

Dynamical response of equatorial waves in four-dimensional variational data assimilation

By NEDJELJKA ŽAGAR*, NILS GUSTAFSSON and ERLAND KÄLLÉN, *Department of Meteorology, Stockholm University, SE-106 91 Stockholm, Sweden*

(Manuscript received 17 January 2003; in final form 27 June 2003)

ABSTRACT

In this study we question the relative importance of direct wind measurements in the tropics by investigating limits of four-dimensional variational assimilation (4D-Var) in the tropics when only wind or mass field observations are available. Typically observed equatorial wave motion fields (Kelvin, mixed Rossby-gravity and $n = 1$ equatorial Rossby waves) are assimilated in a non-linear shallow water model. Perfect observations on the full model grid are utilized and no background error term is used. The results illustrate limits of 4D-Var with only one type of information, in particular mass field information. First, there is a limit of information available through the internal model dynamics. This limit is defined by the length of the assimilation window, in relation to the characteristics of the motion being assimilated. Secondly, there is a limit related to the type of observations used. In all cases of assimilation of wind field data, two or three time instants with observations are sufficient to recover the mass field, independent of the length of the assimilation time window. Assimilation of mass field data, on the other hand, although capable of wind field reconstruction, is much more dependent on the dynamical properties of the assimilated motion system. Assimilating height information is less efficient, and the divergent part of the wind field is always recovered first and more completely than its rotational part.

1. Introduction

The data assimilation problem is a problem of completeness and uniqueness (Bube and Ghil, 1981). The problem of completeness, related to the lack of observational information, is particularly important in the tropics where measurements, especially wind measurements, are sparse. At the same time, measurement of the global wind field is widely acknowledged as a key element for advancement in the understanding and prediction of weather and climate (e.g., Baker et al., 1995). The lack of global tropospheric wind fields has been recognized as a considerable limitation for scientific diagnosis of large-scale processes from the divergent component of the wind field, important in particular for the hydrological cycle.

The lack of large-scale wind information in the current Global Observing System is critical for the application of variational assimilation schemes, in which the data analysis problem is solved globally. As a consequence, significant differences between the National Center for Environmental Prediction–National Center for Atmospheric Research (NCEP/NCAR) and European Centre for Medium Range Weather Forecast (ECMWF) re-analyses

are located in the tropics, throughout the vertical depth of the atmosphere (Kistler et al., 2001). For example, mean re-analysis differences in the wind fields are found to be of the order of their full mean variations. This uncertainty is furthermore transported to ocean models, in which the poorly known surface fluxes usually provided by the atmospheric model, together with the parameterization of physical processes, are main limiting factors for prediction capabilities (Hao and Ghil, 1994).

Besides the imbalance in their content, wind information and mass information are not equally important in the atmosphere. The question of the relative usefulness of mass and wind observations is the problem of adjustment in a rotating fluid, addressed by Rossby (1937), Rossby (1938), Cahn (1945) and many others. During the process of adjustment, energy initially contained in a perturbation is redistributed until ultimately developing into a balanced state, characterized by a particular distribution of potential and kinetic energy. The partition of the energy in the final state between the potential and kinetic energy forms has been shown (e.g., Haltiner and Williams, 1980, Chapter 2) to depend on the initial perturbation size, the average fluid depth and the value of the Coriolis parameter f . The theory shows that wind information, i.e. kinetic energy, dominates the adjusted state for smaller perturbations in the horizontal and for lower latitudes, while potential energy, i.e. mass information, dominates at larger horizontal scales and at high latitudes.

*Corresponding author.
e-mail: nedjeljka@misu.su.se

The present study addresses the question of the relative importance of wind and mass information in four-dimensional variational data assimilation (4D-Var) in the tropics. The main question we ask here is the following one: having access to the measured tropospheric wind field, to what extent do we need data about the mass field? The same question can be addressed in the reverse way: to what extent can the wind field in the tropics be determined from an accurately observed time history of the mass field by the aid of 4D-Var? In other words, how much physically relevant information could be transferred from the mass field to the wind field by using 4D-Var in the tropics?

In spite of the relatively smaller usefulness of a single time insertion of mass field information, one may argue that continuous assimilation of accurate mass field data would force the model's wind field to adjust towards a balance with the mass field. By such an argument one would expect that much more information about the wind field in the tropics can be obtained from the mass field measurements than the simple geostrophic adjustment theory would indicate. The question was considered by Bube and Ghil (1981), Talagrand (1981) and Daley (1991) in the context of continuous data assimilation. Bube and Ghil (1981) provided a mathematical proof that, in principle, a complete time history of mass information (geopotential) ensures knowledge of the wind field for the linearized shallow water system on an f -plane. By numerical experimentation it was found that the non-rotational part of the wind field was almost always reconstructed first (Talagrand, 1981). The reconstruction of the complete wind field from synoptic mass data deteriorated significantly as the equator was approached (Daley, 1991, Chapter 13), i.e. the vorticity reconstruction was very poor for small values of f (Talagrand, 1981). It was also shown (Ghil, 1980) that in the non-linear case the problem leads to equations equivalent to the non-linear balance equation.

These results may not be strictly relevant for the data assimilation problem in the tropics, since the linear problem on the f -plane becomes singular at the equator, while the ellipticity condition for the non-linear balance equation restricts its usefulness at low latitudes. Another deficiency of studies on the f -plane is that an important part of the solution might be missing. Above mentioned studies suffer from the exclusion of two special spherical modes, namely Kelvin and mixed Rossby-gravity (MRG) waves, confined mainly to the tropics (Matsuno, 1966). Both types of motion appear as special solutions of linear shallow water equations on the equatorial β -plane. Depending on the equivalent depth H and the zonal wavenumber, they have characteristics similar to either equatorial Rossby (ER) or inertia-gravity waves (EIG). They can be excited by latent heat release in deep tropical convection. As they propagate through the equatorial waveguide, they communicate effects of the heat sources far away along the equator and through their vertical energy and momentum transport they play an important role for the general circulation (Holton, 1992).

4D-Var, rather than a simple forward continuous data assimilation, is a more appropriate framework to address the question of the relative usefulness of mass and wind data for the purpose of Numerical Weather Prediction (NWP). In the 4D-Var (e.g., Le Dimet and Talagrand, 1986; Talagrand and Courtier, 1987) the influence of observations is spread both forward and backward in time in order to find a model solution that represents an optimal balance between observations and a priori information. Key a priori information is provided by the governing equations. Additional a priori information is the background (first guess) state for 4D-Var and its error statistics. The background error term is not considered in this study. It has important filtering properties; hence, it is not applied in the present study that aims at assessing the basic properties of the adjustment process in 4D-Var in the tropics. Our primary goal is to investigate the relative usefulness of the height and wind field observations in the tropics in more detail compared to previous studies on the f -plane and by using wave motions typical for the tropical atmosphere.

A significant portion of the tropical time-space variability, as found in satellite-observed outgoing longwave radiation (OLR) data, can be described in terms of equatorially trapped wave solutions of linear shallow water theory (Wheeler and Kiladis, 1999; Wheeler et al., 2000). The most frequently identified waves are Kelvin waves, MRG waves, meridional mode $n = 1$ ER wave and $n = 1, 2$ westward-moving EIG waves (WEIG). In this study, we utilize the linear theory of equatorial waves for carrying out "identical twin" Observing System Simulation Experiments (OSSE) in the tropics. We consider equatorial waves as "dry", i.e. freely propagating through the atmosphere governed by dry dynamics only and regardless of the mechanism by which they were generated. Perfect mass and wind observations of the equatorial waves are assimilated by 4D-Var. During the 4D-Var assimilation of either mass or wind field information, the other field is adjusted through the model dynamics. We will examine success or failure of this adjustment for typical equatorial waves by using a non-linear shallow water system.

Further details about the numerical framework are described in Section 2 of the paper. In order to understand aspects of the adjustment process related to interaction between the observations and the model in 4D-Var experiments, certain adjustment experiments in the tropics are described in Section 3. 4D-Var assimilation experiments are described and discussed in Section 4. Further discussion and conclusions are given in Sections 5 and 6, respectively.

2. Numerical model framework

The numerical model solves a system of shallow water equations for the eastward and northward velocity components u and v , respectively, and the height of the free surface (or equivalent depth) h on a limited area of a rotating sphere:

$$\frac{du}{dt} - fv = -g \frac{\partial h}{\partial x} - \epsilon_u u \quad (1)$$

$$\frac{dv}{dt} + fu = -g \frac{\partial h}{\partial y} - \epsilon_v v \quad (2)$$

$$\frac{dh}{dt} + h \nabla \cdot \mathbf{V} = Q - \epsilon_h h. \quad (3)$$

Here, the horizontal wind vector is $\mathbf{V} = (u, v)$, g is gravity, f stands for the Coriolis parameter, ϵ for the friction, and Q is the space and time variable source.

A spectral transform formulation using Fourier series is utilized for numerical solution of the system (1–3). The equations for a single wave component (k, l) read:

$$\begin{aligned} \frac{d\hat{u}_{kl}}{dt} = F_{kl} & \left\{ -F^{-1}(\hat{u})m_x^{-1}F^{-1}(ik'\hat{u}) \right. \\ & - F^{-1}(\hat{v})m_y^{-1}F^{-1}(il'\hat{u}) \\ & \left. + fF^{-1}(\hat{v}) - gm_x^{-1}F^{-1}(ik'\hat{h}) - \epsilon_u F^{-1}(\hat{u}) \right\} \end{aligned}$$

$$\begin{aligned} \frac{d\hat{v}_{kl}}{dt} = F_{kl} & \left\{ -F^{-1}(\hat{u})m_x^{-1}F^{-1}(ik'\hat{v}) \right. \\ & - F^{-1}(\hat{v})m_y^{-1}F^{-1}(il'\hat{v}) \\ & \left. - fF^{-1}(\hat{u}) - gm_y^{-1}F^{-1}(il'\hat{h}) - \epsilon_v F^{-1}(\hat{v}) \right\} \end{aligned}$$

$$\begin{aligned} \frac{d\hat{h}_{kl}}{dt} = F_{kl} & \left\{ -F^{-1}(\hat{u})m_x^{-1}F^{-1}(ik'\hat{h}) \right. \\ & - F^{-1}(\hat{v})m_y^{-1}F^{-1}(il'\hat{h}) \\ & - F^{-1}(\hat{h})m_x^{-1}m_y^{-1} \cdot \left(F^{-1}(\hat{u}) \frac{\partial m_y}{\partial x} \right. \\ & + m_y F^{-1}(ik'\hat{u}) + F^{-1}(\hat{v}) \frac{\partial m_x}{\partial y} \\ & \left. \left. + m_x F^{-1}(il'\hat{v}) \right) - \epsilon_h F^{-1}(\hat{h}) + Q \right\}. \end{aligned}$$

Here $\hat{u}_{kl}$, $\hat{v}_{kl}$ and $\hat{h}_{kl}$ are the wave components for velocity components in the zonal and meridional direction and the height, respectively. A zonal wavenumber is denoted by k , while l stands for a meridional wave number. Parameters m_x and m_y represent the metric coefficients. A transformation between the spectral and physical space is carried out by the Fourier transform operator F : $\hat{a}_{kl} = Fa(x, y)$. For a two-dimensional field $a(x, y)$, F reads:

$$a(x, y) = \sum_{k=-N_K}^{N_K} \sum_{l=-N_L}^{N_L} \hat{a}_{kl} \times \exp \left\{ +2\pi i \left(\frac{kx}{N_X} + \frac{ly}{N_Y} \right) \right\} \quad (4)$$

$$\hat{a}_{kl} = \frac{1}{N_X N_Y} \sum_{x=0}^{N_X-1} \sum_{y=0}^{N_Y-1} a(x, y) \times \exp \left\{ -2\pi i \left(\frac{kx}{N_X} + \frac{ly}{N_Y} \right) \right\}. \quad (5)$$

The Fourier representation includes a so-called extension zone (Haugen and Machenhauer, 1993), used to make the space-dependent variables bi-periodic (Fig. 1). The bi-periodicity is

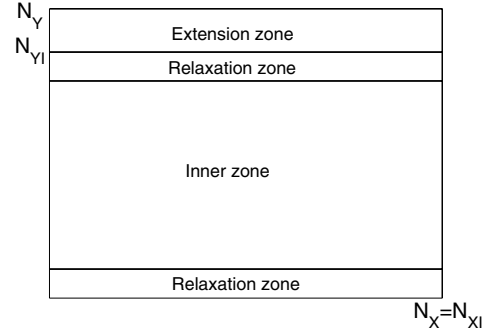


Fig 1. Organization of the integration area of a spectral limited area model.

achieved by extrapolation using sine and cosine functions, as in the spectral HIRLAM model (Gustafsson 1998). The numbers of the domain grid points in the zonal and meridional directions are N_X and N_Y , respectively. An elliptic truncation

$$\left(\frac{k}{N_K} \right)^2 + \left(\frac{l}{N_L} \right)^2 \leq 1$$

ensures a homogeneous and isotropic resolution over the whole model domain. The maximal wavenumbers in the zonal and meridional direction are given by N_K and N_L , respectively.

Fourth-order implicit numerical diffusion is applied to the spectral coefficients as the last operation at each time step to prevent an accumulation of energy at the smallest resolved scales. The adoption of constant map factors in this procedure is assumed to have a negligible effect. For a single wave component $\hat{a}_{kl}$, the equation to be solved is

$$\frac{\partial \hat{a}_{kl}}{\partial t} + K_{kl} \hat{a}_{kl} = 0$$

where

$$K_{kl} = K \left\{ \left(\frac{2\pi}{N_X} \frac{k}{\bar{m}_x} \right)^2 + \left(\frac{2\pi}{N_Y} \frac{l}{\bar{m}_y} \right)^2 \right\}^2.$$

Following Davies (1976), the time-dependent fields are relaxed towards the time-dependent boundary fields across a relaxation zone by using a space-dependent relaxation function. The weighting factor $\alpha(x, y)$ is given as

$$\alpha(x, y) = 1 - \tanh(r/2)$$

where $r = r(x, y)$ is the distance (in grid point units) from the grid point (x, y) to the boundary of the inner integration area, normalized by the width of an extension zone (in grid point units).

Explicit time-stepping is done by using a leap-frog scheme. In the experiments presented in this paper, the extension zone is only applied in the meridional direction while periodic boundary conditions are imposed on the eastern and western model boundaries. Since the equator is located in the centre of the inner

integration area, the number of physically significant points in the meridional direction, N_Y , is an odd number (Fig. 1).

A distinct advantage of the spectral approach for a data assimilation study is that the Fourier transform is a self-adjoint operator. For this purpose, the Fourier transform pair (4–5) is modified by the scale factors $\sqrt{N_X}$, $\sqrt{N_Y}$ in order to make it self-adjoint. In other words, the adjoint of the forward Fourier transform is identical to the inverse Fourier transform and vice versa.

The adjoint version of the shallow water model is needed to minimize a distance function J calculated over a certain time interval with observations distributed among P different times. If $\mathbf{y}_n$, $\mathbf{x}_n$ are the observations and the model state vectors, respectively, at time n , and $\mathbf{R}$ the observation error covariance matrix, then the distance function J to be minimized is

$$J = J_o = \frac{1}{2} \sum_{n=1}^P \left((\mathbf{y}_n - \mathbf{x}_n), \mathbf{R}^{-1} (\mathbf{y}_n - \mathbf{x}_n) \right).$$

The distance function J is usually defined as a sum of two terms $J = J_b + J_o$. Here the background term J_b measures the distance to a background model state. As discussed above, this study takes its value to be zero, which is in agreement with the assumption of perfect observations provided at every grid point. The observational error covariance matrix $\mathbf{R}$ contains three diagonal elements, namely the variance values for the wind components σ_u^2 , σ_v^2 and the height variance σ_h^2 . Since we never consider wind and mass observations simultaneously, the choice for σ_u , σ_v and σ_h is arbitrary. We choose $\mathbf{R}$ to be the identity matrix.

At time step n , the p -th update of the model state variable $\mathbf{x}$ is produced by the non-linear forecast model M , given the state variable at time step $n-1$:

$$\mathbf{x}_n = M(\mathbf{x}_{n-1}).$$

The initial state is given by $\mathbf{x}_0 = \mathbf{x}_b + \delta\mathbf{x}_p$, where $\mathbf{x}_b$ is the background model state and $\delta\mathbf{x}_p$ corresponds to the analysis increment at the p -th step of the minimization.

The gradient of the distance function at the time step n , $\nabla_{\delta\mathbf{x}_n} J$, is computed for each wave component. Using the Fourier series definition and the equality $\hat{a}_{-k-l} = \hat{a}_{kl}^*$ (where $*$ denotes the conjugate operator), contributions to the gradient at time instant n containing the observations are given by

$$\left(\nabla_{\mathbf{h}_n} J_{o,n} \right)_{k,l} = C(\mathbf{y}_n - \mathbf{h}_n) \sigma_h^{-2} e^{-i(kx+ly)}$$

$$\left(\nabla_{\mathbf{u}_n} J_{o,n} \right)_{k,l} = C(\mathbf{y}_n - \mathbf{u}_n) \sigma_u^{-2} e^{-i(kx+ly)}$$

$$\left(\nabla_{\mathbf{v}_n} J_{o,n} \right)_{k,l} = C(\mathbf{y}_n - \mathbf{v}_n) \sigma_v^{-2} e^{-i(kx+ly)}$$

where $C = 2$ for $(k, l) = (0, 0)$, otherwise $C = 4$.

The gradient of J with respect to the model initial condition $\mathbf{x}_0$ is obtained through the backward integration of the adjoint model $\mathbf{M}^T$

$$\mathbf{x}_{n-1}^{ad} = \mathbf{M}_{n-1}^T \mathbf{x}_n^{ad} + \nabla_{\delta\mathbf{x}_n} J$$

starting from the last time step of the assimilation window N and going to 1. At each time step this integration is forced by $\nabla_{\delta\mathbf{x}_n} J$. Here the adjoint variable of $\mathbf{x}$ is denoted $\mathbf{x}^{ad}$.

Minimization of J is carried out by a widely used minimization package M1QN3 (Gilbert and Lemaréchal, 1989). In our application, it uses conjugate gradient iterations to solve the minimization problem by calculating the gradient of the distance function J with respect to the model state increment $\delta\mathbf{x}$ until the predefined convergence criterion is met or until the maximally allowed number of iterations have been performed.

3. Adjustment to local perturbations at the equator

Each update of the model with observations during 4D-Var results in an adjustment between the mass and the wind field. In order to get some insight into the interaction between the observations and the model in 4D-Var, we first present results of adjustment experiments for height and wind perturbations centred at the equator.

Data assimilation schemes are usually formulated in such a way that the structure of an assimilation increment corresponds closely to the result of the adjustment process. For a mid-latitude single geopotential or wind observation, assimilation increments are forced to obey a close geostrophic balance between the mass and the wind field (e.g., Courtier et al., 1998; Gustafsson et al., 2001). The horizontal structure of the increments largely resembles those of the adjusted states from analytical and numerical solutions of the linearized shallow water equation on an f -plane (Barwell and Bromley, 1988). Similar solutions are more complicated for the tropics for which there is no unified theoretical framework such as the quasi-geostrophic theory for mid-latitude motion systems. Besides that, the concept of slow modes, usually applied in data assimilation procedures, is more difficult to apply in the tropics, where Kelvin and MRG waves span the frequency gap between Rossby and inertia-gravity waves, present in mid-latitudes. Finally, equivalent depths of observed large-scale tropical motion systems are small (e.g., Wheeler et al., 2000), so that the dynamics of Rossby and gravity waves are more difficult to separate through a simple frequency criterion.

In order to illustrate the differences to mid-latitude adjustment, we choose to present the model response to the three perturbations considered by Barwell and Bromley (1988), but located at the equator. The perturbations include a height perturbation, a non-divergent wind perturbation and a zonal wind perturbation. The form of all perturbations is based on a Gaussian function. Wave amplitudes have different values as compared to those in Barwell and Bromley (1988), but we are mainly interested in the structures and their differences between the mid-latitudes and the tropics. Adjustment experiments are carried out with the same model setup as used for the assimilation experiments in the next section.

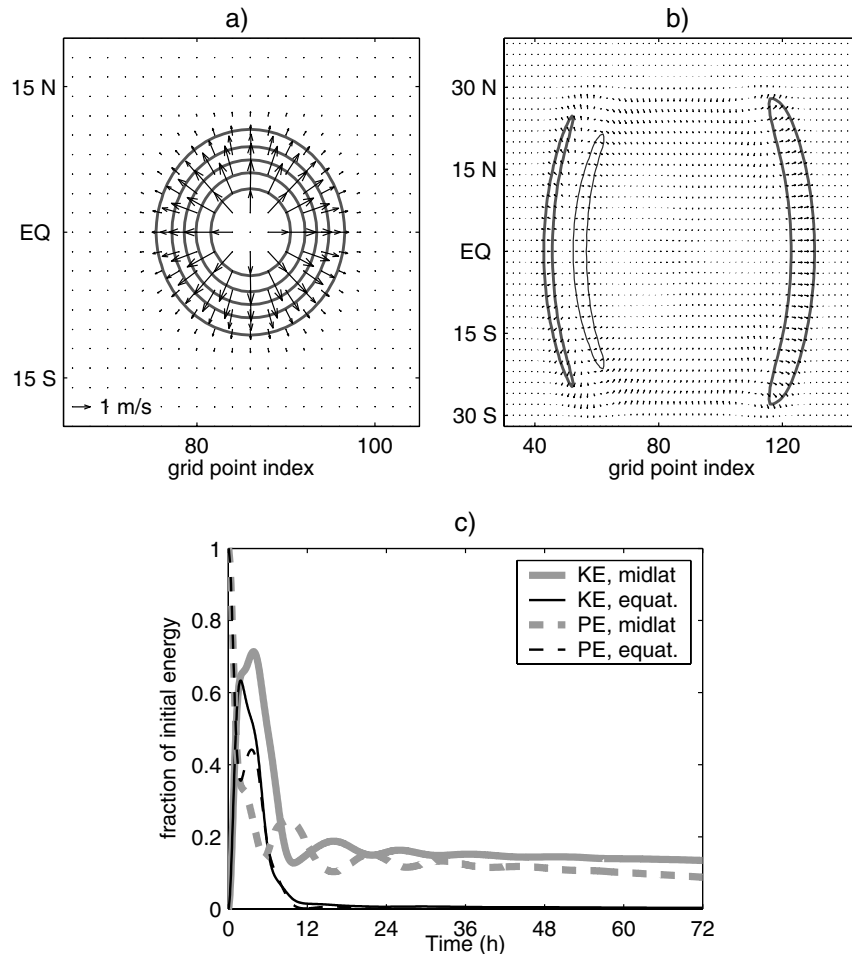


Fig 2. Response of the shallow water model to an idealized, Gaussian shaped mass perturbation located on the equatorial β -plane after 1 h (a) and 12 h (b). The equivalent depth is 1000 m, the magnitude of the perturbation at the centre 100 m and the e -folding distance 6.6° . Isoline spacing is every 4 m, starting from ± 4 m. Thick lines correspond to positive values, while thin lines are used for negative values. Wind vectors are given in every second grid point in both directions. Only a part of the computational model domain is shown. (c) Time evolution of the kinetic and potential energies within the area initially occupied by the height field perturbation, scaled by the initial energy. Thick lines apply to mid-latitude β -plane solution, while thin lines are for the equatorial β -plane solution. Dashed lines apply to the potential and full lines to the kinetic energy.

The difference to the mid-latitude case is mostly seen in the adjustment of a height perturbation (Fig. 2). After 1 h there is a rapid initial change due to dispersion of gravity waves, much the same as that on a mid-latitude β -plane, except for a lack of the Coriolis force effect (Fig. 2a). The fastest equatorial mode, the Kelvin wave, escapes from the other slow modes excited during the adjustment process and propagates eastward (Fig. 2b). The EIG wave is propagating westward. After 12 h, which is the time interval usually used as the 4D-Var assimilation window for NWP, there is basically no mass perturbation left at the place of the initial perturbation. A remaining balanced structure in the wind and mass fields after 48 h (figure not shown) has a shape of the lowest ER mode. However, its amplitudes are small; thus, it is energetically negligible. This is further illustrated in Fig. 2c, which compares the energy redistribution during the adjustment process with that of the same perturbation located on a

mid-latitude β -plane centred on 45°N . The adjusted state on the mid-latitude β -plane has preserved about one-third of its initial energy, about half of that as potential energy. Adjustment of the same mass perturbation at the equator results in practically all potential energy being dispersed away. A negligible amount of kinetic energy is left at the place of the initial disturbance. The energy content of the residual structure depends, of course, on the perturbation size and the equivalent depth used, but in case of a mass perturbation it is always relatively small. With decreasing equivalent depth, the energy content of the weak balanced residual flow increases; in addition, the adjustment process is slower, with energy channeled along the equator by the action of Kelvin and ER waves.

Next we look at the adjustment of a perturbation in the zonal wind at the equator (Fig. 3). Not affected by the Coriolis force, gravity waves induced by the adjustment propagate

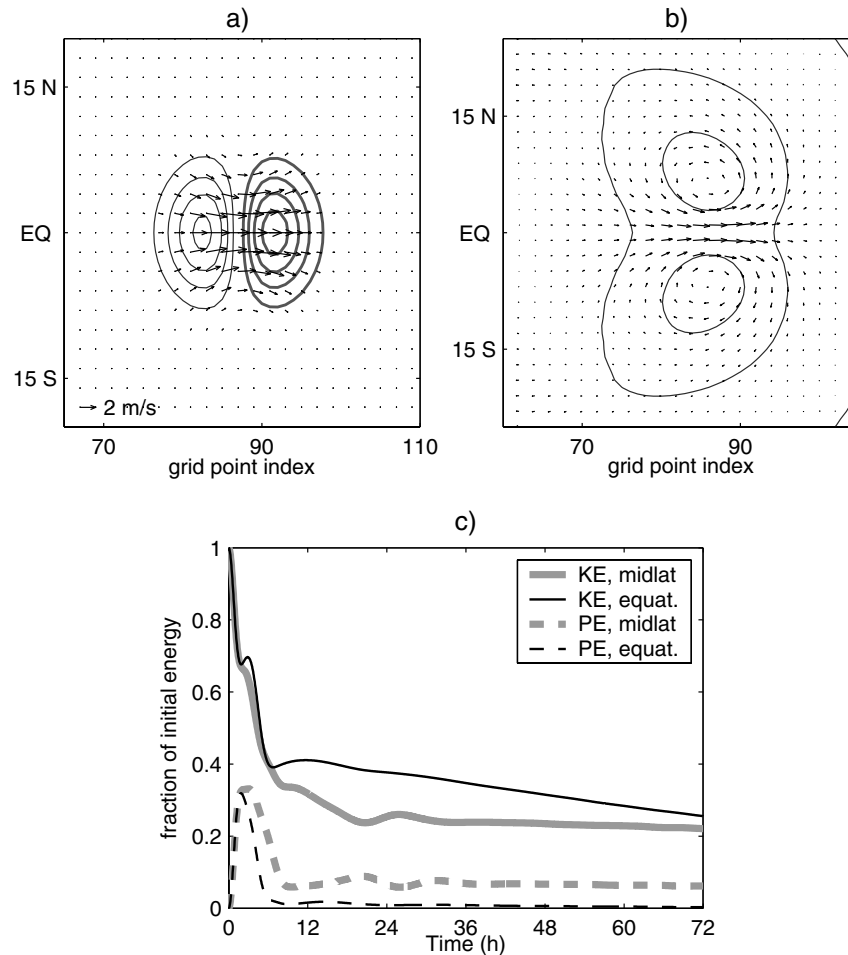


Fig. 3. As Fig. 2, but for an initial zonal wind perturbation. The magnitude of the perturbation at the centre is 5 m s^{-1} . Isolines spacing in (b) is 1 m, starting from $\pm 1 \text{ m}$. (c) As Fig. 2c, but for an initial zonal wind perturbation.

symmetrically eastward and westward (Fig. 3a). After 12 h, the potential energy content of the remaining balanced perturbation is less than 1% of the initial energy. The shape of the balanced structure resembles the $n = 1$ ER wave (Fig. 3b). The structure of the adjusted wind field at the equator looks very much the same as in mid-latitudes, but a larger part of the initial kinetic energy is retained, 40% as compared to about 30% in mid-latitudes (Fig. 3c). While the adjusted mid-latitude β -plane solution remains almost stationary through longer integrations, a slow tropical adjustment continues; the balanced structure of the ER wave propagates slowly westward with its kinetic energy continuously decreasing.

The adjustment of the mass field to the non-divergent wind perturbation (Fig. 4) is fast and efficient in the sense that very little energy is lost due to excitement of gravity waves. In the mid-latitude case, small-amplitude gravity waves propagate away fast, mainly northward and southward, i.e. perpendicular to the direction of the wind at the centre of the perturbation. The same

perturbation placed on the equator only excites eastward and westward propagating gravity waves (Fig. 4a). The balanced structure appears as the lowest ER mode (Fig. 4b). The amplitude of the balanced mass field at the equator is significantly reduced as compared to mid-latitudes (Fig. 4c). In the course of time the balanced states slowly reduce their amplitudes due to horizontal diffusion. Figure 4c also illustrates that the adjustment of a non-divergent wind perturbation is very little affected by the change of geographic latitude. What makes the adjustment process in the tropics “somewhat special” (Gill, 1982) is the change of sign of f across the equator and the equatorial waveguide effect.

4. 4D-Var assimilation experiments

Now we will use the shallow water system described in Section 2 for studying how geopotential and wind observations determine equatorial waves and how the observed information propagates

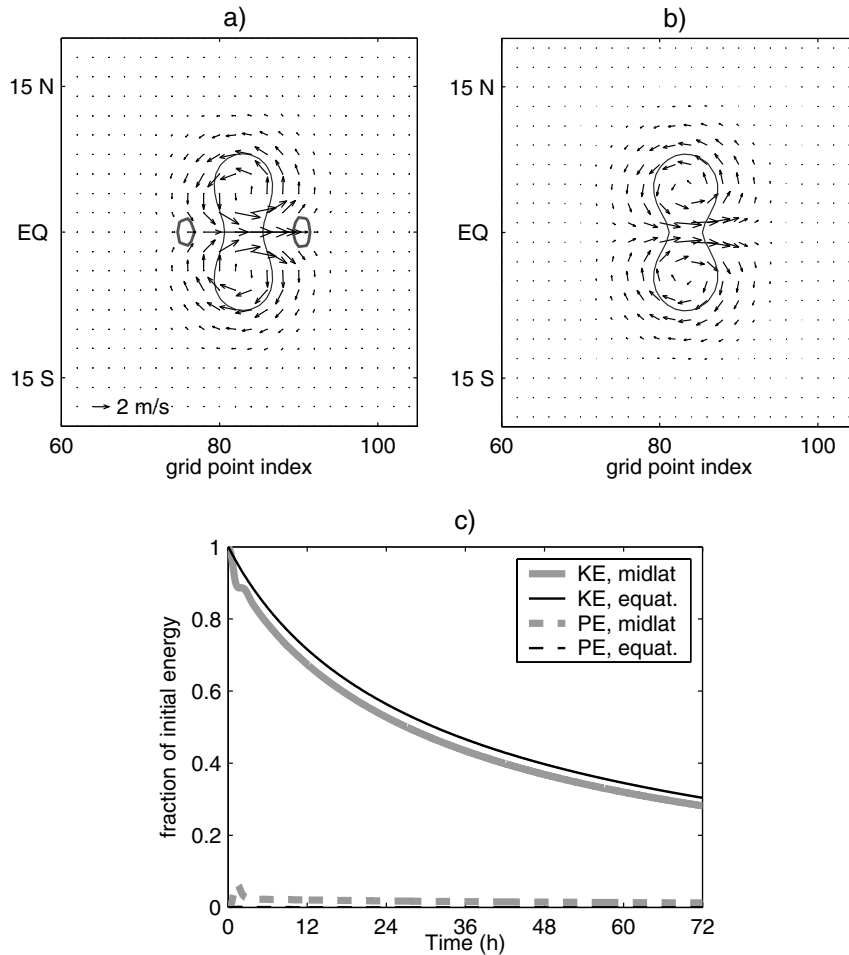


Fig 4. As Fig. 3, but for a non-divergent wind perturbation. The isoline spacing is every 2.5 m, starting from ± 0.5 m. (c) As Fig. 2c, but for a non-divergent wind perturbation.

within the system. A standard “identical twin” approach is used for the experiments. An initial state defined by the analytical expressions for the linear equatorial waves (Gill, 1982, Chapter 11) is propagated in time by a non-linear model to generate simulated observations. Observations of either geopotential or wind generated on the full model grid are then variationally assimilated in the same model to determine how much improvement is brought to the other variable through the internal model dynamical adjustment. During assimilation of only one type of the data, the adjustment between the mass and wind fields is taking place through the model dynamics between each update by observations.

We consider three tropical waves which are most frequently found in observations: the Kelvin wave, the MRG wave and the $n = 1$ ER wave. The equivalent depth, which defines the phase speed and the meridional extent of the considered waves, corresponds to the linear theory estimate for phase speeds, characteristic for observed equatorial waves. The range of equivalent depths found through observations is 12–50 m, and this corre-

sponds to vertical wavelengths in the range 7–13 km; the appropriate horizontal meridional scales are between 5 and 15° latitude (Wheeler and Kiladis, 1999). A common characteristic of the selected equatorial waves is that their energy is mainly kinetic and it is located in the vicinity of the equator. In contrast, their potential energy is more often found off the equator. This important aspect and its consequences for the assimilation process will be discussed in more detail for every wave type below.

4.1. Setup for the assimilation experiments

The domain for the experiments consists of 180 points in the x -direction, and 65 physically useful points in the y -direction. Periodic boundary conditions are applied in the zonal direction, while an extension zone of seven points is added in the meridional direction (Fig. 1). A zonal wavenumber $k = 3$ (corresponds to a global $k = 6$) and a horizontal resolution of 1° latitude/longitude are used in the experiments for all wave types. The equivalent depth is chosen to be $H = 25$ m. Due to the selected H and the

waveguide effect, all motion is confined to the inner model area. Two different lengths of the time window for the data assimilation are used: 12 and 48 h. A 12-h window might seem as a very short time window for assimilating slow equatorial waves, but the intention is to use an assimilation window close to the one used in operational NWP models. On the other hand, a 48-h window, better suited to the dynamics of the assimilated waves, has the disadvantage that the tangent-linear model is very likely not valid in a full NWP system. We will keep this in mind when discussing the relevance of our results for NWP applications in Section 5.

Twelve different experiments with wind and mass observations are performed for each window. In the first experiment observations are provided at the end of the assimilation windows. In each subsequent experiment an additional time instant with observations is added. While time separation between observations is always four hours for the 48-h 4D-Var, observations are evenly distributed through the 12-h window; this has the advantage that the distance between the observations for the first few experiments is similar for two windows. In the case of 12 updates with observations, they are provided every hour for the 12-h window, and every 4 h for the 48-h window. In neither case are observations supplied at the beginning of the assimilation window.

It needs to be said that the 4D-Var solution in the case of one and two updates with height observations and one update with wind observations might not be determined uniquely; in other words, since insufficient information is provided in these three cases, the minimization procedure can provide different solutions. On the other hand, the problem is not serious as a result of application of the horizontal diffusion and the elliptic truncation which reduce the number of effective degrees of freedom.

The background state in all experiments is a motionless fluid of depth H . The minimization procedure is allowed to continue until the norm of the gradient of J is reduced by a factor 1000, or until a maximum of 40 simulations, or 20 conjugate gradient iterations, have been carried out. The verification of the results is presented by the root mean square errors and by the relative recovery of the true energy, divergence and vorticity fields during the assimilation of different types of data.

4.2. Assimilation of the Kelvin wave

Equatorial Kelvin waves are characterized by a geostrophic balance between the zonal velocity and the meridional pressure gradient. They propagate without dispersion eastward relative to the mean flow, and show a meridional decay over a distance corresponding to the equatorial Rossby radius of deformation $R_e = \sqrt{c/2\beta}$ (where $c = \sqrt{gH}$ and $\beta = 2\Omega/r_a = 2.3 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$) (Gill, 1982). The diagnostic equation for the Kelvin wave geopotential and wind is a simple advection equation. The structure of a single Kelvin wave is shown in Fig. 5. The figure also shows the characteristic energy distribution in the Kelvin

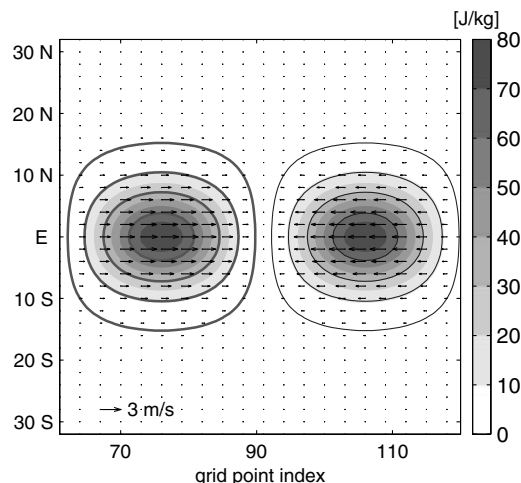


Fig 5. The Kelvin wave solution used in the assimilation experiments of Section 4. Shading is used for energy, with a 10 unit interval between successive levels of shading. This applies to both kinetic and potential energy since they are everywhere locally the same. Contours are used for the perturbation height, with a contour interval 1 m, starting from ± 0.5 m. Thick lines correspond to positive values, while thin lines are used for negative values and the zero contour is omitted. Wind vectors are shown in every second grid point in the meridional and every third point in the zonal direction.

wave; i.e., the energy is everywhere partitioned equally between potential and kinetic energy. This would suggest that mass observations are everywhere as valuable as wind observations for the assimilation of the Kelvin wave. Furthermore, simple arguments from adjustment theory on the f -plane suggest that mass field information becomes of little value as the equator is approached ($f \rightarrow 0$). As this argument ignores the existence of Kelvin waves, it may be that the value of mass field information in the tropics is underestimated, at least to the extent that Kelvin waves are important for tropical dynamics. We investigate this question in what follows.

The validity of the above argument can be checked by inspection of Fig. 6, which presents the root mean square error reduction at the end of the assimilation windows as a function of the number of updates with mass or wind observations. Since perfect observations are provided at the verification time, the assimilation errors for height and wind fields are zero in the case of mass and wind observations, respectively; therefore, these curves are not shown. It is the error reduction in the non-observed field which is of interest. In disagreement with the considerations of the energy distribution in the Kelvin wave, Fig. 6 shows that 4D-Var assimilation is not equally successful with mass observations as with wind observations in reconstructing the Kelvin wave structure. In the case of assimilation of wind field data, a second update with the wind observations is sufficient to restore the mass field. With more updates all errors stay close to zero. Assimilation of mass field data, on the other hand, although capable of wind field

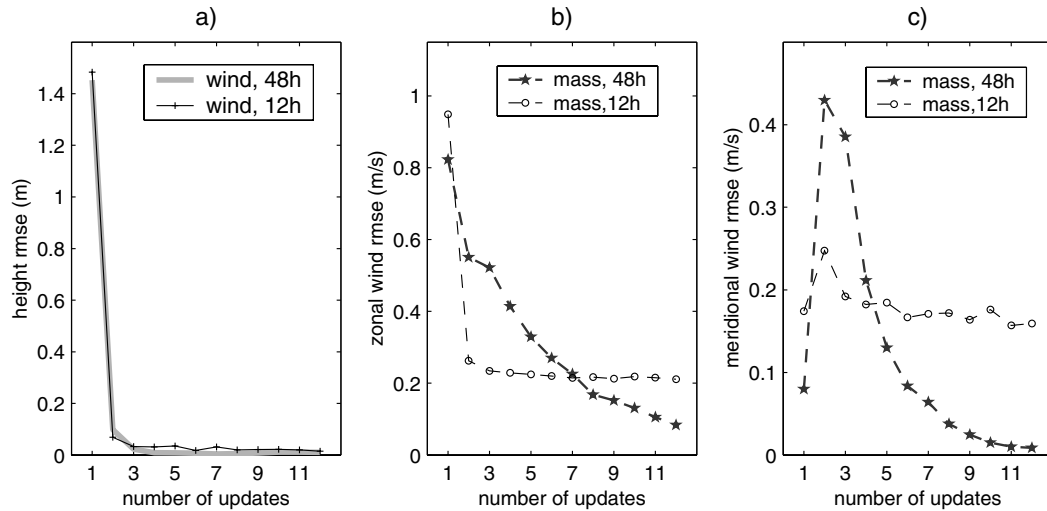


Fig 6. Root mean square error of height (a), zonal wind (b) and meridional wind (c) components for the Kelvin wave at the end of the assimilation windows of lengths 12 and 48 h, as a function of the number of updates with perfect mass or wind observations on the full model grid.

reconstruction, is much less efficient. The degree of reconstruction depends on the length of the assimilation window; in other words, the recovery is determined by the dynamical properties of the assimilated motion system. In the case of a 12-h assimilation window, during which the Kelvin wave did not move much, further error reduction after more than two updates is marginal. With the longer assimilation window, the recovery of the wind field improves as more mass field information is added. Thus, the wind field structure of the Kelvin wave can be recovered from height observations, provided the observations are plentiful and the assimilation window is sufficiently long.

We proceed with a comparison between the assimilated Kelvin wave structure in experiments with one and two updates of the model with mass and wind data, shown in Fig. 7. Presented results are from the long assimilation window. This figure illustrates the difference in the internal model adjustment during 4D-Var in the case of wind and mass field information. First we discuss the adjustment of the height field to the imposed wind observations.

Given the observed wind field, a time tendency of the height field is obtained from the continuity equation. Furthermore, the fields are adjusted to each other during the time integration. The structure of the applied wind observations is that of the zonal wind perturbation discussed in Section 3. According to the same process, EIG waves are generated during 4D-Var and propagate slowly both eastward and westward. With only one update with wind observations, the adjusted height field is presented in Fig. 7a. It has its maxima and minima at the locations of maximal convergence and divergence, respectively. When the wind field information is provided at two time instants, velocity tendencies are known and, therefore, both the height field tendency and the divergence are obtained straightforwardly from the model equations. Hence, the complete wave structure can be

correctly restored after the second update of the model with the wind observations (Fig. 7b).

The internal model adjustment is less simple in the case of height observations of the Kelvin wave. The adjustment process in this case can be thought of in terms of a moving field of successively positive and negative Gaussian height perturbations in accordance with Fig. 2. The wind field, resulting from the adjustment process with a single set of height observations, is presented in Fig. 7c. Since the mass field adjusts to the wind field at the equator, this result appears like a wind field for which the adjusted height field would be the one imposed by the observations. This is a highly unbalanced solution, with an outflow from the low pressure area and inflow into the high pressure. The amplitudes of the wind field, i.e. its kinetic energy, are very small. The imposed height field information is related to wind field divergence and vorticity through their first and second, respectively, time derivatives. Consequently, the recovery of the balanced wind field from a single set of height observations is not straightforward. By applying two updates of mass field information, information on the wind divergence at a single time instant can be obtained. As a result, the recovered balanced wind field structure along the equator is significantly improved (Fig. 7d). Further addition of height data steadily improves the structure of the wind field in the 48-h assimilation, whereas very little improvement is achieved in the 12-h window, regardless of the amount of height observations provided (Figs. 6b and c).

Figure 8 shows the relative vorticity and divergence recovery, calculated as the average absolute value of differences between the analysis and the truth, and normalized by the true values. The results of the assimilation of height observations are shown at the beginning, at the centre and at the end of the assimilation windows. The results of the assimilation of wind data are less interesting because both vorticity and divergence are recovered

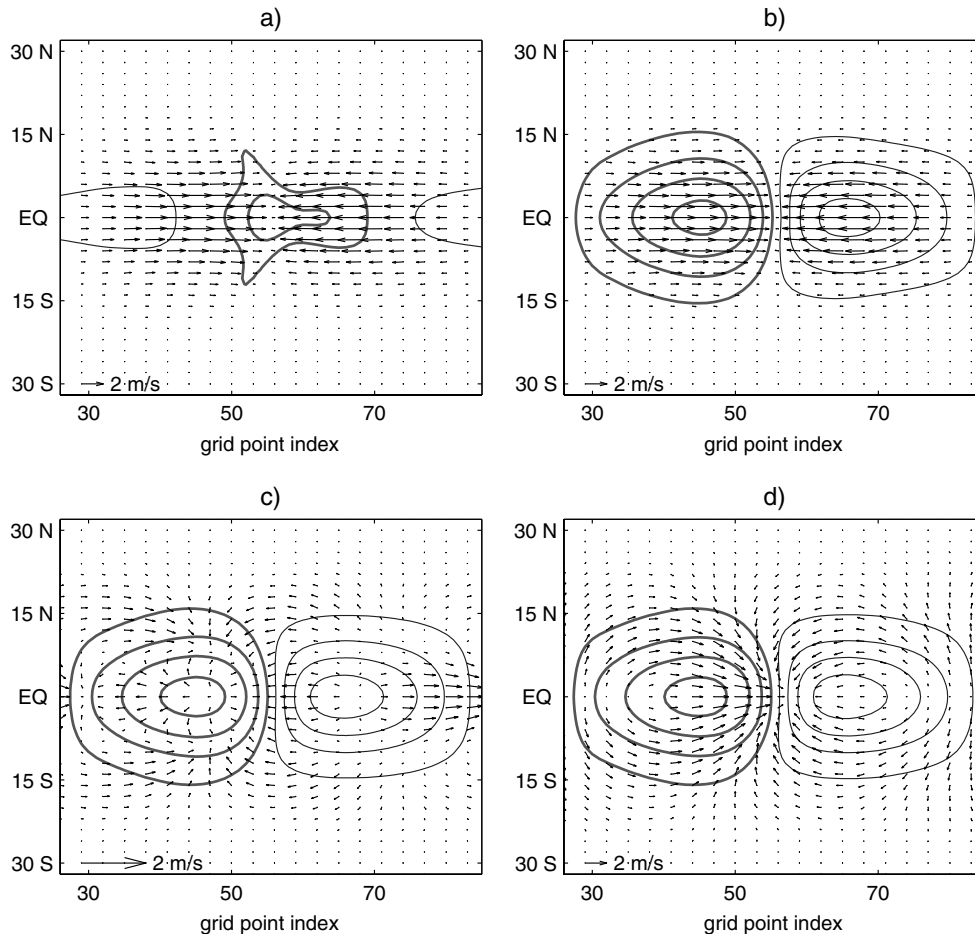


Fig 7. Structure of the Kelvin wave recovered by the assimilation of wind observations supplied at a single time instant (a), at two time instants (b), and by the assimilation of the mass field data supplied at one (c) and two time instants (d). All solutions are valid at the end of a 48-h 4D-Var assimilation window. Thick lines corresponds to positive values, while thin lines are used for negative values with respect to the mean fluid depth. Isoline and arrows spacing is the same as in Fig. 5. The wind field in (c) is multiplied by a factor 3.

well from wind observations; two to three updates with the wind data are almost sufficient for a successful result, independent of the window length. On the contrary, Fig. 8 illustrates the sensitivity of the wind field analysis to the assimilation window length when the mass data are used. It shows that the divergence is always recovered first and more completely than the vorticity. For the 12-h 4D-Var, about 70% (more than 95% at the centre of the window) of the divergence field is recovered from the height observations, as compared to a 20% recovery of the vorticity field. The recovery of vorticity improves little after the third update with data in the case of the 12-h assimilation window. On the other hand, the divergence is restored better as the number of updates with mass field information increases.

A 48-h 4D-Var reveals another interesting feature of the assimilation of height field information; that is, a detrimental effect of increased updates with height data on the recovery of the vorticity field. The negative effect is largest at the beginning of the

48-h window, in which case the vorticity error increases for the first six updates with height observations. In these cases the imposed height observations are still located far in the future so that the internal adjustment process has more “freedom” to select an unbalanced model state at the initial time. Without a background constraint applied, the model picks up a gravity wave solution, whereby EIG waves are channeled along the equatorial belt, and cannot be damped until sufficiently many observation sets are supplied. The analysis solutions for the wind field produced in this case are characterized by strong winds near the equator and a significant wind shear in the meridional direction. The improvement of the divergence recovery at time 00 (after the third update with height data) in the 48-h window is associated with an increased vorticity error over up to six updates with data. This can be explained by looking separately at the divergent and the rotational part of the wind components in the Kelvin wave (Hendon, 1986), which off the equator have opposite signs.

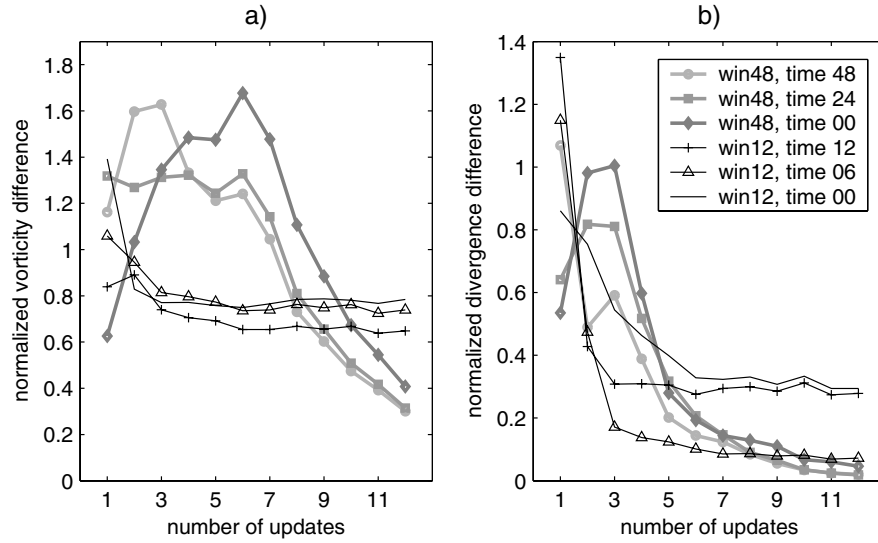


Fig 8. Recovery of the relative vorticity (a) and the divergence (b) for the Kelvin wave at the beginning, at the centre and at the end of the 12- and 48-h assimilation windows as a function of number of updates with height observations. Values are area averages of absolute differences normalized by the average absolute true values.

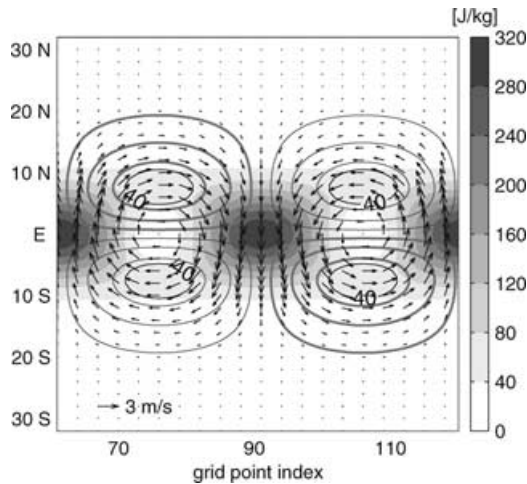


Fig 9. As Fig. 5, but for the mixed Rossby-gravity wave. Shading is used for the kinetic energy, with a 40 unit interval between successive levels of shading. Labeled dashed contours are for the potential energy, with the same contour interval.

4.3. Assimilation of the MRG wave

The MRG waves are dispersive waves, with an eastward group velocity relative to the mean flow. They are antisymmetric relative to the equator. The meridional velocity profile has a Gaussian shape. Thus, the MRG waves are characterized by a strong meridional flow across the equator (Fig. 9). Off the equator, the flow is more geostrophic. The kinetic energy maxima are centred at the equator, while the potential energy maxima are located off the equator. The potential energy constitutes just a small portion of the total energy of the MRG wave; the contribution is about 16% for the parameters used here. This we need to keep in mind

when discussing the wave structure regained by the assimilation. Based on the energy argument we also expect that wind field information is far more useful for the reconstruction of the MRG wave than mass field information.

First we only briefly mention that 4D-Var with wind field data reproduces completely both vorticity and divergence fields of the true MRG wave with two updates with observations, regardless of the length of the assimilation window. This is the reason for concentrating the most of the following discussion on results of the assimilation with height field observations.

In contrast to the Kelvin wave, the assimilation of height observations of the MRG wave results in different wave structures for 12- and 48-h assimilation windows. This can be seen by comparing Fig. 10 with Fig. 11. They display the wave structure which came as output of the assimilation after one update (Fig. 10) and ten updates (Fig. 11) of the model with perfect height observations in 48- and 12-h 4D-Var. The results are valid for the middle time-step of the assimilation windows, where the solution is supposedly the best. A dependence of the results on the assimilation window length is a consequence of a different spatial distribution of the height field, as compared to the Kelvin wave case. The height field of the MRG wave is centred off the equator. Consequently, the adjustment theory implies that the information content of a single height observation is relatively more important for determining the final balanced state than for the Kelvin wave.

In the 48-h 4D-Var the height field information is advected by the model equations both forward and backward in time over a longer time period; thus, the adjustment process takes place more efficiently as compared to the shorter assimilation window. The solution from the 48-h window with a single set of height observations is characterized by a weak cross-equatorial flow

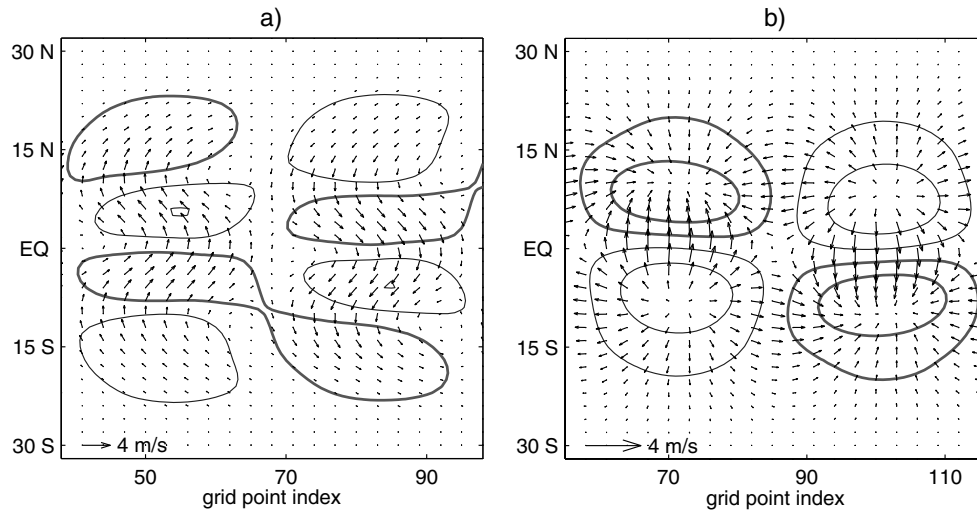


Fig 10. The MRG wave solution at the centres of the assimilation windows of 48 h (a) and 12 h (b), obtained from height observations provided at one time instant. Thick isolines are for the positive and thin for the negative height field perturbation. Isoline and arrows spacing is the same as in Fig. 9. The wind field in (b) is multiplied by a factor 2.

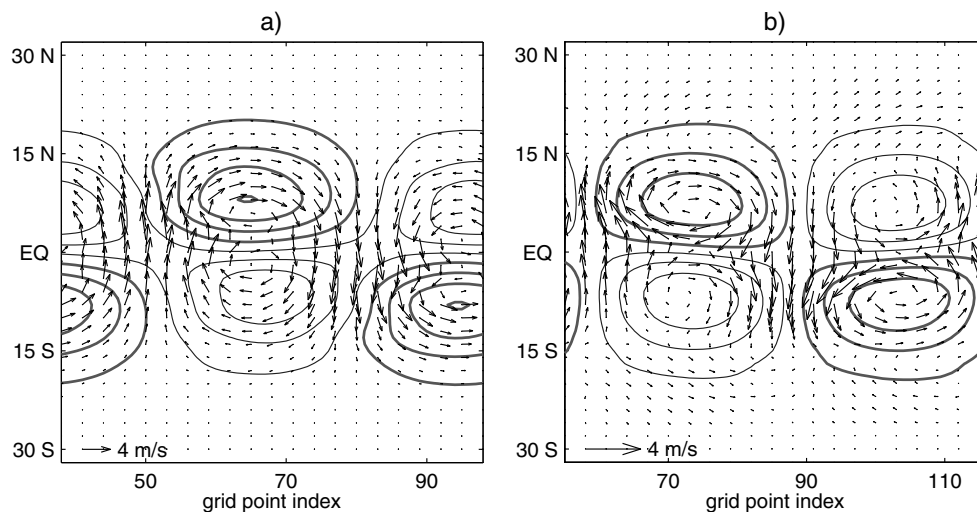


Fig 11. As Fig. 10, but height observations are provided at ten time instants.

from high pressure towards low pressure in the opposite hemisphere (Fig. 10a). In both hemispheres there are corresponding highs and lows to compensate for the flow established across the equator. The solution for the 12-h window is less balanced, with a stronger cross-equatorial flow towards the centres of high pressure (Fig. 10b); it has characteristics similar to that of the Kelvin wave case. The wind field is in both cases almost entirely divergent. One can also notice that a longer adjustment process with height field information destroys the mass field structure of the MRG wave, supplied by observations. In the 12-h window it remains preserved, since there is not sufficient time for the model adjustment to take place.

Inserting height field observations more frequently improves the results, especially in the case of the 48-h assimilation

(Fig. 11). Both the mass and the wind field of the wave are now captured well (Fig. 11a). On the contrary, a significantly smaller part of the wind field is reconstructed by the internal model adjustment when the 4D-Var window is as short as 12 h (Fig. 11b). The wind solution can be compared with the analytical solution for the MRG wave wind field, as derived from stream function and velocity potential by using the Helmholtz theorem (Hendon 1986). A comparison shows that the 12-h 4D-Var reconstructs mainly the divergent part of the wind field near the equator, just as in the case of the Kelvin wave (figure not shown).

For the parameters used here, the rotational part of the wave wind field is larger than the divergent. Hence, the wind field errors remain large through the 12-h window assimilation, as can be seen in the root mean square error scores (Figs. 12b and c). The

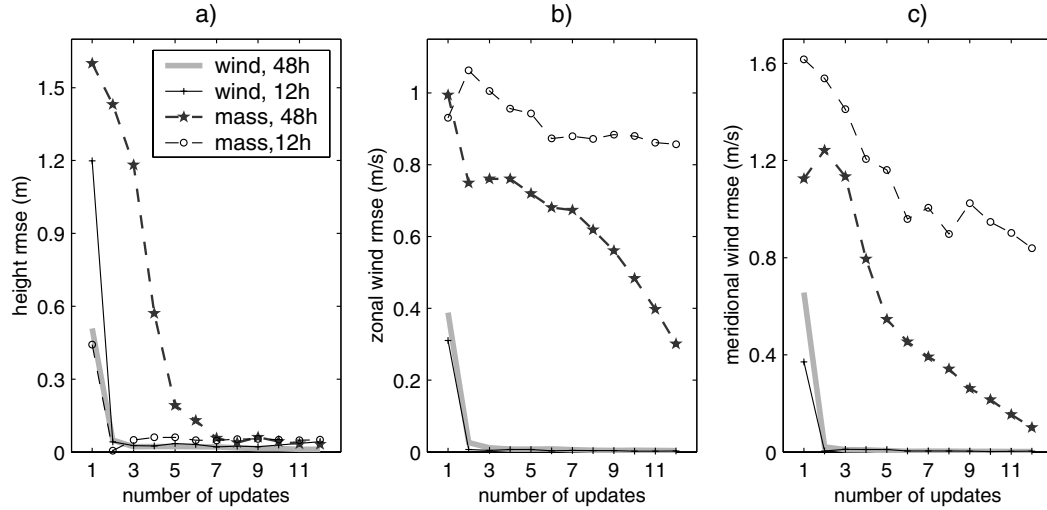


Fig 12. As Fig. 6, but for the MRG wave at the centre of the assimilation windows.

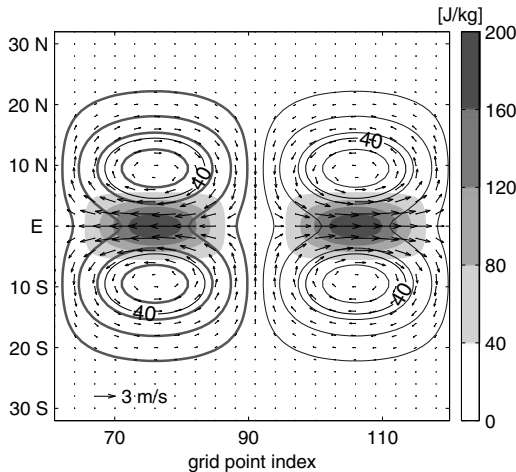


Fig 13. As Fig. 9, but for the $n = 1$ ER wave.

scores also show that the assimilation of increasing amounts of mass field data in a 12-h 4D-Var mainly improves the meridional part of the wind field, i.e. the divergence field. The improvement occurs first at the equator by a removal of the unphysical cross-equatorial flow from the low to high, followed by a geostrophic adjustment of the wind field to the imposed mass observations off the equator. In the 48-h 4D-Var, wind field errors are steadily reduced by adding more height observations, although the zonal wind error remains relatively large even after 12 updates with the height data. Figure 12a illustrates that a longer adjustment process in the height field requires more updates with the height field data in order to reach a good solution.

4.4. Assimilation of the ER wave

The $n = 1$ ER wave mainly consists of kinetic energy which is centred on the equator, while a much smaller part of the en-

ergy in the form of potential energy is located off the equator (Fig. 13). As the value of n increases, the meridional wave structure becomes more complicated and the relative contribution of potential energy increases. As for the MRG wave, wind observations are expected to be more important for a proper description of the wave structure in a numerical model. However, we also expect the mass field information to contribute more than in the case of the MRG waves, since potential energy in the considered ER wave makes a larger contribution to the total energy field as compared to the MRG wave. A relatively larger part of the wind field is balanced with the mass field centred off equator; therefore, the internal model adjustment should contribute relatively more to the recovery of the wind field from a time sequence of height observations.

Verification scores valid at the centres of the 12- and 48-h assimilation windows are shown in Fig. 14. As previously, we discuss first the assimilation of wind observations, results of which display the same features as for the other two wave types.

The assimilation of wind observations supplied at a single time instant restores relatively little mass field. It also contains significant errors in both wind components. With the second update of the model with complete wind field information, the mass field is recovered. While no further observational information is needed for the 12-h assimilation, the 48-h assimilation requires one more data set to reduce the errors to zero. This is a consequence of the longer time allowed for the adjustment process. The rotational part of the wind field is in this case larger as compared to the MRG wave and therefore the mass field information becomes relatively more important.

Similar behaviour of the wind field data is depicted in the energy scores (Fig. 15). The potential energy difference between the recovered and the true fields drops to zero after the second update in the 12-h window, while it takes one more update with observations in the longer window. Although the error in

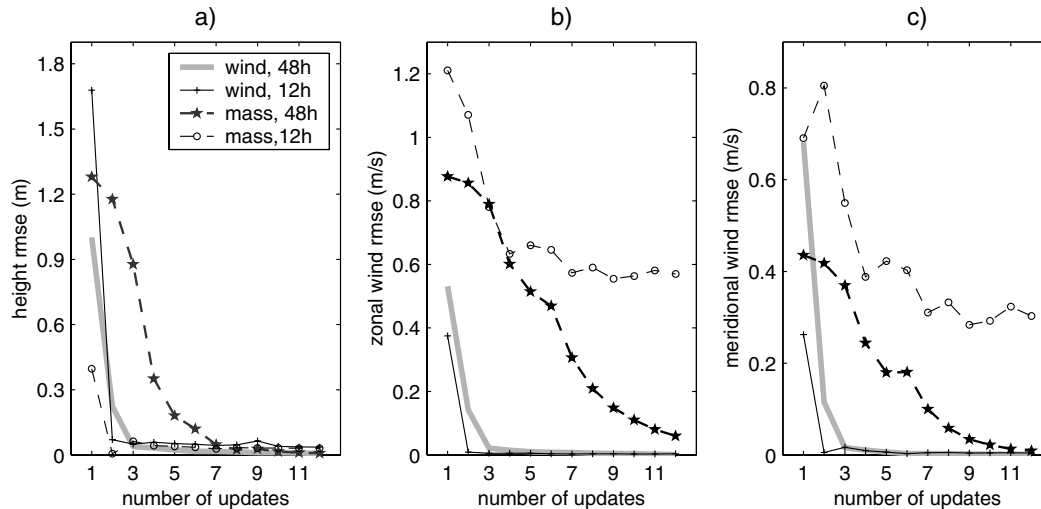


Fig 14. As Fig. 6, but for the ER wave at the centre of the assimilation windows.

geopotential, i.e. in the potential energy, might seem large when only a single wind field is provided, it is relatively small in terms of the total unrecovered wave energy.

On the other hand, assimilation of mass data succeeds in recovering only a minor part of the wind field unless the long assimilation window is used. With four updates of the mass field in a 12-h 4D-Var, the errors in the wind field reduce very little by adding more mass observations (Figs. 14b and c). The 12-h 4D-Var restores about one-third of the wave kinetic energy, as can be seen in Fig. 15a. For a better result, a longer assimilation window is needed, in which the model equations are given sufficient time to adjust appropriately the wind field to the imposed height observations. In the 48-h assimilation, the wind field recovery is continuously improved as the number of updates with the height field information increases (Figs. 14b and c), so the wave kinetic energy (Fig. 15a). A drawback of the longer window is that a larger number of updates with height observations is needed to recover the height field, too. This is also a consequence of the adjustment process; the 12-h adjustment cannot change significantly the imposed mass field observations.

By comparing the structure of the wind field obtained from different experiments it can be seen that two updates with the wind field observations in the short window are equivalent to 12 updates with the height field information in the long window. The assimilation of the mass data over the 12-h window fails to reconstruct the wind field close to the equator. The recovered part of the wind field is that which is geostrophically balanced with the mass observations north and south of the equator.

5. Discussion

It has been shown by numerous studies that the tropics are areas of large uncertainties in the initial data for atmospheric

models. These uncertainties result from a lack of observations and from deficiencies in current global data assimilation systems. The possible reduction of these uncertainties is expected to have a positive impact even on the forecast errors in the extra-tropics, as shown already in a study by Gordon et al. (1972).

In the present study we have addressed the completeness problem of data assimilation in the tropics by investigating the limits of 4D-Var, when either only wind or only mass field observations are available. The internal model adjustment to perfect observations in 4D-Var was investigated in detail by using an “identical twin” OSSE approach. The discussion is simplified by the fact that a shallow water model framework is applied. In spite of their simplicity, it is believed that the experiments presented provide some new insight into the properties of 4D-Var in the tropics; thus, they are physically relevant for atmospheric forecasts based on full scale 3D models.

The relative usefulness of the mass field and the wind field observations has been studied by assimilating simulated perfect observations of typical equatorial waves: the $n = 1$ ER wave, the MRG wave and the Kelvin wave. These waves are selected based on observational evidence; they are frequently identified and re-constructed by linear theory applied to OLR measurements (Wheeler and Kiladis, 1999, and references therein). At the same time, NWP models are less successful in reproducing the observed equatorial waves. Ricciardulli and Garcia (2000), for example, illustrated how the lack of high-frequency heating variability in a climate model convection scheme accounted for an inefficient excitation of the vertically propagating gravity waves; these have been shown to play an important role in the quasi-biennial oscillation (e.g., Scaife et al., 2002).

Details of the adjustment process were studied separately by simulating the model adjustment to isolated perturbations

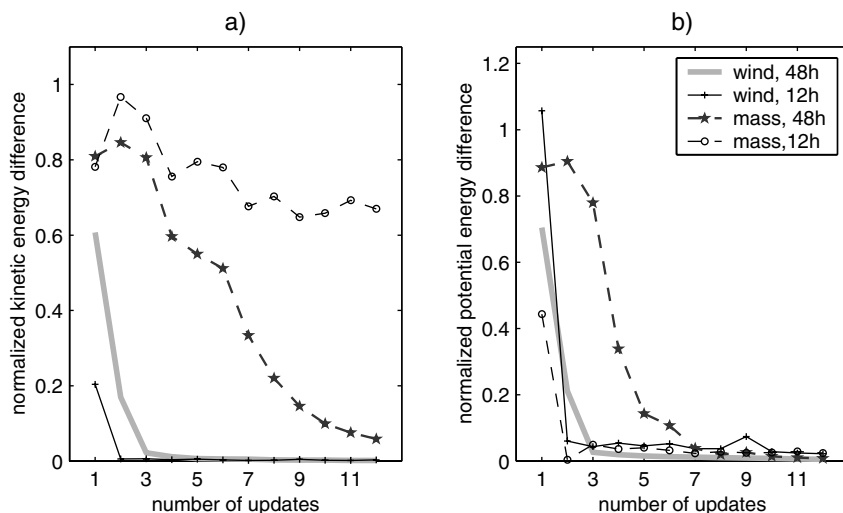


Fig 15. Recovery of the kinetic (a) and the potential (b) energy in the ER wave at the centre of the assimilation windows as a function of number of updates with perfect mass or wind observations. Values are area averages of absolute differences normalized by area averaged true values.

located at the equator. The three types of perturbations included various portions of divergence. Their behaviour during the adjustment process illustrates the importance of vorticity conservation; the more rotation in the initial perturbation, the more energy will be left in the balanced final state. It is the change of the sign of f at the equator, rather than its small value, that makes the tropical dynamics special. One of its particular features, the waveguide effect, channels gravity-waves, generated by heating, along the equator. This waveguide effect is shown to additionally slow down the adjustment process in the case of Kelvin waves. In particular, the vorticity field is reconstructed more slowly. The vorticity field in the ER wave is recovered more efficiently from the height field observations, as compared to the other two waves. This also can be expected from adjustment theory, since the wind field off the equator is in geostrophic balance with the height field.

In agreement with previous studies (Talagrand, 1981; Daley, 1991), the wind field divergence is always restored first in the experiments with an increasing number of height field observations. The assimilation of a single set of mass field data illustrates that the model equations and the 4D-Var are not sufficient to produce a physically balanced solution. More mass field information, i.e. including the time tendency of the mass field, gradually corrects the solution. Nevertheless, 4D-Var assimilation of height data recovers mainly the divergent part of the wind field, unless the assimilation window is sufficiently long for the model adjustment to take place and unless there are many observations. The assimilation cost function always reduced faster and converged to a smaller value for the wind field information, even though in both cases the convergence is fast. The cost function for experiments with a few observation sets was very close to zero at the minimum, i.e. after 20 allowed conjugate gradient iterations. For

several updates with observations, the maximal number of iterations was reached before the gradient test was satisfied. Experiments with a larger allowed number of iterations (50 as compared to 20) demonstrate that J always converges to a value close to zero, especially in case of wind data (figure not shown); the same experiments also confirm that the changes made to the solution after 20 iterations are not significant.

The results of the numerical experiments emphasize two main aspects of 4D-Var near the equator when only one type of observation is available. First, there is a limit of information available through the internal model dynamics. The limit is defined by the length of the assimilation window, in relation to the time-scale characteristics of the motion which is being assimilated. If the motion in question is slow as compared to the length of the assimilation time window, introducing more observational information beyond a certain limit does not improve the analysis. For a 12-h assimilation window, a saturation of the root mean square error is reached after two updates with the height field information; in the 48-h assimilation improvements are obtained for the first five updates with the height field data. The longer the assimilation window, the more useful information is extracted from the observations, since the model dynamics is allowed more time for adjusting those fields, that are not provided by the observations. Many important large-scale motion systems in the tropics are relatively slow as compared to the mid-latitudes; therefore, the length of the useful time window for 4D-Var should be proportionally longer in the tropics. However, a longer time window allows the growth of non-linearities, which is less suitable for NWP applications. The linearity assumption, necessary for the adjoint computations, is especially difficult to achieve with regard to the physical parametrization schemes due to their strongly non-linear nature.

Secondly, there is a limit related to the type of observations used. Assimilation of height field information results in the recovery of the divergence first, while the vorticity field is recovered more slowly or even not at all, if the assimilation time window is too short as compared to the characteristic period of the motion system represented by the observations. Adding more mass field information can even have a detrimental effect unless the observations are plentiful. This has been illustrated in the case of the Kelvin wave; the gravity wave motion, generated during the model adjustment to the height data, propagates along the equatorial waveguide without a chance to get damped until sufficiently much observational information has been introduced.

In up-to-date data assimilation systems used for NWP, an extra term J_c is added to J in order to damp inertia-gravity wave oscillations generated during the 4D-Var. This control term, for example, a digital filter (Gustafsson, 1992), is based on a critical frequency. In the large-scale tropical case, the frequencies of EIG, generated by the assimilation of height field data, are too low for eliminating EIG waves from the solution. On the contrary, EIG waves are important for tropical dynamics. WEIG waves, for example, have been found to explain an important portion of the variance of the OLR data (Wheeler and Kiladis, 1999). The 2–3 days convective variability in the tropical western Pacific is influenced by propagating EIG waves (Clayson et al., 2002), and tropical large-scale heating sources with a timescale $O(1)$ day) generate gravity waves of the same timescale and with a length scale $O(10^5)$ km (Browning et al., 2000).

The numerical experiments presented show that the wind field observations are far more efficient in recovering equatorial waves than mass field observations. After two or three updates of the model with full wind field information, no further information was needed. It has to be remembered, however, that in the case of the shallow water model, complete information of the wind field constitutes two-thirds of the total information. In this sense, the uniqueness problem is better defined when using wind field observations. It may also be mentioned that the 4D-Var assimilation of humidity measurements has a potential of extracting tropical wind field information (Andersson et al., 1994).

We presented results for single values of the zonal wave number ($k = 3$) and the equivalent depth ($H = 25$ m). It is interesting to ask what impact other values of H and/or k would have on the results. With a larger H , the phase speed of the waves is changed, i.e. increased. A Kelvin wave, for example, would move about 40% faster in case of $H = 50$ m, as compared to the $H = 25$ m experiment. This has a positive effect on the wind field restored from height data. Both the divergence and vorticity fields are better recovered in a 12-h experiment with $H = 50$ m; however, there is a negative effect on the vorticity field in case of a 48-h window (figures not shown). Taking another k has little effect on the adjustment properties of the considered waves. In particular, Kelvin waves are non-dispersive while long ER waves are nearly non-dispersive. In contrast, changing k from 3 to 1 (global zonal

wave number 2) would increase the phase speed of a MRG by a factor larger than 3, producing, nevertheless, only insignificant changes in the analysis scores. The main conclusions concerning the behaviour of mass and wind field information in tropical 4D-Var remain unaltered.

All together, the assimilation of mass field data in the tropics demands more care than the assimilation of wind data, in accordance with the basic adjustment process described in Section 2. In the case with only mass field observations, the model is trying to solve for the wind field which has caused the observed mass field distribution. This can result in an unrealistic assimilation output in case of slow dynamics and a short assimilation window. In full scale data assimilation systems, physical balances built into the background error covariance matrix act as filters which prevent such highly unbalanced solutions.

However, current approaches to the tropical data assimilation are far from being optimal. Most data assimilation systems used for NWP rely heavily on the rotational (Rossby) modes in extra-tropics while a univariate analysis is applied in the tropics. In favour of the separate mass and wind analysis in the tropics speak erroneous increments produced by application of a Rossby–Hough expansion in the tropics (Daley, 1996) and the small correlations between the wind and mass fields found in the models (Ingleby, 2001). However, Kelvin and MRG modes, which according to observations play an important role in tropical dynamics (Wheeler and Kiladis, 1999, and references therein) can be used to reduce correlations between the wind field and the mass field (Daley, 1993; Parrish, 1988). We are continuing to work toward developing a tropical J_b term which includes Kelvin and MRG waves. The present study may therefore, be seen as a prelude for a more comprehensive study involving a background error constraint.

6. Conclusions

This study is concerned with the efficiency of information transfer between mass and wind fields in the tropics through internal model adjustments in 4D-Var. Presented 4D-Var assimilation experiments were carried out with perfect observations provided on the full model grid and without a background error term. The results show that wind observations are far more efficient than mass observations in recovering the structure of typical equatorial waves. The comparison of 4D-Var with mass and wind field information highlights two specific aspects of 4D-Var in the tropics when only one type of observation is available: (1) a limit related to the type of observations used, and (2) a limit of information available through the internal model adjustment. Both limits emphasize the importance of direct wind measurements in the tropics. With the wind field information provided at two to three time instants, complete information about the height field of equatorial waves can be reconstructed by using 4D-Var. On the other hand, the assimilation of spatially dense and frequent mass observations is much less efficient in restoring the structure

of the wind field of equatorial waves, with the divergence being always recovered first and more completely than the vorticity. The recovery of the wave structure from mass data is strongly dependent on the length of the assimilation window, in relation to the time-scale characteristics of the motion of interest. On the contrary, the recovery of the wave structure from wind observations is less dependent on the assimilation window length.

7. Acknowledgments

We thank Erik Andersson, Marta Janiskova, Branko Grisogono and two anonymous reviewers for constructive comments on the manuscript.

References

- Andersson, E., Pailleux, J., Thépaut, J.-N., Eyre, J., McNally, A., Kelly, G. and Courtier, P. 1994. Use of cloud-cleared radiances in three/four dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.* **120**, 627–653.
- Baker, W., Emmitt, G. D., Robertson, F., Atlas, R., Molinari, J., Bowdle, D., Paegle, J., Hardesty, R., Menzies, R., Krishnamurti, T., Brown, R., Post, M., Anderson, J., Lorenc, A. and McElroy, J. 1995. Lidar-measured winds from space: a key component for weather and climate prediction. *Bull. Am. Meteorol. Soc.* **76**, 869–888.
- Barwell, B. R. and Bromley, R. A. 1988. The adjustment of numerical weather prediction models to local perturbations. *Q. J. R. Meteorol. Soc.* **114**, 665–689.
- 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.
- Bube, K. and Ghil, M. 1981. Assimilation of asynoptic data and initialization problem. In: *Dynamic Meteorology: Data assimilation methods* (eds L. Bengtsson, M., Ghil and E., Källén), Springer-Verlag, Berlin, 111–138.
- Cahn, A. 1945. An investigation of the free oscillations of a simple current system. *J. Meteorol.* **2**, 113–119.
- Clayson, C. A., Strahl, B. and Schrage, J. 2002. 2–3 day convective variability in the tropical eastern 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). I: Formulation. *Q. J. R. Meteorol. Soc.* **124**, 1783–1807.
- Daley, R. 1991. *Atmospheric data analysis*. Cambridge University Press, Cambridge, UK, 460 pp.
- 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.
- Davies, H. C. 1976. A lateral boundary formulation for multilevel prediction models. *Q. J. R. Meteorol. Soc.* **102**, 405–418.
- Ghil, M. 1980. The compatible balancing approach to initialization and four dimensional data assimilation. *Tellus* **32**, 198–206.
- Gilbert, J. C. and Lemaréchal, C. 1989. Some numerical experiments with variable storage quasi-Newton algorithms. *Math. Prog.* **B25**, 407–435.
- Gill, A. E. 1982. *Atmosphere–ocean dynamics*. Academic Press, New York, 662 pp.
- Gordon, C. T., Umscheid, L. and Miyakoda, K. 1972. Simulation experiments for determining wind data requirements in the tropics. *J. Atmos. Sci.* **29**, 1064–1075.
- Gustafsson, N. 1992. Use of a digital filter as weak constraint in variational data assimilation. *Proc. ECMWF Workshop on Variational assimilation with special emphasis three-dimensional aspects*, Reading, U.K., 9–12 November 1992, 327–338.
- Gustafsson, N. 1998. The numerical scheme and lateral boundary conditions for the spectral HIRLAM and its adjoint. *Proc. ECMWF Seminar on Recent developments in numerical methods for atmospheric modelling*, Reading, UK., 7–11 September 1998, 335–363.
- Gustafsson, N., Berre, L., Hörnquist, S., Huan, X.-Y., Lindskog, M., Navascués, 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.
- Haltiner, G. J. and Williams, R. T. 1980. *Numerical prediction and dynamic meteorology*. John Wiley and Sons, New York, 477 pp.
- Hao, Z. and Ghil, M. 1994. Data assimilation in a simple tropical ocean model with wind stress errors. *J. Phys. Ocean.* **24**, 2111–2128.
- Haugen, J. E. and Machenhauer, B. 1993. A spectral limited-area model formulation with time-dependent lateral boundary conditions applied to the shallow-water equations. *Mon. Wea. Rev.* **121**, 2618–2630.
- Hendon, H. H. 1986. Streamfunction and velocity potential representation of equatorially trapped waves. *J. Atmos. Sci.* **43**, 3038–3042.
- Holton, J. R. 1992. *An introduction to dynamic meteorology*, 3rd edn. Academic Press, New York, 511 pp.
- 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.
- Le Dimet, F.-X. and Talagrand, O. 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* **38A**, 97–110.
- Matsuno, T. 1966. Quasi-geostrophic motions in the equatorial area. *J. Meteorol. Soc. Jpn.* **44**, 25–42.
- Parrish, D. 1988. The introduction of Hough functions into optimum interpolation. *Proc. Eighth Conf. on NWP*, Am. Meteorol. Soc., Boston, Mass., 191–196.
- Ricciardulli, L. and Garcia, R. R. 2000. The excitation of equatorial waves by deep convection in the NCAR community climate model (CCM3). *J. Atmos. Sci.* **57**, 3461–3487.
- Rossby, C. G. 1937. On the mutual adjustment of pressure and velocity distribution in certain simple current systems. *J. Marine Res.* **1**, 15–27.

- Rossby, C. G. 1938. On the mutual adjustment of pressure and velocity distribution in certain simple current systems, II. *J. Marine Res.* **1**, 239–263.
- Scaife, A. A., Butchart, N., Warner, C. and Swinbank, R. 2002. Impact of a spectral gravity wave parameterization on the stratosphere in the Met Office Unified Model. *J. Atmos. Sci.* **59**, 1473–1489.
- Talagrand, O. 1981. A study of the dynamics of four-dimensional data assimilation. *Tellus* **33**, 43–60.
- Talagrand, O. and Courtier, P. 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Q. J. R. Meteorol. Soc.* **113**, 1311–1328.
- 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.
- Wheeler, M., Kiladis, G. N. and Webster, P. J. 2000. Large-scale dynamical fields associated with convectively coupled equatorial waves. *J. Atmos. Sci.* **57**, 613–640.