An idealized study of African easterly waves. I: A linear view

By C. D. THORNicrof*T and B. J. HOSKINS
University of Reading, UK

(Received 7 May 1993; revised 27 October 1993)

SUMMARY

The linear instability problem of the African easterly jet has been investigated using the primitive equations on a sphere. This has included an examination of the linear growth mechanisms and structure using diagnostics traditionally employed for mid latitudes, such as Eiasssen–Palm (EP) fluxes and potential vorticity. It has been shown that a growing normal mode on an African easterly jet is characterized by divergent EP fluxes in the region of the jet, implying both barotropic and baroclinic energy conversions. The linear instability is dominated by the interaction between positive and negative potential vorticity gradients at the level of the jet and, as found in previous studies, the normal modes grow mainly through barotropic energy conversions. Many of the synoptic features associated with the normal modes are in good agreement with those observed, except for the vertical-velocity pattern, which has a 'checkerboard' structure in the vertical and is much weaker.

Changing the jet latitude while keeping the baroclinicity constant changed the growth rates very little, and although barotropic energy conversions remained dominant, the most unstable modes became more baroclinic when the jet was more poleward. The most unstable modes, which grow on a thinner jet, have a larger growth rate, a smaller wavelength and stronger barotropic energy conversions. The 'checkerboard' pattern in the vertical velocity persists with the normal modes which grow on these jets.

The effect of including diabatic effects in the linear problem has also been examined. First, a simple boundary-layer scheme was found to have very little effect on the normal modes. With simply parametrized latent heat release however, the growth rates of the most unstable modes were increased slightly and the modes became less dominated by barotropic energy conversions. Also, an asymmetry is found between the ascent and descent regions in the wave, with the length scale of the updraught contractcd relative to that of the downdraught. The unrealistic 'checkerboard' pattern in the vertical velocity is almost removed and the amplitude is increased. The structure of the normal mode, using a CISK-type scheme for the latent heating, has more in common with the observed structures over west Africa than the structure of the dry modes.

It is suggested that African easterly waves may arise through a mixed barotropic/baroclinic instability mechanism where the role of latent heating is important in increasing the baroclinic energy conversions relative to the barotropic energy conversions, and also in determining the synoptic structure.

1. INTRODUCTION

African easterly waves are westward-travelling waves originating over northern Africa between June and October. They are important features of the summer climate in the African and Atlantic regions but can also reach as far west as the Caribbean (Riehl 1954) and even the west Pacific (Wallace 1970). African easterly waves are particularly important in the Atlantic as they are often transformed into tropical cyclones. Indeed Frank (1970) suggested that African easterly waves lead to about half of the Atlantic tropical cyclones and that they may also be important in triggering east Pacific tropical cyclones. It has also been suggested that in Africa they can provide a favourable environment for the initiation or enhancement of squall lines (Payne and McGarry 1977; Chen and Ogura 1982).

Easterly waves are also initiated in the tropical Pacific (Palmer 1952; Reed and Recker 1971) but, because of the special land–sea configuration which exists over Africa (see section 2 below), the characteristics of the Pacific waves are found to be different from those of the African waves and therefore were not included directly in this study.

The waves over Africa are generally thought to be due to an instability of the mid-tropospheric easterly jet situated over north Africa in the summer. The jet has a maximum wind speed of about 15 m s\(^{-1}\) at about 15°N. Burpee (1972) showed that the jet satisfied the Charney and Stern (1962) instability criterion, and suggested that it could support

* Corresponding author: Department of Meteorology, University of Reading, 2 Earley Gate, Whiteknights, Reading, Berkshire, RG6 2AU, UK.
waves growing both barotropically and baroclinically. A discussion of the zonal flow over north Africa in the summer appears in section 2.

Carlson (1969a,b), using rawinsonde data, was the first to carry out a detailed study of the easterly waves over Africa. They were studied further, using statistical methods, by Burpee (1972, 1974). The most intensive observational program however was the Global Atlantic Tropical Experiment (GATE) in which high time and spatial resolution data of easterly waves was obtained. Summaries of the GATE observations are given by Reed et al. (1977), Norquist et al. (1977) and Thompson et al. (1979). From all these studies it has been shown that African easterly waves occur typically at about 15°N, have a wavelength of between 2000 and 4000 km and a westward phase speed of about 8 m s⁻¹. Further information about the nature of easterly waves is given in section 3.

The initiation, maintenance and development of easterly waves involves a complicated interaction between adiabatic dynamics, boundary layer, moist and radiative processes. Easterly waves have been examined in GCMs and NWP models with these processes included (Walker and Rowntree 1977; Estoque et al. 1983; Reed et al. 1988), however with such complicated models it is difficult to further the fundamental understanding of easterly waves. Some studies using simpler models have also been done to obtain a better understanding of the dynamics of easterly waves, but of these there have only been a few. Rennick (1976), Simmons (1977), Mass (1979) and more recently Kwon (1989) have all carried out linear instability studies of African easterly jets. Rennick (1976) investigated the instability of a perhaps excessively strong (22.6 m s⁻¹) easterly jet, the shape of which was based on Burpee's (1972) observations at 5°E (reproduced here in Fig. 1(a)). Simmons (1977), Mass (1979) and Kwon (1989) studied much simpler analytical flows with jet maxima of about 15 m s⁻¹. Kwon (1989) also studied a second more complicated analytical jet which included other flow features and which had a maximum easterly speed of 17.7 m s⁻¹. All of these authors found unstable waves dominated by barotropic energy conversions and with horizontal wavelengths within the range quoted above, although generally towards the long wavelength end of the range. The waves have some similarities to those observed. The inclusion, by Rennick (1976), Mass (1979) and Kwon (1989) of a simple parametrization of latent heat release made only minor differences to the wave properties, although the last two of these authors suggested that the structure of the waves was improved.

It is important to note that the observations which are usually used for comparison with numerical models are the GATE structures which were observed over west Africa and the east Atlantic. By the time the waves reach this region they have usually reached their maximum amplitude and may therefore be expected to exhibit strong nonlinearity. None of the previous studies has examined a modelled easterly wave growing to finite amplitude. A calculation of this kind is presented by Thorncroft and Hoskins (1994) in Part II of this study. Kwon and Mak (1990), however, have investigated an easterly wave integration over the Atlantic in a study of the transformation into hurricanes. However this was initiated with a large-amplitude easterly wave and may have little relevance to easterly waves growing over north Africa.

The work described here follows Simmons (1977), and uses the Reading University spectral model (see Hoskins and Simmons 1975). After a discussion of the observed mean flows and wave characteristics in sections 2 and 3 the linear study of Simmons is repeated at higher resolution. This is followed by a brief study of the sensitivity to changes in the jet structure and location. Then the effect of including a simple boundary layer and latent heat release is demonstrated. In Part II of this study (Thorncroft and Hoskins 1994) a nonlinear life-cycle of an easterly-wave normal mode is presented together with the effects of simply parametrized diabatic effects.
Figure 1. Mean zonal wind over west Africa from two observational studies: (a) the mean zonal wind for August at 5°E from Burpee (1972), contour interval 5 m s$^{-1}$; (b) the mean zonal wind from GATE between 23 August and 19 September 1974 from Reed et al. (1977), contour interval 2.5 m s$^{-1}$. This is averaged between 10°E and 31°W; the ‘zero’ latitude corresponds to the average latitude of a disturbance path which was 11°N over land and 12°N over the ocean.

2. LARGE-SCALE MEAN CHARACTERISTICS

Meridional cross-sections showing the typical zonal wind structure in the west African region during the northern hemisphere summer months are shown in Fig. 1. Figure 1(a) is from Burpee (1972) derived using radiosonde data at about 5°E, and Fig. 1(b) is from Reed et al. (1977) using GATE data and averaged between 10°E and 31°W. The main features are these:

1. A 600 mb African easterly jet (henceforth denoted AEJ) at about 15°N with a peak strength of 12–15 m s$^{-1}$.
2. An upper-level tropical easterly jet (henceforth denoted TEJ) at about 200 mb, and equatorward of the AEJ.
3. A low-level monsoon westerly flow south of this.
4. A mid-latitude westerly jet to the north.
5. A near-surface easterly flow north of the AEJ, known as the Harmattans (not shown in Fig. 1(b)).
6. A second 600 mb easterly jet in the southern hemisphere (not shown in Fig. 1(b)).
The AEJ, TEJ and low-level westerlies are at their strongest during the northern hemisphere summer months. They are features of the summer monsoon, associated with the poleward shift of the maximum surface heating and the land–sea contrast.

As first suggested by Burpee (1972), the African easterly waves could be associated with a dynamical instability of the AEJ. However, these sections clearly show that the flow in this region is quite complicated. The effect that each of the features indicated above and their associated PV gradients have on easterly waves and their development is not well understood. However, in most of the previous idealized studies of African easterly waves, the AEJ has been considered in isolation and will so be, in this study, also.

3. OBSERVED CHARACTERISTICS OF EASTERLY WAVES

(a) Energetics

With both vertical and horizontal wind shears present in the AEJ, it has been suggested, in previous work, that unstable waves will grow through both baroclinic and barotropic energy conversions (Burpee 1972). A comprehensive study of the energetics of easterly waves was carried out by Norquist et al. (1977), using the composite fields of the GATE data from Reed et al. (1977). It should be remembered, however, that 'the composite' represents a wave as it was seen in the GATE area in the west African region. This is the region where easterly waves generally have their maximum amplitude, and 'the composite' does not necessarily represent the structure of a growing wave. It also represents a smoothed structure, and so some waves will be more energetic than the composite wave. Averaged between 10°E and 31°W and between 1°S and 26°N the composite wave has positive baroclinic conversions (henceforth denoted CE) and barotropic energy conversions (henceforth denoted CK) both with similar magnitudes, 0.043 W m⁻² and 0.053 W m⁻², respectively. There appear however to be marked differences between the waves over the land and over the ocean. The ratio CK/CE, over land is 0.2 and over the ocean is 3.5. This implies that barotropic processes are more important over the ocean and baroclinic processes are more important over the land. The dominance of CK over the ocean was confirmed by Thompson et al. (1979) using the high-resolution ship data in GATE. Burpee (1972) using upper-air station data also found that, over land, CE was larger than CK.

Several GCM case studies of easterly waves over west Africa have also found CE to be larger than CK (cf. Walker and Rowntree 1977; Karyampudi and Carlson 1988; Ross 1991).

Observations and GCM studies of easterly waves over west Africa suggest that baroclinic energy conversions are more important than barotropic energy conversions. However as already indicated, linear studies of the AEJ all suggest that waves grow mainly through barotropic energy conversions. The linear view has more in common with the GATE observations over the east Atlantic where CK appeared to be much larger than CE. However this is a region where the waves are expected to be most nonlinear. Thus there appears to be some confusion about the energy conversions associated with growing easterly waves.

(b) Synoptic structure of easterly waves

The observational studies of Carlson (1969a, b), Burpee (1972, 74) and Reed et al. (1977) are in general agreement about the following basic characteristics. The African waves have their largest amplitude in the mid troposphere at about 650 mb near the jet level. At this level the wave is characterized by a SW to NE tilt mainly equatorward of
the jet, consistent with barotropic growth (i.e. positive CK). At low levels, however, two regions of development have been identified, one between about 5 to 10°N and the other further north at about 20°N at the edge of the Sahara. The poleward region is usually associated with clear skies and larger vorticity anomalies than the more equatorward region, which is usually associated with cloudy skies and rainfall. The vorticity anomalies associated with these two regions tend to have a NW–SE orientation relative to each other and sometimes merge off the west African coast. It is also of interest to note that, from an examination of ECMWF vorticity analyses at 700 mb and 850 mb together with satellite imagery, Reed et al. (1988) located two preferred regions of development on either side of the AEJ, one between 18°N and 25°N, and 10°W and 5°E downwind of the Hoggar mountains in the Sahara, and the second between 8°N and 15°N, and 0 and 10°E in the climatological rainbelt (Thompson 1965). Reed et al. (1988) also noticed that the disturbances originating in these two regions often merge off the west African coast.

Considering the vertical structure, Burpee (1972, 1974) suggested that near the latitude of the AEJ the meridional wind maxima tilt eastward up to the jet level, and then westward above this level, consistent with positive CE above and below the jet and the temperature gradient changing sign at the jet level. At 5°N however, where there is easterly shear throughout the troposphere (see Fig. 1), Burpee (1974) suggested that the meridional wind maxima tilt eastwards up to 400 mb. At 28°N, where there is only westerly shear, the meridional wind maxima tilt westwards. All of these vertical phase tilts are consistent with positive baroclinic energy conversions. Reed et al. (1977) found vertical tilts similar to those found by Burpee (1974) at 15°N; but equatorward of this, and beneath the jet, the phase was more vertical. They suggested that this was due to the weaker baroclinicity and more important role of latent heating effects at lower latitudes.

(c) The observed life cycle of an easterly wave

The observational papers referred to above suggest that easterly waves probably form somewhere between 15°E and 30°E, and reach their largest amplitude at or near the west coast of Africa with most development occurring between 10°E and the west coast (Carlson 1969b; Albignat and Reed 1980). Passing from the African continent to the relatively cool east Atlantic, waves generally decay (Carlson 1969a, b) but can remain as ‘debris’ to reach the western Atlantic, Caribbean and even the Pacific, where they can regenerate into tropical storms (Frank 1970).

Suppose that easterly waves are initiated at about 25°E and have a westward phase speed of 8 m s⁻¹ at 15°N, then it will take about 6½ days for a wave to reach the west coast. Dugdale (personal communication) suggests that 6 to 8 days is a realistic estimate for this time. A question of interest is why the waves reach their peak amplitude near the west coast. In mid latitudes the observed decay time is similar to that produced in idealized life-cycles (Simmons and Hoskins 1980; Thornicroft and Hoskins 1990). It is of interest to know if 6–8 days is a natural life-cycle time for easterly waves or whether the waves decay near the west coast because the basic flow can no longer support growing waves.

Norquist et al. (1977) using GATE data suggested a doubling time for easterly waves of about 4 days; although the EKE doubling time implied by energy conversions was about 2 days. Burpee (1972) suggested a doubling time of 2 days for the total eddy energy, based on data for Niamey. An EKE doubling time of 2 days implies a doubling time of nearer 3 days for the meridional wind, equivalent to a linear growth rate of about 0.4 d⁻¹. The observed meridional wind speed over western Africa from GATE is about
5 m s\(^{-1}\). If this growth rate had occurred 7 days earlier the meridional wind speed would have been about 0.3 m s\(^{-1}\).

Given the small growth rates and limited longitudinal extent of the jet the initial perturbation may be very important in determining the structure of the waves when they reach west Africa. 'Debris' from the east, orographic forcing, differential surface heating, or perhaps a more organized perturbation in the form of a squall line could lead to different structures. The availability of moisture will also vary in time and space. In the moist monsoon winds, between 0 and 20\(^o\)W, latent heat release may be expected to be important, whereas in the initiation region, dry dynamics may be more important. This could influence the relative magnitudes of CK and CE over land and sea.

4. NORMAL-MODE STUDY BASED ON SIMMONS (1977)

(a) The zonal mean flow

The analytically defined flow used by Simmons (1977), has been re-examined in this study. The zonal wind and potential temperature of this initial state are shown in Fig. 2(a). The zonal wind which has an easterly maximum of about 15 m s\(^{-1}\) at 600 mb was based on the flow analysed by Reed et al. (1977) and is shown here in Fig. 1(b). Note that, for simplicity, features such as the TEJ and low-level westerlies are absent. The temperature field, in balance with the zonal wind, has the characteristic reverse temperature gradient below the jet maximum, with warmer air to the north and cooler air at the equator. Above the jet, the temperature decreases with increasing latitude. The temperature gradient at the surface is about 5 K 1000 km\(^{-1}\), similar to that observed by Reed et al. (1977).

The Charney–Stern instability criterion states that for a flow to be unstable either (a) the quasi-geostrophic potential vorticity gradient (henceforth denoted \(\overline{q}_y\)) on a constant pressure surface must change sign in the fluid interior, or (b) \(\overline{q}_y\) in the fluid interior must have the opposite sign to that of the surface temperature gradient (henceforth denoted \(\overline{\theta}_y\)). The \(\overline{q}_y\) for the basic state used here is shown in Fig. 2(b). It is characterized by negative values in the jet core and slightly smaller positive values on the flanks. Thus, the Charney–Stern instability criterion (a) above is satisfied. Instability criterion (a) is also satisfied with the positive \(\overline{q}_y\) above and below the jet and the negative \(\overline{q}_y\) in the jet, although this will probably be much weaker since the positive \(\overline{q}_y\) is much weaker in these regions. Also present, however, is a positive \(\overline{\theta}_y\) at the surface (see Fig. 2(a)) which, together with the negative \(\overline{q}_y\) associated with the jet, satisfies the Charney–Stern instability criterion (b) above. There are therefore several possible interactions which could result in linearly unstable growth. The strongest interactions, expected to be associated with the largest \(|\overline{q}_y|\) and surface \(|\overline{\theta}_y|\) are therefore:

- between \(-ve\ \overline{q}_y\) (jet) and \(+ve\ \overline{\theta}_y\) (poleward flank)
- between \(-ve\ \overline{q}_y\) (jet) and \(+ve\ \overline{\theta}_y\) (equatorward flank)
- between \(-ve\ \overline{q}_y\) (jet) and \(+ve\ \overline{\theta}_y\) (surface)

Also, for instability to be present, the Fjørtoft condition states that the mean zonal wind, \(\overline{u}\), must be positively correlated with the \(\overline{q}_y\) and surface \(\overline{\theta}_y\) gradients. An examination of the fields in Fig. 2 shows that this condition is indeed satisfied.

In terms of interacting Rossby waves on these gradients (cf. Hoskins et al. 1985), unstable growth implies that waves on the positive \(\overline{q}_y\) gradients on the flanks of the jet, and waves on the positive surface \(\overline{\theta}_y\) gradient should tilt downstream (i.e. towards the west) relative to waves on the negative \(\overline{q}_y\) gradient. Generally the expectation is that the
interaction between the negative $\tilde{q}_y$ and positive $\tilde{q}_y$ should be associated mainly with barotropically unstable growth or positive CK, and that the interaction between the negative $\tilde{q}_y$ and surface $\tilde{\theta}_y$ should be associated mainly with baroclinically unstable growth or positive CE.

It should be noted that, as the zonal wavelength becomes smaller, one expects its vertical scale to decrease. When the depth scale of the wave becomes too small, compared with the height of the jet, the $\tilde{q}_y$-surface $\tilde{\theta}_y$ interaction will become impossible. It can be shown, using simple scaling arguments, that this will occur for wavenumbers greater than about $m_e = \pi a f \cos \phi / h N$, where $h$ is the height of the AEJ, $a$ is the earth's radius, $\phi$ is the latitude, $N$ is a typical Brunt–Väisälä frequency and $f$ is the Coriolis parameter. Around $15^\circN$, $m_e$ is about 15. For waves of this order only the interaction between the
negative and positive $\tilde{q}_s$ will be expected. Since the meridional scales of high wavenumbers will also probably be smaller, one also expects that one side the jet will be preferred.

In the atmosphere one would expect the relative magnitudes of the $\tilde{q}_s$ extrema to vary, resulting in different types of instabilities. In Fig. 2(b) of Reed et al. (1977) the vorticity gradient appears to be stronger on the equatorward flank than on the poleward flank. This is perhaps consistent with the fact that their vorticity pattern at 700 mb showed a trough predominantly on the equatorward flank, tilting NE–SW. Burpee's (1972) analysis of $\tilde{q}_s$ also had larger positive values on the equatorward flank. These observations all look qualitatively like Fig. 2 and are in agreement that the necessary conditions for mixed barotropic/baroclinic instability are satisfied.

Also shown here in Fig. 2(c) is the zonal mean Ertel potential vorticity for the basic state used here. Comparing this with Fig. 2(a) we can see that a PV maximum exists on the isentropes passing through the jet, so confirming (using this more exact measure) the linear instability of the jet. The Ertel potential vorticity associated with the jet shown in Fig. 1(a) above (Burpee 1972) also showed the reverse PV gradient on isentropes required for instability.

(b) The method and model

The baroclinic spectral model developed by Hoskins and Simmons (1975) was used, as by Simmons (1977), but with the horizontal resolution, rhomboidal truncation 57, improved to triangular truncation 95 (T95), giving a resolvable length scale (wavelength/2$\pi$) of about 67 km. The vertical resolution is similar with 15 sigma levels at 0.033, 0.113, 0.216, 0.326, 0.432, 0.524, 0.600, 0.662, 0.714, 0.759, 0.803, 0.848, 0.894, 0.940 and 0.982.

The same initial value method as used by Simmons was used here. The most unstable mode for a particular wavenumber is found by integrating the primitive equations on a hemisphere, using as initial conditions the zonal flow and a small perturbation confined to the wavenumber of interest. The perturbation is allowed to grow while keeping the zonal mean constant. So as to keep the integration linear, the perturbation amplitude was reduced when it became too large. The growth of the perturbation amplitude is followed until it is exponential to within a specified accuracy. The growth rate, phase speed and structure of the most unstable mode can then be examined directly. The model was integrated until the relative change in growth rate between one instant and half a day later was less than $10^{-3}$. A comparison of the present results with those of Simmons shows negligible differences for the most unstable modes (less than 5%). Similar negligible differences were also found when comparing the most unstable modes with those obtained using a matrix inversion method with a rhomboidal truncation at wavenumber 57 and the same 15 levels.

The initial value method was also used for the moist instability analysis in section 6(b) with a resolution of T95 and the same 15 levels. Here the model was integrated until the relative change in growth rate between one instant and half a day later was less than $10^{-2}$. This is because with latent heat included convergence was found to be less rapid.

It should be noted however that for the dry weakly-growing high-wavenumber modes (roughly 14 and greater), the growth rate decreased slightly with increased vertical resolution. For these wavenumbers, the vertical resolution in the jet region probably needs to be enhanced, because, as noted above for high wavenumbers, all of the amplitude is expected to be concentrated in the jet region and probably on one flank. Examination of these modes (not shown) shows this to be the case. Indeed, the vertical resolution used in the previous linear studies of Rennick (1976), Mass (1979) and Kwon
(1989) is much poorer than that used here and so their short wavelengths, and perhaps longer ones also, may be inadequately represented. For the ‘moist’ study in section 6(b), the vertical resolution problem for high wavenumbers is probably less important, since \( m_c \) is much larger, owing to a smaller equivalent \( N \). The high-wavenumber moist modes are deeper and do not become so concentrated at the jet level in the same way as do the dry modes. An examination of the normal modes obtained (not shown) confirms this.

(c) Results

(i) Growth rates and phase speeds. The variation of growth rate and phase speed with zonal wavenumber is shown in Fig. 3 for the most unstable modes of the jet shown in Fig. 2. The growth-rate curve is very similar to that of Simmons (1977) and is a maximum at about zonal wavenumbers 10 and 11. At 15°N, \( m = 11 \) is equivalent to a wavelength of about 3500 km, towards the large wavelength end of the observed range. The growth rate for \( m = 11 \) is 0.28 d\(^{-1}\) and it has a westward phase speed of 6.8 deg d\(^{-1}\), implying a period of about 4.8 days. The phase speed decreases with increasing wavenumber, consistent with the modes being concentrated on one flank of the jet, and also with the discussion above in subsection (a).

(ii) Global energy conversions. The global energy conversions of the most unstable mode for wavenumber \( m = 10 \) will now be studied. This wavenumber was chosen rather
than \( m = 11 \) because it was the one considered by Simmons (1977), and because the vertical resolution for this wavenumber is almost definitely adequate. Throughout this discussion the mode is scaled to give a meridional wind maximum of about 5 \( \text{ms}^{-1} \), similar to the composite easterly waves in the GATE study (Reed et al. 1977).

The global energy conversions for the \( m = 10 \) mode are presented here in the form of a Lorenz energy box diagram in Fig. 4. The conversion terms contributing to the growth of eddy kinetic energy (henceforth denoted EKE) have been normalized to unity.

![Diagram](image)

Figure 4. Normalized energy conversions for the most unstable mode at wavenumber 10. AZ and AE are the zonal and eddy available potential energies, and KZ and KE are the zonal and eddy kinetic energies. The eddy energy conversions are indicated by the arrows and the normalized magnitudes of these are also included. The conversions are normalized by the total contribution to the growth of eddy kinetic energy, which is equal to \( 8.38 \times 10^{-2} \text{W m}^{-2} \) when the mode is scaled to give a maximum meridional wind of 5 \( \text{ms}^{-1} \).

In agreement with previous studies the dominant energy conversion is through the barotropic term, the ratio CK to CE being 5.6 to 1. From a comparison with Fig. 1 of Norquist et al. (1977) and section 3(a) above, it is apparent that the energetics of this normal mode have a closer resemblance to those calculated over the ocean than to those over the land. It is found that the ratio CK/CE increases with increasing wavenumber, consistent with the modes having more amplitude near the jet and less near the surface.

It is clear that the barotropic positive-\( \dot{q}_z \), negative-\( q_z \) internal interaction is dominating this mixed barotropic/baroclinic instability.

(iii) Zonal mean eddy kinetic energy and energy conversions. The zonal mean EKE is shown in Fig. 5(a). In the region of the jet, the structure is dominated by two EKE maxima on the poleward and equatorward flanks. A third much weaker maximum is at the ground level in the region of the \( \theta_0 \) gradient. The EKE on the poleward flank, although slightly smaller in magnitude, is deeper than the equatorward flank maximum and appears to mirror the vertical spreading of the isentropes.

Also shown in Fig. 5(a) are small non-zero values of EKE at the equator, which suggests that the boundary in this hemispheric situation may be affecting the results slightly. When the normal-mode calculation was repeated for this mode on the whole sphere, non-zero values of EKE were found on the equator even when there was no wall there. We conclude from this, and because the values of EKE are small compared with the maxima near the jet, that the effect of the equatorial boundary is small.

The horizontal eddy momentum flux (Fig. 5(b)) is dominated by the flux of easterly momentum away from the jet on both flanks of the jet, associated with barotropic energy conversions and positive CK. The maxima values on the equatorward and poleward
flanks are +8.6 and -10.7 m²s⁻², respectively. As with the EKE, the structure on the northern flank is deeper.

Burpee (1972), averaging between 15°E and 15°W at 14°N on the equatorward flank found a positive momentum flux of 2.7 m²s⁻² at 700 mb, while Albignat and Reed (1980) found the largest momentum fluxes on the equatorward flank at 700 mb equal to around 10 m²s⁻². This maximum was situated just off the west coast of Africa.

The horizontal eddy heat flux (Fig. 5(c)) is dominated by two maxima with a negative flux below, and a positive flux above, the jet. These fluxes are both in the right sense for baroclinic growth. It is hypothesized that the fluxes above the jet arise solely from the internal $\overline{q}$ interactions but that the deeper, slightly weaker region of fluxes below the
jet is associated with the negative $\overline{\bar{q}_\sigma}$—surface $\overline{\bar{\theta}}$, interaction as well. The observed heat flux (Norquist et al. 1977; Burpee 1972) is dominated by a negative heat flux below the jet at about 850 mb.

The vertical eddy heat flux (Fig. 5(d)), has two extrema above and below the jet, both corresponding to upward heat fluxes, consistent with baroclinic growth and the horizontal heat fluxes shown in Fig. 5(c). The vertical eddy heat flux shown by Norquist et al. (1977) has a more complicated structure. The main regions of upward flux were beneath the jet, as here, and also around 300 mb and 500 mb, equatorward of the jet. They also found a sloping region of downward flux from below and equatorward of the jet to above and poleward of the jet. As suggested by Norquist et al. (1977), this could be due to evaporative effects.

(iv) **EP flux and its divergence.** The horizontal momentum fluxes and horizontal heat fluxes shown above may be summarized by constructing EP (Eliassen–Palm) cross-sections. For $\beta$-plane geometry and using height coordinates, the quasi-geostrophic EP flux vector is given, in standard notation, in the form:

$$ F = \left( \overline{-u'v'} ; \frac{f}{\overline{\theta'}} \right), $$

where the overbar denotes a zonal average, and the primes deviations from it; $u$ and $v$ are zonal and meridional wind components; $\theta$ is the potential temperature; $\overline{\theta}$ is a standard static stability and $f$ is the Coriolis parameter. Edmon et al. (1980) show how this diagnostic can be related to wave-action and group-velocity concepts. In the present context it is clearly a very useful way of summarizing the barotropic and baroclinic energy fluxes.

The EP flux and its divergence are shown in Fig. 6 for the $m = 10$ mode. At the jet level the EP vectors are mainly horizontal and pointing equatorwards on the equatorward flank of the jet and polewards on the poleward flank of the jet, consistent with the horizontal momentum fluxes shown above in Fig. 5(b). Above and beneath the jet, the EP vectors point upwards and downwards, respectively, consistent with the horizontal

![Figure 6. Latitude–height section of EP flux and its divergence for the most unstable mode at wavenumber 10 which grows on the basic state in Fig. 2. The mode has been scaled to give a maximum meridional wind of 5 m s$^{-1}$. The zero contour is dotted and the negative contours are dashed. The contour interval is $1 \times 10^{15}$ m$^2$.](image)
heat fluxes shown in Fig. 5(c). It is therefore clear that for an easterly jet such as this, the EP flux signature for a growing wave is characterized by divergence in the region of the jet and convergence on the flanks of the jet at around 8°N and 20°N. The weaker convergence region at the surface is consistent with the large value of the ratio CK/CE.

It can be shown (Edmon et al. 1980) that F also gives information about how the waves feed back onto the mean flow. For example, the divergent EP flux signature implies that the $\bar{q}_v$ in the jet will increase, thus stabilizing the flow. If we were to specify constant linear growth and start with a mode scaled to give a meridional wind maximum of 1 m s$^{-1}$, as in the life cycle examined by Thorncroft and Hoskins (1994), it would take about 9 days to remove the negative $\bar{q}_v$—in reasonable agreement with the numerical results.

In summary, the EP fluxes indicate clearly how the normal mode grows at the expense of the jet, and also how the dynamical instability is removed.

(v) Horizontal structure. Perturbation fields of PV on the 315 K surface and at the lowest-level temperature are shown to illustrate the horizontal structure. Figure 7(a) shows the perturbation PV on the 315 K surface. This surface is close to 650 mb and was chosen because it transects the jet (Fig. 2(a)). PV anomalies are present in the three different regions of PV gradient (cf. Fig. 2(b,c)), on the poleward flank, on the equatorward flank and in the jet itself, consistent with the findings above in section 4(a). The largest anomalies are in the jet while the weaker ones on the flanks have the expected westward displacement relative to these, implying barotropic growth. The relative vorticity anomalies near the level of the jet (not shown) have the same pattern as the PV.

(a)

(b)

Figure 7. The horizontal structure of the most unstable mode at wavenumber 10, scaled to give a maximum meridional wind of 5 m s$^{-1}$. The perturbation Ertel potential vorticity on the 315 K surface is shown in (a) with a contour interval of 0.05 PV units. The perturbation wind vectors are also included. The perturbation temperature at $\sigma = 0.982$ is shown in (b) with a contour interval of 0.4 K. In each case the zero contour is dotted and negative contours are dashed.
The temperature perturbations near the surface, shown in Fig. 7(b), have a simple sinusoidal variation centred at about 17°N with the pattern tilted westwards relative to the PV pattern in the jet, again consistent with the findings in section 4(a) above. The relative vorticity perturbations at low levels (not shown) are also consistent with this, with cold and warm low-level temperature anomalies corresponding to positive and negative vorticity anomalies, respectively.

Observed easterly waves (see section 3 above) have shown two maxima at low levels, poleward and equatorward of the jet. It is hypothesized that the surface maximum seen here in the normal mode corresponds to the observed poleward maximum, but that the observed equatorward maximum is probably produced through nonlinear or diabatic effects.

(vi) **Vertical structure.** Vertical longitudinal cross-sections at 15°N are now presented for relative vorticity, meridional wind and vertical velocity. The relative vorticity in Fig. 8(a) has a maximum of $3.1 \times 10^{-5}$ s$^{-1}$ just above the jet. The GATE composite (Reed et al. 1977) also had a maximum of about $3 \times 10^{-5}$ s$^{-1}$ just below and equatorward of the jet. In Fig. 8(a) at 15°N at the surface can be seen a secondary, much weaker, maximum from which there is a clear eastward tilt with height towards the maximum at the jet level; above the jet there is a westward tilt with height. The GATE composite of relative vorticity is dominated by cyclonic vorticity around the jet level, with the largest anticyclonic vorticity above at about 200 mb. This is clearly not a simple sinusoidal structure, which suggests that moist or nonlinear processes are important.

For this mode at 15°N, the vertical tilts are significant, whereas poleward and equatorward of the jet (not shown) they become small. It should also be noted that at 15°N the vertical tilts are quite complicated. The structure resembles the letter 'W' on its side with the upper and lower tilts in the right sense for positive baroclinic energy conversions. The central tilts in the region of the jet imply negative but small baroclinic energy conversions because of the small horizontal temperature gradients there (cf. Fig. 5(c, d) and Fig. 4 of Norquist et al. 1977). For small wavelengths which cannot interact with the surface $\theta_v$, these jet-level-damping baroclinic energy conversions, although still small, will tend to be the dominant baroclinic signature. This again draws attention to the importance of having high resolution in the jet region when studying the smaller wavelengths.

The meridional wind shown in Fig. 8(b), with a maximum at the jet level and a weaker maximum at the surface, has the same vertical tilts as the relative vorticity. The meridional wind observed by Reed et al. (1977) also had two $\omega$ maxima, one just below jet level and the other higher up, just above 200 mb, in disagreement with this normal mode. The temperature field (not shown) tilts in the opposite sense to the meridional wind, consistent with a growing baroclinic structure.

The 'checkerboard' pattern in the vertical velocity, $\omega$, shown in Fig. 8(c) is consistent with quasi-geostrophic theory. A vorticity anomaly on a jet with opposite baroclinicity, above and below, immediately gives this checkerboard pattern (Hoskins and Pedder 1980). Note that the upper-level part of the $\omega$ pattern has a larger amplitude than the lower part, consistent with the maximum vorticity anomaly (see Fig. 8(a)) located slightly above the jet. This is also consistent with the slightly larger horizontal and vertical heat fluxes seen above, rather than below, the jet in Fig. 5(c, d). The dry-dynamics 'checkerboard' pattern looks very different from the observed structure of Reed et al. (1977); also, the amplitude is much weaker—0.7 mb h$^{-1}$ compared with the observed 5 mb h$^{-1}$. These major differences are probably due to latent heating effects. The
Figure 8. Longitude–height sections at 15°N looking polewards: (a) perturbation relative vorticity with a contour interval of $7 \times 10^{-6} \, \text{s}^{-1}$; (b) meridional wind with a contour interval of 1 m s$^{-1}$; (c) vertical velocity, $\omega$, with a contour interval of 0.1 mb h$^{-1}$. In each case the zero contour is dotted and negative contours are dashed.

The observed structure does however have a double maximum in the vertical, unlike the Pacific composite of Reed and Recker (1971). The fact that this double maximum is seen in the African wave, but not in the Pacific wave, is very likely to be due to the presence of the AEJ.
5. NORMAL-MODE SENSITIVITY TO CHANGING JET CHARACTERISTICS

(a) Motivation

The jet examined in section 4 was based on the GATE data which were averaged over the period 23 August to 19 September 1974. However, the observed AEJ characteristics, such as the latitude of the wind maximum and strength of the horizontal shear, may be expected to vary interannually, and even during the season. This may result in significant changes in the synoptic climatology and, in particular, in the waves growing on these jets. For example, two Meteorological Office GCM simulations for the summers of 1983 and 1950 (Rowell et al. 1992) show differences between the two years, the former having anomalously dry conditions in the Sahel region, and the latter anomalously wet conditions in that region. The AEJ averaged between July and September in the two simulations were also quite different (Rowell, personal communication). In the anomalously dry season the jet had a maximum of about 12 m s\(^{-1}\) at 14°N, whereas in the anomalously wet season the jet had a maximum of about 8 m s\(^{-1}\) at 18°N. Fontaine and Janicot (1992) also found from observations that in drought years the AEJ is stronger.

The fact that the 1950 summer was the wettest summer in the Sahelian region this century might be associated with the corresponding poleward shift of the moist monsoonal air. However changes in the waves growing on the jet might also be expected; although it has been suggested (Avila and Clarke 1989) that drought years, such as 1983, have similar numbers of waves.

Regarding interseasonal variations, Mass (1979) suggested that the jet used by Simmons (1977), which was based on the GATE jet of Reed et al. (1977), was oversmoothed because of the large time-averaging period. It was suggested that a meridionally thinner jet obtained by averaging over a shorter period would be more relevant for instability studies.

(b) Sensitivity to changes in jet structure

To test the sensitivity to the jet structure, three different versions of the AEJ have been examined (see Fig. 9): (a) a jet shifted poleward to 20°N, (b) a jet shifted equatorward to 13°N and (c) a thinner jet at 15°N.

For the shifted jets the strength of the jet was varied inversely as the Coriolis parameter, thereby maintaining the baroclinicity. The 20°N jet then has a wind maximum of about 11 m s\(^{-1}\) and the 13°N jet a wind maximum of 17 m s\(^{-1}\). As noted above, the poleward-shifted jet of 1950 also had a weaker jet maximum than the more equatorward 1983 jet. The \(\bar{q}_v\) contrast for the 13°N, 15°N and 20°N jets decreases with increasing latitude owing to the weakening horizontal curvature in \(\bar{u}\). The expectation must be that the barotropic component of the instability will weaken with increasing latitude.

For the thinner jet the zonal wind has the same analytic form as that of the basic Simmons jet except that the sine function is raised to the tenth power rather than the sixth power. The consequent jet width is decreased by about 20%. The negative \(\bar{q}_v\) in the jet has a very similar magnitude to that for the basic jet.

Table 1 displays, for each of the flows and including the basic jet, the most unstable wavenumber, its growth rate and its phase speed. The normalized CE and CK energy conversions are also presented for the most unstable \(m = 10\) modes. Figure 9 shows the zonal mean eddy kinetic energy associated with each of these \(m = 10\) modes.

The growth rate of the most unstable mode does not change much if the jet is shifted in latitude while maintaining the temperature gradient. There is a slight shift towards shorter wavelengths as the jet is moved polewards. The most striking difference, however, is in the energetics. As the jet is moved polewards the baroclinic contribution CE
Figure 9. Latitude–height sections of mean zonal wind and potential temperature for different easterly jets together with the eddy kinetic energy of the corresponding most unstable mode at wavenumber 10: (a, b) are for the jet shifted polewards by 5° of latitude; (c, d) are for the jet shifted equatorwards by 2° of latitude; (e, f) are for the thinner jet (see text for details). The zonal wind is shown with a contour interval of 2 m s⁻¹ with dashed contours negative. The potential temperature is shown with a contour interval of 5 K and with the 300 K contour dotted. For the EKE pictures the modes have been scaled to give a maximum meridional wind perturbation of 5 m s⁻¹. The contour intervals in (b), (d) and (e) are 1, 2 and 2 J kg⁻¹, respectively.
TABLE 1. The most unstable wavenumber, $m$, its growth rate, $\sigma$, and phase speed, $C_{ph}$, in deg d$^{-1}$ corresponding to the jets indicated in the first column

<table>
<thead>
<tr>
<th>JET</th>
<th>$m$</th>
<th>$\sigma$</th>
<th>$C_{ph}$</th>
<th>CK</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>20°N</td>
<td>12</td>
<td>0.27</td>
<td>-5.6</td>
<td>0.61</td>
<td>0.39</td>
</tr>
<tr>
<td>Basic</td>
<td>10-11</td>
<td>0.28</td>
<td>-7.0</td>
<td>0.86</td>
<td>0.14</td>
</tr>
<tr>
<td>13°N</td>
<td>10</td>
<td>0.27</td>
<td>-7.6</td>
<td>0.99</td>
<td>0.01</td>
</tr>
<tr>
<td>Thin</td>
<td>11-12</td>
<td>0.35</td>
<td>-6.2</td>
<td>0.98</td>
<td>0.02</td>
</tr>
</tbody>
</table>

For two wavenumbers with the same growth rate the average phase speed of the two modes is given. Also included in the table are the normalized energy conversions CK and CE for the most unstable $m = 10$ mode for each jet. The normalizations used for the 20°N jet, basic jet, 13°N jet and thin jet are 8.19, 8.38, 6.67 and $8.82 \times 10^{-2}$ W m$^{-2}$, respectively, for modes scaled to give a maximum meridional wind of 5 m s$^{-1}$.

increases relative to the barotropic contribution CK. Consistent with the energetics, the zonal mean EKE sections show that the more baroclinic mode associated with the 20°N jet has more amplitude at the surface than the more barotropic mode associated with the 13°N jet, which has most of its amplitude at the jet level.

The relative insensitivity of both growth rate and frequency to changes in the jet latitude is perhaps consistent with the observation that the number of waves in a season is similar if the jet is more poleward or equatorward than normal. These results strongly suggest that for a more poleward jet, as in the wet 1950 season, one would expect the waves to be more baroclinic, and that for a more equatorward jet, as in the dry 1983 season, they should be less baroclinic. This suggestion should be examined in future work.

For the thinner jet the most unstable mode has a slightly smaller wavelength and a larger growth rate, than for the basic jet, in agreement with Nitta and Yanai (1969) and with simple barotropic instability theory (Kuo 1949), which states that the growth rate is proportional to the shear. The zonal mean EKE associated with this normal mode is dominated by just two maxima on either side of the jet, and the energetics are indeed almost entirely dominated by barotropic conversions.

6. Sensitivity to the inclusion of diabatic processes

(a) Boundary-layer damping

(a) The parametrization. A very simple boundary layer has been included in the instability calculations for the basic jet. It uses Rayleigh friction on vorticity and divergence and the same Newtonian damping on temperature. The damping rate $\lambda$ was chosen to vary linearly from 0.4 d$^{-1}$ at $\sigma = 1.0$ to zero at $\sigma = 0.9$. A similar dissipation, but with $\lambda$ vanishing at $\sigma = 0.8$, was used by Valdes and Hoskins (1988) and resulted in the growth rates of their baroclinic modes being halved. The results discussed below were found to be insensitive to which $\lambda$ distribution was used.

(ii) The results. The growth-rate and phase-speed curves with this boundary-layer damping are not shown because the curves are very similar to those of the adiabatic modes shown in Fig. 3. There is only slight stabilization for wavenumbers greater than 9; for smaller wavenumbers there is even slight destabilization. The EKE cross-section for zonal wavenumber 10 shown in Fig. 10(a) no longer exhibits a surface maximum, and the EP flux structure shown in Fig. 10(b) is very similar to that of the adiabatic mode shown in Fig. 6, except for weaker fluxes below the jet level indicating a damped interaction between the surface $\vec{\theta}_y$ and interior $\vec{q}_y$.

The wavenumber-10 energetics (not shown) are also very similar to those for the
frictionless mode, with the normalized CK equal to 0.86. This reinforces the conclusion reached above that the internal $\bar q_y$ interactions dominate the linear instability of the AEJ and not the interactions with the surface $\bar \theta_y$.

(b) **Effect of including a simple latent heating parametrization**

(i) **Previous studies.** In the previous studies of Rennick (1976), Mass (1979) and Kwon (1989), latent heat was included using a wave-CISK-type parametrization. The growth rates and phase speeds were found to change very little with this latent heating included, although Mass (1979) and Kwon (1989) suggested that a more realistic structure was obtained. These studies examined only the effect of latent heating on single zonal wavelengths. However, previous work on moist baroclinic instability in mid latitudes (e.g. Emanuel et al. 1987) has found that, with latent heating, the width of the ascent region decreases relative to the descent region. This asymmetry has also been found in
the present study. The results presented here allow for this by including the higher-order harmonics of the wavenumber examined at T95.

(ii) *Observed convection in easterly waves.* Observational studies have often reached contrasting conclusions regarding the position of convection in easterly waves. For example Carlson (1969a) suggested that the main convective area is west of the trough, in agreement with Aspliden et al. (1976), Burpee and Dugdale (1975) and Payne and McGarry (1977), whereas, in a different study, Carlson (1969b) suggested it was in the trough. Burpee (1972) initially concluded that there was no preferred region, in agreement with Bolton (1984) and Rowell and Milford (1993), whereas later Burpee (1974) found it to be mainly west of the trough south of 12.5°N and west of the ridge north of 12.5°N. More recently, Duvel (1990) has suggested that at around 10°N there is increased convection west of the trough, whereas at around 20°N increased convection is more likely east of the trough, in some agreement with Burpee (1974).

Before examining the 'moist' normal modes, it is first worth considering the effect of moisture on a dry normal-mode-type structure like that in Fig. 8. This exercise is perhaps quite pertinent since the easterly waves are thought to be initiated at around 20–30°E (see section 3(c)) where the moisture supply is small. It is only when they reach about 0° that the moisture supply becomes large and the effects of latent heat release more likely to be significant.

Figure 11(a) shows a schematic of a cyclonic vorticity anomaly centred at the level of an easterly wind maximum. The vertical-velocity field associated with this anomaly in the case of dry dynamics is indicated by the small arrows, and corresponds to the checkboard pattern described above (see section 4(d)). Thus, below the jet, warm northerly air rises ahead of the vortex and cooler southerly air sinks behind the vortex in the right sense for baroclinic growth. Note that, since the specific humidity around 15°N over west Africa decreases from large values on the equatorward side of the jet to very small values on the poleward side, the southerlies are expected to be moist and the northerlies dry. Clearly, the baroclinic structure of the easterly wave with descending southerlies is not conducive to saturation of the moist air. This is in contrast to baroclinic growth in mid latitudes where latent heat release is usually in the same sense as the dry baroclinic heat fluxes and results in a positive feedback. The different behaviour here is clearly due to the fact that over west Africa the meridional temperature and humidity gradients have opposite signs, whereas in the usual mid-latitude situation they have the same sign.

Insight can be gained regarding the observed convection patterns if we consider the situation depicted in Fig. 11(a), but at different latitudes. For example, Fig. 11(b) indicates the situation around 20°N in the Saharan region where the low-level air is very dry. This is clearly not conducive to deep convection. The only position in the wave where moist effects may be important is above the jet in the southerlies. Here air is also rising and so saturation may occur if the air is moist enough.

Figure 11(c) indicates the situation around 10°N. Here the air at low levels is very moist, and it is suggested here that the low-level ascent ahead of the trough may be an important trigger for deep convection, despite this air being northerly and despite there being descent above the jet. Above the jet level to the rear of the trough again, large-scale condensation may occur in the southerlies, as in the schematic at 20°N. This schematic depicting the situation poleward and equatorward of the jet may help to explain some of the observed convection patterns, in particular those of Duvel (1990) and Burpee (1974).

It is also worth noting here that part of the wave where squall-line development may
be expected to be most likely. One condition for squall-line development is the presence of a convectively unstable lapse rate. Looking at Fig. 8(b) showing the meridional wind, one can see that, just west of the trough at low levels, the meridional wind becomes more northerly with height—a consequence of the baroclinic nature of the wave. Therefore, if we consider the observed meridional $\theta_e$ distribution with highest values south of $15^\circ$N and at the ground, and with lower values north of $15^\circ$N and above the ground, the differential meridional $\theta_e$ advections west of the trough in the sloping baroclinic region will tend to decrease the quantity $\partial \theta_e / \partial z$ making it more convectively unstable. Squall-
line development may therefore be expected to be most likely in this region ahead of the trough, and to be a direct consequence of the baroclinic nature of the easterly wave.

(iii) **Latent heat parametrizations.** Two simple parametrizations have been used here and will be referred to as the $\theta_c$-scheme and the CISK scheme. In the $\theta_c$-scheme, as first proposed by Bennets and Hoskins (1979), the ascending air is supposed to be saturated at all levels. The heating rate is then calculated given that the rising air conserves a particular value of $\theta_c$. In this case a value of 345 K was chosen for $\theta_c$ to represent the low-level moist air over west Africa. The heating rate per unit mass is then given by the expression $-\omega C_p A(p)$, where $A(p)$ is an analytical approximation to the heating gradient in the presence of moist adiabatic ascent with constant $\theta_c$ equal to 345 K. The analytical profile is given in Fig. 12; it has a maximum at around 500 mb.

![Figure 12. The heating gradient profile used in the $\theta_c$-scheme, based on saturated ascent whilst conserving $\theta_c = 345$ K.](image)

The CISK scheme is a wave-CISK scheme similar to that of Kwon (1989) and is based on the supposition that all the air rising at $\sigma = 0.894$ has a prescribed specific humidity and that the heating associated with the saturation of this 'moist' air is distributed instantaneously throughout the depth of the column according to a prescribed weighting function. Here the heating rate per unit mass is $Lq\omega^* \eta(p)/\Delta p$, where $L$ is the latent heat of vaporization, $q$ is the specific humidity, $\omega^*$ is the vertical velocity at $\sigma = 0.894$, $\Delta p$ is the depth of the atmosphere above this, and $\eta(p)$ is the normalized weighting function which distributes the heating in the vertical. The parametrization is kept as simple as possible by using a constant value of $q$ of 14 g kg$^{-1}$ from the equator up to 15$^\circ$N. Polewards of 15$^\circ$N, $q$ decreases to zero by about 20$^\circ$N, mimicking the observed gradients over west Africa (Reed et al. 1977). Rather than choosing $\eta(p)$ to give the observed heating distribution as in previous studies, $\eta(p)$ is allowed to vary without bias according to the heating distribution obtained through saturated ascent whilst conserving $\theta_c$ (as in Fig. 12). This gives a maximum heating rate at about 500 mb.

In approximate terms the $\theta_c$-scheme may be thought of as representing large-scale latent heat release, while the CISK scheme represents low-level driven convection. With this in mind, an extra constraint is included in the $\theta_c$-scheme which takes into account the observed meridional humidity gradients. First, the amplitude $A(p)$ is damped with
increasing latitude in the same way as \( q \) in the CISK scheme, and secondly, latent heating is allowed only when the meridional wind is polewards. This removes the unrealistic possibility of large-scale latent heat release in air moving equatorwards, which, in reality, is likely to be very dry. The CISK scheme and \( \theta_e \)-scheme described above therefore roughly represent the two types of convection discussed in section 6(b)(ii) above, the first driven by low-level ascent and the second by slantwise ascending poleward-moving air.

(iv) Growth-rate and phase-speed curves. The growth-rate and phase-speed curves for the \( \theta_e \)-scheme are shown in Fig. 13. The growth-rate curve peaks at \( m = 11 \), but is extremely flat with all wavenumbers examined having growth-rates greater than 0.3 d\(^{-1}\).

![Growth-rate and phase-speed curves](image)

Figure 13. The growth-rate (above) and phase-speed (below) curves as a function of wavenumber for the most unstable modes which grow on the basic state in Fig. 2, with the \( \theta_e \)-scheme (see text for details). Filled circles indicate modes which did not converge to within the specified accuracy of 10\(^{-2}\) (see text) but still converged to less than 0.03.

The growth rate maximum is 0.34 d\(^{-1}\)—a little larger than for the dry modes. For wavenumbers less than 9 the method of determining the normal modes fails because wavenumber \( 2m \) has a similar growth rate to \( m \), which makes convergence too slow or impossible. In fact the \( m = 9 \) value shown here is probably the least accurate. The phase speeds are larger than for the dry modes, with values around \(-10.7^\circ \text{d}^{-1}\) for \( m = 11 \), equivalent to about \(-13 \text{ m s}^{-1}\).
The growth-rate and phase-speed curves for the CISK scheme are shown in Fig. 14. The growth rate has a maximum at around \( m = 11 \) and 12 with a value of \( 0.30 \) d\(^{-1}\)—only a little higher than for the dry modes. As with the \( \theta_c \)-scheme, the phase speeds are larger than for the dry modes. For \( m = 11 \) the phase speed is \(-8.3\) deg d\(^{-1}\)—equivalent to about \(-10\) m s\(^{-1}\).

Thus, both schemes increase the growth rate by only a small amount, in agreement with previous linear studies with parametrized latent heating. Both schemes however result in a slight shift of the most unstable mode to smaller wavelengths and a marked increase in phase speed. With latent heating, the phase speed of the easterly wave exceeds two thirds of the jet speed, whereas in the dry case it is about half.

(v) Global energetics. The global energetics for the \( \theta_c \) and CISK \( m = 10 \) modes are shown in Fig. 15. The contribution of the baroclinic conversion CE increases relative to the barotropic conversion CK for both schemes, but more so for the \( \theta_c \)-scheme. The barotropic energy conversion remains the most dominant energy source, even with latent heating included. For both schemes the direct input into eddy available potential energy due to the latent heating is much smaller than the other energy conversions.

(vi) Zonal EKE and EP fluxes. The zonal EKE and EP fluxes are shown for the \( \theta_c \)
mode in Fig. 16(a, b). Consistent with the discussion above in section 6(b)(ii), the \( \theta_c \) mode has maximum amplitude above the jet. The increased vertical component of the EP vectors is clear, consistent with increased positive heat fluxes.

The zonal EKE and EP fluxes are shown for the CISK mode in Fig. 16(c, d). The EKE for the CISK mode is very similar to the dry mode except that the main EKE maximum on the poleward flank of the jet is lower down at around 700 mb as opposed to near 500 mb, and the surface amplitude is now larger. The EP fluxes and their divergence also have more amplitude below the jet than the dry normal mode. As with the \( \theta_c \) mode, the vertical component of the EP vector below the jet is larger, implying stronger heat fluxes.

A comparison of these figures with the corresponding figure for the dry mode (see Fig. 6) suggests that the \( \theta_c \)-scheme tends to invigorate the dry mode structure above the jet, whereas the CISK scheme tends to invigorate the dry mode structure below the jet.

(vii) **Vertical structures.** Longitudinal cross-sections of vertical velocity at 15°N are shown in Fig. 17 for the \( \theta_c \) and CISK, \( m = 10 \), modes to indicate the different structures. The modes have been scaled to give a maximum meridional wind perturbation of 5 m s\(^{-1}\).
Figure 16. Latitude–height sections showing the structure of the most unstable modes at zonal wavenumber 10 growing on the basic state shown in Fig. 2: (a, b) the $\theta$-scheme; (c, d) the CISK scheme (see text for details). (a) and (c) show the EKE both with a contour interval of 1 J kg$^{-1}$; (b) and (d) show the EP flux vectors and their divergence with a contour interval of 6 and $8 \times 10^{14}$ m$^3$, respectively. The modes have been scaled to give a maximum meridional wind perturbation of 5 m s$^{-1}$. In each case the zero contour is dotted and negative contours are dashed.
Figure 17. Longitude–height sections at 15°N looking polewards, of vertical velocity for the m = 10 normal modes using: (a) the $\theta_e$-scheme; (b) the CISK scheme. The modes have been scaled to give a maximum meridional wind of 5 m s$^{-1}$ and the contour intervals in (a) and (b) are 0.8 and 0.2 mb h$^{-1}$, respectively. In each case the zero contour is dotted and negative contours are dashed.

For both modes the maximum meridional wind is poleward of the jet, whereas the maximum vertical velocity is near the jet. At 15°N the $\theta_e$ mode has a maximum vertical velocity of −4.1 mb h$^{-1}$ at about 450 mb, consistent with the zonal EKE section. The most striking feature however is the absence of the ‘checkerboard’ pattern and the contraction in scale of the ascent region with respect to the descent region. This type of structure has not hitherto been discussed in the literature with regard to easterly waves. Indeed most of the observational work on easterly waves has been unable to show this, owing to the form of presentation of composites with the longitudinal scale given by position within the wave rather than by length scale. The ratio of the ascent scale to the descent scale, $L_a/L_d$ is about 0.3 for the $\theta_e$ mode. This result shows that the previous work on the inclusion of a latent heat parametrization is inaccurate owing to the absence of higher harmonics.

The vertical velocity for the CISK mode has a similar amount of scale contraction with the ratio $L_a/L_d$ equal to about 0.4. At 15°N the maximum amplitude of −1.4 mb h$^{-1}$ is at about 800 mb but it also has significant amplitude up to about 300 mb. The perhaps unrealistic ‘checkerboard’ pattern seen in the dry modes is much weaker, although at 20°N where the latent heating is zero this pattern returns. This CISK structure has more
in common with the observations (Thompson et al. 1979) than that of the \( \theta_e \) mode. The observed ascent pattern also has a maximum at about 800 mb and a weaker one at about 350 mb.

7. Conclusion

Following on from the work of Simmons (1977), we have re-examined the linear instability problem of the African easterly jet using the primitive equations on a sphere. Insight into this problem has been gained by using diagnostics traditionally used for mid-latitude studies; this has included an examination of the African easterly jet from a potential vorticity perspective, and also of the nature of the energy fluxes associated with the growing normal modes using Eliassen–Palm fluxes. For example it has been shown that a growing normal mode on an AEJ is characterized by divergent EP fluxes in the region of the jet, implying both barotropic and baroclinic energy conversions. This also implies the removal of the negative zonal-mean quasi-geostrophic potential-vorticity gradient, \( \vec{q}_y \), which gives rise to the linear instability in the first place.

In agreement with previous linear studies of African easterly jets the normal modes are seen to grow mainly through barotropic energy conversions. The linear instability is dominated by the interaction between positive and negative \( \vec{q}_y \) at the level of the jet. Many of the synoptic features associated with the normal modes are in good agreement with those observed. The largest discrepancy, however, is with the vertical-velocity field which, consistent with dry dynamics, has a 'checkerboard' pattern in the vertical and is also much weaker than observed.

Changing the jet latitude while keeping the baroclinicity constant showed that, although barotropic energy conversions remain dominant, the most unstable modes become more baroclinic when the jet is more poleward. A thinner jet results in a larger growth rate, a smaller wavelength for the most unstable mode and reduced importance in baroclinic energy conversions. The 'checkerboard' pattern in the vertical velocity persists with the normal modes which grow on these jets.

The effect of including diabatic effects in the linear problem has also been examined. A simple boundary-layer scheme had very little effect on the normal modes, reducing the growth rates only slightly. This is consistent with the fact that the dry instability mechanism is dominated by the internal \( \vec{q}_y \) interactions at the level of the jet maximum. Two simple parametrizations of latent heating have also been used here, one representing low-level driven convection (CISK-type) and the other representing large-scale ascent (\( \theta_e \)-type). In agreement with previous linear studies which included the effect of latent heating, the growth rates found here using the two different schemes were increased only slightly. However, the schemes were found to change the normal-mode structures substantially. The most unstable modes became less dominated by barotropic energy conversions. Also, the most significant change to the synoptic structure is the asymmetry found between the ascent and descent regions in the wave. In agreement with previous moist baroclinic instability studies the length scale of the updraught, \( L_u \), is less than that of the downdraught, \( L_d \), the ratio \( L_u/L_d \) being about 0.3 and 0.4 for the \( \theta_e \)-scheme and CISK scheme, respectively. This property of easterly waves has not been taken into consideration in many of the observational studies of easterly waves. Also, with latent heating included, the perhaps unrealistic 'checkerboard' pattern in the vertical velocity is almost removed and the amplitude increased.

In conclusion, it appears that African easterly waves arise through a mixed barotropic/baroclinic instability mechanism where the role of latent heating is important in increasing the baroclinic energy conversions relative to the barotropic energy con-
versions and also in determining the synoptic structure. However, even with latent heating effects included, barotropic energy conversions still dominate the instability, in disagreement with observations over west Africa and with GCM results.

As noted in the introduction, most of the observations of easterly waves are from western Africa and the eastern Atlantic. In this region the waves may be expected to be nonlinear, and so it is suggested that, to better understand the easterly wave developments and to make possible a more relevant comparison with the observations, nonlinear integrations of easterly waves should be examined. A study to this end is described in Part II of this paper (Thornicroft and Hoskins 1994). Further studies should take account of the longitudinal varying environment on which the waves grow, and also consider the more realistic effect of the inclusion of diabatic processes.

REFERENCES


Chen, Y.-L. and Ogura, Y. 1982 Modulation of convective activity by large-scale flow patterns observed in GATE. *J. Atmos. Sci.*, 39, 1260–1279


Thompson, R. M., Payne, S. W., Recker, E. E. and Reed, R. J. 1979 Structure and properties of the synoptic-scale wave disturbances in the intertropical convergence zone of the eastern Atlantic. *J. Atmos. Sci.*, 36, 53–72


