

QUARTERLY JOURNAL
OF THE
ROYAL METEOROLOGICAL SOCIETY

Vol. 123

JULY 1997 Part B

No. 541

Q. J. R. Meteorol. Soc. (1997), **123**, pp. 1445–1482

The role of inertial instability in determining the location and strength of
near-equatorial convection

By ROBERT A. TOMAS and PETER J. WEBSTER*

University of Colorado, USA

(Received 23 October 1995; revised 23 December 1996)

SUMMARY

There are two major organized cloud configurations in the vicinity of the equator. Where there is a small cross-equatorial surface pressure gradient, convection is close to the equator and is generally tied to the location of the lowest sea-level pressure (SLP) and warmest sea-surface temperature (SST), in agreement with arguments based upon simple thermodynamical considerations. However, when there is a substantial cross-equatorial pressure gradient, such as occurs in the monsoon regions, organized convection appears off the equator in the summer hemisphere, equatorward of the SLP minimum and not necessarily collocated with the warmest SSTs. Thus, in this instance, simple thermodynamical considerations alone cannot explain the location of the convection. In this situation, the zero absolute vorticity contour ($\eta = 0$) also lies in the summer hemisphere. Therefore, between the equator and the $\eta = 0$ contour is a region of locally-anticyclonic absolute vorticity and an inertially unstable regime. It is argued that the convection results from the low-level divergence–convergence doublet centred about the $\eta = 0$ contour which is the mitigating response to the inertial instability. The associated latitude–height secondary circulation should provide subsidence (suppressed convection) over the equator and rising motion (enhanced convection) to the north of the zero absolute vorticity contour.

Signatures of the inertial instability predicted by theory are found in observations supporting the hypothesis. Wherever a strong cross-equatorial pressure gradient exists, the $\eta = 0$ contour bisects a maximum in the divergent wind field. Divergence is found equatorward of the zero contour and convergence on the poleward side. Latitude–height cross sections show strong local meridional circulations with maximum rising motion on the poleward side of $\eta = 0$. As the regions where the rising motions occur are conditionally unstable, there is deep convection and the vertical circulations extend throughout the troposphere. It is noted that the intensity of the off-equator convection is deeper (and probably stronger) than convection located at the equator. This is probably because the convection associated with the inertial instability is more efficient. Necessary conditions for the location of near-equatorial convection are listed.

Arguments are presented whereby inertial instability is established as the cause, rather than an effect, of off-equatorial convection. These include an outline of the sequence of processes leading up to the convection. The factors that limit the encroachment of the $\eta = 0$ contour into the summer hemisphere are discussed and an explanation for the existence of the low-level westerly monsoon wind maximum is suggested. The possible role played by the instability mechanism (or the lack of it) in coupled model simulations that produce seasonally migrating and/or double ITCZs in the eastern Pacific Ocean is discussed. Finally, it is proposed that the instability mechanism is important in the initiation of westward-moving disturbances found in the eastern Pacific and in determining active and break periods in the summer Indian monsoon.

KEYWORDS: Absolute vorticity advection Inertial instability Intertropical convergence zone Symmetric instability Tropical convection

1. INTRODUCTION

It has been long realized that near-equatorial convection possesses a global climate significance (see, for example, Riehl and Malkus 1958; Riehl 1979; Hastenrath 1988). The zonally-oriented bands define the region of confluence of the trade-wind regimes and are

* Corresponding author: Program in Atmospheric and Oceanic Sciences, University of Colorado, Boulder, CO 80309, USA.

the point of coalescence, in a mean sense, of the local Hadley circulations. Furthermore they harbour the deepest convection ('hot towers') which has been claimed to account for most of the necessary vertical heat-transport in the tropics to satisfy a global heat-balance (Riehl and Malkus 1958). Whereas the importance of deep tropical convection is readily accepted, there still appears to be considerable argument regarding the nature of the so-called ITCZ and the mechanisms that define its location.

Figures 1(a) and 2(a) show the distribution of the mean sea-surface temperature (SST) and sea-level pressure (SLP) for February 1992. Superimposed on the SST diagrams, are contours of outgoing longwave radiation (OLR) (W m^{-2}) which are indicators of deep penetrative convection and, consequently, significant precipitation. Superimposed on the SLP diagrams are hatched and stippled bands indicating regions where the cross-equatorial surface pressure gradients are strong and weak, respectively. The strongest and most persistent precipitation in the near-equatorial latitudes occurs within the zonally oriented cloud bands often referred to as the Intertropical Convergence Zone (ITCZ). The heavy bold line represents the zero absolute vorticity contour (i.e. $\eta = (\partial v/\partial x - \partial u/\partial y + f) = \zeta + f = 0$) at 925 hPa, the relevance of which will be discussed later. Figures 1(b) and 2(b) show the same fields for July 1992. When viewed from a mean perspective on timescales of weeks to months, the zonally oriented convective features appear as coherent entities around the globe. On shorter timescales, the ITCZ is made up of a series of propagating disturbances and, in the major warm pools, aperiodic conglomerations of mesoscale convective clusters (Webster and Lukas 1992). During the boreal summer, the regions of organized convection are generally found in the northern hemisphere (NH) displaced 4° to 12° north of the equator in certain regions, especially in the Atlantic Ocean and eastern Pacific Ocean. Over the Indian Ocean, they are found at a similar distance to the north of the equator in the boreal summer and to the south of it during the boreal winter.

(a) *The relative location and intensity of convection, SST and surface pressure*

The broadest dynamic feature of the equatorial regions is the near-equatorial trough (Hastenrath 1988). This surface pressure minimum is generally collocated with maxima in the SST field or with the monsoon regions over the land areas. Embedded within the low-pressure trough are regions of wind convergence, the axis of confluence between the trade-wind regimes of the two hemispheres and maxima in convection and precipitation (Hastenrath 1988). Hastenrath points out that there is a general misconception that all these features, often collectively referred to as the ITCZ, are of similar scale and are necessarily collocated. Careful regional analysis (for example by Ramage 1974) has shown that collocation does not occur, at least in the regions considered. In particular, the lateral scales of convergence, cloudiness and convection are much smaller than the broad extrema in surface pressure and SST. Indeed, Hastenrath (1988) raises the important point that if the maximum convection were simply tied to the maxima in the SST, how could this extremum be maintained in a region of minimum insolation? Figures 1 and 2 show that there are two modes of organization. Collocation of minima in surface pressure, maxima in SST, and convection or convergence does not occur where the cross-equatorial pressure gradient is strongest (red bands e.g., in the eastern Pacific and eastern Atlantic Oceans during the boreal summer and the Indian Ocean during both summer and winter). On the other hand, where the cross-equatorial pressure gradient is weakest (stippled bands), the collocation of convection, SST maxima and surface pressure minima is more readily observed.

Some comments on the strength of the convection in the different regions need to be made. Convection of nearly equal intensity, as indicated by comparably low values of OLR ($< 200 \text{ W m}^{-2}$), occurs in both regions, despite the seas being about 2 deg C

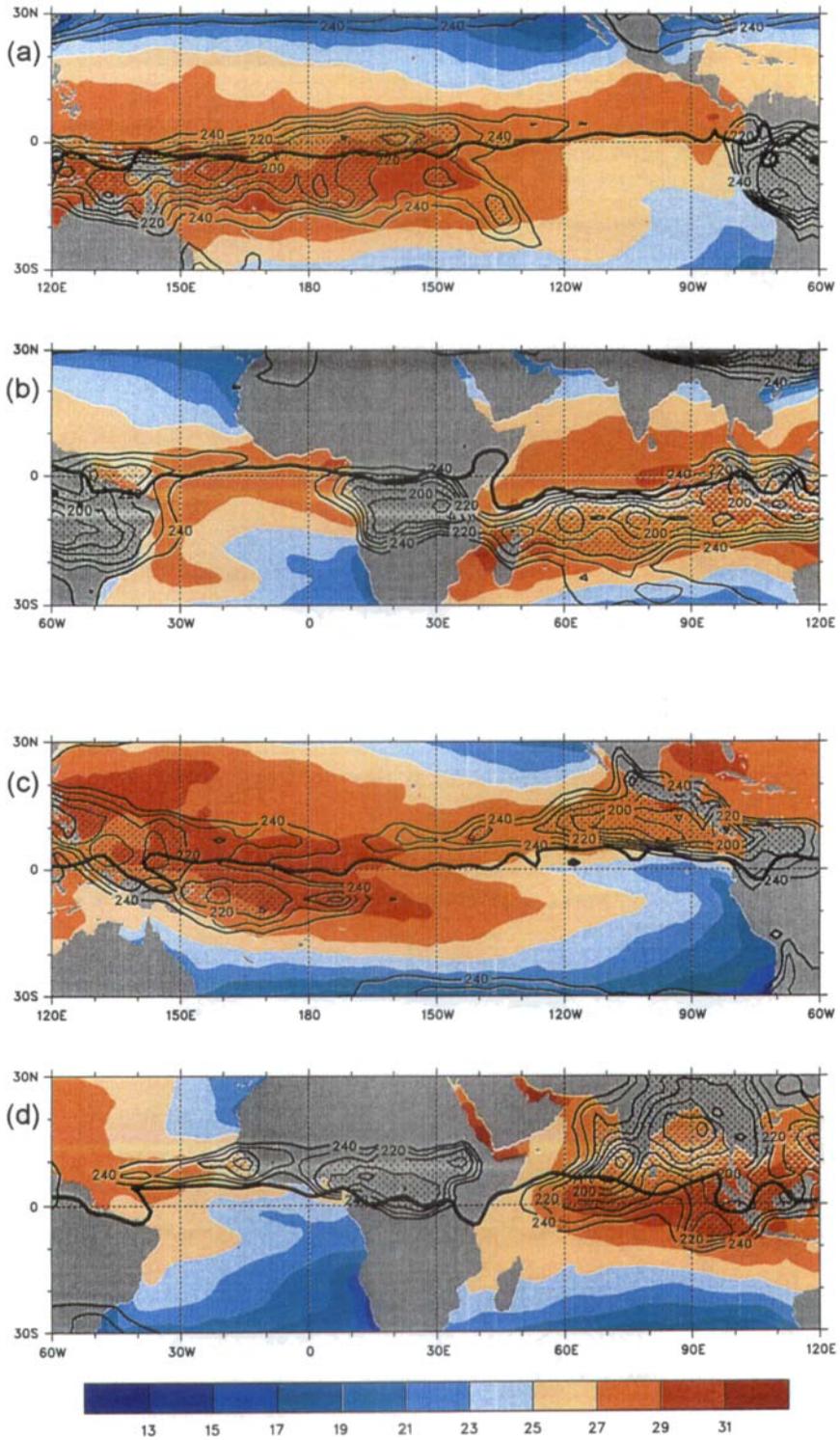


Figure 1. Distributions of outgoing long-wave radiation measured from satellite (OLR) (W m^{-2}) and sea-surface temperature (SST) ($^{\circ}\text{C}$): (a) and (c) Pacific Ocean, 120°E to 60°W ; (b) and (d) Atlantic and Indian Oceans, 60°W to 120°E ; (a) and (b) February 1992; (c) and (d) July 1992. Bold lines denote the zero absolute vorticity contours ($\eta = \zeta + f = 0$) at 925 hPa; stippled areas denote OLR $< 240 \text{ W m}^{-2}$; only areas where OLR $< 240 \text{ W m}^{-2}$, indicating deep convection, are shown. Values of SST ($^{\circ}\text{C}$) are given by the scale at the bottom of the figure.

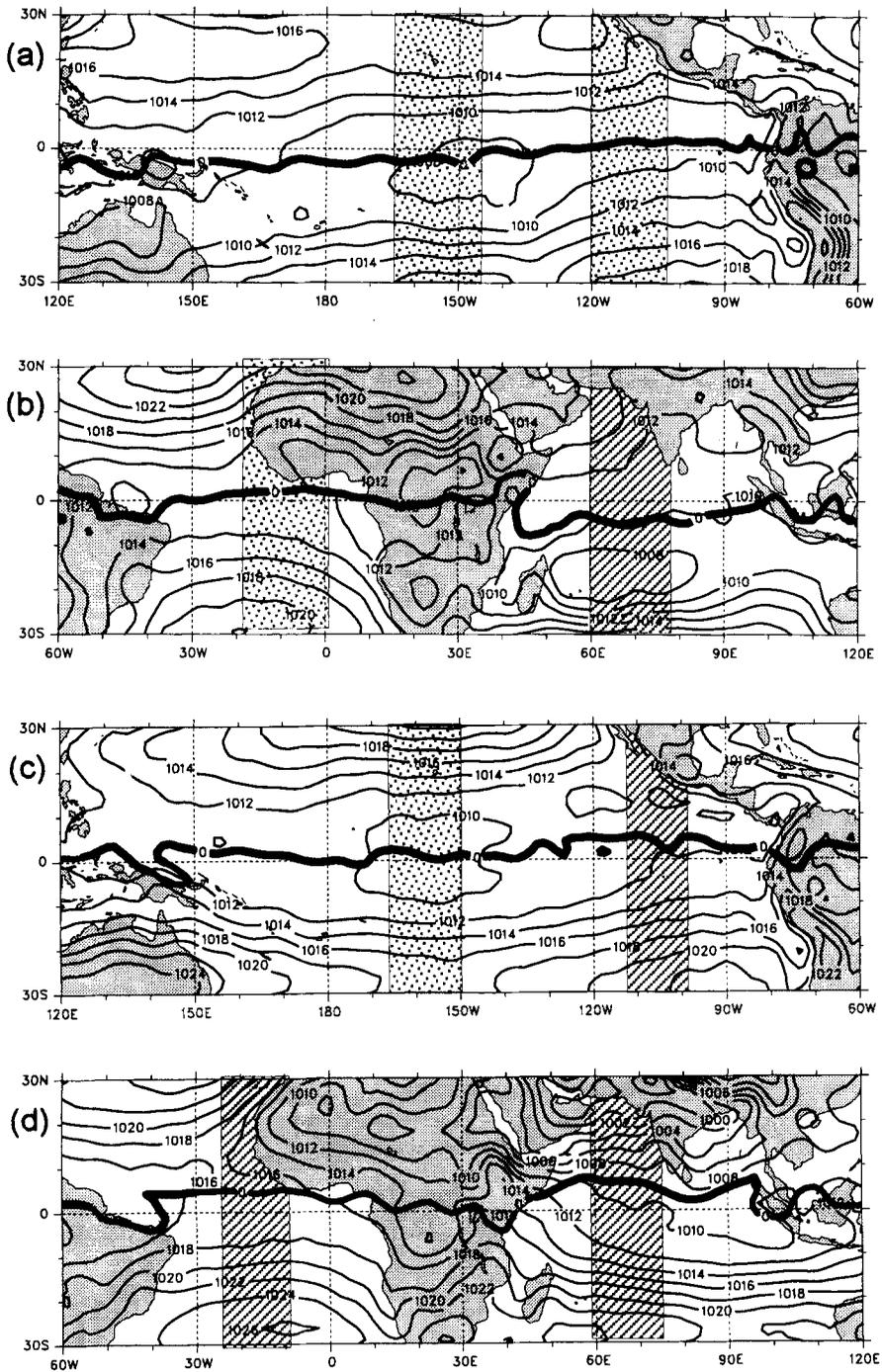


Figure 2. Distributions of sea-level pressure (SLP): (a) and (c) Pacific Ocean, 120°E to 60°W; (b) and (d) Atlantic and Indian Oceans, 60°W to 120°E; (a) and (b) February 1992; (c) and (d) July 1992. Bold lines denote the zero absolute vorticity contours ($\eta = \zeta + f = 0$) at 925 hPa; hatched and stippled areas indicate where the cross-equatorial gradient of SLP is strong and weak respectively.

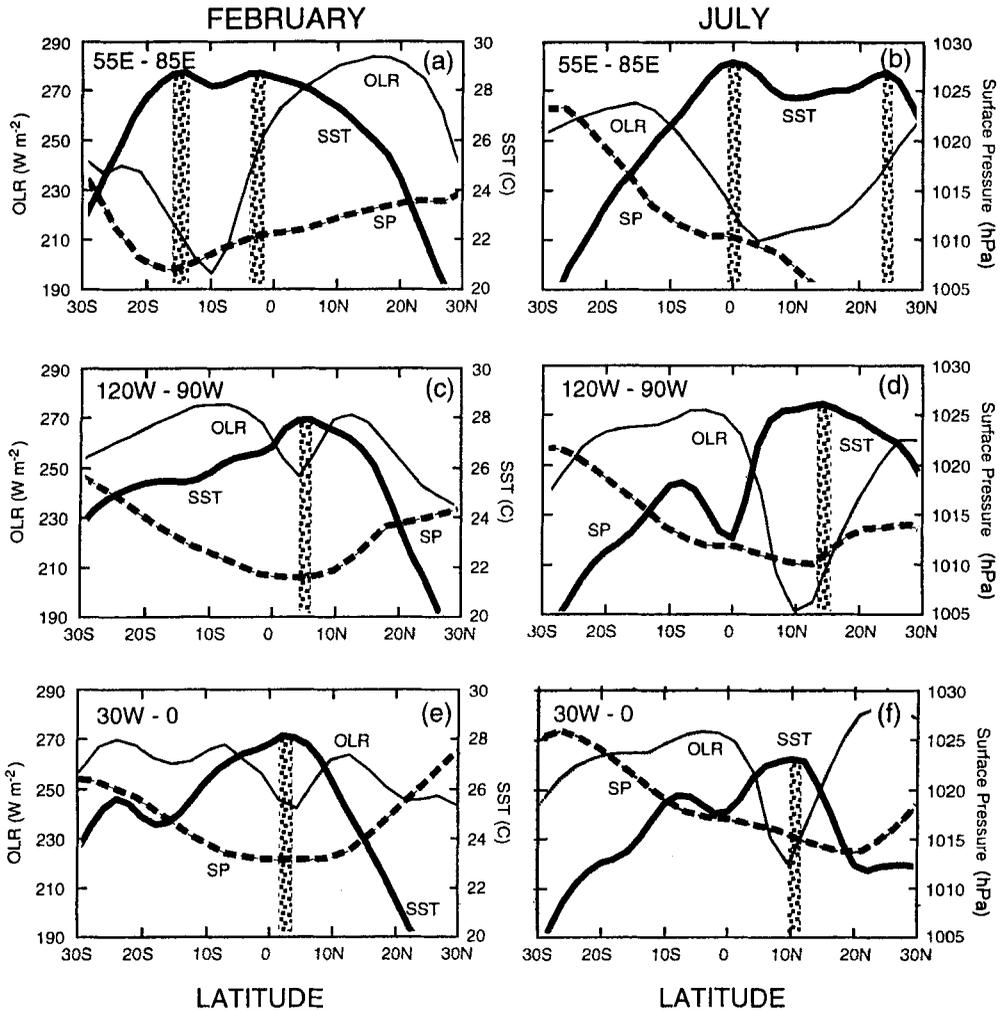


Figure 3. Distributions of outgoing long-wave radiation measured from satellite (OLR) ($W m^{-2}$) (thin solid lines), sea-surface temperature (SST) ($^{\circ}C$) (heavy solid lines) and mean sea-level pressure (SP) (hPa) (heavy pecked line) for three longitude bands: (a) and (b) Indian Ocean, 55°E to 85°E; (c) and (d) eastern Pacific Ocean, 120°W to 90°W; (e) and (f) eastern Atlantic Ocean, 30°W to 0°; (a), (c) and (e) February 1992; (b), (d) and (f) July 1992. Stippled areas denote the areas of maximum SST.

warmer in the near-equatorial regions. The presence of equally strong convection in both regions suggests that some mechanism is acting to increase the strength and/or efficiency of convection in the off-equatorial locations, to compensate for the lower SSTs there.

Figure 3 shows the distributions of monthly-mean OLR measured from satellite, SST and surface pressure for longitudinal bands in the Indian Ocean for February 1992 and July 1992. In the Indian Ocean, during the boreal summer (Fig. 3(b)), the maximum SST is nearly centred over the equator. Maximum convection occurs at 5°N but equatorwards of the low pressure minimum. Summer convection in the Atlantic is slightly equatorwards of the SST maximum and well equatorwards of the surface pressure minimum. In the eastern Pacific Ocean section (Fig. 3(d)), boreal summer convection occurs equatorward (10°N) of the SST maximum (14°N) and the surface pressure minimum (14°N). During the boreal winter similar characteristics occur in the Indian Ocean (Fig. 3(a)). The maximum

TABLE 1. NEAR-EQUATORIAL CONVECTION, SST AND SURFACE-PRESSURE EXTREMA

Location	Period (1992)	OLR minimum (W m^{-2})	SST maximum ($^{\circ}\text{C}$)	SLP minimum (hPa)
Indian Ocean (55°E – 85°E)	July	210, 5°N	29.2, 0°	<1005, > 14°N
	February	196, 10°S	28.8/28.8, $4^{\circ}\text{S}/14^{\circ}\text{S}$	1007, 17°S
Pacific Ocean (120°W – 90°W)	July	192, 10°N	28.5, 16°N	1010, 14°N
	February	250, 4°N	28, 5°N	1008, 5°N
Atlantic Ocean (30°W – 0°)	July	220, 10°N	27, 11°N	1014, 20°N
	February	244, 5°N	28.2, 4°N	1012, 5°N

SST occurs near 4°S (although there is a second maximum near 14°S) while convection is maximized at 10°S . In the other two boreal winter sections (Fig. 3(c,e)) there is a closer correspondence between extrema in the SST and OLR. But, in these two cases the convection is very weak, with OLR values of 250 W m^{-2} compared to values between 190 W m^{-2} and 220 W m^{-2} in all other cases shown. The locations of the extrema are summarized in Table 1.

Figure 3 may be summarized as follows.

(i) The equatorial trough, convection and the SST maxima appear collocated when there is little or no cross-equatorial pressure gradient. These collocations occur in the mid-oceans, well away from continental land masses and monsoon regions.

(ii) In the monsoon regions, where there is a strong cross-equatorial pressure gradient, convection is located on the low-pressure side (i.e. summer hemisphere) of the equator. The convection is not collocated with the minimum surface pressure but occurs equatorward of it. This is important as it means that the location of the ITCZ is not responsible for the existence of the broad-scale latitudinal pressure gradient. The broad pressure gradient in the monsoon regions is the result of differential heating of land and ocean.

(iii) Most of the time, the convective maxima are located equatorward of the surface pressure minima and either equatorwards or polewards of the SST maxima. Furthermore, when the convection is weak (i.e. values of OLR are larger), there is a closer collocation of extrema in convection, SST and surface pressure.

Hastenrath (1988) also notes that there is a strong correspondence between the annual variation of the trough and the SST maximum which indicates that the trough is a heat-induced phenomenon defining an annually varying thermal equator. In that sense, the reason for the existence of the broad thermal trough is basically understood. Thus, the primary question is why the maxima in convection and precipitation reside in a particular part of the trough and the SST distribution? Point (iii) above, raises another point. Why is off-equator convection stronger than near-equator convection?

(b) *Theories of the location of near-equatorial convection*

Two main classes of theories can be identified. In the first, the position of the near-equatorial convection is assumed to be determined almost exclusively by the SST distribution. Manabe (1969) and Manabe *et al.* (1974), using general circulation models, theorized that the ITCZ would be located where the near-equatorial local SST maximum existed. If the SST maximum were at the equator, the convection would lie at the equator. If equatorial upwelling occurred, convection could lie on either side of the equator producing a single ITCZ or possibly on both sides, producing a double ITCZ. Based on results from a zonally symmetric coupled ocean–atmosphere model, Pike (1971) suggested that cold equatorial

upwelling resulting from equatorial easterlies was responsible for the off-equator location of convection. In this theory, deep convection is associated only with SSTs that are greater than or equal to 27°C. The second class of theories depend on the dynamics of near-equatorial disturbances (e.g. Holton *et al.* 1971; Lindzen 1974). An earlier hybrid dynamic–thermodynamic theory had been proposed by Charney (1971) in which the off-equatorial deep convection was controlled by an optimization of Ekman pumping (which decreased towards the equator) and the availability of moisture (which decreased away from the equator). Mitchell and Wallace (1994) hypothesized the existence of a robust cold tongue–ITCZ complex resulting from interaction between the atmosphere and the ocean in the eastern Pacific and Atlantic regions. A positive feedback between upwelling (which influences the SST distribution and thus sea-level pressure distributions) and the low-level winds is central to their argument. They suggest that relatively warm sea surfaces off Central America and West Africa lead to enhanced northward flow out of the cold tongues and a westward extension of continental convection resulting in the oceanic ITCZs. The idea of feedbacks between the atmosphere and ocean and the resulting pressure distribution are not inconsistent with the central hypothesis of this paper, which focuses on the atmospheric response to the pressure distribution near the equator.

Lindzen and Hou (1988), Hou and Lindzen (1992) and Plumb and Hou (1992) attempted to explain the location and magnitude of the tropical Hadley cell. These studies considered circulations relative to fixed and imposed heating distributions and only dry thermodynamics was considered. Despite these simplifications and lack of feedbacks, there are some interesting conclusions that pertain directly to the question at hand. Lindzen and Hou (1988) found that for the same magnitude of heating, locating the heating incrementally off the equator led to considerable increases in the magnitude and asymmetry of the winter hemisphere circulation. Hou and Lindzen (1992) found that concentrating the off-equator heating in the north–south direction increased the magnitude of the circulation considerably even if the total heating were not increased. The procedure was justified on the basis that condensation (if it had been included in the model) would tend to concentrate the heating in this way. They gave no specific reasons as to why the heating would ‘choose’ to be off the equator. However, recalling the results of Lindz and Nigam (1987), Hou and Lindzen (1992) suggested that an off-equator maximum of the SST would possess even greater significance by producing very strong symmetric circulations. This conclusion is not completely in keeping with the observations cited above, which show that strong convection is often not associated with maximum SST, either on the equator or off it. Although the three studies made the important point that off-equator heating leads to stronger responses, none of them presented a consistent argument about why the maximum heating occurred where it did.

In a similar vein, Schubert *et al.* (1991), Hack and Schubert (1990) and Hack *et al.* (1989) investigated the response of the tropical meridional (Hadley) and zonal circulations over a range of specified diabatic forcing fields. Their results demonstrate how the position and intensity of convection (i.e. heating) leads to very different large-scale circulation responses and how the inertial stability effects the poleward extent of the Hadley cell. However, they too did not address the question of why the heating occurred where it did.

Walliser and Somerville (1994) sought a dynamic basis for the off-equator location of convection. In developing their theory, they concluded that neither zonally propagating disturbances nor an off-equatorial maximum in the sea surface temperature (SST) is required for convection to be located off the equator. In this sense, they corroborated the observational results of Ramage (1974), Sadler (1975a, b) and Hastenrath and Lamb (1977a, b). They suggested that convection occurs some 4 to 12 degrees off the equator because, at these latitudes, low-level convergence is dynamically maximized, so that the convergence

of moist static energy is enhanced. These two processes would then conspire to produce the strongest convection. Walliser and Somerville (1994) argued that these collocations could be understood through an analysis of the low-level momentum budget. In a shallow-water system, an imposed low-latitude heat source induces a pressure-gradient anomaly across the equator which is balanced by a combination of friction and geostrophy. The friction is associated with low-level convergence and, because the atmosphere is conditionally unstable, convection. If the heat source is moved some small distance off the equator, the role played by friction increases on the equatorward side of the pressure anomaly, where the Coriolis force is small. Continued poleward displacement of the heat source leads to decreased low-level convergence because the Coriolis force on its equatorial side increases. Extending their analysis to a more complex, moist, primitive-equation model, Walliser and Somerville (1994) showed that if the convection is allowed to be determined internally, a positive feedback develops between the mid-tropospheric latent heating and the low-level convergence, with the effect of enhancing the organization of convection at latitudes between 4° and 12° .

Walliser and Somerville provided convincing evidence that a dynamical mechanism is responsible for the location of the off-equator convection and that location is where the low-level convergence is maximized. Indeed, it would be difficult to contemplate an explanation for the features shown in Fig. 3 without some type of dynamic influence or control. However, it is believed that the physical processes responsible for the convergence are somewhat different from those suggested by Walliser and Somerville. The explanation proposed here involves circulations resulting from inertial (symmetric) instability of the low-level flow which arise because of the displacement of the zero absolute (potential) vorticity contour from the equator by strong cross-equatorial pressure gradients.

A hypothesis that inertial instability in the lower troposphere determines the location of off-equatorial convection in the region of strong surface cross-equatorial pressure gradients is presented in section 2. In section 3, the parcel argument is used to illustrate qualitatively several important aspects of inertial instability. These qualitative arguments lead to a schematic template that may be used to recognize the manifestation of the instability. In section 4, monthly-mean horizontal fields showing the divergent wind and divergence are presented. Several features in these fields are interpreted as resulting from inertial instability. These features in the horizontal fields suggest the presence of secondary circulations in the height–latitude plane (section 5) which act to reduce the instability of the background flow. In section 6, the effect of atmospheric static stability on near-equatorial inertial instability is examined as well as the thermodynamic consequences of the interactions between the motion and the moisture fields. Summaries and concluding remarks are made in section 7. In addition, arguments are presented that suggest why the $\eta = 0$ contour is located where it is. Furthermore, a physical explanation for the monsoon low-level westerly wind maximum is presented. An explanation of why some coupled ocean–atmosphere models simulate the ITCZ poorly is also proposed. A description of the data and methodology used in processing the data is found in an appendix.

2. HYPOTHESIS: INERTIAL (SYMMETRIC) INSTABILITY AND EQUATORIAL CONVECTION

(a) Hypothesis

In discussing Figs. 1 and 2, it was noted that the off-equatorial ITCZ, which occurs in regions of strong cross-equatorial pressure gradient, was located poleward of the zero absolute vorticity (AV) contour. As a $\eta = 0$ contour displaced from the equator signifies inertial instability of the low-level flow, it is proposed that three conditions are necessary to determine the location of near-equatorial convection:

(i) the flow is locally inertially (symmetrically) unstable. The instability results from the advection of absolute (potential) vorticity across the equator by strong cross-equatorial pressure gradients arising either from the SST distribution about the equator and/or the unevenly heated distribution of land and ocean which occur, for example, in the Asian and African regions. The convection itself is produced by secondary circulations in the height–latitude plane that render the system stable;

(ii) the SST should be close to the maximum tropical SST but it need not be the maximum overall SST nor even the local SST maximum; and,

(iii) the atmosphere must be locally conditionally unstable poleward of the $\eta = 0$ contour. Conditional instability is necessary so that the atmosphere does not impede the development of the rising branch of the secondary circulations. In conjunction with the secondary circulations, conditional instability promotes deep penetrative convection. Convective intensity is enhanced through the convergence of moist static energy in the boundary layer. The downward branch of the circulations must be located where the atmosphere is relatively dry, so that descending air can cool radiatively, thus compensating for any adiabatic warming.

If the cross-equatorial pressure gradient is weak, so that there is no cross-equatorial advection of AV and the flow is inertially stable, the location of convection is determined by the magnitude of the SST and the thermodynamic state of the environment. That is, conditions (ii) and (iii) determine the location of convection.

An excellent example illustrates why condition (ii) is not sufficient to determine the location or even the existence of deep convection. During the boreal spring and early summer, the highest SSTs of the entire globe exist in the northern Indian Ocean with values between 29 °C and 30 °C. Yet, the region possesses an OLR maximum (convection minimum). The reason that convection is not collocated with this maximum in SST is that the atmosphere above the Indian Ocean warm pool is statically stable because of the large-scale, strong subsidence in the region caused by the landlocked nature of the ocean basin (Webster 1994; Loschnigg and Webster 1995).

Interactions and feedbacks between convection and circulations are notorious for making it difficult to distinguish cause from effect. However, it will be argued that the inertial-instability processes determine the location of convection only where there is a large-scale pressure gradient. Indeed, it is important to note that the strong off-equator convection occurs only in the monsoon regions, where differential heating over land and ocean produce a planetary-scale pressure gradient. Indeed, reference to Fig. 3 shows that in the monsoon region, the pressure gradient is on a much larger scale than the convection. That is, the ITCZ is not collocated with minimum pressure in these regions.

In the absence of convection, a cross-equatorial pressure gradient would still exist in the monsoon regions because of differential surface heating. Cross-equatorial flow and AV advection would ensue, setting up the inertially unstable regime. If sufficient moisture were present, secondary circulations would develop and convection would follow. Thus, it is appropriate to think of the inertial instability as a cause of off-equatorial convection rather than an effect resulting from the convection. This is not to say that the convection does not play a significant role in determining the observed circulation. Clearly, the response of a moist atmosphere will be very different from that of a dry atmosphere.

(b) *Inertial instability: atmospheric context*

The concept of inertial instability has a long history in the atmospheric sciences. Solberg (1936) introduced the concept of dynamic instability (now more frequently referred to as symmetric instability) into meteorology. The name is derived from the observation that

circulations resulting from the instability are symmetric when viewed along the direction of the basic-state flow (e.g. Emanuel 1988). Solberg made use of the parcel method, allowing symmetric instability to be seen as the result of two forces that are nearly perpendicular to each other, one acting in the horizontal direction (the imbalance between the pressure gradient and Coriolis forces) and the other in the vertical (the imbalance between the vertical pressure gradient and gravitational forces, defining buoyancy). Inertial instability is closely related to symmetric instability but only involves the horizontal forces. For motion along a constant potential temperature surface, the buoyancy force is zero and symmetric instability reduces to inertial instability. For motion along a surface of constant angular momentum, symmetric instability reduces to convective instability (e.g. Emanuel 1988).

The relevance of symmetric instability in meteorological problems was addressed by Sawyer (1949) a little over a decade after Solberg's work was published. Sawyer examined the possible role of symmetric instability in the formation of extratropical cyclones, limits on wind shear on the south side of jet streams, the formation of tropical cyclones and frontal structure. He concluded that, although it may not be important in the formation of extratropical cyclones, symmetric instability may play some role in the other phenomena listed above.

Hoskins (1974), in comments on the relevance of symmetric instability to fronts and jet streams, noted that symmetric instability could be triggered by the effective decrease in static stability due to latent-heat release and proposed that this might account for observations of rain bands made by Browning *et al.* (1973). Bennetts and Hoskins (1979) extended this argument by introducing the concept of conditional symmetric instability in their study of the formation of frontal rain bands. This form of symmetric instability includes the effects of latent heating in buoyancy forcing. Later observational work showed that frontal regions occur in a state that is near neutral with respect to conditional symmetric instability (Emanuel 1988). Conditional symmetric instability is also considered to be important in explosive cyclogenesis in the mid-latitudes (Shutts 1990).

(c) *Inertial instability and the tropics*

Consequences of symmetric instability for tropical flows were investigated by Stevens (1983) who focused on understanding the reasons for the lack of cross-equatorial wind shear in zonally symmetric global fields. Stevens concluded that symmetric instability is a major homogenizer of flow near the equator. Although the study presents a sound review of the importance of symmetric instability in the zonally averaged tropics, the restriction to zonally averaged flows precluded the consideration of strong regional cross-equatorial flows such as in the monsoon regions noted in the introduction.

Krishnamurti *et al.* (1983) studied low-level flow in the Indian monsoon region using a high-resolution, limited-area, planetary boundary-layer model forced by pressure fields obtained from observations. Of particular relevance, they noted that in the boundary layer there is geostrophic balance to the north of the ITCZ and an advective balance (i.e. parcels accelerating) to the south of the ITCZ (see also Krishnamurti and Wong 1979). As part of the BALSAMINE experiment, Reverdin and Sommeria (1981) analysed the trajectories of balloons at a height of 1 km over the Arabian Sea. They found that the balloons experienced a component of acceleration opposing their motion when they were located south of the equator. After crossing the equator, however, this component of acceleration reversed and pointed in the direction of the balloon's trajectory. In collaboration, Krishnamurti *et al.* (1983) noted that '... just north of the equator the flow is accelerated, with little impediment, toward lower pressure . . .'. The analyses in this study will show that the zero

AV contour in the lower troposphere is located north of the equator in the western Indian Ocean in summer indicating that the flow is inertially unstable.

Mapes and Houze (1992) noted that the $\eta = 0$ contour was '*... displaced from its resting latitude (i.e. the equator) into the southern hemisphere ...*' in the Indonesian–Australian sector during the boreal winter monsoon. The advection of the zero contour was noted to precede the advection of locally large values of positive AV on timescales of days. They noted that on the poleward side of the $\eta = 0$ contour there was considerable convergence which they attributed to convection. A different interpretation and sequence of events will be postulated here, where it will be argued that the convection is the result of the cross-equatorial advection of AV.

The modelling results of Rodwell and Hoskins (1995) show that for summertime conditions, parcels crossing the equator in the Indian monsoon region accelerate significantly when they are in a region that is symmetrically unstable (defined as where the product of potential vorticity (PV) and the Coriolis parameter is negative), thus suggesting symmetric (inertial) instability. In addition, the authors emphasize the importance of PV modification in allowing parcels that cross the equator to remain there. They argue that if the PV is not modified sufficiently, parcels will return to the hemisphere in which they originate.

(d) *Criteria for instability*

There are two equivalent criteria that may be used to determine whether or not a particular flow is symmetrically unstable. Firstly, if, on an isentropic surface, the absolute vorticity of the flow has the opposite sign to the local planetary component of vorticity (i.e. the absolute vorticity is 'locally anticyclonic'), the flow is symmetrically unstable.

In the lower tropical troposphere, the potential temperature variation on a constant pressure surface is small on large spatial scales (thousands of kilometres), at most about 5 K in the regions examined. That is, the isentropic surfaces are nearly parallel to the isobaric surfaces at low latitudes (Tomas and Webster 1994). Thus, for the case considered, symmetric instability reduces to inertial instability and the instability criterion for the near-equatorial region becomes the second criterion. This is, if, on a pressure surface, the absolute vorticity of the flow is of opposite sign to the local planetary component of vorticity (i.e. that the absolute vorticity is 'locally anticyclonic') the flow is inertially unstable.

Thus, the very similar structure of isentropic and isobaric surfaces near the equator renders the interpolation of data to isentropic surfaces unnecessary. This has the advantage that interpolation errors are not introduced into the data (or extrapolation errors when the desired constant potential temperature surface lies below the lowest pressure surface). Nevertheless, it will become apparent later that if the interpolation is made, the criteria for symmetric instability would also be met. Viewing the problem as composed of separate components, an inertial instability *and* convective instability, instead of a single process (i.e. moist symmetric instability), emphasizes the different roles played by the dynamics and the thermodynamics. This makes it easier to understand the underlying physical processes that conspire to produce off-equatorial convection.

(e) *Association between the location of the $\eta = 0$ contour and the large-scale surface pressure field*

Figure 4 shows the mean monthly AV at 925 hPa during February and July 1992. For the most part, values are negative in the southern hemisphere and positive in the northern hemisphere. However, this rule is broken in several locations where the $\eta = 0$ contour strays either southward and northward of the equator a few degrees into the summer

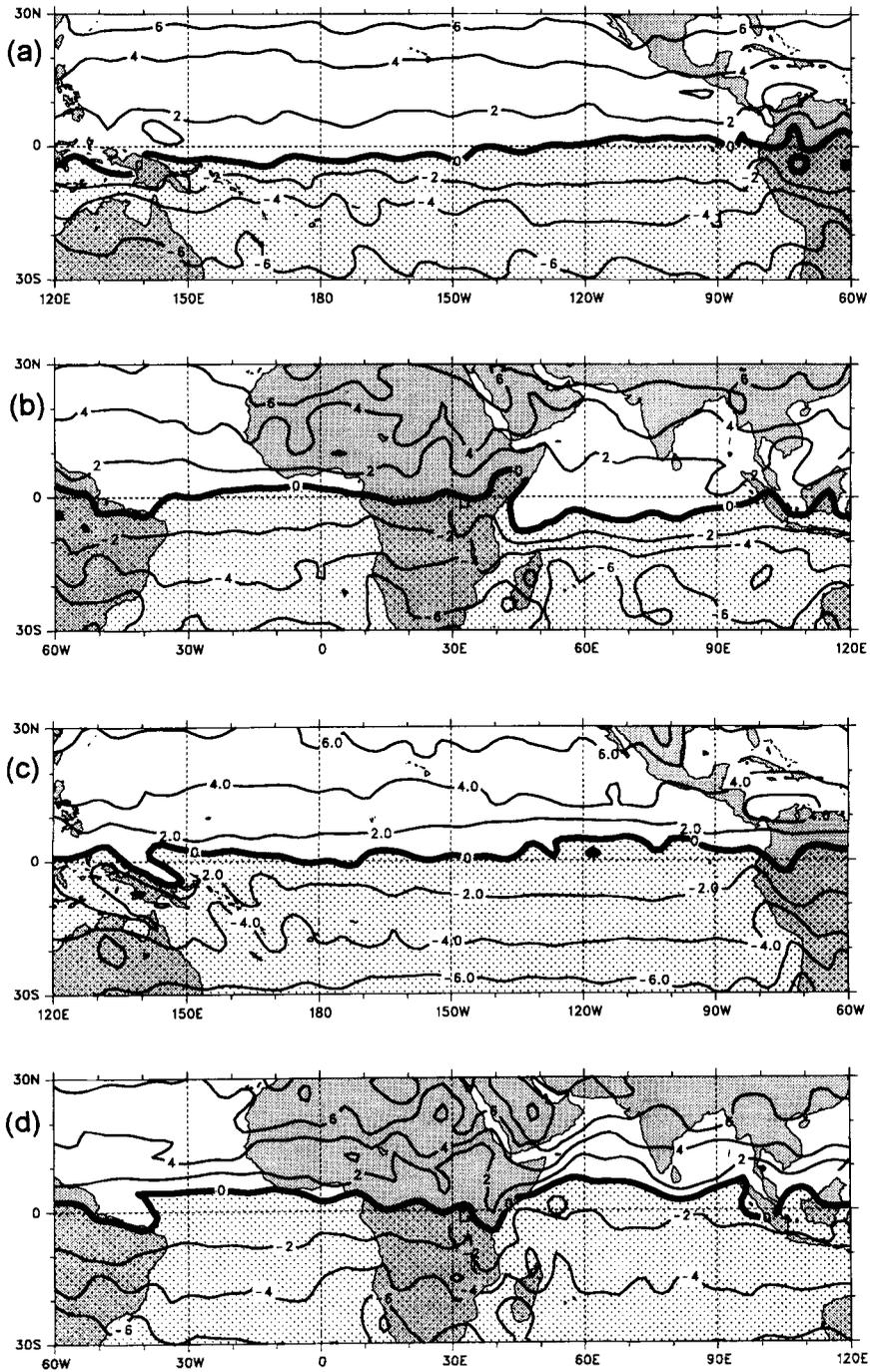


Figure 4. Distributions of monthly-mean absolute vorticity ($\eta = \zeta + f$) (10^{-5} s^{-1}) at 925 hPa: (a) and (c) Pacific Ocean, 120°E to 60°W; (b) and (d) Atlantic and Indian Oceans, 60°W to 120°E; (a) and (b) February 1992; (c) and (d) July 1992. Bold lines denote the zero absolute vorticity contours ($\eta = \zeta + f = 0$); areas of negative absolute vorticity are stippled; contour intervals $2 \times 10^{-5} \text{ s}^{-1}$.

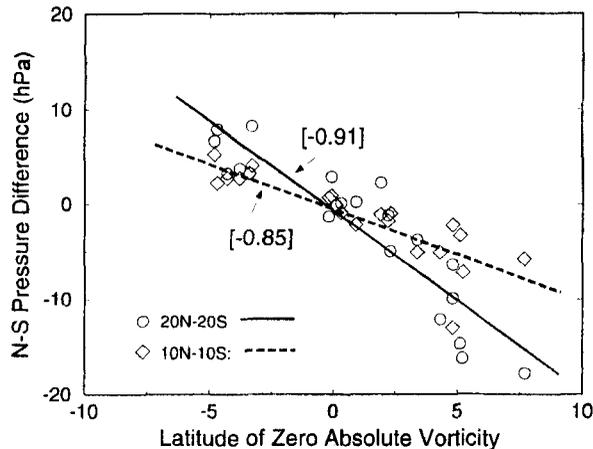


Figure 5. Latitude of zero absolute vorticity contour ($\eta = \zeta + f = 0$) at 850 hPa and the cross-equatorial surface-pressure difference: diamonds 20°N–20°S; circles 10°N–10°S. Regression lines and correlation coefficients for the two latitude zones are also shown.

hemisphere. The greatest deviations away from the equator occur in the regions of strongest convection (Figs. 1 and 2).

During February 1992, (Fig. 4(a)) the greatest departure of the $\eta = 0$ contour from the equator is to the south in the longitude band between 60°E and 90°E (southern Indian Ocean). In this region, there is strong southward monsoonal cross-equatorial flow. Positive AV extends as far south as 8°S and across the Indian Ocean with a mean latitude of about 5°S. During July 1992 (Fig. 4(b)) there are three regions where the $\eta = 0$ contour is located north of the equator. These are roughly between 120°W and 90°W (eastern Pacific), between 30°W and the Greenwich Meridian (eastern Atlantic), and between 60°E and 90°E (northern Indian Ocean). Negative AV is found to extend to about 5°N in the first two locations and to almost 10°N in the last.

Within the relationships discussed above, there is a strong association between the local cross-equatorial pressure gradient and the position of the $\eta = 0$ contour in the lower troposphere. Figure 5 shows the correlation of the latitude of the $\eta = 0$ contour with the cross-equatorial pressure difference between 20°N and 20°S and 10°N and 10°S. The correlations were -0.85 and -0.91 respectively, although the number of degrees of freedom is relatively small. It is important to note that the degree of incursion of the $\eta = 0$ contour is much the same whether the pressure gradient is calculated for $\pm 10^\circ$ about the equator or $\pm 20^\circ$. That is, the location of the $\eta = 0$ contour is determined by the macro-scale cross-equatorial pressure gradient.

Other months during 1992 have been examined and it was found that the February excursions (Fig. 4(a)) are representative of the period January to March, although the extent of the excursions is found to be less in other months. Similarly, Fig. 4(b) is representative of the period June to August. The excursions in the eastern Pacific and eastern Atlantic are also indicative of those found during September to November 1992. Figure 6 shows the monthly-mean positions of the $\eta = 0$ contour in the Indian Ocean–western Pacific Ocean sector based on a 15-year data base (1980–1994) at the 850 hPa level. The mean locations are similar to those found in February and July 1992 at the 925 hPa level. Thus, examination of the two months of 1992 should allow general conclusions to be drawn.

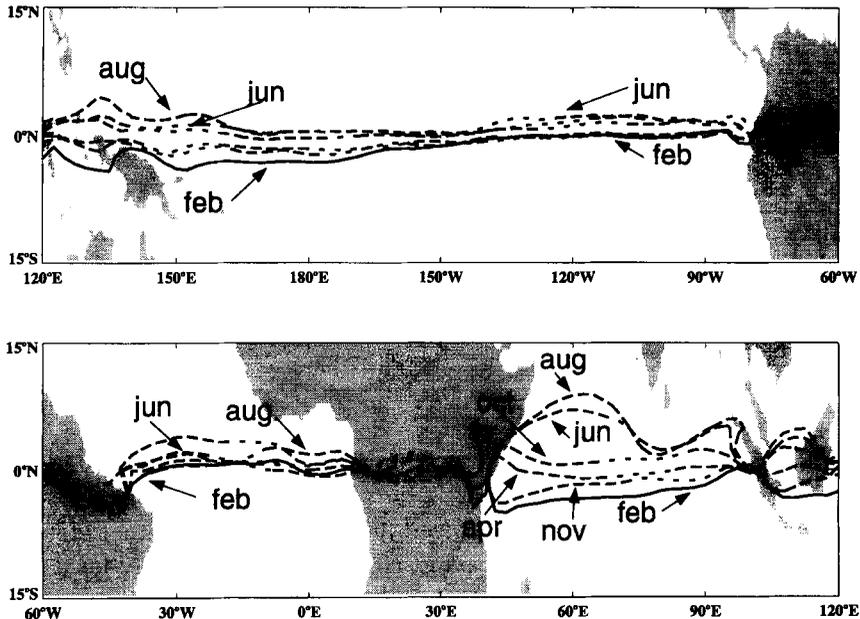


Figure 6. Long-term (1980–1994) monthly-mean positions of the zero absolute vorticity contour ($\eta = \zeta + f = 0$) at 850 hPa, from European Centre for Medium-Range Weather Forecasts (ECMWF) data.

In summary, in all regions where there is an excursion of the $\eta = 0$ contour across the equator, the criteria for inertial instability are met in the manner defined above. As these criteria are based upon a rather simple theoretical analysis, some scepticism is justified given the complexity of the real atmosphere. Nevertheless, the analysis presented so far is sufficiently interesting for us to seek consistencies with the concept of the instability.

3. INTERPRETATIONS

The parcel method, which has been used quite successfully to provide qualitative and quantitative insight into atmospheric processes (e.g. Sawyer 1949; Thorpe *et al.* 1989; Holton 1992), is employed to suggest the atmospheric response to the existence of locally-anticyclonic absolute vorticity in the lower troposphere. The most familiar context in which this method has been applied is with vertical displacements from hydrostatic equilibrium to test for convective instability.

It is possible to derive a criterion for inertial instability using similar reasoning, and, at the same time, shed some light on the physics that produces the instability. Choosing to look at inertial instability, rather than symmetric instability (see discussion in section 2), simplifies the analysis while retaining the physical processes important for our hypothesis. Further simplifications result from assuming that the basic state is purely zonal and in geostrophic balance. Such a derivation is well known and may be found in text books such as Holton (1992). For brevity, the details of the derivation are omitted.

The equation governing the non-disruptive motion of a parcel moving through a zonal basic-state field that is in geostrophic balance, is

$$\frac{dv}{dt} = \frac{d^2(\delta y)}{dt^2} = -f \left(f - \frac{\partial u_g}{\partial y} \right) \delta y, \quad (1)$$

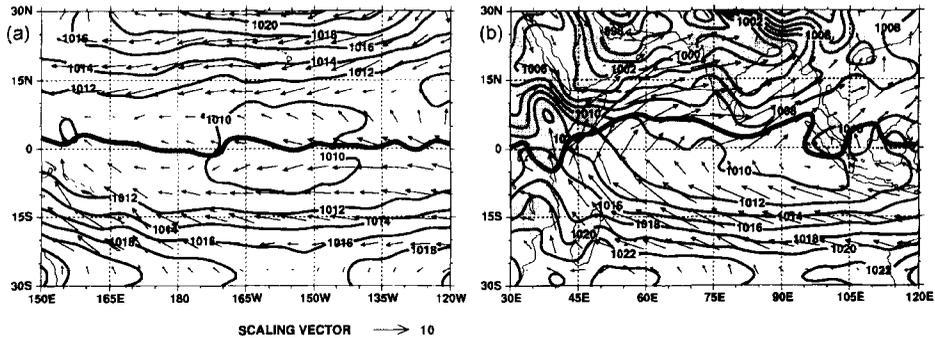


Figure 7. Distributions of sea-level pressure (SLP) (hPa) and vector total wind (m s^{-1}) commonly encountered about the equator, July 1992: (a) Pacific Ocean 150°E to 120°W ; (b) Indian Ocean 30°E to 120°E . Bold lines denote the zero absolute vorticity contours ($\eta = \zeta + f = 0$) at 925 hPa. Scaling vector for winds is shown.

(Holton 1992). Depending on the sign of the acceleration dv/dt , the parcel will experience a force that acts either to return it to its original position or to accelerate it in the direction of the displacement. In (1), v is the meridional velocity component, δy is the meridional displacement of the parcel, f is the Coriolis parameter and u_g is the zonal geostrophic velocity component. The critical factor is whether $f \leq \partial u_g / \partial y$ in (1). There are three cases:

(i) $f < \partial u_g / \partial y$. Locally anticyclonic AV. A parcel displaced polewards from its equilibrium position will be moving sub-geostrophically (slower than the basic state) and will thus experience a net force in the direction of the displacement. This case is inertially unstable.

(ii) $f = \partial u_g / \partial y$. Zero AV. A parcel displaced polewards from its equilibrium position will be moving geostrophically (at the same speed as the basic state) and will thus experience no net force. This case is inertially neutral.

(iii) $f > \partial u_g / \partial y$. Locally cyclonic AV. A parcel displaced polewards from its equilibrium position will be moving super-geostrophically (faster than the basic state) and will thus experience a net force opposite the direction of the displacement. This case is inertially stable.

The distribution of the geostrophic wind u_g is given by the background basic-state pressure gradient $\partial p / \partial y$. In near-equatorial regions, the pressure gradient zonally averaged around the globe is very small (i.e. $\partial p / \partial y \approx 0$), which corresponds to the case studied by Stevens (1983). However, regionally there are strong cross-equatorial pressure gradients (see Figs. 1, 2 and 3). Thus, a sufficiently large geostrophic shear may be produced that could lead to inertial instability. Indeed, the location of the $\eta = 0$ contour away from the equator indicates that this is true (Fig. 6).

Figure 7(a, b) shows the pressure distributions commonly encountered about the equator. The two panels show the wind and pressure fields in the northern Indian Ocean and the mid-Pacific Ocean during July. In Fig. 7(a) the pressure gradient across the equator over the Pacific Ocean is weak. Here, the subtropical high-pressure regions and the near-equatorial trough produce easterly geostrophic flow but with little or no cross-equatorial component. Without a strong cross-equatorial pressure gradient, the $\eta = 0$ contour resides very close to the equator (Fig. 4(a, b)). Figure 7(b), by contrast, shows a strong cross-equatorial gradient produced either by the surface fluxes associated with differential heating of the land and the ocean or by a strong cross-equatorial SST difference. Both the eastern Pacific and Indian Ocean possess similar configurations during the boreal summer. In all of these locations the

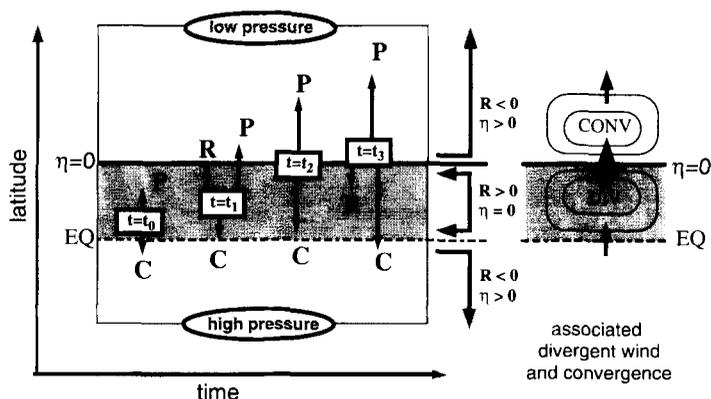


Figure 8. Schematic diagram showing the forces acting on a parcel of air under the influence of a cross-equatorial pressure gradient. Flow in the region is assumed to be inertially unstable. Axes representing time and latitude are drawn with orientation appropriate for northern hemisphere (boreal) summer. P, C and R represent the pressure gradient force, Coriolis force and residual force respectively. The region where the local absolute vorticity is negative ($\eta = \zeta + f < 0$) is stippled. On the right-hand side is the divergence-convergence distribution consistent with an inertially unstable flow.

$\eta = 0$ contour resides in the summer hemisphere well away from the equator. Thus, there are regions where the AV is locally anticyclonic and, therefore, locally inertially unstable.

The ideas discussed above are applied to those near-equatorial regions where the cross-equatorial pressure gradient is substantial and the local AV is anticyclonic, and defer discussion of those regions with weaker cross-equatorial gradients until later. Figure 8 shows the forces acting on a parcel as it passes through a region that is inertially unstable. Letters 'P' and 'C' represent the pressure gradient and Coriolis forces, respectively and 'R' represents the residual force. The axes on the diagram represent time (abscissa) and latitude (ordinate). The panel is drawn with orientation appropriate for northern summer conditions with low pressure to the north of the equator and high pressure to the south.

Consider a parcel at time $t = t_0$ that is located just north of the equator, in geostrophic balance, in a region where the AV is negative, and that it is then displaced northward at time $t = t_1$. The zonal velocity changes as a result of this displacement as dictated by (1) but, because the environment is inertially unstable, the new zonal velocity of the parcel is sub-geostrophic (case (i): $R > 0$). The unbalanced pressure gradient force R acts to accelerate the parcel northward ($R > 0$). At time $t = t_2$, the parcel reaches the zero AV contour ($\eta = 0$). At this stage, the acceleration is zero (i.e. case (ii): $R = 0$). Further northward motion (time $t = t_3$) results in the parcel experiencing a southward restoring force (case (iii): $R < 0$).

Thus, when the zero AV contour is found some distance from the equator, the wind will be accelerated on the equatorward side and decelerated on the poleward side of the $\eta = 0$ contour. If this interpretation is applicable to the real atmospheric system, the signal of these processes should be evident in the divergent wind field: a local maximum in the divergent wind field should be centred on the $\eta = 0$ contour; divergence should be found on the equatorial side of the $\eta = 0$ contour and convergence on the poleward side. The divergent pattern consistent with the existence of inertial instability is shown on the right-hand side of Fig. 8. Before examining the characteristics of the lower-tropospheric divergence fields, the limitations of the simple theoretical analysis are addressed and an interpretation of features in the divergent wind is provided.

So far, only parcels initially displaced polewards have been considered. In the absence of any processes that impart a directional bias to the system, however, equatorward

displacement is equally likely. Since the acceleration is proportional to the direction and the magnitude of the displacement (Eq. (1)), a parcel displaced equatorwards in the inertially unstable region experiences an equatorward acceleration. Clearly, the response of an inertially unstable system that does not contain any mechanisms that would introduce a directional bias would be very different from the one proposed here.

An example of the former system and the ensuing response has been observed by O'Sullivan and Hitchman (1992). In their modelling study of the mesosphere, they find horizontal convergence and divergence maxima stacked in an alternating fashion over the $PV = 0$ contour when it is located away from the equator. The horizontal scale of the divergence–convergence maxima is similar to the scale of the locally anticyclonic PV but the vertical scale is the smallest scale resolved by the model.

It is speculated that there are two mechanisms in the lower troposphere, absent elsewhere in the atmosphere, that induce a one-sided response. First, δy will be exclusively positive (during boreal summer) or negative (during boreal winter) because parcels entering the unstable region are moving towards the summer pole within the cross-equatorial flow under the influence of the macro-scale pressure field and surface friction. This idea is supported by the observations of Reverdin and Sommeria (1981) and modelling work by Rodwell and Hoskins (1995). These workers found that, upon crossing the equator in the Indian Ocean during boreal summer, parcels in the lower troposphere accelerated in the direction they possessed before crossing the equator.

Directional bias may also be accounted for by interaction between the motion and moisture fields, resulting in convection (latent heating) in the rising branch and radiative cooling in the subsiding branch (see section 6). These thermodynamic processes, which are absent in the mesosphere, would be expected to increase the vertical scale of the response.

4. LOW-LEVEL DIVERGENT WINDS

Figure 9 shows the 925 hPa vector divergent wind fields for February 1992 and July 1992. The contours represent the magnitude of the divergent wind and the $\eta = 0$ position is plotted as the bold contour. In all regions where the $\eta = 0$ contour is found some distance away from the equator, there are three consistent relationships between the wind field and the vorticity field:

(i) The divergent wind is always directed so that it advects absolute vorticity across the equator responding to the macro-scale cross-equatorial monsoonal pressure gradient force discussed earlier. Presumably, this advection is responsible for the locally anticyclonic absolute vorticity.

(ii) There is a maximum in the divergent wind speed that is bisected by the zero absolute vorticity maximum especially in regions where $\eta = 0$ away from the equator. This relationship is more complicated in the northern Indian Ocean region during July 1992 (Fig. 9(d)). This will be discussed later.

(iii) Three contributions to the generation of absolute vorticity exist in the tropics. The vorticity equation scaled for tropical motions is given by

$$\frac{\partial \eta}{\partial t} = -\mathbf{V} \cdot \nabla \eta - \eta \nabla \cdot \mathbf{V} - \alpha \zeta, \quad (2)$$

such that the local tendency of absolute vorticity is determined by the combination of absolute vorticity advection, vortex stretching and dissipation. The parameter α is a dissipative coefficient. Note that (2) differs from the form found by Holton (1992) in that the advection term involves the total wind field, not just the rotational component, and

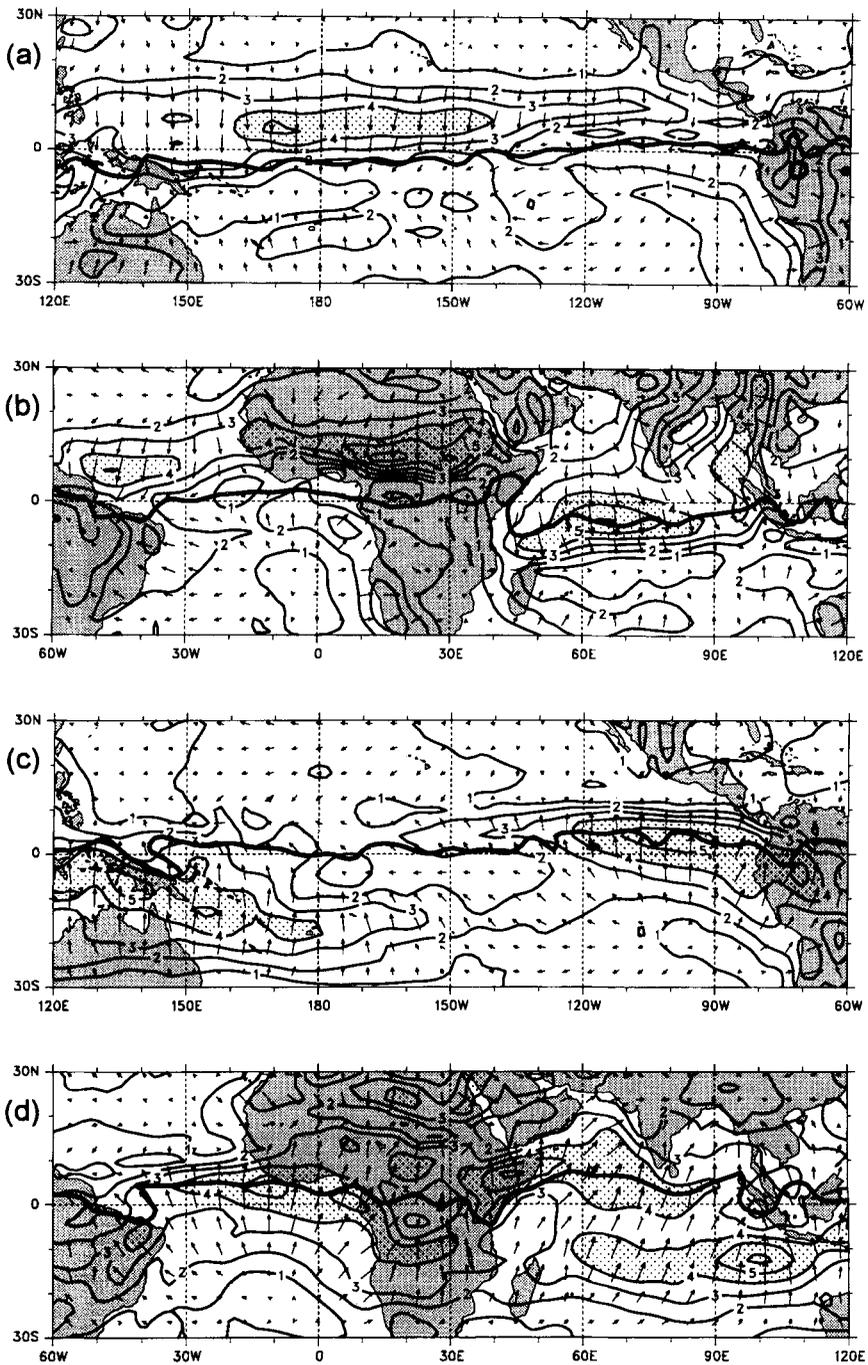


Figure 9. Distribution of mean vector divergent wind at 925 hPa (m s^{-1}): (a) and (c) Pacific Ocean, 120°E to 60°W; (b) and (d) Atlantic and Indian Oceans, 60°W to 120°E; (a) and (b) February 1992; (c) and (d) July 1992. Bold lines denote the zero absolute vorticity contours ($\eta = \zeta + f = 0$) at 925 hPa; areas of divergent winds greater than 4 m s^{-1} are stippled.

it includes dissipation. The use of the total wind field is supported by observations that the cross-equatorial advection of absolute vorticity is accomplished primarily by the divergent wind, except in the western Indian Ocean where the advection by the rotational component is roughly equivalent to that by the divergent component. Thus, from a regional perspective, neither the divergent part of the wind field nor the rotational can be ignored.

Considering Eq. (2), and starting with an AV field that is inertially stable (and thus locally cyclonic everywhere), it is clear why cross-equatorial advection of AV is the mechanism responsible for the observed regions of locally anticyclonic AV in Fig. 4. The first two terms on the right-hand side of (2) are capable of making negative contributions to the local tendency. However, in the absence of cross-equatorial vorticity advection, they can contribute only to cyclonic vorticity generation. Without cross-equatorial advection, the advection term acts only to redistribute cyclonic AV. The divergence term (or vortex stretching term) can make a locally anticyclonic contribution to the tendency if there is divergence. However, as the AV approaches zero, this mechanism becomes less and less efficient. That is, as $\eta \rightarrow 0$, $\eta \nabla \cdot \mathbf{V} \rightarrow 0$. The balances that exist near the $\eta = 0$ contour are discussed in sections 7(d) and 7(e).

It is argued that the general collocation of divergent wind maxima and the zero AV contour away from the equator is precisely the distribution expected to result from inertial instability (see Fig. 8). That is, the divergent wind is being accelerated on the equatorward side of the $\eta = 0$ contour and decelerated on the poleward side as hypothesized. In contrast, if the mechanism proposed by Walliser and Somerville (1994) were acting (see section 1), one would expect to observe the maximum divergent wind speeds on the equator, where the Coriolis force is zero.

In summary, a dual role for the divergent part of the wind field is claimed:

- (i) the divergent wind is responsible for creating the conditions that lead to near-equatorial inertial instability via interhemispheric advection of absolute vorticity; and,
- (ii) the divergent wind is accelerating locally in an effort to mitigate the inertially unstable conditions caused by the advection of absolute vorticity across the equator.

The dual role of the divergent wind distribution can be seen through an examination of the 925 hPa divergence fields for February and July 1992 shown in Fig. 10. In regions where the AV is locally anticyclonic, divergence is located on the equatorward side of the $\eta = 0$ contour whereas convergence is located on the poleward side. It is claimed that this particular configuration of divergence, resulting from the acceleration of the divergent wind across the region where the AV is locally anticyclonic, acts to reduce the local inertial instability. The relaxation of the instability is accomplished by making a locally cyclonic contribution to the AV tendency via the divergence generation term in the vorticity equation (2). Furthermore, the strong convection which is collocated with the maximum convergence is a direct consequence of the convergence of moist static energy in the boundary layer and the upward vertical velocity accompanying the divergence. That is, ITCZ off-equator convection is the result of the stabilizing secondary circulation resulting from the inertial instability.

Although this discussion concerning the role of the divergence–convergence doublet in mitigating the instability has merit because of its simplicity, this simplicity comes at the cost of ignoring the vertical structure and thermodynamics of the real atmosphere. To include these aspects, it is useful to expand the arguments presented thus far in several ways, including rephrasing the problem in terms of potential vorticity. When this is done, one must reconcile the idea of processes that modify the AV distribution with processes that modify the potential-vorticity distribution. For example, divergence and convergence *by themselves* do nothing to change the PV distribution: PV is modified only by diabatic

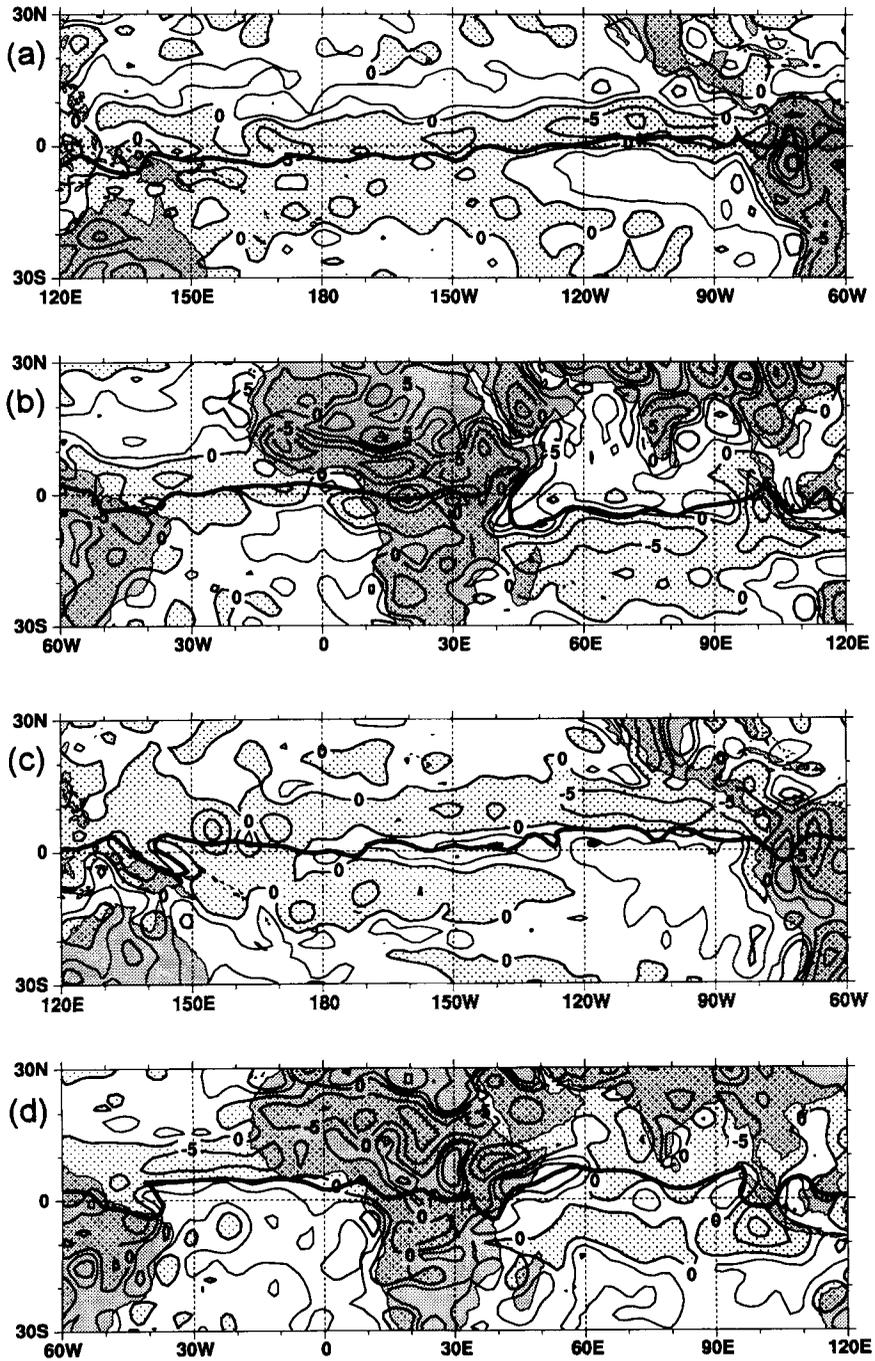


Figure 10. Distribution of monthly-mean divergence at 925 hPa (10^{-6} s^{-1}): (a) and (c) Pacific Ocean, 120°E to 60°W; (b) and (d) Atlantic and Indian Oceans, 60°W to 120°E; (a) and (b) February 1992; (c) and (d) July 1992. Bold lines denote the zero absolute vorticity contours ($\eta = \zeta + f = 0$) at 925 hPa; contour interval is $5 \times 10^{-6} \text{ s}^{-1}$.

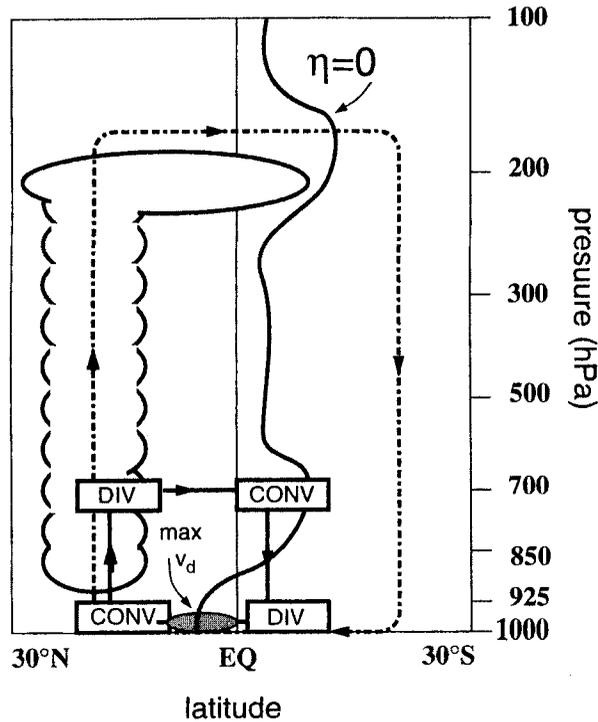


Figure 11. Schematic diagram showing the expected circulation in the latitude–height plane in regions where near-equatorial inertial instability is occurring. In a conditionally unstable atmosphere, a deeper convective circulation may be expected. The bold line denotes zero absolute vorticity ($\eta = \zeta + f = 0$), with negative absolute vorticity to the right (south) of the line.

processes such as friction and vertical gradients in heating. Thus one may be led to conclude that the arguments concerning the secondary circulations mitigating the instability of the basic state are spurious. In sections 5 and 6, however, it is shown that the divergence–convergence doublet is associated with vertical circulations in the latitude–height plane, which in turn are associated with vertical gradients in the heating that act to modify the PV distribution in a way that reduces the instability, thus reconciling the AV and PV viewpoints. In fact, consistency between the dynamics and thermodynamics requires that the divergence and convergence that modify the AV distribution be associated with vertical gradients in heating that modify the PV distribution in a similar way.

5. THE LATITUDE–HEIGHT CIRCULATIONS

One of the shortcomings of the parcel argument is that it fails to explain how environmental air is replaced in the space vacated by the parcel and how it disappears from the space where the parcel eventually comes to rest. The characteristics of the divergent wind fields shown in Figs. 9 and 10, along with mass-conservation arguments, imply that this is accomplished by the existence of closed circulations in the height–latitude plane. Figure 11 shows a schematic diagram generalized from features apparent from the plots of the divergent wind fields about the equator. This diagram is oriented to represent advection of negative AV into the northern hemisphere as occurs regionally in the July 1992 cases. The orientation of the circulation would be reversed for the boreal winter case. Two circulations dominate the longitudinal sections. There is a deep, large-scale divergent circulation that

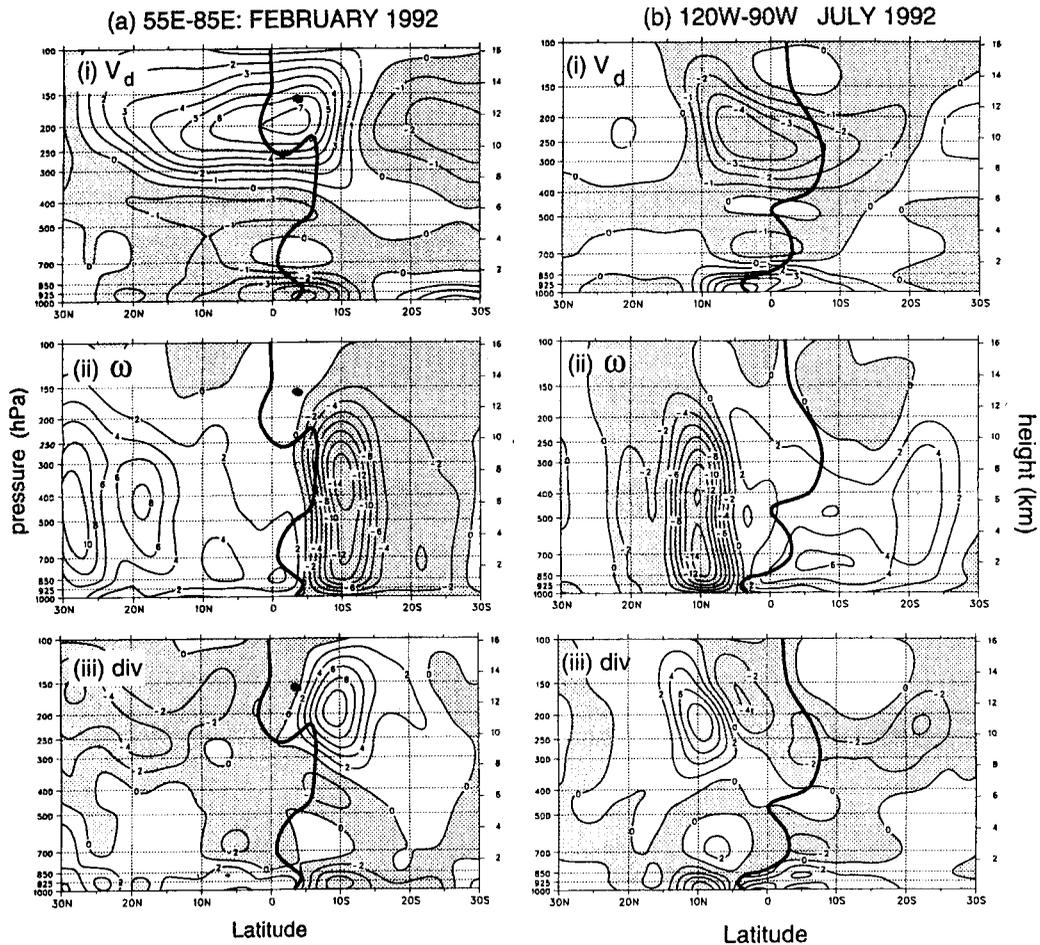


Figure 12. Latitude–height cross-sections of circulations: (i) northward divergent wind V_d (m s^{-1}); (ii) vertical velocity $\omega(dp/dt)$ ($10^{-2} \text{ Pa s}^{-1}$); (iii) divergence (10^{-6} s^{-1}); (a) Indian Ocean, 55°E to 85°E , February 1992; (b) eastern Pacific Ocean, 120°W to 90°W , July 1992; (c) central and eastern Atlantic Ocean, July 1992; (d) Indian Ocean, 55°E to 85°E , July 1992. The bold line denotes zero absolute vorticity ($\eta = \zeta + f = 0$), with negative absolute vorticity to the right (south) of the line. Stippling denotes negative values, including upward vertical velocities.

is responsible for advecting negative AV across the equator and into the northern hemisphere. The circulation extends over the entire depth of the troposphere and from 30°S to 20°N . The circulation is driven by the planetary-scale cross-equatorial pressure gradient associated with the distribution of land and ocean in the Indian Ocean sector and the SST distribution in the eastern Pacific. Superimposed on the northern branch of this circulation, there is divergent circulation closer to the equator that is associated with the instability and the location of the near-equatorial convection.

If the near-equatorial latitude–height circuit is a secondary circulation, it must conspire to reduce the condition that leads to the instability. Viewing the problem as one of an inertial instability, this is accomplished by making positive contributions to the local AV tendency where the AV is negative north of the equator. In other words, the locally anticyclonic AV decreases as a result of the near-equator circulation rendering the system

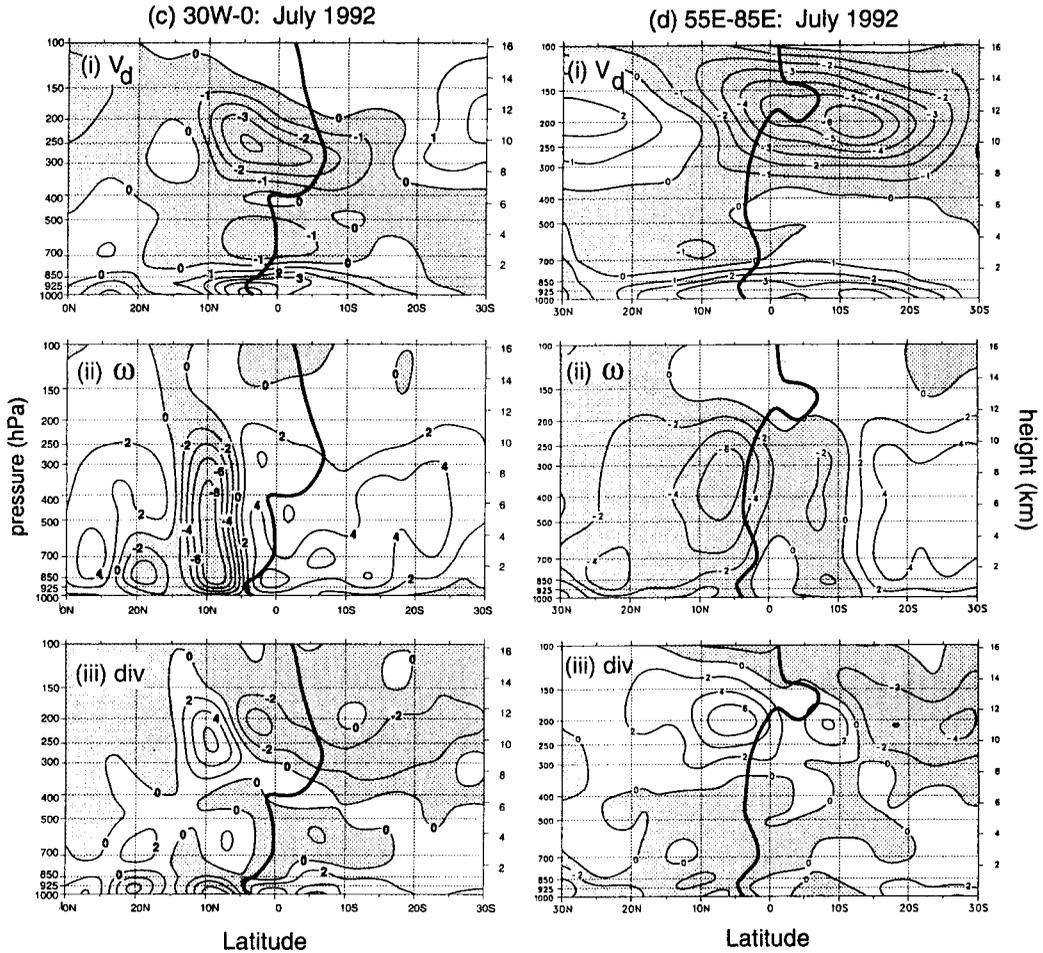


Figure 12. Continued.

towards inertial stability. On the other hand, the gross-scale circuit, resulting from the very-large-scale pressure gradient continues to advect locally anticyclonic AV across the equator moving the system towards inertial instability. The near-equatorial circulation, and the resulting convection, appear as the mean stabilizing features of the system as long as a gross-scale cross-equatorial pressure gradient persists. The inertial instability viewpoint is reconciled with that of symmetric instability in section 6.

If the atmosphere in Fig. 11 is statically stable, one might expect an impediment to both the upward and downward motion of the near-equatorial circulation that results from the inertial instability. Focusing on the upward branch first, if there is enough moisture at low levels, so that lifted parcels become saturated during forced ascent, it is possible that they may reach their level of free convection and continue to rise on their own. That is, in an atmosphere that is conditionally unstable, the circulations will be allowed by, and, in fact, be enhanced by, convection, attaining depths much greater than they would in a dry atmosphere. The latent-heat release from the convection counters, and greatly exceeds, the adiabatic cooling resulting from the upward motion in the rising branch

of the secondary circulation. Similarly, the subsiding branch of the circulation must be accompanied by cooling. If the downward motion transports drier, middle tropospheric air into the lower tropical troposphere, radiative effects can account for this cooling. The heating and cooling associated with the upward and downward branches of the secondary calculations is examined in section 6.

Figure 12 shows height–latitude cross-sections of the zonally averaged divergent flow and the divergence for the Indian Ocean sector (55°E to 85°E) during February 1992 and, for July 1992, the eastern Pacific Ocean sector (120°W to 90°W), the central to eastern Atlantic Ocean sector (30°W to 0°) and the Indian Ocean sector (55°E to 85°E). The top panels (i) show the northward divergent wind with negative (southern) winds stippled. The middle panels (ii) show the vertical velocity component (specifically the ω -field dp/dt) with negative (ascending) values stippled. The bottom panels (iii) show the divergence. The zero AV contour is represented by the bold contour in all panels. Figure 12 shows the characteristics depicted in the schematic Fig. 11 allowing, with the exception of the July Indian Ocean sections (Fig. 12(d)), a general description of the three sets of sections to suffice. The complexity of the northern Indian Ocean section requires a separate discussion.

The poleward branch of the divergent flow is constrained to the lower troposphere, generally below 850 hPa, with maxima occurring at the position of the zero AV contour. The low-level convergence/divergence doublet straddles the $\eta = 0$ contour. Extrema in the vertical motions concur with the centres of these fields. In the upper troposphere, return flows occur near 200 hPa and also near 700 hPa. Overall magnitudes of the divergent horizontal flow are generally between 4 m s^{-1} and 6 m s^{-1} .

The divergent flow in the upper troposphere is directed from the summer hemisphere to the winter hemisphere, the reverse of the flow in the lower troposphere. Commensurate with this change in direction of flow, there is a reversal in the advection of absolute vorticity and the $\eta = 0$ contour is found in the winter hemisphere in the upper troposphere. The maximum in the upper-level divergent wind is not collocated with the $\eta = 0$ contour as it was in the lower troposphere; this point will be discussed later. In combination, the vertical and horizontal flows define strong latitude–height circulations with the ascending branches tightly bounded just polewards of the low level $\eta = 0$ contour. A comparison of the latitudes of maximum ascent in Fig. 12 with the OLR distributions (Fig. 1) reveals that they match well with the regions of low OLR and thus deep convection.

The displacement of the $\eta = 0$ contour means that local anticyclonic vorticity exists in the upper troposphere of the winter hemisphere. Thus, by the same criteria, the upper troposphere must also be inertially unstable. Yet, the pattern of divergence in the upper troposphere does not seem the same as in the lower troposphere. Instead of a tight local convergence-maximum attending the $\eta = 0$ contour, the convergence is more spread out with weaker local extrema. The probable reason for the difference is that the upper troposphere is statically very stable so that vertical motions are inhibited. Also, there are no mechanisms whereby, if secondary circulations were to develop, they could produce the necessary thermodynamic response to counter this impediment, as occurs in the lower troposphere through latent-heat generation. Radiative cooling occurs in the stratosphere but this is slow, with a timescale of the order of days. The absence of heating processes with short timescales, and the stronger static stability, result in a much broader descending region which extends as far polewards as 30° of latitude.

It was noted that in Fig. 9 the July divergent wind fields over the northern Indian Ocean were somewhat different from those in the other regions. There was a general correspondence between the zero AV contour and extrema in the divergent wind but the divergent wind maximum is to the north of the $\eta = 0$ line and not quite bisected by it. Figure 12(d) contains features similar to the other vertical cross-sections (Fig. 12(a, b, c))

but with less clear detail. For example, a very broad lower-tropospheric divergent wind maximum spans the equator and the vertical velocity shows three regions of ascent. The first upward maximum exists just to the north of the $\eta = 0$ contour in a position consistent with inertial instability. A weak ascending region occurs to the south of the equator and a third over the south Asian land mass. In the mean, the entire section polewards of the equator appears to be ascending. A number of studies (e.g. Sikka and Gadgil 1980; Webster 1983; Srinivasan *et al.* 1993) have indicated that convection in the northern Indian Ocean is very transient, oscillating aperiodically between locations just north of the equator and the south Asian land mass. Whether the system showed propagation between the two limits or possesses multiple solutions has not been determined (Webster *et al.* 1997). However, one can only speculate that the more complex aspects of the divergent flow found in Fig. 12(d) are the result of the transience occurring on sub-monthly timescales. This region will be discussed further in section 7.

6. LATENT HEATING, RADIATIVE COOLING AND PV MITIGATION BY SECONDARY CIRCULATIONS

For deep convection to be generated, the lower troposphere must be conditionally unstable. To test for conditional instability, calculations were made of the equivalent potential temperature θ_e of the atmospheric column, and the equivalent potential temperature of a hypothetically saturated atmosphere, θ_e^* , which has the same thermal structure as the real atmosphere, following Ooyama (1969). Figure 13 shows the differences between θ_e and θ_e^* for a parcel displaced from 1000 hPa to 700 hPa in the mean for February and July 1992 atmospheres. Only positive (conditionally unstable) contours are plotted. It is seen that the lower atmosphere is conditionally unstable over much of the tropics, including all the regions of interest to this investigation. Thus, the secondary circulations can lead to convection in these regions.

It also appears that interaction between the motion and moisture fields occurs in association with the subsiding branch of the secondary circulations and that this interaction also has important thermodynamic consequences. In this case, the sinking motion brings dry air downwards, resulting in a humidity profile that is conducive to low- and mid-level radiative cooling. The cooling balances the adiabatic warming in the subsiding branch. To investigate this, cross-sectional diagrams of the mixing ratio, the material tendency of the potential temperature and the partial derivative with respect to pressure of the latter quantity (Fig. 14) were constructed for the eastern Pacific for July 1992. The discussion that follows also applies to the other regions discussed in the present paper, but, in the interests of brevity, these diagrams have not been included.

The mixing-ratio field (Fig. 14(a)) shows that, at any pressure level, the atmosphere is generally much drier to the south of the $\eta = 0$ contour than to the north. This is in agreement with the vertical motion field which shows subsidence to the south and rising motion to the north. The dry air to the south of the $\eta = 0$ contour permits efficient radiative cooling, as is seen in the mean material change in potential temperature shown in Fig. 13(b). Away from the lower boundary, there is cooling, with a maximum located near 700 hPa. Of particular interest, there is a small, local, minimum in the mixing ratio and maximum in the cooling located just equatorward of the $\eta = 0$ contour, between 850 hPa and 925 hPa (Fig. 14(a, b)). This is where the weak, low-level return circulations are located (Fig. 12). North of the $\eta = 0$ contour, the heating field is generally positive and is dominated by extremely strong latent heating that increases in magnitude abruptly between 1000 hPa and 700 Pa.

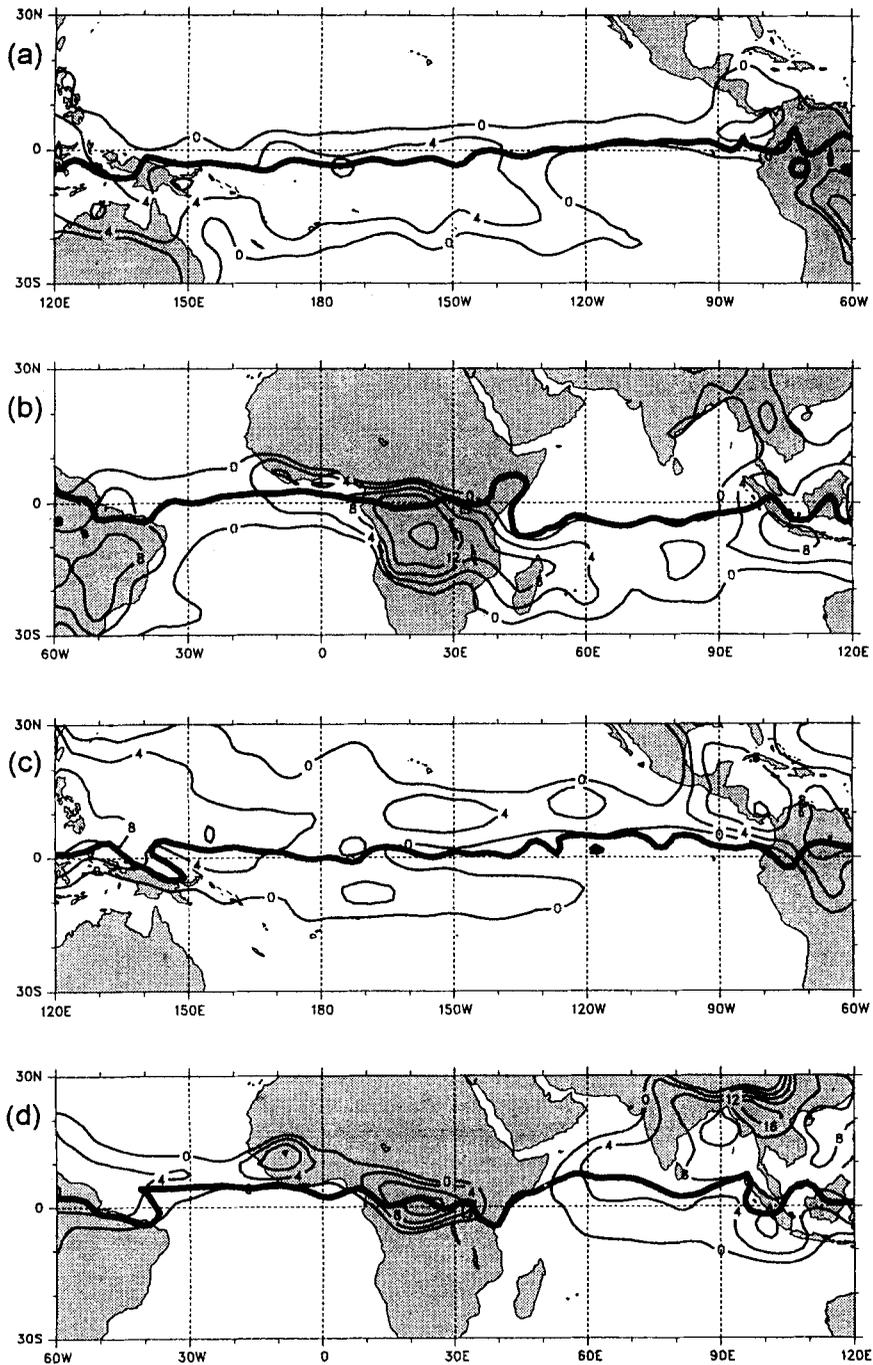


Figure 13. Distribution of the difference between the mean equivalent potential temperature of air at 1000 hPa from the equivalent potential temperature of hypothetically saturated air at 700 hPa ($\theta_e - \theta_e^*$): (a) Pacific Ocean 120°E to 60°W, February 1992; (b) Atlantic and Indian Oceans 60°W to 120°E, February 1992; (c) Pacific Ocean 120°E to 60°W, July 1992; (d) Atlantic and Indian Oceans 60°W to 120°E, July 1992. Bold lines denote the zero absolute vorticity contours ($\eta = \zeta + f = 0$) at 925 hPa.

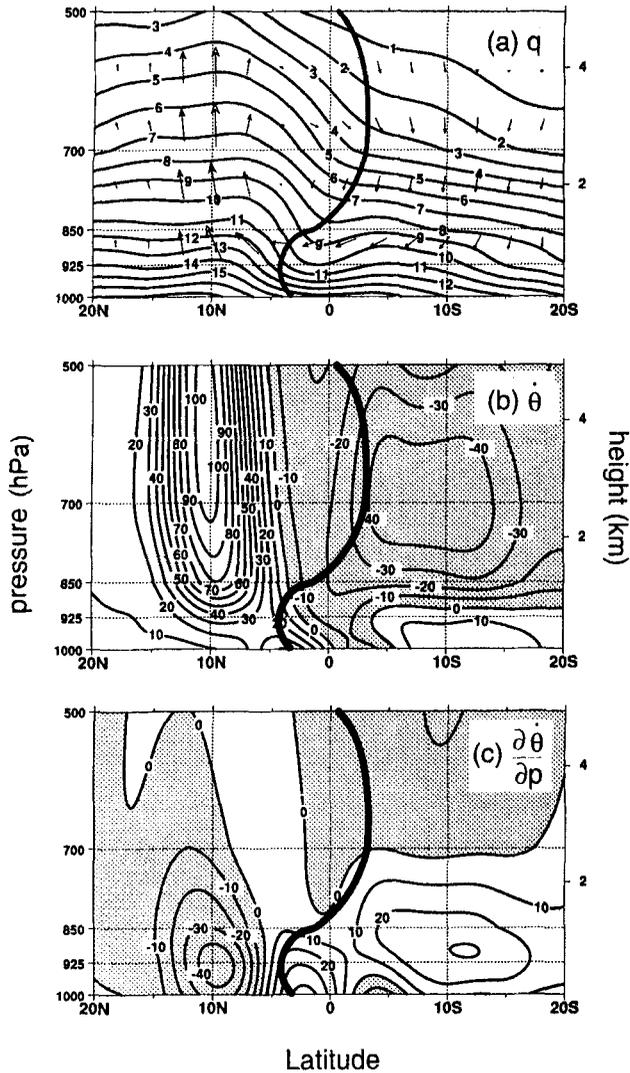


Figure 14. Latitude–height cross-sections of moisture and thermodynamic quantities over the eastern Pacific Ocean 120°W to 90°W, July 1992: (a) humidity mixing ratio q (g kg^{-1}) and wind vectors denoting the divergent part of the circulation (the vertical component of each vector is scaled by a factor of 50 relative to its horizontal component); (b) material tendency of the potential temperature $\dot{\theta}$ (K s^{-1}); (c) partial derivative of the material tendency of the potential temperature with respect to pressure $\partial\dot{\theta}/\partial p$ ($\text{K s}^{-1}\text{Pa}^{-1}$). The bold line denotes zero absolute vorticity ($\eta = \zeta + f = 0$), with negative absolute vorticity to the right (south) of the line. Stippling denotes negative values.

Near 1000 hPa, surface processes appear to affect the heating field strongly. Generally, warming is observed as air moves northward, up the gradient of potential temperature. An exception occurs near the equatorial cold tongue, where air initially cools and then warms rapidly as it leaves the cold-water region.

The vertical gradient of the heating (Fig. 14(c)) is of interest when considering the PV balance since its product with the local AV is proportional to the material source of

PV. For large-scale flows, the vertical components dominate in the PV equation so that

$$\frac{d(\text{PV})}{dt} = \underline{-g\eta \frac{\partial \dot{\theta}}{\partial p}} - g \frac{\partial \theta}{\partial p} \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right). \quad (3)$$

The first term on the right-hand side of (3) (underlined) is the diabatic source of PV and the second is the frictional dissipation. $\text{PV} = -g\eta \partial \theta / \partial p$ is the potential vorticity on an isobaric surface; g is the gravitational acceleration; $\partial \dot{\theta} / \partial p$ is the partial derivative of the material tendency of the potential temperature with respect to pressure; $\partial \theta / \partial p$ is the partial derivative of the potential temperature with respect to pressure and $(\partial F_y / \partial x - \partial F_x / \partial y)$ is the curl of the frictional force.

Below 700 hPa the partial derivative of heating with respect to pressure is positive to the south of the $\eta = 0$ contour and negative to the north (Fig. 14(c)). On the other hand, absolute vorticity is negative to the south of the $\eta = 0$ contour and positive to the north. Thus the diabatic source term, which is the product of these two quantities and a negative definite quantity (minus the gravitational acceleration) is positive everywhere and the diabatic source acts to oppose the low-level advection of potential vorticity.

The extrema in Fig. 14(c) merit some discussion. The large negative maximum centred near 10°N at 925 hPa is associated with latent heating increasing with height. To the south, there is a maximum lying between 925 hPa and 850 hPa, extending from the equator to southward of 20°S . There is another maximum that is stronger, although smaller, centred near 2°N , at 1000 hPa. The large magnitude of the latter results from warming near the surface as air moves off the equatorial cold tongue, in conjunction with cooling aloft. Recall that the weak low-level return branch of the secondary circulations is located here (Fig. 12).

Thus, from the PV perspective, the secondary circulations act to mitigate the symmetrically unstable large-scale flow by inducing the necessary heating distributions through interaction with the moisture field. This is analogous to the AV perspective, where the secondary circulations mitigate the inertially unstable large-scale flow by producing the necessary divergence and convergence distributions. This agreement is not surprising since, in the tropics, heating tends to be balanced by vertical motion. Thus, through continuity, heating increasing with height is associated with convergence and cooling increasing with height with divergence (cf. Figs. 14(c) and 12(b)(iii)).

Since it is reasonable to expect that frictional dissipation usually opposes the local relative vorticity, then the dissipation term also acts to oppose the cross-equatorial advection of PV. This can be seen by noting that, for linear dissipation, the distribution of the dissipation mirrors the distribution of relative vorticity, i.e. $-(\partial F_y / \partial x - \partial F_x / \partial y) = -\alpha \zeta$. Maps of the relative vorticity (not shown), reveal locally anticyclonic relative vorticity where the $\eta = 0$ contour is located off the equator. Furthermore, these maxima are bisected by the $\eta = 0$ contour. Thus, whereas diabatic processes are important in opposing advection to the north and south of the $\eta = 0$ contour, frictional dissipation is important in opposing advection in its immediate vicinity. This subject is considered further in subsections 7(d) and 7(e).

7. SUMMARY AND CONCLUSIONS

An attempt has been made to define the physical processes that determine where mean convection will be located about the equator. To do this, a very simple theory has

been developed and the signatures predicted by the theory have been compared with observations. The following conclusions may be drawn.

(a) *Location of near-equatorial convection*

Depending on the macro-scale pressure-gradient field, two major classes of large-scale convective organization can be defined. From Figs. 1 and 2, either the zone of maximum convection is relatively close to the equator (when the macro-scale cross-equatorial pressure gradient is small) or substantially removed from the equator (when the macro-scale cross-equatorial pressure gradient is large).

(i) *Regions of weaker cross-equatorial pressure gradient.* In regions where the pressure gradient is small (e.g. the western and central Pacific Ocean during summer or winter and the western Atlantic Ocean during winter) the AV zero contour appears very close to the equator. As a result, the circulation on both sides of the equator is inertially stable. Figures 1 and 2 show that in these regions the convection appears in closer association with the maximum SST distribution and the distribution of conditional instability rather than with any specific dynamical process (Figs. 2, 13). The convection in these regions is quite intense ($OLR < 200 \text{ W m}^{-2}$) and the sea surface is quite warm ($SST > 29 \text{ }^\circ\text{C}$).

(ii) *Regions of strong cross-equatorial pressure gradient.* In regions where there is a substantial lower-tropospheric cross-equatorial pressure gradient, such that there is a divergent wind flow from the winter to the summer hemisphere, the zero AV contour is found between 5° and 10° from the equator in the summer hemisphere. The displacement of the contour results from the advection of AV from one hemisphere to the other, producing a region of anticyclonic (or 'wrong signed') AV. Major displacements of the $\eta = 0$ contour occur in the eastern Atlantic Ocean, the eastern Pacific Ocean and the Indian Ocean in association with the local summer and winter monsoon circulations. In these regions, the maximum zone of convection is not necessarily associated with the warmest sea-surface nor with the lowest surface pressure. Indeed, maximum convection usually occurs on the equatorward side of the latter and either equatorward or poleward of the former. SSTs are approximately 2 deg C lower than those observed in the near-equatorial regions (discussed above), yet the convection is comparable in intensity ($OLR < 200 \text{ W m}^{-2}$).

Simple theoretical considerations indicate that the local anticyclonic circulation associated with the displacement of the $\eta = 0$ contour is inertially unstable. The parcel method was used to predict the character of the circulations arising from the instability. Consistently, the predicted signal of the unstable circulations is found in monthly-mean fields of the AV, divergent wind, divergence and vertical motion. The intense near-equatorial convection arises from a secondary circulation that acts to stabilize the region by the local creation of cyclonic vorticity. In addition, there is a thermodynamic constraint that must be overcome. At first glance, it appears that the static stability of the atmosphere with respect to dry adiabatic ascent and descent should impede or preclude the development of convection. However, it was shown that the near-equatorial regions of ascent are conditionally unstable so that low-level convergence readily results in deep convection and heating. Additionally, the convergence appears to promote strong convection through the convergence of moist static energy in the boundary layer. Where there is sinking motion, the circulations produce a local dry anomaly which is conducive to anomalous radiative cooling. That is, through interactions between the circulations and the moisture field, the circulations induce ameliorating heating and cooling responses. Furthermore, if the problem is viewed as one of symmetric instability in which the dynamics and thermodynamics are combined into a single quantity, then the appropriate instability criterion is that the PV

is locally anticyclonic. The mitigating effect of the circulations is then realized as a result of vertical gradients in the heating (cooling) that produce material tendencies of PV on both sides of the $\eta = 0$ contour that act to relax its location back to the equator (Fig. 14).

(b) *Criteria for the existence of off-equatorial organized convection*

The analysis allows the listing of conditions and processes necessary for the formation of off-equator convection and for the determination of the location and strength of the convection. The necessary conditions are the following.

- (i) *A cross-equatorial pressure gradient.* Organized zonally-oriented convection occurs off the equator when there is a north–south pressure gradient. The strongest gradients occur over the Indian Ocean in both the boreal summer and the boreal winter, and over the Atlantic and the eastern Pacific Ocean during the boreal summer.
- (ii) *A warm sea surface.* The sea surface must be sufficiently warm for the atmosphere on the summer side of the equator to be conditionally unstable.

Clear evidence that the second criterion alone does not necessarily produce deep convection has been noted (see Fig. 3). This failure is especially evident in the northern Indian Ocean during the period from March to June when the warmest ocean temperatures on the planet are not matched by deep convection.

(c) *Sequence of processes creating off-equatorial organized convection*

The simple instability theory and the corroborating evidence from observations suggest a sequence of processes following the establishment of the necessary conditions listed above. The sequence is as follows.

- (i) *A cross-equatorial advection of absolute vorticity.* The macro-scale cross-equatorial pressure gradient must be sufficiently strong to produce a divergent wind field that advects absolute vorticity across the equator. The divergent wind field is strong enough in the summer hemisphere over all three oceans, but, during the boreal winter, only over the Indian Ocean and the Indonesian–Australian sector.
- (ii) *An inertially unstable state.* The advection of absolute vorticity across the equator renders air in the near-equatorial summer hemisphere to be inertially unstable.
- (iii) *An ameliorating secondary circulation.* In order to relax the inertial instability, a divergence–convergence doublet occurs in the vicinity of the zero AV contour, producing local cyclonic vorticity to counter the advection of locally anticyclonic vorticity. The local convergence (divergence) requires upward (downward) vertical motion for mass continuity. The convergence within the boundary layer acts to increase the intensity of convection in these regions.
- (iv) *A thermodynamical conditional instability.* If the atmosphere is conditionally (statically) unstable, the vertical velocity associated with the convergence will allow the development of deep convection. Conditional instability depends on the magnitude of the SST and the structure of the atmosphere above.
- (v) *Subsequent feedbacks.* The ensuing convection feeds back positively into (i) by increasing the cross-equatorial pressure gradient and cross-equatorial advection, and feeds back negatively into (iii) by increasing low-level convergence and thus ameliorating the circulation. The vertical heating distribution acts to reduce the instability by the generation of oppositely signed PV.

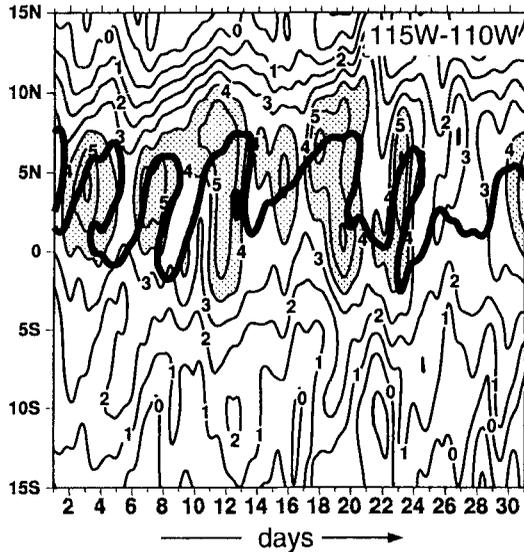


Figure 15. Latitude–time depiction of the northward component of the divergent wind V_d (m s^{-1}) at 925 hPa in the longitude band 115°W to 110°W (eastern Pacific Ocean) for August 1992. The bold line denotes zero absolute vorticity ($\eta = \zeta + f = 0$), with negative absolute vorticity below (south of) the line. Stippling denotes values greater than 4 m s^{-1} .

Further support for the theory comes from matching the theoretical time- and space-scales with observations. To date only monthly-mean distributions have been considered. Figure 15 shows a latitude–time section of the divergent part of the wind field for August 1992, averaged between 115°W and 110°W . The bold line denotes the $\eta = 0$ contour which, during the month, oscillates between the equator and 8°N . As in the monthly distributions, the contour also bisects the divergent wind maxima even on sub-monthly timescales. Clear regions of divergence and convergence can be inferred from the divergent wind distribution on the equatorward and poleward sides of the contour respectively.

Rather clear temporal and spatial scales can be inferred from Fig. 15. There appear to be seven periods of maximum northward divergent wind straddling the $\eta = 0$ contour, suggesting a period of about 4.5 days. The inertial period goes as f^{-1} which for 6° of latitude is also about 4.5 days. Emanuel (1979) suggested that the space scale of inertial instability goes as l_z/s_θ where l_z is the vertical scale of the unstable domain and s_θ is the slope of the isentropic surfaces. From the cross-sections (Fig. 12), one can estimate that $l_z \approx 2 \text{ km}$ and $s_\theta \approx 1.5 \text{ km}/2000 \text{ km}$ which produces a scale of about 2500 km. From Fig. 12 or from Fig. 15, the scale appears to be about 20° of latitude or 2200 km.

In the Indian Ocean during July (the region of strongest cross-equatorial pressure gradient), the divergent maxima appear more complicated than in other regions. It was speculated that this region is one in which sub-monthly variability is very strong as the monsoon goes through active and break periods. At these times, the major locus of convection appears to move from the 5° – 10°N location, predicted by the instability argument, to the monsoon trough region. However, periodically the convection returns to the location determined by the inertial instability.

(d) *Location of the zero AV contour*

It is clear that the reason the $\eta = 0$ contour lies in the summer hemisphere is the advection of AV by the flow associated with the cross-equatorial pressure gradient. However,

even where the surface pressure gradient is very strong (e.g. over the Indian Ocean), the $\eta = 0$ contour is located within $5\text{--}10^\circ$ of the equator. Two basic questions arise. Firstly, what limits the poleward advection of the zero AV contour? For example, why does the $\eta = 0$ contour not lie at 30°N in the same regions? Secondly, why is the $\eta = 0$ contour lying off the equator a feature of the monthly-mean circulation even though the timescale of inertial instability goes as f^{-1} ? Clearly, the 5–6 day timescale is much shorter than the monthly average.

The answer to the first question comes from noting that there are two competing processes that provide for the location of the contour. Consider the steady-state version of (2). Such a situation will arise if the advection of AV is balanced by the generation of AV by divergence and dissipation. Schneider (1987) noted that, for the shallow-water system, dissipation is required if the flow is both non-divergent and the absolute vorticity is non-zero. That is

$$-\mathbf{V} \cdot \nabla \eta = \eta \nabla \cdot \mathbf{V} + \alpha \zeta. \quad (4)$$

The velocity \mathbf{V} is determined by the pressure gradient force. Thus, the stronger the pressure gradient force the greater the advection of wrong-signed AV across the equator. However, the divergence produced by the secondary circulation and the association convection in the conditionally unstable atmosphere, are magnified in importance as latitude increases. At the same time, as latitude increases, the advection term becomes dominated by the advection of the planetary vorticity. That is, as latitude increases,

$$\begin{aligned} -\mathbf{V} \cdot \nabla \eta &\rightarrow -\beta v \\ -\eta \nabla \cdot \mathbf{V} &\rightarrow -f \nabla \cdot \mathbf{V}. \end{aligned} \quad (5)$$

Thus, the limits of the unstable regime (and thus near-equatorial convection) are set at the point where the cyclonic generation of AV by convergence matches the advection of anticyclonic AV, less frictional dissipation. The advection has latitudinal limitations because the impact of the cyclonic-vorticity generation is increasing as $\sin \varphi$ whereas the advection of vorticity is decreasing as $\cos \varphi$. In other words, as the advection decreases polewards, the efficiency of vortex-tube stretching increases. Thus, irrespective of the strength of the cross-equatorial pressure gradient there will be a limit to the extent of the unstable regime and thus, the convection will not occur too far from the equator.

With regard to the second question, there appears to be a simple explanation as to why the zero AV contour remains in the summer hemisphere on timescales that are very much longer than the timescale of the instability relaxation itself. The system is rendered continually inertially unstable because the large-scale cross-equatorial pressure gradient exists on very long timescales compared to the instability timescale. Other examples of this type of relationship exist in the atmosphere. The timescale of baroclinic instability in the extratropics is a few days. However, it exists as a fairly continual process because the ultimate driving mechanism, latitudinal differential radiational heating, persists on a much longer timescale. Similarly, Bernard convection, the secondary circulation systems relaxing convective instability, exists as long as the unstable vertical heating gradient is maintained. Finally, the near-surface daytime lapse rate may continually be superadiabatic because the timescale of the solar heating of the surface is much longer than that of the convective mixing. In this sense, near-equatorial inertial instability, as it is forced by the long-lived cross-equatorial pressure gradient force, is another example of a forced instability system.

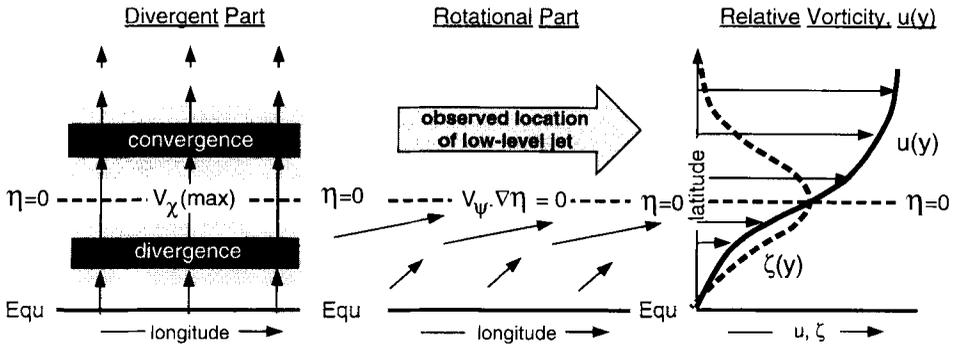


Figure 16. Schematic diagram showing the balances in the vicinity of the zero absolute vorticity contour ($\eta = \zeta + f = 0$).

(e) *The low-level westerly monsoon wind maximum*

In regions of strong cross-equatorial pressure gradient, a low-level westerly wind maximum in the rotational part of the flow is often observed situated off the equator and in the vicinity of the convection. Specifically, this low-level wind maximum is located just poleward of the $\eta = 0$ contour (Fig. 7). This location turns out to be consistent with the AV balance that occurs in the vicinity of the $\eta = 0$ contour.

Figure 10 shows a doublet of divergence and convergence about the $\eta = 0$ contour. Thus, in (4), $\eta \nabla \cdot \mathbf{V} \rightarrow 0$ in the vicinity of $\eta = 0$ since $\eta = 0$ and $\nabla \cdot \mathbf{V} \rightarrow 0$, leaving the following terms to balance

$$-\mathbf{V}_\chi \cdot \nabla \eta - \mathbf{V}_\psi \cdot \nabla \eta = \alpha \zeta|_{\eta=0}, \tag{6}$$

where we have used $\mathbf{V} = \mathbf{V}_\chi + \mathbf{V}_\psi$, the sum of the divergent and rotational parts of the wind vector, respectively.

At $\eta = 0$, $\nabla \eta$ is finite and \mathbf{V}_χ is at a local maximum (see Fig. 9), but \mathbf{V}_ψ turns to become orthogonal to $\nabla \eta$ so that $\mathbf{V}_\psi \cdot \nabla \eta \rightarrow 0$. Thus, only the divergent part of the wind is responsible for the advection of AV at $\eta = 0$ and this advection must be balanced by dissipation. That is,

$$-\mathbf{V}_\chi \cdot \nabla \eta = \alpha \zeta|_{\eta=0}. \tag{7}$$

As \mathbf{V}_χ is a local maximum at $\eta = 0$, ζ must also be a local maximum to satisfy the balance. In a zonally axisymmetric context, this relative vorticity maximum must result from a maximum in zonal shear. That is, as η tends to zero, $|\partial u / \partial y|$ tends to a maximum. The sign of the shear must be consistent with the sign of the advection, which changes between the northern and southern hemisphere cases.

Consider the northern hemisphere case. At $\eta = 0$, $-\mathbf{V}_\chi \cdot \nabla \eta < 0$, so that $\zeta|_{\max} < 0$ and $\partial u / \partial y|_{\max} > 0$ for balance. Typically, easterlies are observed in the tropics. However, in regions of strong cross-equatorial pressure gradient, the shear in the zonal wind necessary to balance the advection of AV is sufficient to turn the wind to westerlies. Furthermore, westerly winds must rapidly increase in strength to the north of the $\eta = 0$ contour. Figure 16 shows a schematic of the balances in the vicinity of the $\eta = 0$ contour. In the southern hemisphere, similar arguments produce a low-level westerly maximum to the south of the $\eta = 0$ contour.

The momentum balance obtained by integrating (7) with respect to y , assuming $\mathbf{V}_\chi = v_\chi = \text{constant}$ and $\zeta = -\partial u_\psi / \partial y$, is

$$v_\chi \eta = \alpha u_\psi.$$

For the case of a northwards directed v_x maximum, bisected by the $\eta = 0$ contour, this balance requires easterlies to the south of the $\eta = 0$ contour, zero zonal velocity at $\eta = 0$ and westerlies to the north. The westerlies observed at, and to the south of, $\eta = 0$ in the Indian Ocean region during boreal summer are an indication that the west to east pressure gradient is an important term in the momentum balance. This raises the question of what the impact is of omitting it in the vorticity balance and in the discussion above.

Figure 7(b) helps to answer the question. It shows that the zonal pressure gradient decreases with latitude between the equator and the $\eta = 0$ contour. Such a distribution results in less shear being needed to balance vorticity advection by the divergent wind. Despite this moderating influence of the zonal pressure gradient, strong shear is observed in association with the low-level monsoon jet, supporting the idea that the discussion above, neglecting the zonal pressure gradient term, is qualitatively correct.

(f) *Double ITCZs simulated in coupled models*

There has been considerable discussion in the modelling community as to why some coupled models simulate the ITCZ over the Pacific Ocean only poorly. Observations show that, over the eastern Pacific, the ITCZ is never located south of the equator, yet some models simulate a double ITCZ in the eastern Pacific or an ITCZ that migrates seasonally about the equator. One feature common to all the models that produce unrealistic simulations of this convection is the widespread and systematic overestimation of the SST over a broad region west of Peru and in a narrow band at 10°S (Mechoso *et al.* 1995). In the models that simulate a double ITCZ, the warm band near 10°S persists as a strong feature throughout the year. This error in the SST simulation might be expected to result in a reduction or elimination of the cross-equatorial sea-level pressure gradient. Consequently, the region would be inertially stable rather than inertially unstable. For thermodynamic reasons, a double ITCZ would form in the regions having maxima in the SST; over the maximum north of the equator and over the spurious maximum to the south of the equator. In the models that simulate a seasonally migrating ITCZ, the warm band near 10°S is stronger during boreal winter than it is during boreal summer. The seasonally migrating ITCZ can be accounted for as resulting from a seasonally reversing cross-equatorial pressure gradient produced by the oscillation of the warmest sea surface about the equator. Recall that observations presented here indicate that a seasonal oscillation in the surface temperature and pressure gradients results in the seasonal migration of the ITCZ about the equator in the Indian Ocean.

(g) *Extensions and future work*

An attempt has been made to provide consistent explanations for the position of monthly-mean locations of the near-equatorial distribution of convection. Considerable work remains to extend the analyses to shorter timescales, especially in the eastern Pacific during boreal summer and associated with the Asian–Australasian monsoon system in both summer and winter.

In the eastern Pacific, there is evidence that westward-moving disturbances are produced in association with the development of strong secondary circulations. These disturbances have a typical length of about 5000 km and a period of about 5 days, which is close to the inertial time scale 5° from the equator. Indeed, the formation of westward-moving disturbances with these length and temporal scales, and in association with cross-equatorial flow, has been documented by Fujita *et al.* (1969) using data derived from satellite photographs. It appears that the disturbances Fujita *et al.* (1969) observed were of the type seen here and that inertial instability could possibly explain their observations.

Together with suggestions from other studies that indicate that the monsoon has very strong variance on sub-monthly timescales (e.g. Sikka and Gadgil 1980), it is believed that the more complicated relationships in the northern Indian Ocean during boreal summer, between monthly-mean convection, divergence and the location of the inertially unstable regions, is due to vigorous sub-monthly variability. Within this hypothesis, monsoon convection oscillates between two locations. The first is near-equatorial convection determined by the processes discussed in the present paper. The second state is convection occurring over southern Asia (e.g. Webster 1983; Srinivasan *et al.* 1993). Webster *et al.* (1997) described two potential solutions as placing the monsoon in 'dynamic tension'. An investigation into this higher-frequency regime, with the anticipation of learning more about the transient structure of the monsoon itself, is intended.

It appears that there may be other tropical phenomena that may be associated with inertial instability. In the current study, focus has been given to the shear in the zonal wind component. However, the relative vorticity may also change if there is zonal shear in the meridional component of the basic flow and if that shear is negative (i.e. $\partial v/\partial x < 0$). Here, planetary-scale flows have been considered, where $\partial v/\partial x$ is very small. However, there are locations where the magnitude of $\partial v/\partial x$ may be relatively large so that $|\partial v/\partial x| > f$ close to the equator. In this circumstance, the local anticyclonic AV would be anticyclonic. The first location is in the atmosphere in the western Indian Ocean where a very sharp north-south Somali jet stream crosses the equator in both summer and winter. The second location is in the ocean Somali current, below the Somali jet, which also changes direction between the seasons (Knox 1985). Furthermore, it is associated with recurring gyres (referred to as the Great Swirl) that to date have defied explanation.

ACKNOWLEDGEMENTS

Support for this work was provided by National Science Foundation Grant ATM-9526030 through the Climate Dynamics Division. R. Tomas was supported by the Advanced Study program at the US National Center for Atmospheric Research (NCAR). Computations were carried out at the computer facilities of the National Center for Atmospheric Sciences and the Program in Atmospheric and Oceanic Sciences at the University of Colorado. We appreciate the insight provided through interesting discussions with Professor J. R. Holton (particularly for bringing the implications of the momentum balance to our attention), and Dr. Chidong Zhang of the University of Washington and Dr. H. R. Chang of the Program in Atmospheric and Oceanic Sciences of the University of Colorado during the completion of this work. Finally, we thank Christian Guillemot of the Climate Analysis Section (CAS) at NCAR for providing the software (and technical assistance) used to produce many of the diagrams.

APPENDIX

The primary source of dynamical data used in this study is the European Centre for Medium-Range Weather Forecasts (ECMWF) Tropical Ocean Global Atmosphere-World Climate Research Program (TOGA-WCRP) archive at NCAR. A reduced version of this data set was utilized which had been truncated to T42 resolution (Trenberth and Solomon 1993) and made available by the CAS at NCAR. All analyses were available at 14 or 15 levels (1000, 925, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 10 hPa) with the 925 hPa level becoming available after 1992. The investigation used 1992 data as, at that time, additional-level lower-tropospheric data were available. Monthly-mean values of the National Oceanic and Atmospheric Administration outgoing longwave radiation (Gruber

and Krueger 1984) and monthly-mean values of Reynolds SST (Reynolds and Marsico 1993), both also made available by the CAS at NCAR, were also used.

REFERENCES

- Bennetts, D. A. and Hoskins, B. J. 1979 Conditional symmetric instability—a possible explanation for frontal rain bands. *Q. J. R. Meteorol. Soc.*, **105**, 945–962
- Browning, K. A., Hardman, M. E., Harrold, T. W. and Pardoe, C. W. 1973 The structure of rainbands within a mid-latitude depression. *Q. J. R. Meteorol. Soc.*, **99**, 215–231
- Charney, J. G. 1971 ‘Tropical cyclogenesis and the formation of the Intertropical Convergence Zone’. Pp. 355–368 in Mathematical Problems of Geophysical Fluid Dynamics, W. H. Reid, Ed., *Lectures in Applied Mathematics*, **13**, Am. Math. Soc.
- Emanuel, K. A. 1979 Inertial instability and mesoscale convective systems. Part I: Linear theory of inertial instability. *J. Atmos. Sci.*, **36**, 2425–2449
- 1988 Observational evidence of slantwise convective adjustment. *Mon. Weather Rev.*, **116**, 1805–1816
- Fujita, T. T., Watanabe, K. and Izawa, T. 1969 ‘Formation and structure of equatorial anticyclones caused by large-scale cross equatorial flows determined by ATS-1 photographs’. Satellite and Mesometeorology Research Project Research Paper 78
- Gruber, A. and Krueger, A. F. 1984 The status of the NOAA outgoing longwave radiation data set. *Bull. Am. Meteorol. Soc.*, **65**, 958–962
- Hack, J. J. and Schubert, W. H. 1990 Some dynamical properties of idealized thermally-forced meridional circulations in the tropics. *Meteorol. Atmos. Phys.*, **44**, 101–117
- Hack, J. J., Schubert, W. H., Stevens, D. E. and Kuo, H.-C. 1989 Response of the Hadley Circulation to convective forcing in the ITCZ. *J. Atmos. Sci.*, **46**, 2957–2973
- Hastenrath, S. 1988 *Climate and circulation in the tropics*, Atmospheric Sciences Library, Reidel
- Hastenrath, S. and Lamb, P. J. 1977a *Climatic atlas of the tropical Atlantic and Eastern Pacific Oceans*, University of Wisconsin Press
- 1977b Some aspects of the circulation and climate over the Eastern Atlantic. *Mon. Weather Rev.*, **105**, 1019–1023
- Holton, J. R. 1992 *An Introduction to Dynamic Meteorology, 3rd Edition*, Academic Press
- Holton, J. R., Wallace, J. M. and Young, J. A. 1971 On boundary layer dynamics and the ITCZ. *J. Atmos. Sci.*, **28**, 275–280
- Hoskins, B. J. 1974 The role of potential vorticity in symmetric stability and instability. *Q. J. R. Meteorol. Soc.*, **100**, 480–482
- Hou, A. Y. and Lindzen, R. S. 1992 The influence of concentrated heating on the Hadley Circulation. *J. Atmos. Sci.*, **49**, 1233–1241
- Knox, R. 1985 ‘The Indian Ocean: Interaction with the monsoon’. In *Monsoons*. Ed. J. S. Fein and P. L. Stephens
- Krishnamurti, T. N. and Wong, V. 1979 A simulation of cross-equatorial flow over the Arabian Sea. *J. Atmos. Sci.*, **36**, 1895–1907
- Krishnamurti, T. N., Wong, V., Pan, H.-L., Pasch, R., Molinari, J. and Ardanuy, P. 1983 A three-dimensional planetary boundary layer model for the Somali jet. *J. Atmos. Sci.*, **40**, 894–908
- Lindzen, R. S. 1974 Wave-CISK in the tropics. *J. Atmos. Sci.*, **31**, 156–179
- Lindzen, R. S. and Hou, A. Y. 1988 Hadley circulations for zonally averaged heating centered off the equator. *J. Atmos. Sci.*, **45**, 2416–2427
- Lindzen, R. S. and Nigam, S. 1987 On the role of the sea surface temperature gradients in forcing low level winds and convergence in the tropics. *J. Atmos. Sci.*, **44**, 2418–2436
- Loschnigg, J. P. and Webster, P. J. 1995 ‘SST Regulation in the Indian Ocean: A Contrast to the Pacific Warm Pool’. In proceedings of the 21st Conference on Hurricanes and Tropical Meteorology, April 24–28 1995, Miami, Florida, USA
- Manabe, S. 1969 Climate and the ocean circulation, Part 2: The atmospheric circulation and the effects of heat transfer. *Mon. Weather Rev.*, **97**, 775–805
- Manabe, S., Hahn, D. G. and Holloway, J. L. Jr. 1974 The seasonal variation of the tropical circulation as simulated by a global model of the atmosphere. *J. Atmos. Sci.*, **31**, 775–805

- Mapes, B. and Houze, R. 1992 An integrated view of the 1987 Australian monsoon and its meso-scale convective system. I: Horizontal structure. *Q. J. R. Meteorol. Soc.*, **118**, 927–963
- Mechoso, C. R., Robertson, A. W., Barth, N., Davey, M. K., Delecluse, P., Gent, P. R., Ineson, S., Kirtman, B., Latif, M., LeTreut, H., Nagai, T., Neelin, J. D., Philander, S. G. H., Polcher, J., Schopf, P. S., Stockdale, T., Suarez, M. J., Terray, L., Thual, O. and Tribbia, J. J. 1995 The seasonal cycle over the tropical Pacific in coupled ocean–atmosphere general circulation models. *Mon. Weather Rev.*, **123**, 2825–2838
- Mitchell, T. P. and Wallace, J. M. 1994 The annual cycle in equatorial convection and sea surface temperature. *J. Clim.*, **5**, 1140–1156
- Ooyama, K. 1969 Numerical simulation of the life cycle of tropical cyclones. *J. Atmos. Sci.*, **26**, 3–40
- O’Sullivan, D. J. and Hitchman, M. H. 1992 Inertial instability and Rossby wave breaking in a numerical model. *J. Atmos. Sci.*, **49**, 991–1002
- Pike, A. C. 1971 Intertropical convergence zone studied with an interacting atmosphere and ocean model. *Mon. Weather Rev.*, **99**, 469–477
- Plumb, R. A. and Hou, A. Y. 1992 The response of a zonally symmetric atmosphere to subtropical thermal forcing: Threshold behavior. *J. Atmos. Sci.*, **49**, 1790–1799
- Ramage, C. S. 1974 Structure of an oceanic near-equatorial trough deduced from research aircraft traverses. *Mon. Weather Rev.*, **102**, 754–759
- Reynolds, R. W. and Marsico, D. C. 1993 An improved real-time global sea surface temperature analysis. *J. Clim.*, **6**, 114–119
- Reverdin, G. and Sommeria, G. 1981 ‘Dynamics of the tropical boundary layer deduced from constant level balloon trajectories over the Indian Ocean’. Pp. 12.12–12.16 in Proceedings of the International Conference on Early Results of FGGE and Large Scale Aspects of its Monsoon Experiments, Tallahassee, WMO
- Riehl, H. 1979 *Climate and Weather of the Tropics*, Academic Press, London
- Riehl, H. and Malkus, J. S. 1958 On the heat balance of the equatorial trough zone. *Geophysica.*, **6**, 503–537
- Rodwell, M. J. and Hoskins, B. J. 1995 A model of the Asian Summer Monsoon. Part II: Cross-Equatorial flow and PV behavior. *J. Atmos. Sci.*, **52**, 1341–1356
- Sadler, J. 1975a The monsoon circulation and cloudiness over the GATE area. *Mon. Weather Rev.*, **103**, 369–387
- 1975b The upper tropospheric circulation over the global tropics. Department of Meteorology, University of Hawaii, UHMET-75-05
- Sawyer, J. S. 1949 The significance of dynamic instability in atmospheric motions. *Q. J. R. Meteorol. Soc.*, **75**, 364–374
- Schneider, E. K. 1987 A simplified model of the modified Hadley circulation. *J. Atmos. Sci.*, **44**, 3311–3328
- Schubert, W. H., Ciesielski, P. E., Stevens, D. E. and Kuo, H.-C. 1991 Potential vorticity modeling of the ITCZ and the Hadley Circulation. *J. Atmos. Sci.*, **48**, 1498–1509
- Shutts, G. J. 1990 Dynamical aspects of the October storm of 1987: A study of a successful fine mesh simulation. *Q. J. R. Meteorol. Soc.*, **116**, 1315–1347
- Sikka, D. R. and Gadgil, S. 1980 On the maximum cloud zone and the ITCZ over the Indian longitude during the southwest monsoon. *Mon. Weather Rev.*, **108**, 1840–1853
- Solberg, P. H. 1936 ‘Le mouvement d’inertie de l’atmosphère stable et son rôle dans la théorie des cyclones’. U.G.G.I. Assoc. Meteorol. Edinburgh., 66–82
- Srinivasan, J., Gadgil, S. and Webster, P. J. 1993 Meridional propagation of large-scale monsoon convective zones. *Meteorol. Atmos. Phys.*, **52**, 15–35
- Stevens, D. E. 1983 On symmetric stability and instability of zonal mean flows near the equator. *J. Atmos. Sci.*, **40**, 882–893
- Thorpe, A. J., Hoskins, B. J. and Innocentini, V. 1989 The parcel method in a baroclinic atmosphere. *J. Atmos. Sci.*, **46**, 1274–1284
- Tomas, R. A. and Webster, P. J. 1994 The horizontal and vertical structure of cross-equatorial wave propagation. *J. Atmos. Sci.*, **51**, 1418–1430

- Trenberth, K. E. and Solomon, A. 1993 Implications of global atmospheric spatial spectra for processing and displaying data. *J. Clim.*, **6**, 531–545
- Walliser, D. E. and Somerville, R. C. J. 1994 Preferred latitudes of the intertropical convergence zone. *J. Atmos. Sci.*, **51**, 1619–1639
- Webster, P. J. 1983 Mechanisms of monsoon low-frequency variability: Surface hydrological effects. *J. Atmos. Sci.*, **40**, 2110–2124
- 1994 The role of hydrological processes in ocean–atmosphere interaction. *Rev. Geophys.*, **32**, 427–476
- Webster, P. J. and Lukas, R. 1992 TOGA-COARE: The Coupled Ocean–Atmosphere Response Experiment. *Bull. Am. Meteor. Soc.*, **73**, 1377–1416
- Webster, P. J., Palmer, T. N., Skukla, J., Magaña, V., Yanai, M. and Yasumari, T. 1997 The Monsoon: Processes and Prediction. *J. Geophys. Res.* In press