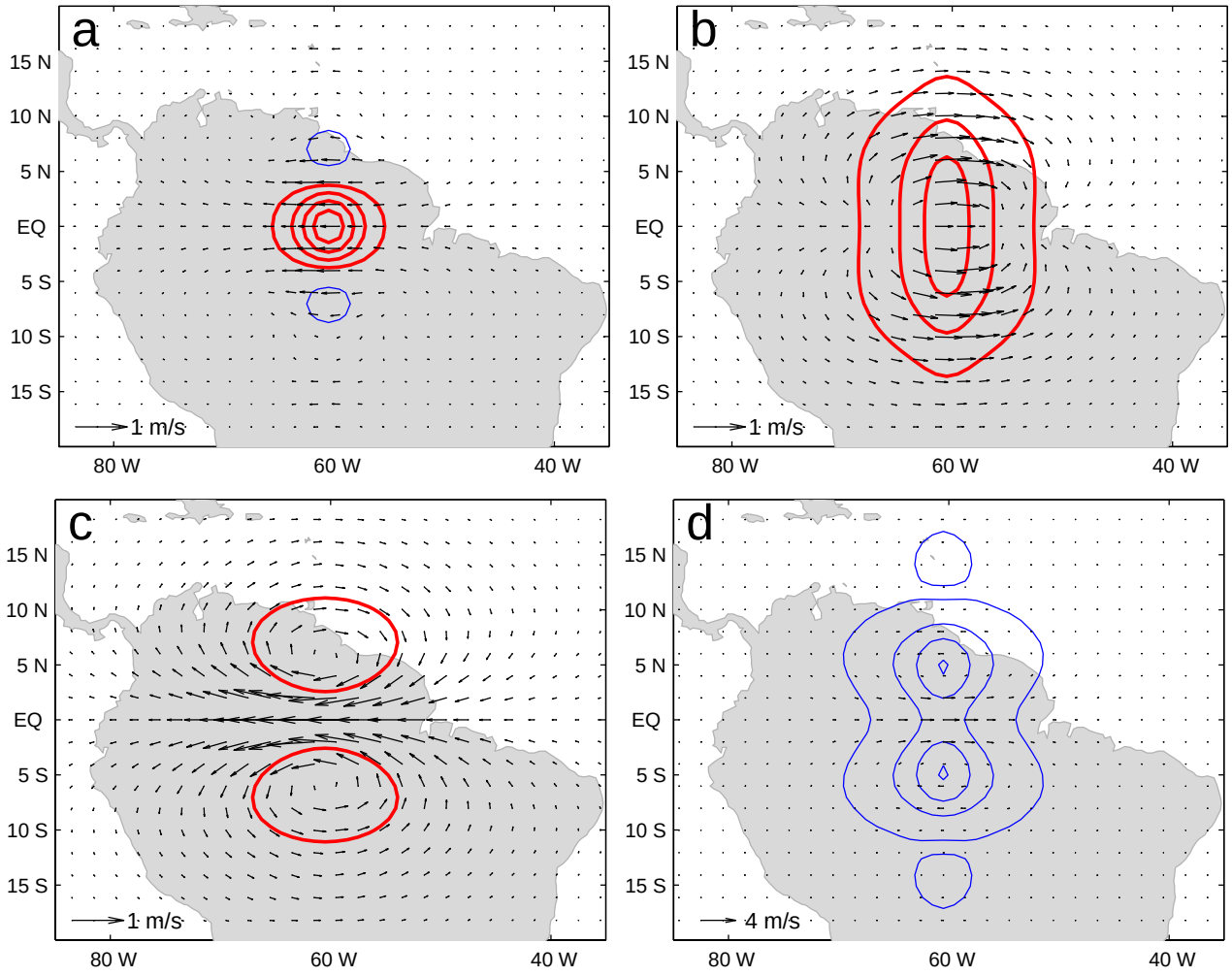

Figure 12: As Fig. 9, but for a height observation (a-c) and a zonal wind observation (d) at model level 39. (a,d) without Kelvin wave variance, (b) without EIG wave variance, (c) without Kelvin and EIG wave variance. Isoline spacing is ±0.5 m in (a-c), and ±0.2 m in (d).

this experiment, EIG variance in each *k* and *n* has been removed and distributed to the corresponding ER mode. In the case of a height field observation, the meridional scale of increments is increased 2-3 times at all altitudes. Balanced zonal winds are stronger and display maxima off the equator. The prevailing structure is still a Kelvin wave. When both Kelvin and EIG wave are removed from the spectrum, analysis increments become rotational (Fig. 12c). The amplitude of height-field increments is reduced by factor three and increments are placed off the equator. Along the equator, strong easterly winds are produced, compared to Fig. 9b.

Without the Kelvin-wave balance in the background-error term, increments due to a zonal wind observation (Fig. 12d) have a structure of an n=1 ER wave throughout the atmosphere. The horizontal scale of balanced height-field increments increases upwards more than in Fig. 10. In other words, 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 influence of the EIG modes is similar to that of the Kelvin modes, but weaker.

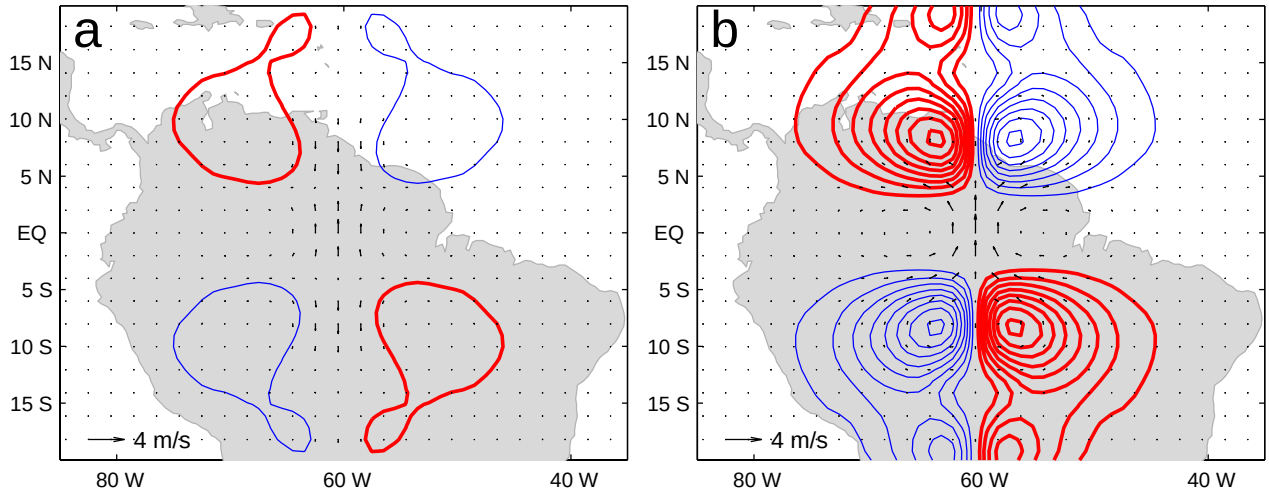

Figure 13: As Fig. 11, but for model level 39. (a) without MRG wave variance, (b) without EIG wave variance. Isoline spacing is ±0.05 m.

Removal of the MRG wave variance from the spectrum does not produce a dramatic change similar to that 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 counter flow produced off the equator (Fig. 13a, to be compared with Fig. 11b). However, EIG modes exhibit a major impact by reducing the meridional correlation scales and the magnitudes of the balanced height-field increments (Fig. 13b).

Having seen these large sensitivities, it is not surprising that the progress with balanced tropical data assimilation has been slow, and that the best treatment so far in global NWP systems has been to consider it a univariate problem!

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 with an ensemble of data assimilation. It was found that the equatorial waves (Equatorial Rossby, Equatorial Inertio-Gravity, Kelvin and Mixed Rossby-Gravity) explain a substantial fraction (60-70%) of the tropical large-scale variability, which is of consequence

for tropical data assimilation. Using a mode-based background-error covariance model, idealized single-observation analysis experiments were performed, which illustrate the main features of balanced analysis increments at various horizontal levels in the tropics.

Among the various tropical modes, the largest percentage of variance pertains to ER waves (30-40%) and to EIG waves (20-30% each for the westward- and eastward-propagating modes). The vertical correlations for ER waves display characteristics similar to their extra-tropical counterparts: correlations narrow towards shorter scales and in the stratosphere. On the contrary, EIG mode correlations do not display vertical correlation broadening with increased zonal scale. Kelvin and MRG modes explain a relatively small fraction of the variance in the troposphere. However, 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 in determining the characteristics of the horizontal correlations near the equator. This was suggested on the basis of idealized experiments (ŽGK) and here it has been confirmed 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 show a marked reduction in the stratosphere, which is most likely related to the strong eastward phase of the quasi-biennial oscillation (QBO) in the study period. The easterly winds suppress the upward propagation of westerly modes to the stratopshere. The impact of QBO on stratospheric correlations as well as the fact that the largest part of tropospheric errors is associated with the intertropical convergence zone (ITCZ) suggests that the background-error statistics should vary in the course of the year. The concentration of errors along the ITCZ, i.e. asymmetric with respect to the equator, has also been found to be the main problem of a methodology based on the 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 like the Rossby deformation radius. However, quasi-geostrophic considerations do not take into account the possible influence of non-rotational tropical modes on the horizontal correlations near the equator. The covariance model derived in this paper 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 smaller than that which would be obtained from the quasi-geostrophic approach or from globally averaged error statistics. It was also found that correlations are vertically asymmetric. Both features can be associated with deep tropical convection acting as a generator of equatorial wave motion.

The background error covariance model in the ECMWF's 4D-Var scheme is effectively univariate near the equator, as are many other global data assimilation schemes. The 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 over-estimated Hadley circulation observed in ECMWF operational analysis and in ECMWF's 40-year re-analysis (ERA-40) (Uppala, 2001). Excessive analysed amounts of gravity waves in the lower stratosphere/upper troposphere may contribute to an over-estimation of the Brewer-Dobson circulation also observed in operational analyses and in ERA-40. An approach that combines tropical and mid-latitude dynamics and produces balanced increments globally is required for further progress.

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 and Adrian Simmons (ECMWF) for reading the paper and their relevant comments. We also thank Agathe Untch (ECMWF) for discussions of the results, and for pointing out the role of the QBO.

Appendix A: 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 $\mathbf{x}^a$ given a background $\mathbf{x}^b$ and a set of observations represented by the vector $\mathbf{y}$:

$$\mathbf{x}^a = \mathbf{f}(\mathbf{x}^b, \mathbf{y}). \tag{5}$$

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:

$$\mathbf{x}^f(T) = \mathscr{M}_T(\mathbf{x}^a) \tag{6}$$

The analysis, background and observation errors are

$$\varepsilon^a = \mathbf{x}^a - \mathbf{x}^t \tag{7}$$
$$\varepsilon^b = \mathbf{x}^b - \mathbf{x}^t \tag{8}$$
$$\varepsilon^o = \mathbf{y} - \mathbf{y}^t \tag{9}$$

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

$$\varepsilon^f(T) = \mathbf{x}^f(T) - \mathbf{x}^t(T) \tag{10}$$

where $\mathbf{x}^t(T)$ 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 $\varepsilon^m(T)$ the error of a forecast made with perfect initial conditions:

$$\varepsilon^s = \mathbf{f}(\mathbf{x}^t, \mathbf{y}^t) - \mathbf{x}^t \tag{11}$$
$$\varepsilon^m(T) = \mathscr{M}_T(\mathbf{x}^t) - \mathbf{x}^t(T) \tag{12}$$

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

$$\varepsilon^a = \frac{\partial \mathbf{f}(\mathbf{x}^t, \mathbf{y}^t)}{\partial \mathbf{x}^b}\varepsilon^b + \frac{\partial \mathbf{f}(\mathbf{x}^t, \mathbf{y}^t)}{\partial \mathbf{y}}\varepsilon^o + \varepsilon^s + O(\varepsilon^2). \tag{13}$$

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

$$\varepsilon^f(T) = \frac{\partial \mathscr{M}_T(\mathbf{x}^t)}{\partial \mathbf{x}}\varepsilon^a + \varepsilon^m(T) + O(\varepsilon^2). \tag{14}$$

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

$$\widehat{\mathbf{x}^a} = \mathbf{f}(\mathbf{x}^b + \zeta, \mathbf{y} + \eta) + \omega. \tag{15}$$

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

$$\widehat{\varepsilon^a} = \frac{\partial \mathbf{f}(\mathbf{x}^t, \mathbf{y}^t)}{\partial \mathbf{x}^b}(\varepsilon^b + \zeta) + \frac{\partial \mathbf{f}(\mathbf{x}^t, \mathbf{y}^t)}{\partial \mathbf{y}}(\varepsilon^o + \eta) + (\varepsilon^s + \omega) + O(\varepsilon^2). \tag{16}$$

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

$$\delta\widehat{\mathbf{x}}^a = \frac{\partial \mathbf{f}(\mathbf{x}^t, \mathbf{y}^t)}{\partial \mathbf{x}^b}\delta\zeta + \frac{\partial \mathbf{f}(\mathbf{x}^t, \mathbf{y}^t)}{\partial \mathbf{y}}\delta\eta + \delta\omega + O(\varepsilon^2). \tag{17}$$

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 $\left(\varepsilon^b, \varepsilon^o, \varepsilon^s\right)^{\mathrm{T}}$. This may be achieved by perturbing each analysis with different random errors drawn from the distribution of $\left(\varepsilon^b, \varepsilon^o, \varepsilon^s\right)^{\mathrm{T}}$. Comparing Eq. (13) and Eq. (17), we see that, to first order in ε, the covariance matrix for $\delta\widehat{\mathbf{x}}^a$ is twice the covariance matrix for the analysis error ε^a of the unperturbed analysis.

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

$$\widehat{\mathbf{x}}^f = \mathscr{M}_T(\widehat{\mathbf{x}}^a) + \xi \tag{18}$$

The error of this forecast is:

$$\widehat{\varepsilon}^f(T) = \frac{\partial \mathscr{M}_T(\mathbf{x}^t)}{\partial \mathbf{x}}\widehat{\varepsilon}^a + \xi + \varepsilon^m(T) + O(\varepsilon^2). \tag{19}$$

By a similar argument to that gven above, the difference between two forecasts made from differently-perturbed analyses is:

$$\delta\widehat{\mathbf{x}}^f(T) = \frac{\partial \mathscr{M}_T(\mathbf{x}^t)}{\partial \mathbf{x}}\delta\widehat{\mathbf{x}}^a + \delta\xi + O(\varepsilon^2). \tag{20}$$

Comparing Eq. (14) and Eq. (20), we see that if the vector $\left(\delta\widehat{\mathbf{x}}^a, \delta\xi\right)^{\mathrm{T}}$ has covariance matrix equal to twice that of $(\varepsilon^a, \varepsilon^m(T))^{\mathrm{T}}$, then to first order, $\delta\widehat{\mathbf{x}}^f(T)$ will have 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 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.* 2004).

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 6 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 observed 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., 1992: A continuous three-dimensional model of short-range forecast error covariances. *Tellus,* **44A,** 217–235

Browning, G.L., Kresiss, 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

Clayson, C. A., Strahl, B. and Schrage, J., 2002: 2-3 day convective variability in the tropical Western Pacific. *Mon. Wea. 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). Part I: Formulation. *Q. J. R. Meteorol. Soc.,* **124,** 1783–1807

Daley, R., 1991: *Atmospheric data analysis.* Cambridge University Press, Cambridge

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

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

Derber, J. C. and Bouttier, F., 1999: Formulation 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. Wea. 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, U.K.

Fisher, M., 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, U.K.

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, U.K.

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

Lindzen, R.S. and Holton, J.R., 1968: A theory of the quasi-biennial oscillation. *J. Atmos. Sci.*, **25,** 1095–1107

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. Amer. Meteor. 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. Wea. Rev.*, **117,** 1941–1953

Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. *J. Meteor. Soc. Japan*, **44,** 25–42

Parrish, D., 1988: 'The introduction of Hough functions into optimum interpolation'. Pp. 191–196 in Proceedings of the Eighth Conf. on NWP, Am. Meteorol. Soc., Boston, Mass.

Parrish, D. F. and Derber, J. C., 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. 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.P., 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). Part 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, Amer. Meteor. Soc., Ft. Lauderdale, FL

Untch, A. and Hortal, M., 2003: A finite-element scheme for the vertical discretization in the semi-Lagrangian version of the ECMWF forecast model. *ECMWF Tech. Memo.*, **382,** Pp. 27

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

Vialard, J., Vitart, F., Balmaseda, M.A., Stockdale, T.N. and Anderson, D.L.T., 2004: An ensemble generation method for seasonal forecasting with an ocean-atmosphere coupled model. Submitted to Mon. Wea. Rev.

(accepted pending minor revisions)

Wergen, W., 1988: The diabatic ECMWF normal mode initialization scheme. *Beitr. Phys. Atmosph.*, **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., 2004a: Variational data assimilation in the tropics: the impact of a background error constraint. *Q. J. R. Meteorol. Soc.*, **130,** 103–125

Žagar, N., Gustafsson, N. and Källén, E., 2004b: Dynamical response of equatorial waves in four-dimensional variational data assimilation. *Tellus*, **56A,** 29–46