LARGE-SCALE AND MESOSCALE CIRCULATIONS IN CONVECTIVELY ADJUSTED ATMOSPHERES

Kerry A. Emanuel
Center for Meteorology and Physical Oceanography
Massachusetts Institute of Technology
Cambridge, Massachusetts, USA

Summary. There is mounting evidence that the state of conditional instability, corresponding to finite reservoirs of convective available potential energy, is a highly anomalous circumstance in the atmosphere and is confined to certain middle latitude continents at particular times of the year. More frequently, the atmosphere is very nearly neutral to moist convection—upright in the tropics and slantwise elsewhere. In view of this circumstance I argue that Conditional Instability of the Second Kind (CISK) is a false hypothesis.

Given this condition, it is of interest to examine the nature of large- and mesoscale circulations that occur in moist-neutral atmospheres. We begin by examining what is known from both observations and theory about frontal circulations and baroclinic development under conditions of moist slantwise neutrality. We particularly emphasize that the structure and evolution of moist-neutral baroclinic circulations is in far better accord with observations than is the classical "dry" theory. We then proceed to demonstrate that many cyclonic storms previously attributed to CISK (e.g., hurricanes and polar lows) are more profitably viewed as air-sea interaction instabilities that arise from the wind-dependence of surface heat fluxes. We also present evidence that certain global and mesoscale circulations arise from air-sea interaction as well. New results pertaining to the effects of wind-dependent surface heat fluxes on maritime cyclogenesis are presented, and we make several recommendations pertaining to possible improvements in large-scale numerical models in light of these recent findings.

1. INTRODUCTION

It has been recognized for some time that phase changes of water represent an important source of energy for atmospheric motions; indeed some of the early ideas on middle latitude cyclogenesis attributted it to latent heat release (Espy, 1841). While the equations governing condensation and, to a lesser extent, precipitation and ice formation are well known, the means of handling such phase changes in large-scale atmospheric simulation models remain controversial. A glance at any satellite photograph or radar display immediately reveals that clouds and precipitation exhibit far finer structure than does the pressure field, which has been used historically to define many large-scale systems. The fact that latent heat release generally occurs within very small structures presents an enormous challenge to both conceptual understanding and numerical simulation.

Among the most trying problems is that of understanding large- and mesoscale circulations in conditionally unstable atmospheres, for here virtually all latent heat release occurs within convective towers that cover a fractional area of only a few per cent. Attempts to understand the "interaction" of convective-scale circulations with those on the mesoscale and larger scales have frustrated the best minds. Today it must be admitted that our sketchy conceptual models of convective effects on the larger scales help us very little in understanding what we observe in nature or even in numerical models.

The novice acquainted with classical dry convection might at first be tempted to deny the existence of any fundamental scale interaction problem for, as he would point out, convective motions occur on such short space and time scales compared to those associated with large-scale circulations that there would be little interaction; instead, the convective motions would maintain a state that the large-scale views as statically neutral. Indeed, when it comes to dry convection, we are content to enforce a dry convective adjustment in our models and very little is heard in the way of scientific gnashing of teeth about this. How do we explain to the novice how moist convection differs?

We might first introduce the concept of conditional instability and point out

that phase changes of water, among other things, allow Convective Available Potential Energy (CAPE) to be built up over long periods of time and then released suddenly when and where the larger scale permits it. So as not to leave our novice friend in doubt we pull from our files a sounding from Texas in the spring and, sure enough, we see CAPE in great abundance, from which we and he are tempted to conclude that conditional instability is a common occurence. A glance at virtually any conventionally plotted tropical sounding supports our generalization. Once the concept of conditional instability is realized, the idea of scale interaction is revitalized and we are led to wonder about what circumstances conduce to releasing CAPE and whether the resulting large-scale flow feeds back positively on the convection. We are soon well down the garden path to Conditional Instability of the Second Kind (CISK).

But suppose our novice happens to be particularly gifted. He first points out that none of the large-scale circulations we attribute to CISK (e.g., tropical cyclones) occur in Texas in the spring, where CAPE is most abundant. He then inquires into the exact definition of CAPE and discovers that we have entirely neglected the effects of water substance (gaseous and condensed) on density. We point out defensively that after all, many clouds precipitate and he responds that this may *increase* the condensed water content in some parts of clouds (e.g., cloud base) even while decreasing it elsewhere. He has rediscovered Betts (1982). Figure 1 shows the "Jordan sounding" (Jordan, 1958) plotted both conventionally and in the form of virtual temperature of a reversibly lifted parcel against environmental virtual temperature. The latter contains almost no CAPE. Figure 2 shows the buoyancy of a reversibly lifted sub-cloud layer parcel averaged over many tropical soundings (as opposed to the buoyancy computed from an average sounding); this too shows that tropical soundings are neutral to within the instrumental error of standard rawinsondes. This state of neutrality is maintained by the convection itself against larger-scale destabilizing processes.

The near neutrality of the tropical atmosphere to reversible ascent from the sub-cloud layer is surprising, in view of the well-known highly dilute character of cumulus

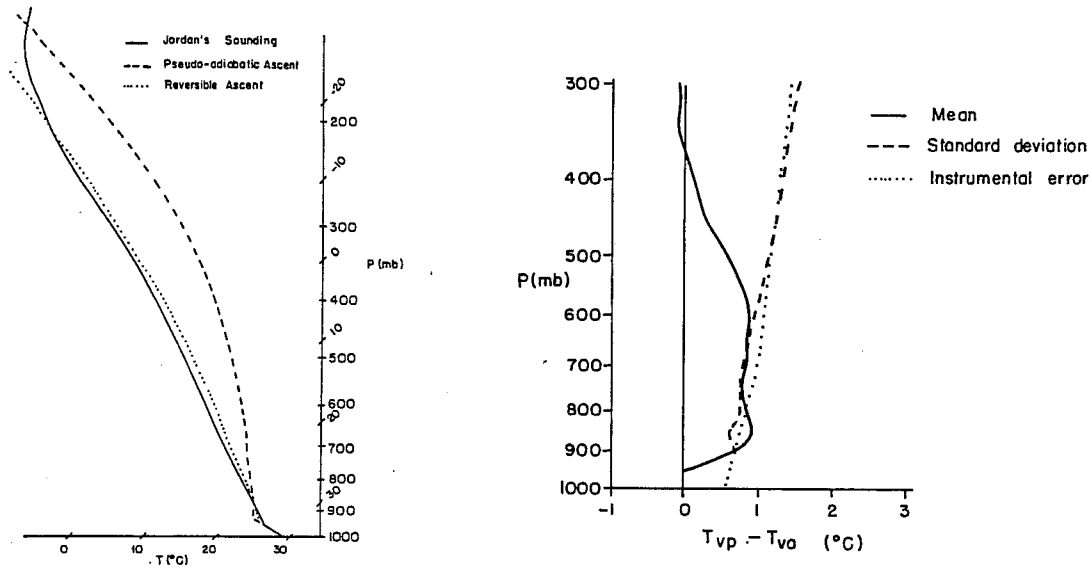

Fig. 1. (left) Tephigram showing Jordan's (1958) composite hurricane season virtual temperature sounding from the Carribean together with the virtual temperature of a parcel lifted pseudo-adiabatically from the sub-cloud layer (dashed) and a parcel lifted reversibly from the sub-cloud layer (dotted). The latter includes condensate loading in the definition of virtual temperature.

Fig. 2. (right) The difference between the virtual temperature of a parcel lifted reversibly from the top of the sub-cloud layer and the environmental virtual temperature averaged over a larger number of summertime soundings made at Truk. The dashed line shows the standard deviation of the same quantity while the estimated instrumental error is shown by the dotted line. Courtesy of K. Xu.

clouds. Observations (e.g., Warner, 1955) show, however, that such clouds are extremely inhomogeneous, and recent data (Paluch, 1979) shows that small samples of undilute sub-cloud layer air can usually be found. This observed neutrality to undilute parcels presents a challenge to cloud physicists.

The state of neutrality to moist convection is evidently not confined to the tropical atmosphere. The parcel theory of slantwise moist convection (Emanuel, 1983a, b) demonstrates that the stability to a combination of buoyant and centrifugal convection can be assessed by constructing soundings along sloping "M" surfaces rather

than in the vertical. Here M is defined

$$M \equiv V_g + fx, \tag{1}$$

where V_g is the component of geostrophic wind in the direction of the thermal wind, f is the Coriolis parameter and x is distance orthogonal to the thermal wind. The stability of soundings along M surfaces can be assessed the same way as in the case of upright soundings; in fact, the latter is a special case of the former.

Several attempts have been made to measure the slantwise moist stability in highly baroclinic ascent regions of middle latitude cyclones. All of these have found that such regions are slantwise neutral. Figure 3 presents a sounding made by flying an instrumented aircraft along an M surface, together with a (vertical) rawinsounding. Clearly the former is neutral and the latter is stable. A cross-section (Figure 4) constructed from closely spaced rawinsonde observations taken during a major field experiment at about the same time as the sounding in Figure 3 gives us additional information. Note that while lapse rates along M surfaces are moist adiabatic in many regions (Figure 4a), θ_e decreases upward along such surfaces in several places (Figure 4b). As the sounding was made in a general region of ascent, the implication of this is that while *potential* slantwise instability is present, it is released so quickly upon saturation that soundings are always approximately neutral. It is likely that this release of potential instability results in mesoscale precipitation and cloud bands in middle latitude cyclones (Seltzer et al., 1985).

It appears that the condition of moist neutrality is very general in the earth's atmosphere; large reservoirs of CAPE are as anomalous as are the very severe thunderstorms and tornadoes that result from it. Thus we shall focus on the implications of convective neutrality for understanding and modelling large-scale and mesoscale circulations.

The first and perhaps most important conclusions that follows is that in the absence of conditional instability, CISK (as defined by Charney and Eliassen, 1964) is impossible. This is because in a neutral atmosphere there is no physical way of lowering the center of gravity (except by evaporation of rain, as will be discussed in Sec-

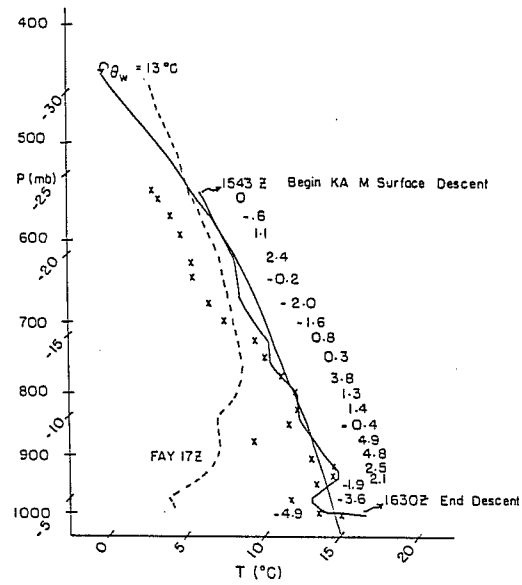

Fig. 3. M-surface sounding by NCAR King Air between 1543 GMT and 1630 GMT on 26 January 1986. Solid line shows temperature while x's denote dewpoint. Numbers to right show departures of M (m s^{-1}) from initial value. Dashed line is temperature from the 17 GMT CLASS sounding at Fayetteville, North Carolina (FAY).

tion 5) and thus there is no source of energy. This is true no matter how great the "moisture convergence" and latent heat release in a system. Likewise, very small conditional instability represents a correspondingly small energy source, so the state of neutrality is in no sense a singular limit. The only times and places where CISK is possible is when and where there is large CAPE and here CISK-like systems are not observed. We submit that CISK is a false hypothesis.

It also follows that in a convectively neutral atmosphere there is a one-to-one relationship between the density of the free atmosphere and the moist entropy of the sub-cloud layer. The only way to generate potential energy in such an atmosphere is to either change the moist entropy content of the sub-cloud layer or submit the atmosphere to a large-scale process that *increases* the stability of the atmosphere (for which there is no restoring convective adjustment). The latter process always represents a sink of energy, since work must be done against gravity, and thus no self-

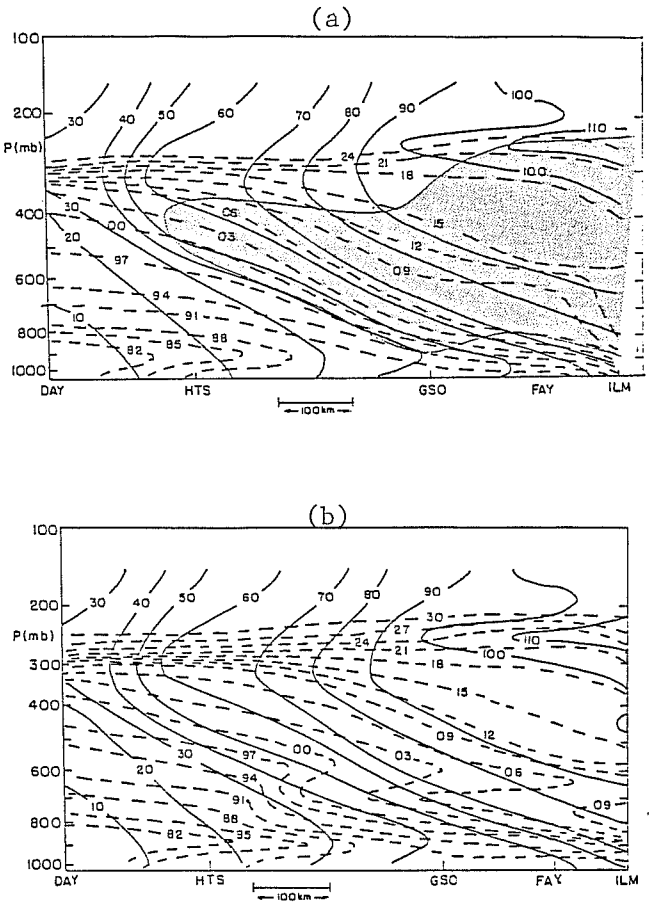

Fig. 4. a. Cross-section from Dayton, Ohio (DAY) to Wilmington, North Carolina (ILM) constructed from National Weather Service and CLASS soundings made at DAY, Huntington, West Virginia (HTS), Greensboro, North Carolina (GSO), Fayetteville, North Carolina (FAY) and ILM at 15 GMT 26 January 1986. Solid lines denote M (m s^{-1}) while dashed lines show θ_e^* (K; first digit omitted and values above 324 K not contoured). Stippling denotes regions where lapse rate along M-surfaces is greater than or equal to moist adiabatic rate. b. Same as Fig. 4a but showing θ_e (K) and M (m s^{-1}). Values of θ_e above 330 K not contoured.

exciting systems are possible by this route. In Section 3 we review the evidence that systems formerly attributed to CISK (such as tropical cyclones) result instead from an unstable feedback between large-scale circulations and the moist entropy content of the sub-cloud layer. The effects of surface entropy fluxes on baroclinic middle latitude cyclones are explored in Section 4.

There is one other potential energy source in convectively neutral barotropic atmospheres: this results from the cold source available by evaporation of rain. Even though the atmosphere is conditional neutral, its θ_e often decreases with height so that the introduction of liquid water at low θ_e results in the realization of the potential buoyant energy. There is an increasing tendency to look to evaporation of rain as the major source of energy for certain types of mesoscale convective systems. These will be discussed briefly in Section 5.

Finally, the condition of neutrality by itself has important implications for the development and structure of baroclinic systems, which are sensitive to the effective static stability. We discuss these implications in the following section.

2. BAROCLINIC CIRCULATIONS IN SLANTWISE NEUTRAL CONDITIONS

2.1 Frontal circulations

To form a conceptual picture of the effects of slantwise neutrality on baroclinic circulations it is first necessary to determine where the atmosphere is saturated (slantwise neutral) and where it is not. To date, most inquiries have made the simplifying postulate that *where there is ascent the atmosphere is saturated and slantwise neutral while the atmosphere is subsaturated and stable in descent regions.* Emanuel (1985) solved the Sawyer–Eliassen equation for specified frontogenetical forcing and showed that this results in extremely thin sloping sheets of ascent and broad, weak downdrafts. These effects were treated somewhat more elegantly by Thope and Emanuel (1985), who solved the classical semi-geostrophic frontogenesis model of Hoskins and Bretherton (1972) for small (but, for numerical reasons, finite) moist slantwise stability. Some of their results are presented in Figure 5. Note that frontal collapse occurs first at the surface and that the ascent is very strong and narrow. Also, dry potential vorticity (PV) is generated near the surface by condensation, giving rise to a narrow strip of high PV that may be unstable to internal barotropic/baroclinic instabilities. The features displayed in Figure 5 are in somewhat better agreement with observations than is the classical dry model.

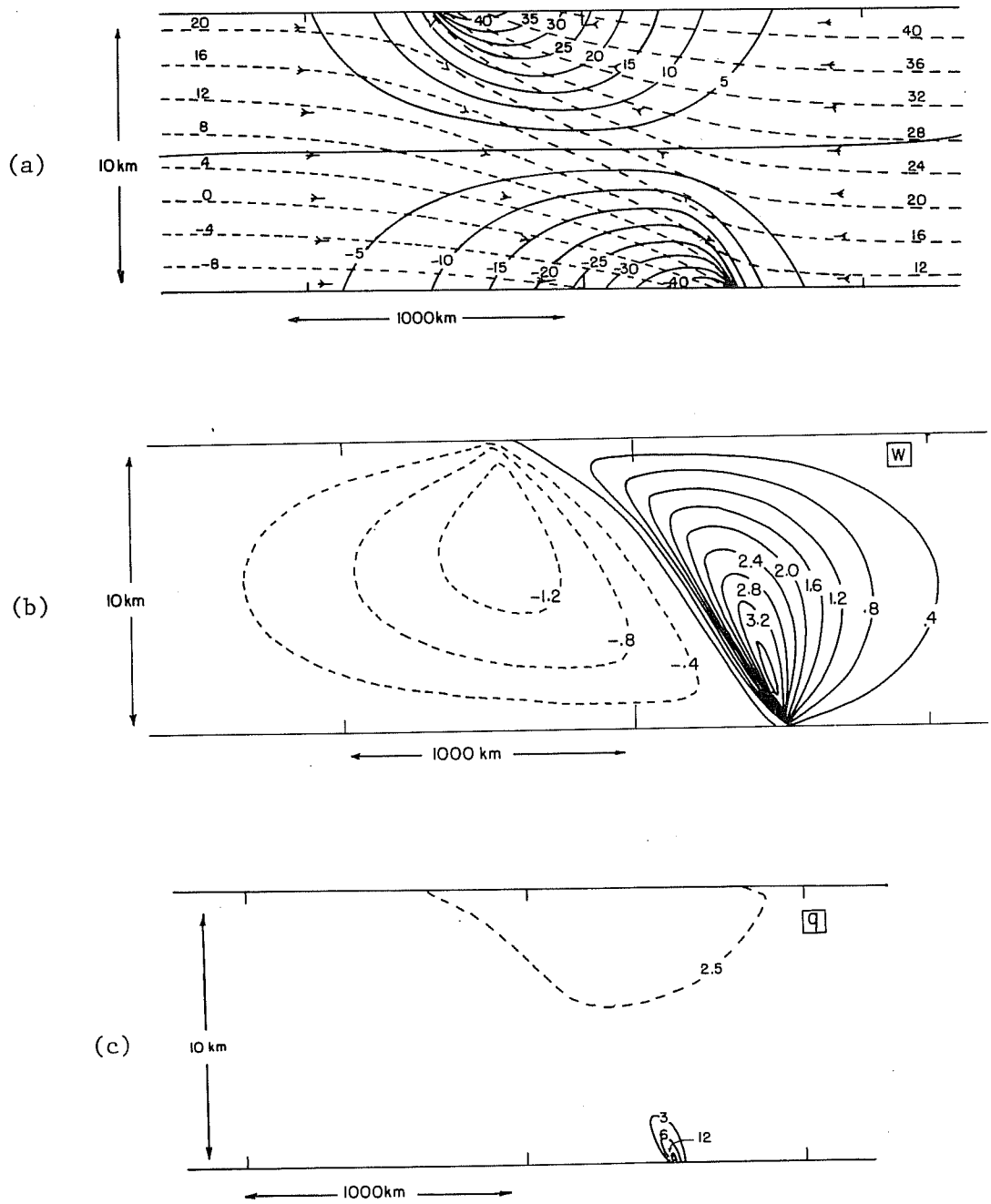

Fig. 5. a. Meridional velocity (solid, m s^{-1}) and potential temperature (dashed, K-300) after 24 hrs of simulated frontogenesis in a semi-geostrophic model forced by constant stretching deformation and with specified moist potential vorticity of 0.1 times the initial dry potential vorticity. From Thorpe and Emanuel (1985). b. Same as 5a but for vertical velocity (cm s^{-1}). c. Same as 5a but for potential vorticity (10^7 m^2 K s^{-1} kg^{-1}).

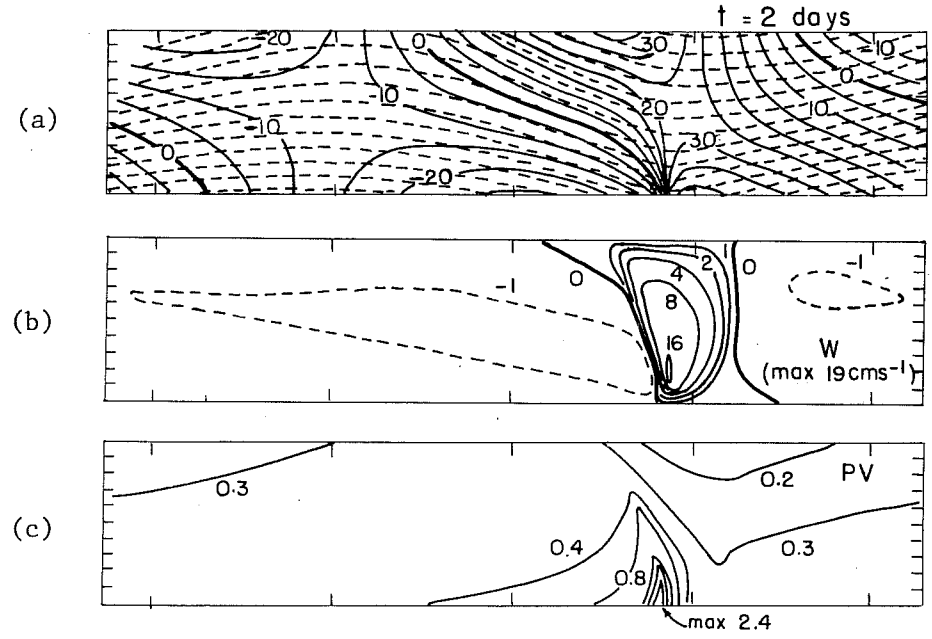

Fig. 6. a. Meridional velocity (solid, m s^{-1}) and potential temperature (dashed with contour interval of 4K) after 2 days of simulated 2-dimensional Eady wave evolution in a numerical semi-geostrophic model with moist potential vorticity equal to 0.1 times the initial dry value. Domain is periodic in x with total wavelength = 4554 km, and is 10 km deep. From Emanuel et al. (1987). b. Same as 6a but for vertical velocity (cm s^{-1}). c. Same as 6b but for potential vorticity (10^{-6} K m^2 Kg^{-1} s^{-1}).

2.2 Baroclinic instability

Progress is being made in understanding baroclinic instability in slantwise neutral conditions. The simplest approach is to examine the stability of a two-layer semi-geostrophic model to two-dimensional Eady-type waves, once again assuming neutral conditions in the ascent regions. Analytic solutions were obtained by Emanuel, Fantini and Thorpe (1987) who showed that for zero slantwise stability the growth rate is about 2.5 times the classical dry value and the total wavelength of maximum growth is about 60% of the dry wavelength. Once again, the updraft collapses into an infinitesimally thin sloping sheet and frontogenesis occurs first at the surface. The phase speed is not altered. Numerical solutions of the semi-geostrophic equations for two-dimensional moist Eady waves (once again using small but finite moist slantwise

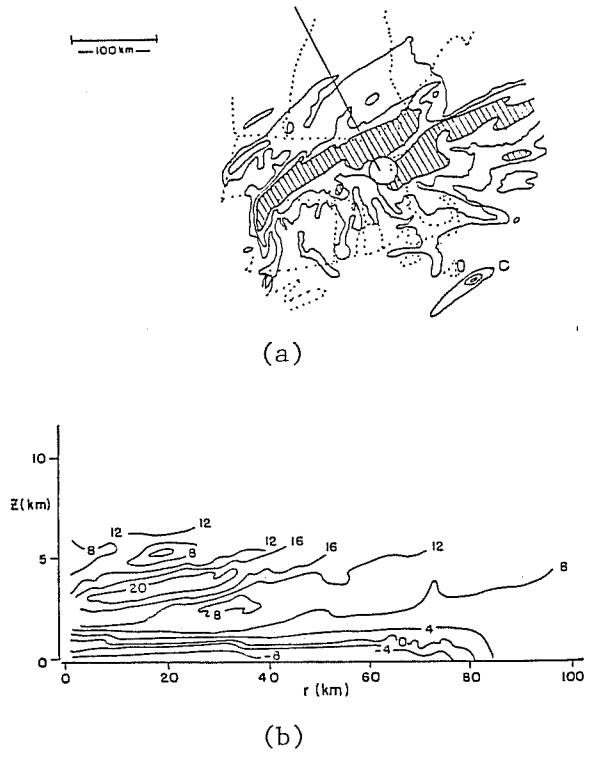

Fig. 7. a. Constant altitude display of radar reflectivity (dbz) in the layer from 2-4 km from M.I.T. 10 cm radar located at Cambridge, Massachusetts, at 7.29 GMT on 12 February 1983. Hatched regions show reflectivity greater than dbz. Precipitation was in the form of snow. Heavy bar shows cross-section line of Fig. 7b. b. Radial velocity (m s^{-1}) along cross-section show in Fig. 7a. Positive values indicate flow away from radar.

stability) show the structure in greater detail (Figure 6). Thorpe (personal communication) has performed numerical simulations of three-dimensional moist baroclinic waves that show some of the same features; i.e., faster growth, frontal formation first at the surface and sloping sheets of ascent. Fantini (personal communication) has demonstrated that non-semi-geostrophic effects cause the updraft to have finite width even at zero moist slantwise stability, so that dissipative processes need not be invoked to explain the finite width of updrafts in nature.

The structures seen in simulations at neutral slantwise stability bring to mind the "conveyor belts" of Browning and Harrold (1969). Indeed, there is some observational evidence that the ascent regions of baroclinic cyclones do indeed take the

form of thin sheets. Figure 7 shows a Doppler radar cross-section through the major regions of ascent in an intense winter storm, which was assessed by Sanders and Bosart (1985) to be nearly slantwise neutral. The cross-section was made almost perpendicularly to the geostrophic wind at all levels and shows extraordinary sloping ageostrophic flow of close to 30 m s^{-1}. The author has seen very similar, though less intense, examples of such structures in other storms as well. Perhaps we are seeing Browning's conveyor belt in Figure 7.

Many avenues of further research along these lines suggest themselves. How does non-modal baroclinic development (of the type discussed by Farrell, 1984) proceed in a slantwise neutral environment? What are the effects of evaporating precipitation? These and others constitute important scientific questions of direct consequence for understanding and predicting middle latitude weather systems.

3. CIRCULATIONS DRIVEN BY SURFACE SENSIBLE AND LATENT HEAT FLUXES

As stated in the Introduction, no self-exciting systems are possible in barotropic conditionally neutral atmospheres without evaporation of rain or actual non-moist adiabatic processes that change the sub-cloud layer moist entropy. There does, however, appear to be a class of air-sea interaction instabilities in which surface winds associated with large-scale or mesoscale circulations induce anomalous surface latent and/or sensible heat fluxes that increase the sub-cloud layer entropy and thereby increase the temperature aloft, resulting in further intensification of the circulation and so on. Examples and possible examples of such instabilities are discussed below.

3.1 Tropical cyclones

Observations (e.g., Riehl, 1954) leave little doubt that tropical cyclones are associated with dramatic radial gradients of sub-cloud layer moist entropy. Emanuel (1986) showed that there is a direct relationship between the total inward increase of entropy in such storms and their intensity, and that the mature storm can be viewed as

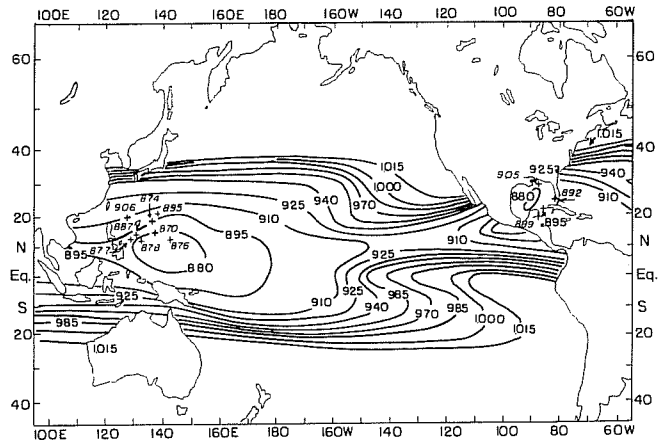

Fig. 8. Minimum sustainable pressures (mb) under September mean climatological conditions, assuming ambient pressure of 1015 mb. Crosses mark positions and intensities of some of the most intense tropical cyclones on record (Anthes, 1982).

a simple Carnot engine, with air extracting (latent) heat from the ocean at high temperature and losing it by export or radiation at the cold temperatures of the lower tropical stratosphere. (The conversion of latent to sensible heat by condensation is incidental and does not influence the overall energetics.) Once again, we emphasize that the energy source for such storms is the entropy difference between the sub-cloud layer and the ocean, and *not* the entropy difference between the sub-cloud layer and the free atmosphere. The latter difference, when and where it briefly occurs, is used to drive cumulus clouds, not the cyclone.

Because the entropy difference between the atmosphere and ocean is finite, there is an upper bound on the intensity of tropical cyclones. This has been calculated by Emanuel (1986) for September climatological conditions and is reproduced in Figure 8 together with locations and intensities of some of the most intense storms on record (from Anthes, 1982). Evidently, the most intense storms approach the upper bound on intensity.

The evolution of tropical cyclones is controlled by the feedback of surface winds on sub-cloud layer moist entropy and by the dynamical interaction of the storm with its environment. Both of these are complex and poorly understood processes. Ro-

tunno and Emanuel (1987) showed that in an axisymmetric primitive equation numerical model (in which interactions with the environment are necessarily excluded) the moist entropy budget of the sub-cloud layer is highly complex. In the outer regions of the storm, the excess surface entropy flux is nearly balanced by a convective flux out the top of the sub-cloud layer so that the sub-cloud layer entropy does not increase there. Thus surface fluxes are a necessary but not sufficient condition for increasing the sub-cloud layer moist entropy. In the inner region, the powerful surface fluxes are not thus balanced and entropy increases. In the mature stage, the entropy fluxes in the inner region are balanced by inward advection of low entropy air.

It is apparent that, aside from environmental interactions, the sub-cloud layer entropy budget exercises crucial control of the storm evolution. This is difficult to model. For example, precipitating downdrafts in nature are known to import low θ_e air into the sub-cloud layer, possibly "over-correcting" the initial small entropy surfeit that led to the precipitating cloud. If this were to occur extensively in a weak tropical cyclone, the resulting net decrease in sub-cloud layer entropy might very well kill the cyclone.

Even aside from these considerations, it has become apparent to us that the air-sea interaction instability that drives the cyclone is finite-amplitude in character; that is, sufficiently small initial disturbances decay while larger ones amplify. This was apparent in the simulations by Rotunno and Emanuel (1987) which showed that under identical environmental conditions weak initial cyclones decayed while strongers ones intensified. We are just beginning to understand the reasons for this. In the first place, the aforementioned simulations showed that, to a good approximation, the cyclone can be regarded as passing through a sequence of equilibrium states, each characterized by slantwise neutrality along angular momentum surfaces. Then, according to the theory of Emanuel (1986), the intensity and structure of the cyclone at any time is determined by the distribution of sub-cloud layer moist entropy with respect to angular momentum (M). Moreover, the Sawyer–Eliassen equation in p–M space is separable so that the following heuristic picture emerges (refer to Figure 9). In simplest terms, there are two effects on the distribution of sub-

cloud layer entropy with respect to M. The first is the all-important surface source of entropy, which by the classical aerodynamic flux formulae is proportional to the surface wind speed and the difference between the saturation entropy of the ocean surface and the entropy of the air. The second is the advection of entropy across M surfaces which, since M is conserved, must be accomplished by dissipative or non-axisymmetric processes. In an axisymmetric model the latter is absent so we focus on the former. The flow across M surfaces is proportional to the surface stress, which scales as V^2, and the radial gradient of θ_e. Thus we have, symbolically,

$$\begin{aligned} \text{surface entropy flux} &\sim V\Delta \ln \theta_e \\ \text{advection of entropy across } M \text{ surfaces} &\sim rV^2 \frac{d \ln \theta_e}{dM}. \end{aligned} \tag{2}$$

Now consider two scenarios. In the first (Figure 9a), we have a very weak vortex and weak radial entropy gradients in the sub-cloud layer. According to (2) the surface entropy flux scales as the first power of the velocity ($\Delta \ln \theta_e$ is finite) while the advection scales as the product of V^2 and the perturbation radial gradient of $\ln \theta_e$. Thus in a linear system the surface flux dominates by two orders over the cross-M advection. The resulting streamfunction in the p–M plane (from the Sawyer–Eliassen equation) shows upward motion in phase with the heating (i.e., with the surface V) and, more importantly, *no radial velocity at the location of maximum* V. Since the spin-up of V is proportional to the radial velocity u, no spin-up can occur and in fact friction will spin down the system. This constitutes a heuristic proof that the linear system is stable (this is borne out by quantitative analysis).

Now consider a vortex of sufficient amplitude that the surface fluxes and cross-M advection are of comparable magnitude (Figure 9b). The latter process displaces the maximum local time rate of change of sub-cloud layer entropy (in p-M space) radially inward of V_{max}. Once again, the Sawyer–Eliassen equation shows that the streamfunction dipole is centered above the maximum of $\partial \ln \theta_e / \partial \tau$, but this time it is radially inward of V_{max}, giving rise to radially inward flow which, if sufficiently strong, can spin-up V_{max} against friction. Thus amplification is conceivable provided the initial vortex is strong enough that the two processes given by (2) are of comparable magnitude. Note that this simple picture yields a finite-amplitude criterion

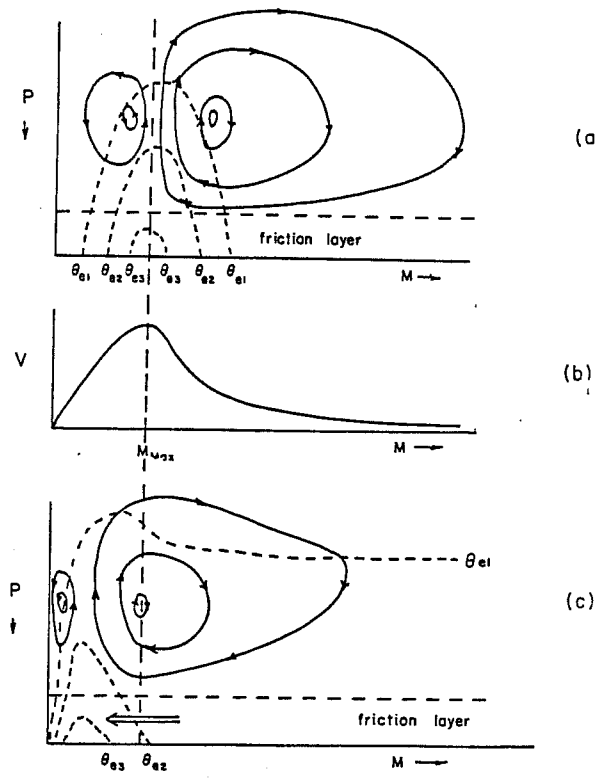

Fig. 9. Heuristic description of tropical cyclone development. Surface wind speed profile given in (b) as function of angular momentum (M). In (a), disturbance is weak so that local change of sub-cloud layer entropy is dominated by surface fluxes. Streamfunction dipole centered at M (V_{max}). In (b), disturbance is strong enough that cross-M entropy advection by friction-driven wind displaces sub-cloud layer entropy tendency radially inward of V_{max}.

even in the absence of dynamic interactions with the environment and of cold downdraft effects on the sub-cloud layer.

The finite-amplitude nature of this instability is reassuring in light of observations which shown that tropical cyclones never seem to form "out of thin air" but always arise from pre-existing disturbances such as easterly waves or cyclones on old fronts. Obviously we have a long way to go to determine sufficient conditions for amplification.

3.2 Polar lows

There is increasing evidence that polar lows operate at least partially on the same principle as tropical cyclones. Auer (1986) and S. Businger (personal communication) have correlated polar low occurrence and intensity with an index defined as the difference between the 500 mb temperature of a parcel lifted from saturation at sea surface temperature and ambient temperature at 500 mb. Emanuel and Rotunno (1988) have shown that the thermodynamic conditions that arise when arctic air flows out over relatively warm water (such as occurs, for example, over the Norwegian and Barent Seas in winter) are capable of supporting intense hurricanes, and were able to spin up an intense polar low in an axisymmetric tropical cyclone model (but with ice physics) starting from observed initial conditions. The structure of one such simulated polar low is shown in Figure 10. (In the case of polar lows, most of the surface flux comes from sensible rather than latent heat transfer.) Once again, the instability is finite-amplitude in nature and requires a starting vortex which in this case is probably provided by baroclinic instability. In Section 4 we shall explore the possibility of cooperation between sea-air heat fluxes and baroclinic instability.

3.3 The 40-60-day oscillation

The most popular explanation of this phenomenon is CISK and there is some evidence that the modes observed in certain General Circulation Models (GCM's) are caused by CISK (I. Held, personal communication). Yet these may be artifacts of cumulus parameterization schemes. The Kuo scheme, for example, requires moisture convergence for adjustment so that slight large-scale moisture divergence accompanied by cooling aloft will lead, in the model, to build-ups of CAPE that are probably artificial. The problem here is a chicken-and-egg one: without moisture convergence the model cannot convect, but in nature random noise might well set off convective clouds that generate their own moisture convergence. The issue might be settled by examining vertical soundings taken from the models to see whether CAPE exists at various times and places and then comparing the results to nature.

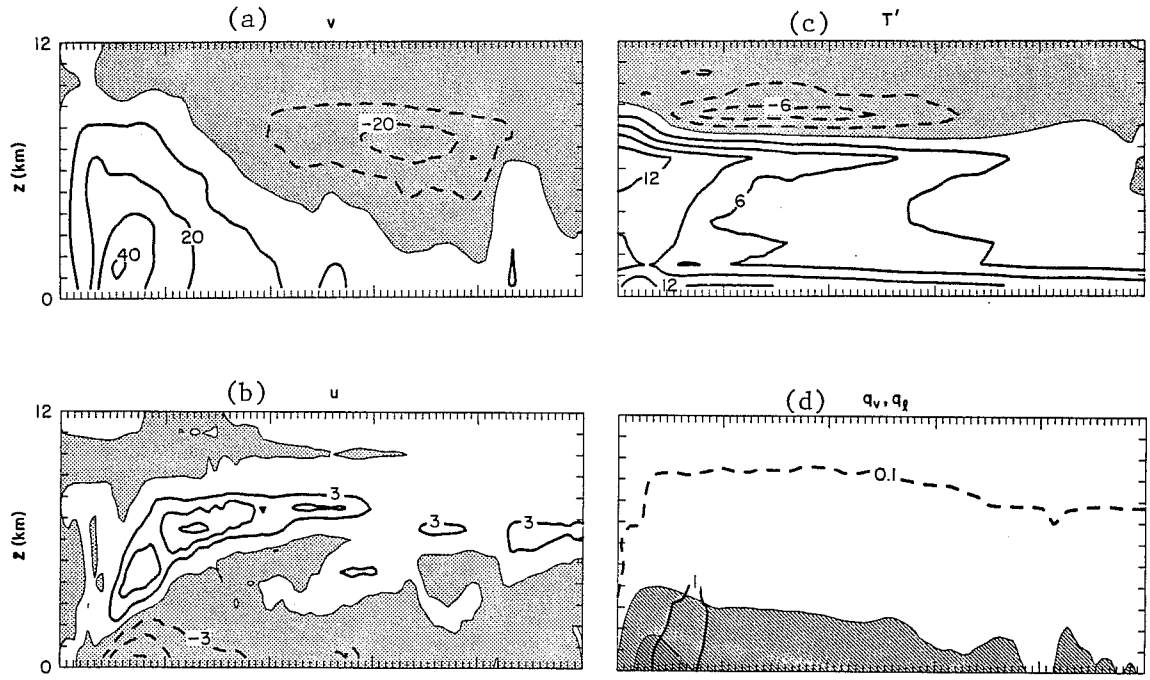

Fig. 10. Radius-height cross-sections of various quantities averaged over 20 hrs of a numerical simulation of a polar low described by Emanuel and Rotunno (1988). The simulated low was in a quasi-steady condition. The quantities are (a), azimuthal velocity (m s^{-1}); b, radial velocity (m s^{-1}); (c), temperature perturbation from initial condition (K); and (d), cloud water content greater than 0.1 g kg^{-1} (dashed line), snow content in g kg^{-1} (solid line) and water vapor content (shading).

There is some tenuous evidence that the 40–60-day oscillation may be due to an air-sea interaction instability which, as in the case of tropical cyclones, would be due to the wind dependence of surface latent heat fluxes. Investigations by Neelin et al. (1987) and Emanuel (1987) show that linear models based on evaporation-wind feedback in neutrally stratified atmospheres give phase speeds in the right ball park, but the scale selection is in the wrong direction. A heuristic picture of the hypothetical dynamics is shown in Figure 11: In a mean easterly flow, anomalous easterlies give rise to anomalous surface latent heat fluxes which, when distributed upward to maintain convective neutrality, warm the atmosphere and drive the wave eastward. Measurements of surface relative humidity are probably not good enough to test this hypothesis, and a clear-cut experiment to do so has not yet been proposed.

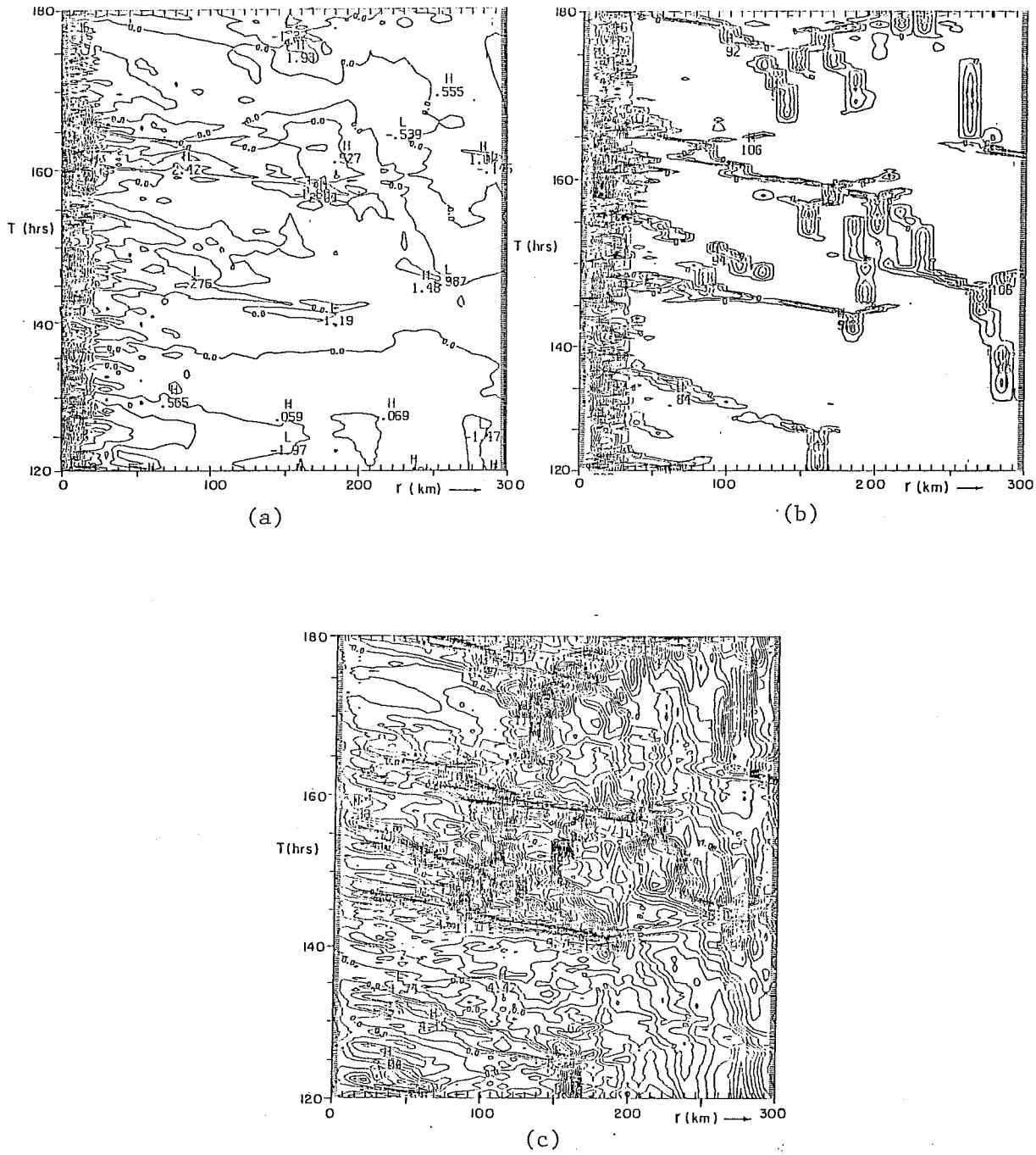

Fig. 11. Time-radius diagrams of various quantities from the lowest grid points of a high resolution simulation of a tropical cyclone (Experiment "D" from Rotunno and Emanuel, 1987), beginning at 120 hrs (bottom) and ending at 180 hrs (top). Quantities shown represent departures from the 120-180 hr time average time average fields of azimuthal velocity (a, m s^{-1}); rain water (b, g kg^{-1}) and equivalent potential temperature (c, °K).

3.4 Concentric eye walls in tropical cyclones

These have been described in some detail by Willoughby et al. (1984). In very intense tropical cyclones, multiple eyewalls are often observed. These generally contract inward and are accompanied by a transient intensification of winds and low pressure. After an eyewall moves inward of some critical radius, it begins to decay and is replaced by another convective ring at larger radius.

Concentric eyewalls have been found in numerical models as well. I have recently discovered them in a numerical simulation reported by Rotunno and Emanuel (1987). They use a primitive equation, axisymmetric nonhydrostatic model with 7.5 km radial resolution and no convective parameterization. Figure 11 displays time-radius diagrams of departures from time averages of rainwater (ℓ), azimuthal velocity (V) and equivalent potential temperature (θ_e) at the lowest model grid point. These plots were made from 60 hours of simulated time during which the model storm was in a quasi-steady state.

Inward propagating convective rings are clearly indicated in the rainwater field. These are moving inward at about 6 m s^{-1}. They are clearly associated with positive anomalies in both the V and θ_e fields; careful analysis shows that all three anomalies are nearly exactly in phase. The existence of high θ_e anomalies associated with the convective rings suggests that these may be examples of a mesoscale air-sea interaction instability, though more analysis will be necessary to test this supposition.

4. EFFECTS OF SURFACE HEAT FLUXES ON MARITIME BAROCLINIC CYCLOGENESIS

Meteorologists have long suspected that the physics of certain extreme maritime cyclogenesis events is in some ways different from ordinary baroclinic development. An examination of the distribution of a large number of cyclogenesis events compiled by Roebber (1984) shows that significantly more extreme events occur than would characterize a normal distribution. This raises the suspicion that the physics of the non-Gaussian events departs from ordinary baroclinic dynamics.

This suspicion is deepened by the observation that certain operational numerical models systematically underestimate the intensity of maritime "bombs" even while doing well with extreme continental events. The National Meteorological Center's Limited Area Fine Mesh (LFM) model is among those with this characteristic. One deficiency of the model's physics that may contribute to the problem is its use of fixed (non-wind-dependent, non-time-dependent) heat fluxes over the ocean. MIT graduate student Chris Davis and I sought to test the hypothesis that deficient surface fluxes are responsible for the systematic underforecast of the intensity of bombs.

We developed an easily evaluated quantity that measures the susceptibility of the atmosphere to surface heating. This quantity is the difference between the "potential thickness" and the actual 1000–500 mb thickness, where potential thickness is defined as the thickness that would obtain were the atmosphere heated to a moist adiabatic lapse rate representing ascent of a parcel saturated at sea water temperature at 1000 mb. This potential thickness is the maximum the atmosphere could attain by heating from below and the difference between it and the actual thickness thus measures the susceptibility of the atmosphere to heating.

Figure 12 shows a scatter plot of the LFM 24-hr central pressure forecast error against the average difference between potential and actual thickness along the 24-hr storm track. The high correlation coefficient (0.81) between these quantities supports the hypothesis that the underforecast is due to insufficient heating from the ocean.

Preliminary theoretical analysis shows that while sensible heat fluxes may impede the linear growth phase of baroclinic cyclones, latent fluxes accelerate it since latent heat is released only within relatively warm, ascending air. In the mature, nonlinear phase, however, it is possible that both fluxes contribute to development. Here again, theoretical understanding is currently in a very primitive state.

5. MESOSCALE CIRCULATIONS DRIVEN BY EVAPORATION OF RAIN

One of the most prominent and long-lasting features of even weak precipitating convection is the rain-cooled air that spreads out at the surface, sometimes covering ar-

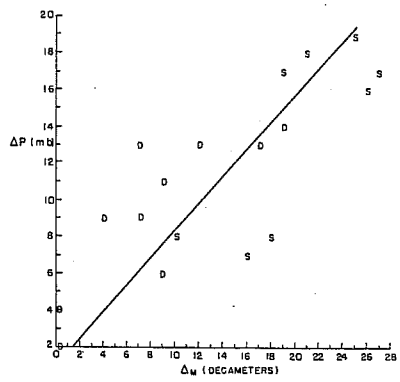

Fig. 12. Difference between LFM 24-hr central pressure forecast and observed central pressure against mean difference between potential and actual thickness. D's mean moderate bombs (1.2 Bergeron $< \Delta P <$ 1.8 Bergeron) while S's denote strong bombs ($\Delta P >$ 1.8 Bergeron). Best fit line with correlation coefficient = 0.81. Courtesy of Christopher Davis, M.I.T.

eas of up to 10^4–10^5 km^2. The sinking of the rain-cooled air is often a major source of kinetic energy in convective systems, and increasing attention is being focused on this process by modellers and theoreticians seeking to understand and predict mesoscale convective systems.

A relatively clear-cut example of a mesoscale system driven perhaps entirely by evaporative cooling was described by Carbone (1982). In this case, an intense narrow precipitation band advanced into a conditionally neutral air mass somewhat in advance of a cold front. The structure of the advancing cold air associated with the line strongly resembled a density current, with the cold air presumably provided by evaporation and melting of hydrometeors.

The possibility of self-exciting instabilities driven by evaporative cooling is an important unsolved theoretical problem. Clearly, in nature, such an instability would necessarily have a finite-amplitude character, since finite displacements are necessary to create rain. Yet not all "gust fronts" advancing into conditionally neutral or unstable air are observed to lead to convection, perhaps because the vertical displacements associated with them are too small to lead to rain-producing clouds having the correct phase relationship with the advancing cold air.

The conditions under which density currents lead to self-sustaining systems are

being explored by Richard Rotunno, Joseph Klemp, David Parsons and Morris Weisman of the National Center for Atmospheric Research. Preliminary numerical modeling efforts by this group indicate that the persistence of a convective system depends sensitively on the low-level component of vertical shear normal to the density current front. They suggest that persistent systems occur when the shear-related component of vorticity along the gust front is approximately canceled by baroclinic generation at the front, so that air entering cloud base has relatively small vorticity. Understanding the factors affecting the persistence of squall lines and other mesoscale convective systems remains an exciting challenge.

6. CONCLUSIONS AND RECOMMENDATIONS

The theme I have emphasized throughout this report is that, in general, the atmosphere is either stable or quasi-neutral to moist processes; the state of conditional instability in which non-trivial CAPE is present is anomalous. Recent work by Betts (1982) suggests that the tropical atmosphere is nearly neutral to *reversible* parcel ascent, and observations of middle-latitude cyclonic systems consistently show slantwise neutrality.

These observations tell us what factors are most important in simulating diabatic effects on large-scale and mesoscale circulations. First, parcel neutrality implies a one-to-one relationship between sub-cloud layer moist entropy and free atmosphere temperature. Therefore, *processes determining the moist entropy content of the sub-cloud layer, such as surface fluxes, entrainment and convective downdrafts, are of crucial importance in convective atmospheres.* These processes are presently handled crudely or not at all in most large-scale models. Secondly, *we should abandon "moisture convergence" as a criterion for latent heat release* even as we would never think of using potential temperature convergence as a criterion for dry adiabatic adjustment. If the atmosphere becomes conditionally unstable and the "negative area" is not large, the instability will be released regardless of the presence or absence of moisture convergence. Of course, if moisture convergence is not present and does not

ensue as a result of the convection, the sub-cloud layer will dry out and convection will cease. But a consistent model should be capable of simulating this process without the crutch of demanding moisture convergence; to do so may well result in an artifical build-up of CAPE.

The importance of surface fluxes cannot be emphasized enough. They are the source of energy in tropical cyclones and probably some polar lows, and there is some evidence that they make the difference between extratropical bombs and garden variety cyclones. The theory relating surface fluxes to wind, waves and stability is not well-developed and modellers should closely follow and encourage progress in understanding such relations. Long-term forecasts must certainly be sensitive to surface fluxes as well.

Finally, existing models should be examined to see whether they are capable of resolving the types of structures that are predicted and observed when baroclinic processes operate in slantwise neutral environments. We must face the possibility that much of the ascent in extratropical cyclones takes place in thin mesoscale sheets and take steps to insure that these are adequately resolved or represented.

7. REFERENCES

Anthes, R. A., 1982: Tropical Cyclones: Their Evolution, Structure and Effects. Boston, Amer. Met. Soc.

Auer, A. H., 1986: A observational study of polar air depressions in the Australasian region. Preprints of the Second Int. Conf. on South. Hem. Met., Boston, Amer. Met. Soc.

Betts, A. K., 1982: Saturation point analysis of moist convective overturning. J. Atmos. Sci., 39, 1484–1505.

Browning, K. A. and T. W. Harrold, 1969: Air motion and precipitation growth in a wave depression. Quart. J. Roy. Meteor. Soc., 95, 288–309.

Carbone, R., 1982: A severe winter squall line. Part I: Stormwide hydrodynamic structure. J. Atmos. Sci., 39, 258–279.

Charney, J. G. and A. Eliassen, 1964: On the growth of the hurricane depression. J. Atmos. Sci., 21, 68–75.

Emanuel, K. A., 1983a: The Lagrangian parcel dynamics of moist symmetric instability. J. Atmos. Sci., 40, 2368–2376.

Emanuel, K. A., 1983b: On assessing local conditional symmetric instability from atmospheric soundings. Mon. Wea. Rev., 111, 2016–2033.

Emanuel, K. A., 1985: Frontal circulations in the presence of small moist symmetric stability. J. Atmos. Sci., 42, 1062–1071.

Emanuel, K. A., 1986: An air–sea interaction theory for tropical cyclones. Part I: Steady-state maintenance. J. Atmos. Sci., 43, 585–604.

Emanuel, K. A., 1987: An air-sea interaction model of intraseasonal oscillations in the tropics. J. Atmos. Sci., 44, 2324–2340.

Emanuel, K. A., M. Fantini and A. J. Thorpe, 1987: Baroclinic instability in an environment of small stability to slantwise convection. Part I: Two-dimensional models. J. Atmos. Sci., 44, 1559–1573.

Emanuel, K. A. and R. Rotunno, 1988: Polar lows as arctic hurricanes. Submitted to Tellus.

Espy, J. P., 1841: The Philosophy of Storms. Boston, Little and Brown.

Farrell, B., 1984: Modal and non-modal baroclinic waves. J. Atmos. Sci., 41, 668–673.

Hoskins, B. J., and F. P. Bretherton, 1977: Atmospheric frontogenesis models: mathematical formation and solution. J. Atmos. Sci., 29, 11–37.

Jordan, C. L., 1958: Mean soundings for the West Indies area. J. Meteor., 15, 91–97.

Neelin, J. D., I. M. Held and K. H. Cook, 1987: Evaporation–wind feedback and low-frequency variability in the tropical atmosphere. J. Atmos. Sci., 44, 2341–2348.

Paluch, I. R., 1979: The entrainment mechanism in Colorado cumuli. J. Atmos. Sci., 36, 2462–2478.

Riehl, H., 1954: Tropical Meteorology. New York, McGraw–Hill, 392 pp.

Roebber, P. J., 1984: Statistical analysis and updated climatology of explosive cy-

clones. Mon. Wea. Rev., 112, 1577–1589.

Rotunno, R. and K. A. Emanuel, 1987: An air–sea interaction theory for tropical cyclones. Part II: Evolutionary study using a nonhydrostatic axisymmetric numerical model. J. Atmos. Sci., 44, 542–561.

Sanders, F. and L. F. Bosart, 1985: Mesoscale structure in the megalopolitan snowstorm of 11–12 February 1983. Part I: Frontogenetical forcing and symmetric instability. J. Atmos. Sci., 42, 1050–1061.

Seltzer, M. A., R. E. Passarelli and K. A. Emanuel, 1985: The possible role of symmetric instability in the formation of precipitation bands. J. Atmos. Sci., 42, 2207–2219.

Thorpe, A. J. and K. A. Emanuel, 1985: Frontogenesis in the presence of small stability to slantwise convection. J. Atmos. Sci., 42, 1809–1824.

Warner, J., 1955: The water content of cumuliform cloud. Tellus, 7, 449–457.

Willoughby, H. E., H.–L. Jin, S. J. Lord and J. M. Piotrowicz, 1984: Hurricane structure and evolution as simulated by an axisymmetric, nonhydrostatic numerical model. J. Atmos. Sci., 41, 1169–1186.

SURFACE DRAG AND THE CIRCULATION OF A BAROCLINICALLY UNSTABLE ATMOSPHERE

I.N. James
Department of Meteorology, University of Reading
READING RG6 2AU, U.K.

1. FEEDBACKS BETWEEN RESOLVED AND PARAMETRIZED PROCESSES

The theme of this paper is that feedbacks exist between the various diabatic and frictional processes, which are generally parametrized in numerical models, and also with the explicitly resolved large scale circulations. Models of various levels of simplicity will be presented to illustrate such feedback mechanisms.

The classical Lorenz energy cycle, exemplified by Fig. 1 taken from Kung and Tanaka (1983) makes this contention plain. Differential heating generates available potential energy. Conversions transform this available potential energy to eddy and eventually to zonal kinetic energy. Dissipation, especially at the surface, destroys kinetic energy and must eventually balance the energy input. Also noted on Fig. 1 are the residence times of energy in each of the boxes, calculated by dividing the energy content by the total flux through the box due to conversions and source/sink terms. The timescales are short; except for the zonal available potential energy, they do not exceed 10 days. Even the zonal available potential energy has a lifetime of 30-40 days, a value which is of course a global measure of the radiative time constant for the atmosphere. These short timescales mean that an error in, say, the dissipation rates will have a significant impact on the entire global energy budget during the course of a 10-day forecast. There would be least impact on the zonal available potential energy, but the zonal winds and steady and transient eddies could respond quite substantially.

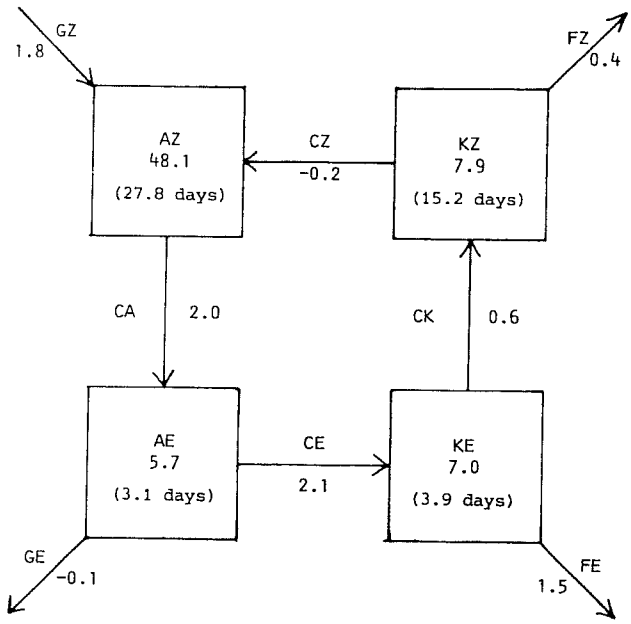

Fig. 1. A Lorenz energy diagram for the FGGE first special observing period, computed by Kung and Tanaka (1983) from ECMWF level IIIb analyses.

An alternative view is provided by potential vorticity considerations. I am indebted to Professor Brian Hoskins for this argument. Fig. 2 shows an isentrope S which intersects the surface of the Earth along the circuit L. The rate of change of potential vorticity is

$$\frac{D}{Dt} P = \frac{\dot{\theta}\zeta + \mathcal{F} \wedge \nabla\theta}{\rho} \quad . \tag{1}$$

Integrating over the surface S and using the circulation theorem gives

$$\frac{D}{Dt} \int_S \rho P dS = \oint_L \left(\mathcal{F} \frac{\partial\theta}{\partial n} - \dot{\theta}\zeta \right) d\ell \tag{2}$$

and so in the time average we must have

$$\overline{\oint \mathcal{F} \frac{\partial\theta}{\partial n} d\ell} = \overline{\oint_L \dot{\theta}\zeta \, d\ell} \quad . \tag{3}$$

Thus, if there is a net diabatic cooling around L the friction $\mathcal{F}$ must be negative, implying surface winds must be positive or westerly. Similarly, diabatic heating around L must imply easterly flow. The integral of P over S is closely related to the circulation around L, and this can be used to estimate a timescale for the modification of P by surface diabatic processes, and hence of the timescale needed for $\mathcal{F}$ and $\dot{\theta}$ to come into balance. Taking $\dot{\theta}$ as around 1 K day^{-1}, $\zeta = f = 10^{-4}$ s^{-1} and a typical surface wind as 5 m s^{-1}, the relaxation time for potential vorticity is again around 40 days, comparable to the radiative time constant.

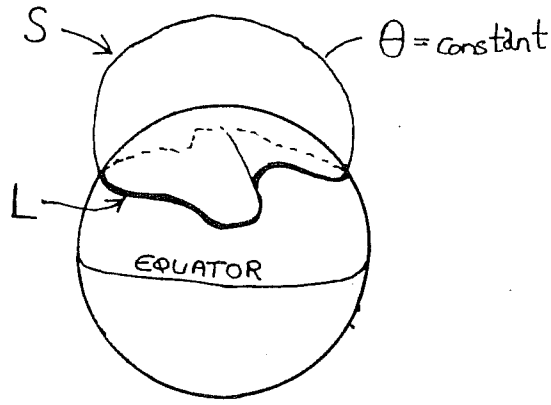

Fig. 2 Illustrating the isentropic surface S which intersects the Earth's surface around L.

Both these arguments serve to demonstrate the coupling between friction and diabatic processes on timescales no longer than that of the radiative time constant, and possibly a good deal shorter for certain specific aspects of the flow. In the succeeding sections, this coupling will be demonstrated using simple models of baroclinic flow on the sphere.

2. THE TROPICAL HADLEY CIRCULATION

Although described by Halley in 1685 and Hadley in 1735, simple quantitative accounts of axisymmetric circulations at low latitudes on a rotating planet have only been published relatively recently. The arguments of Held and Hou (1980) are particularly elegant and are surprisingly successful in accounting for the observed strength and extent of the tropical Hadley circulation.

Their model can be regarded as a two level model of the atmosphere. The surface layer experiences friction which reduces it to rest with respect to the ground on some unspecified but short timescale. The upper tropospheric flow, at height H, experiences no friction and so air conserves angular momentum as it moves away from the equator at upper levels. Conservation of angular momentum leads to a simple formula for the upper level zonal wind which, via the thermal wind equation and an integration with respect

to latitude leads to an expression for the variation of mid tropospheric temperature with latitude. Assuming that the temperature of the tropical atmosphere relaxes radiatively to some simple function symmetric about the equator, a condition of no net heating of parcels circulating in the Hadley cell leads to an expression for the width of the cell: the vertical and meridional velocities can also be calculated if a value for the radiative relaxation time is assumed. Fig. 3 illustrates their model.

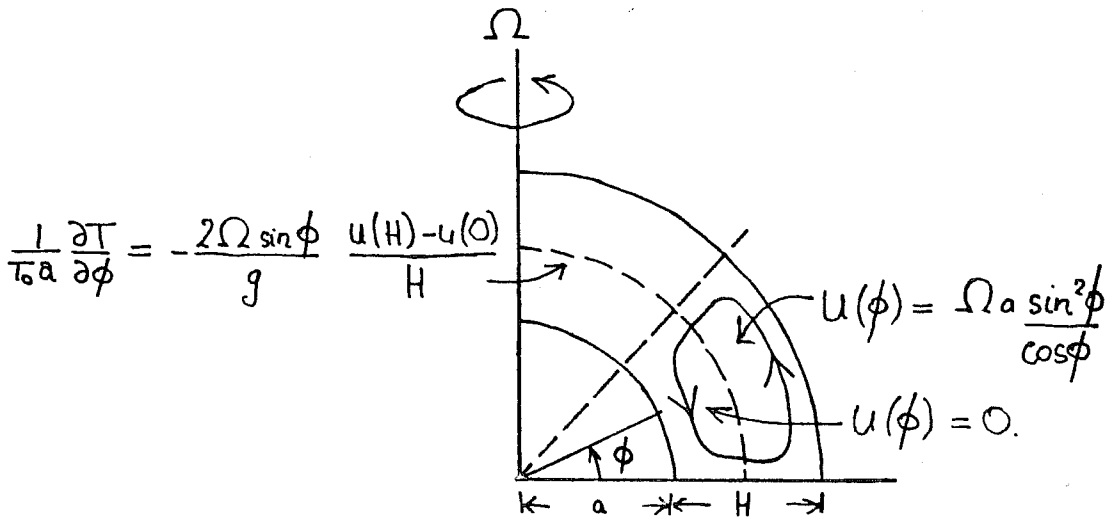

Fig. 3. The Held and Hou (1980) model of the Hadley circulation.

The model illustrates how the friction controls the large scale circulation, though, in this case, the upper level dynamics depend only weakly on the surface friction. The essential Held and Hou results are confirmed by a more elaborate axisymmetric primitive equation model. The thermodynamic equation includes relaxation on a 15 day timescale to a stably stratified basic state. The vorticity and divergence equations include vertical diffusion of the form $K\,\partial^2(\ )/\partial z^2$, which will set up a simple Ekman boundary layer in the presence of interior relative vorticity. The model can be integrated to a steady state, which is achieved after some 20-40 days of

integration, starting from a state of no motion. However, if the Ansatz coefficient K is reduced below some critical value (on the order of 1 m^2 s^{-1}), the flow remains unsteady. In full 3-dimensional models, this unsteadiness is much less pronounced, and it has been suggested that the eddies transport momentum at a rate which is equivalent to an eddy viscosity in excess of 1 m^2 s^{-1} in these integrations. But down to this critical diffusion, the flow changes only rather slowly as K is varied.

The simple arguments of Held and Hou assumed that the radiative equilibrium distribution of temperature was symmetric about the equator. It is interesting to extend the theory to the case when the radiative equilibrium temperature distribution reaches a maximum away from the equator, a situation which is more relevant to the solsticial seasonal mean. Such a modification has been made by Lindzen and Hou (personal communication) and reproduced by Paul James and the author at Reading University. As the maximum is moved away from the equator, the cells become highly asymmetric. The circulation is dominated by a major cell with ascent in the summer hemisphere and descent in the winter hemisphere, with only an extremely weak cell in the summer hemisphere. Paul James has prepared Figs, 4a and b using this modification of the theory; the results are confirmed using the same primitive equation model mentioned above (see Fig. 4c).

An important point of considerable generality is revealed by these calculations. The response, in the form of the Hadley cells, to the diabatic forcing, represented by the radiative equilibrium temperature field, is highly nonlinear. It follows that the mean meridional circulation through an annual cycle is much stronger than the response to the annual mean diabatic forcing. This latter is of course symmetric about the equator. Similar effects arise whenever the response to diabatic forcing is a nonlinear function of the forcing: thus any parametrization of such forcing must necessarily depend upon the space and time scales implicit in the parametrization.

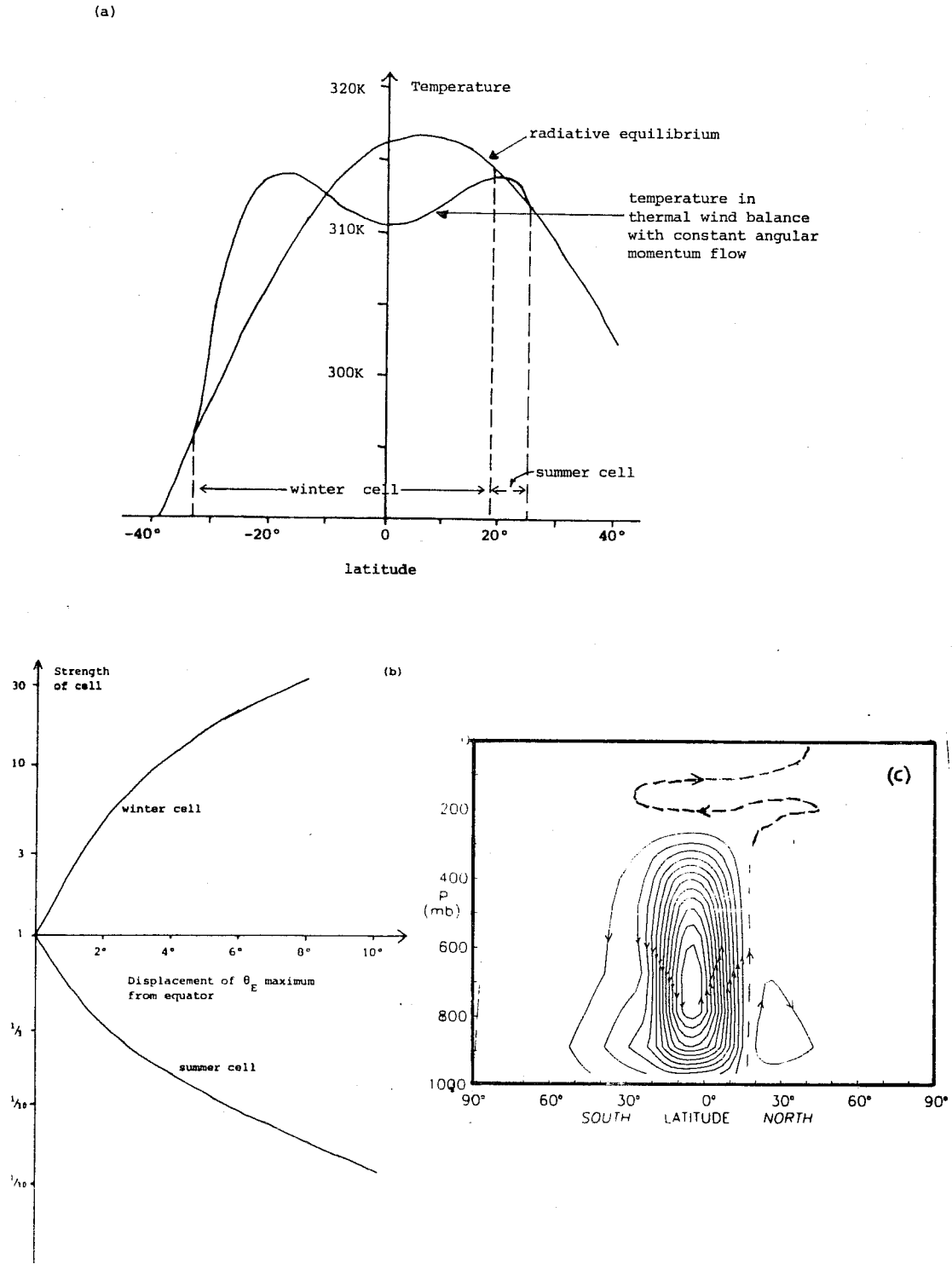

Fig. 4. The results of applying the Held and Hou model. (a) The thermal balance and width of the cell. (b) The strengths of the cells are a function of the latitude of the heating. (c) Axisymmetric primitive equation results for heating centred at 9°N.

3. BAROCLINIC MODELS WITH DRAG

The dynamics of an axysimmetric atmosphere are relatively straightforward, at least in as far as the flow will, with moderate dissipation rates, settle down to a unique steady state. When the axisymmetric constraint is relaxed, much greater complexity is possible. James and Gray (1987) carried out such integrations and this section is based on their paper.

The model used was similar to that described in the previous section, with Newtonian cooling towards some prescribed simple "radiative equilibrium" thermal state on a timescale of 15 days. Mechanical dissipation was introduced in the form of a simple Rayleigh friction (i.e., a term of the form $\underset{\sim}{u}/\tau_D$ on the right-hand side of the momentum equation) at the lowest model level. A rather coarse resolution with triangular truncation to wavenumber 21 and 5 equispaced levels in the vertical was employed, enabling long runs (of order 500 days) to be repeated with different imposed parameters. The drag timescale τ_D was varied between these runs.

When the drag timescale was short (i.e., large mechanical dissipation), the flow was baroclinically unstable and highly unsteady. In the midlatitudes, baroclinic waves developed and collapsed in an irregular fashion, resulting in large fluctuation of the poleward heat transport and the eddy kinetic energy on timescales up to 10-20 days. Fig. 5 shows the time mean zonal flow and meridional circulation for this case. It is a surprisingly faithful simulation of the gross aspects of the observed general circulation, with a low latitude Hadley cell and mid-latitude Ferrel cell driven by transient eddies associated with baroclinic instability of the strongly sheared subtropical jet.

The middle latitudes of this high drag run are filled with vigorous baroclinic waves; the fields of stream function at $\sigma = 0.3$ and temperature at $\sigma = 0.9$ are shown in Fig. 6. The dominant disturbances have zonal wavenumbers less than 8 and the characteristic baroclinic phase shift with height can be discerned.

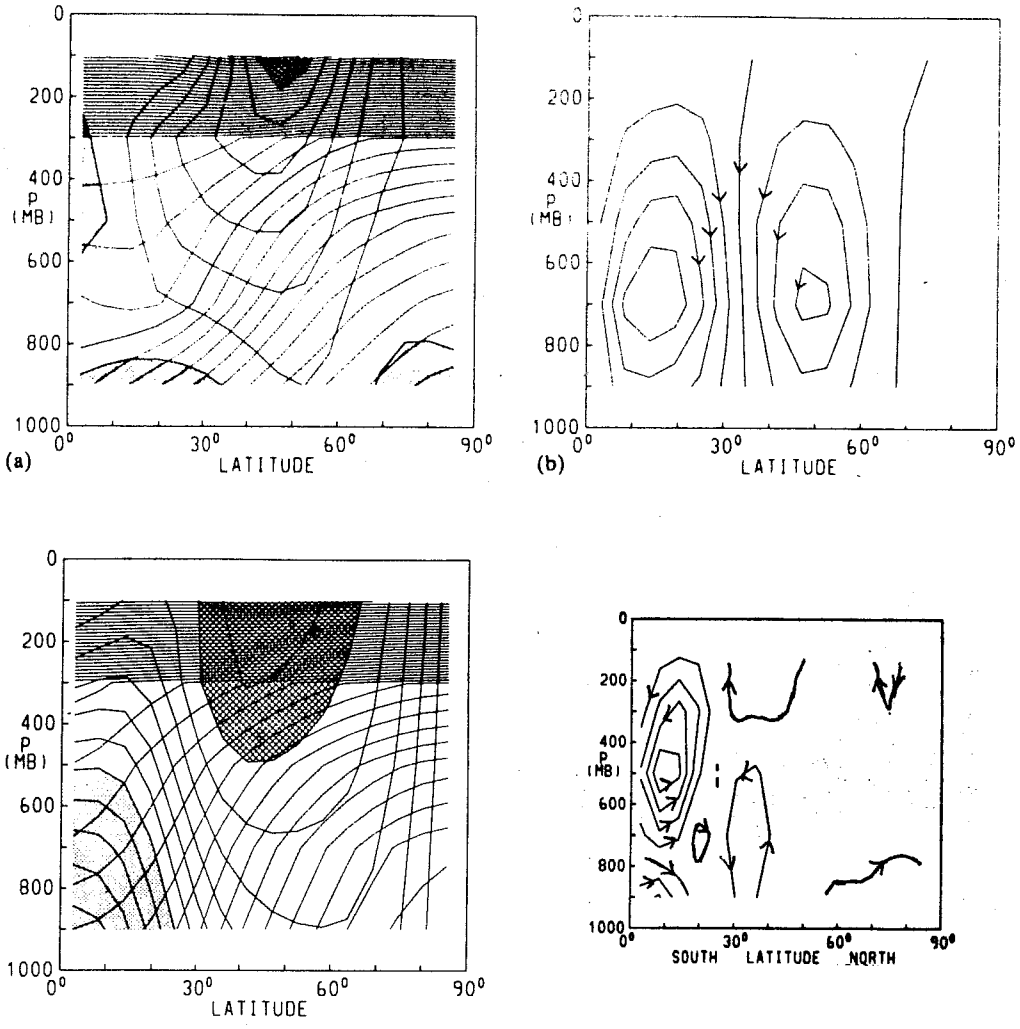

Fig. 5. The zonal mean climate of the simple model for extreme values of the drag.

(a) High drag case, [u] and [θ] (contour intervals 5 m s^{-1} and 5K).

(b) High drag case, meridional stream function, contour interval 22.4 mb s^{-1}.

(c) Low drag case, [u] and [θ] (contour intervals 5 m s^{-1} and 5K).

(d) Low drag case, meridional stream function, contour interval 1.44 mb s^{-1}.

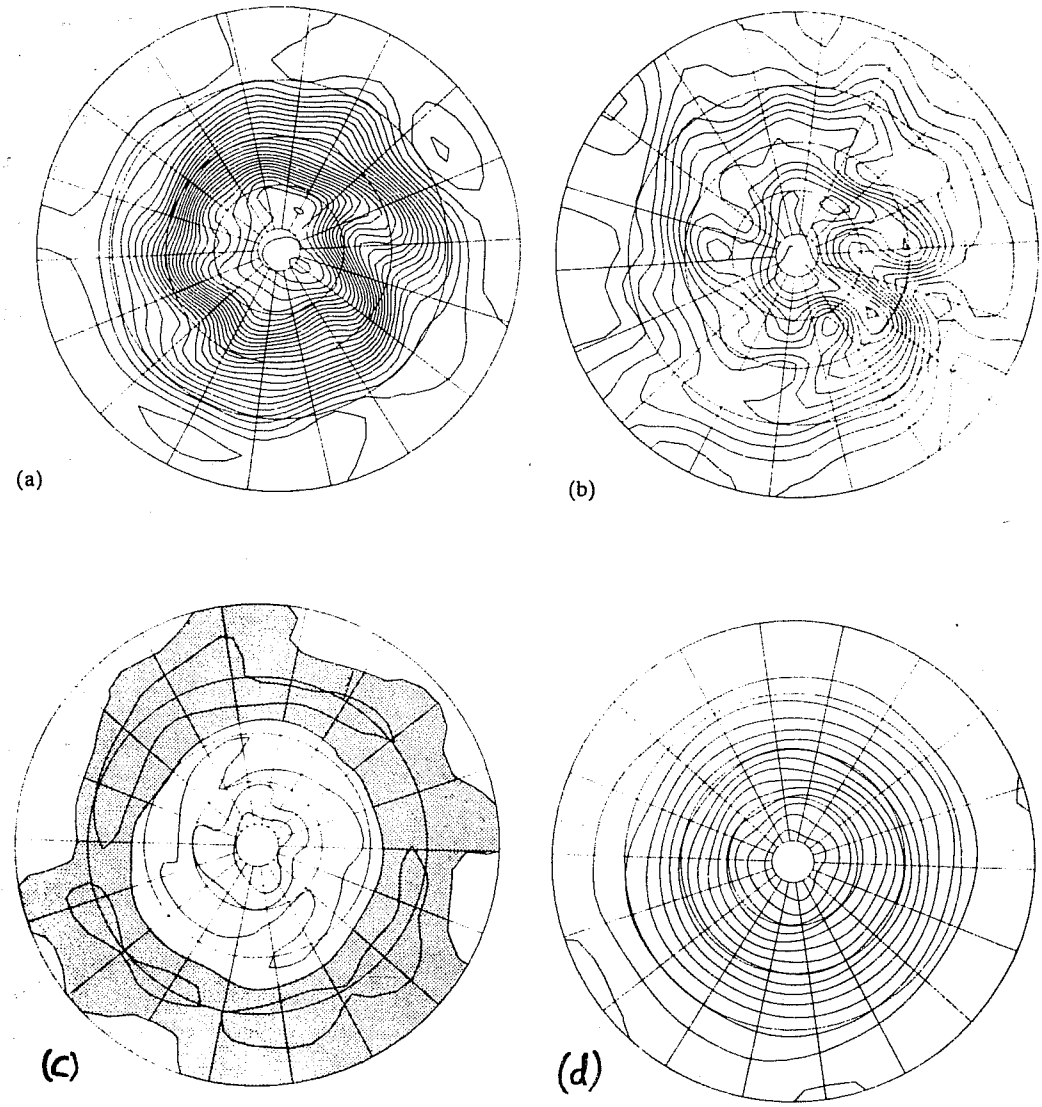

Fig. 6. Instantaneous flow in the simple model for extreme values of the drag.

(a) High drag case, stream function at $\sigma = 0.3$ (contour interval 1.5×10^7 m^2 s^{-1}).

(b) High drag case, temperature at $\sigma = 0.9$ (contour interval 4K).

(c) Low drag case, relative vorticity at $\sigma = 0.3$ (contour interval 7.29×10^{-6} s^{-1}).

(d) Low drag case, temperature at $\sigma = 0.9$ (contour interval 4K).

A dramatic change resulted from the reduction of the surface drag. The case when τ_D was 250 days, a very large value, is presented in this paper. Eddy activity is almost entirely absent and so the eddy transports are negligible. The temperature field differs only very slightly from the "radiative equilibrium" state, and the mean meridional circulation is weak. During the initial, transient phase of the integration, there was a period of 30-50 days when baroclinic activity was vigorous. During that phase, the temperature gradients were eroded and a strong barotropic component to the zonal flow developed. Once the eddies died away, the temperature relaxed back to the radiative equilibrium distribution on the 15 day radiative timescale, but the surface wind remained strong with only very limited dissipation by the small drag term. The final steady state contains a very low level of eddy activity which balances the slight dissipation of the surface wind.

The climate of the model changed smoothly between these extremes for intermediate values of the drag. The variation is most conveniently summarized in the change of the global energetics, shows in Fig. 7. As the drag is reduced, KE falls, rapidly at first, then more slowly. At the same time, AZ increases consistently with the reduced heat fluxes, and KZ increases as the surface winds build up at low drags. The baroclinic conversion CE is always in the direction of increasing KE, but becomes smaller as the drag is reduced. CK is related to the momentum fluxes and acts to increase KZ at the expense of KE. It increased slightly as drag was reduced at large drags but decreased steadily for lower values of drag, when the eddy activity became very small.

At first sight, these results appear paradoxical. The low drag runs have large baroclinicity and little stabilizing influence from dissipation. Yet they are nearly stable to baroclinic disturbances. The high drag runs are baroclinically very active, but have weaker temperature gradients, larger static stability and the large drag would be expected to reduce baroclinic growth rates. The usual criteria for baroclinic instability (Charney and Stern, 1962; see also the summary in James and Hoskins, 1985) are strengthened by the large temperature gradients and smaller static stability of the lower drag runs. The paradox is of course illusory; the Charney-Stern criterion is merely a <u>necessary</u> condition for instability which states that

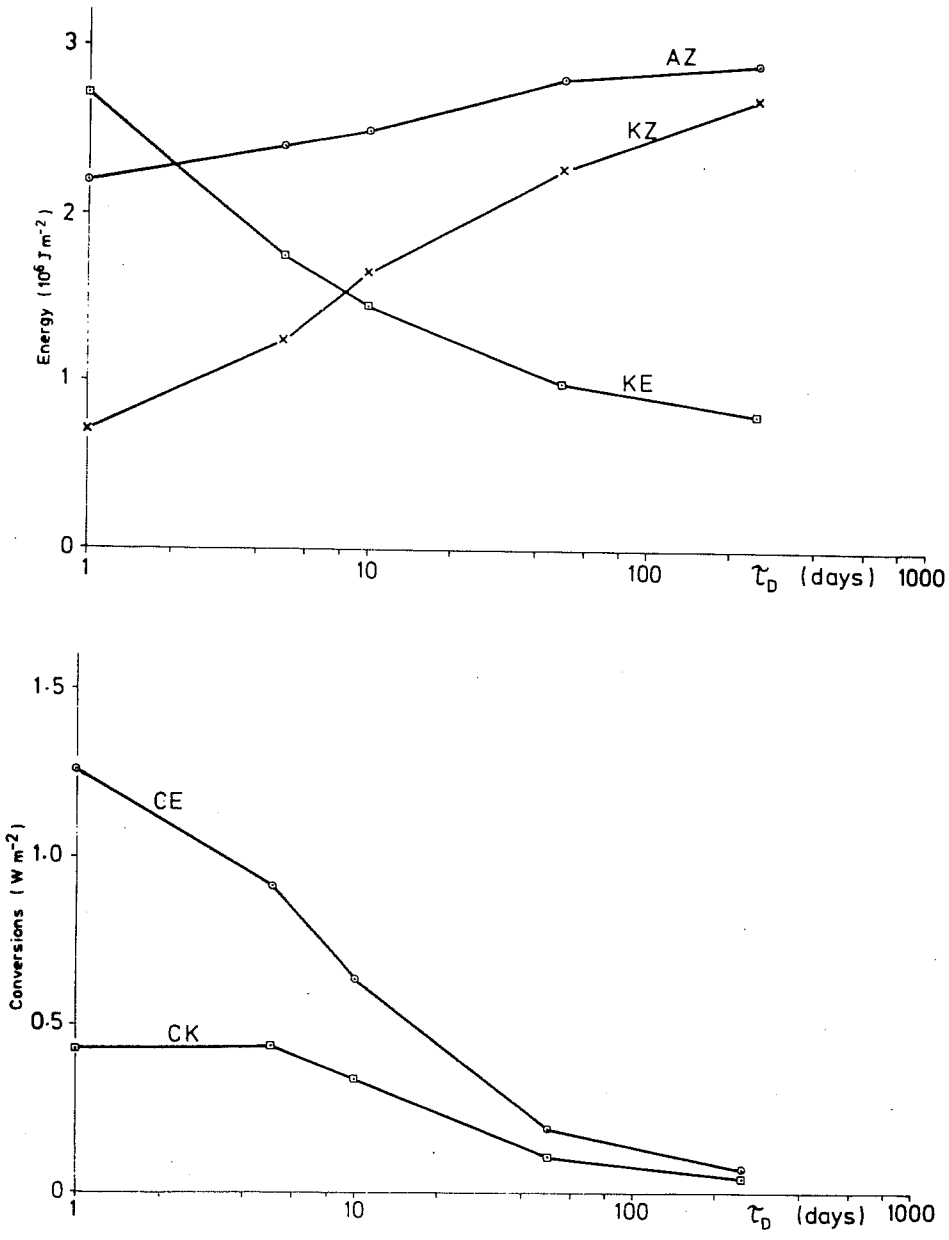

Fig. 7. Variation of energy and conversions in the simple model as a function of the drag timescale.

available potential energy exists. There is no guarantee that it is possible to sustain disturbances whose structure can exploit that available potential energy for substantial periods in any given flow. The classical model of baroclinic instability (those due to Eady, Charney and the two layer system) assume a basic flow with shear in the vertical but which is uniform in the horizontal. We may suspect that the greater baroclinic stability in the low drag runs reported in James and Gray (1986) is associated with the large horizontal shears which developed. I shall demonstrate that this was indeed the case in the next section.

4. BAROCLINIC INSTABILITY WITH HORIZONTAL SHEAR

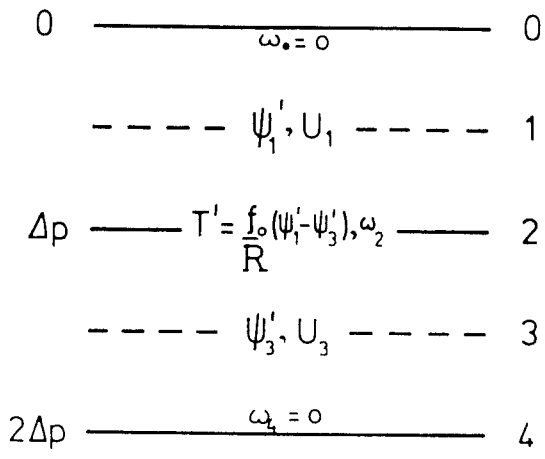

Fig. 8. Illustrating the nomenclature used to describe the two level model discussed in the text.

For simplicity, consider the two layer quasi-geostrophic system, linearized about zonal flow. This is described by the equations:

$$\frac{\partial}{\partial t} q_i' + ik\left[U_i q_i' + \left(\beta - \partial^2 U_i/\partial y^2\right)\psi_i' + \left(2 - i\right)\left(U_2 - U_1\right)\psi_i'\right] = 0 \quad (4)$$

where

$$q_i' = \frac{\partial^2}{\partial y^2}\psi_i' - \left(K_R^2 + k^2\right)\psi_i' + K_R^2\ \psi_{4-i}' \quad , \quad i = 1,3. \quad (5)$$

where q_i' is the perturbation potential vorticity, ψ_i' is the geostrophic stream function, and i = 1 refers to the upper level, i = 3 to the lower level. The parameter $K_R = (f/\sigma\Delta p)$, σ being a static stability parameter. Fig. 8 illustrates the notation. The mean zonal wind is given by

$$U_1 = \Delta U + Ay \quad , \quad U_3 = Ay \quad (6)$$

and for simplicity $\beta = 0$. A measures the horizontal, barotropic shear to which the growing instabilities are subject. Note that the addition of this simple linear horizontal shear does not modify the potential vorticity gradients or the vertical shear. It therefore has no impact on the criteria for the existence of baroclinic instability. James (1987) has solved the eigenvalue problem associated with Eqs. (4) and (5) to determine the structure and growth rates of unstable normal mode disturbances for various values of A. K_R was taken to be 500 km and ΔU to be 20 m s^{-1} for the calculations described here.

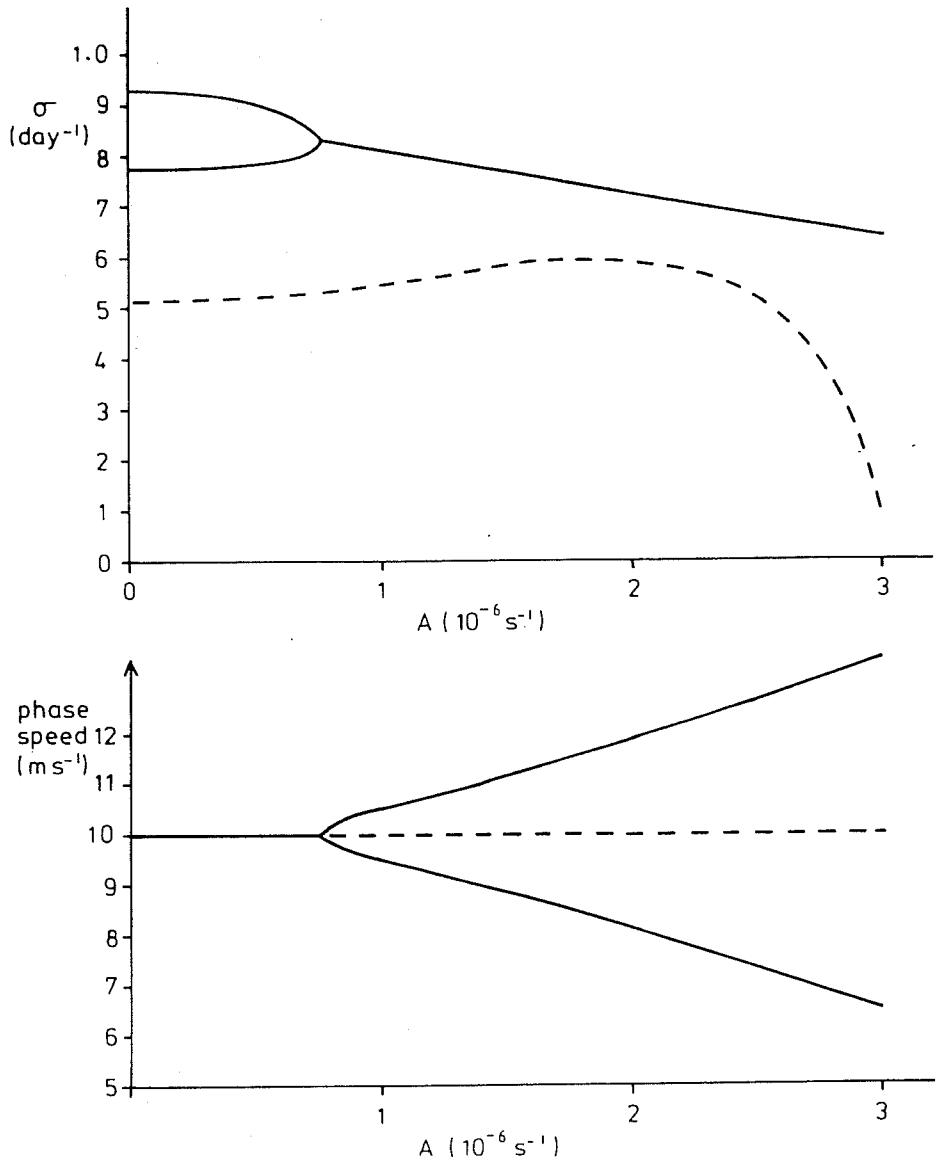

Fig. 9. Showing (a) growth rate and (b) phase speed of baroclinically unstable wavenumber 7 normal modes in the two layer system as a function of the horizontal shear A.

The results for A = 0 are of course the analytic text-book results, showing a most unstable mode with a growth rate of 0.92 day^{-1} and a phase speed of 10 m s^{-1} (i.e., the mean flow speed along the channel). As A approaches 0.7×10^{-6} s^{-1}, the growth rate of the most unstable mode drops slightly, while that of the next mode increases a little. The phase speed remains constant, at 10 m s^{-1}. But for larger values of A, the growth rates are reduced further and the phase speeds begin to change.

These changes are related to changes in the structure of the unstable normal modes, illustrated in Fig. 10. The essential effect of the shear is to

confine the normal modes in the meridional direction. At first, they are confined to the centre of the channel, with symmetry with respect to reversal of y and p; as might be expected from symmetry considerations, the steering level passes through the centre of the channel and so the phase speed is 10 m s^{-1}. For A in excess of 0.7×10^{-6} s^{-1}, the mode splits into two modes, each with the same growth rate, but confined to opposite sides of the channel. Their phase speed reflects the mean flow speed in that region. Fig. 10 shows just the modes on the southern boundary of the channel. The steering level intersects the lower layer at the point

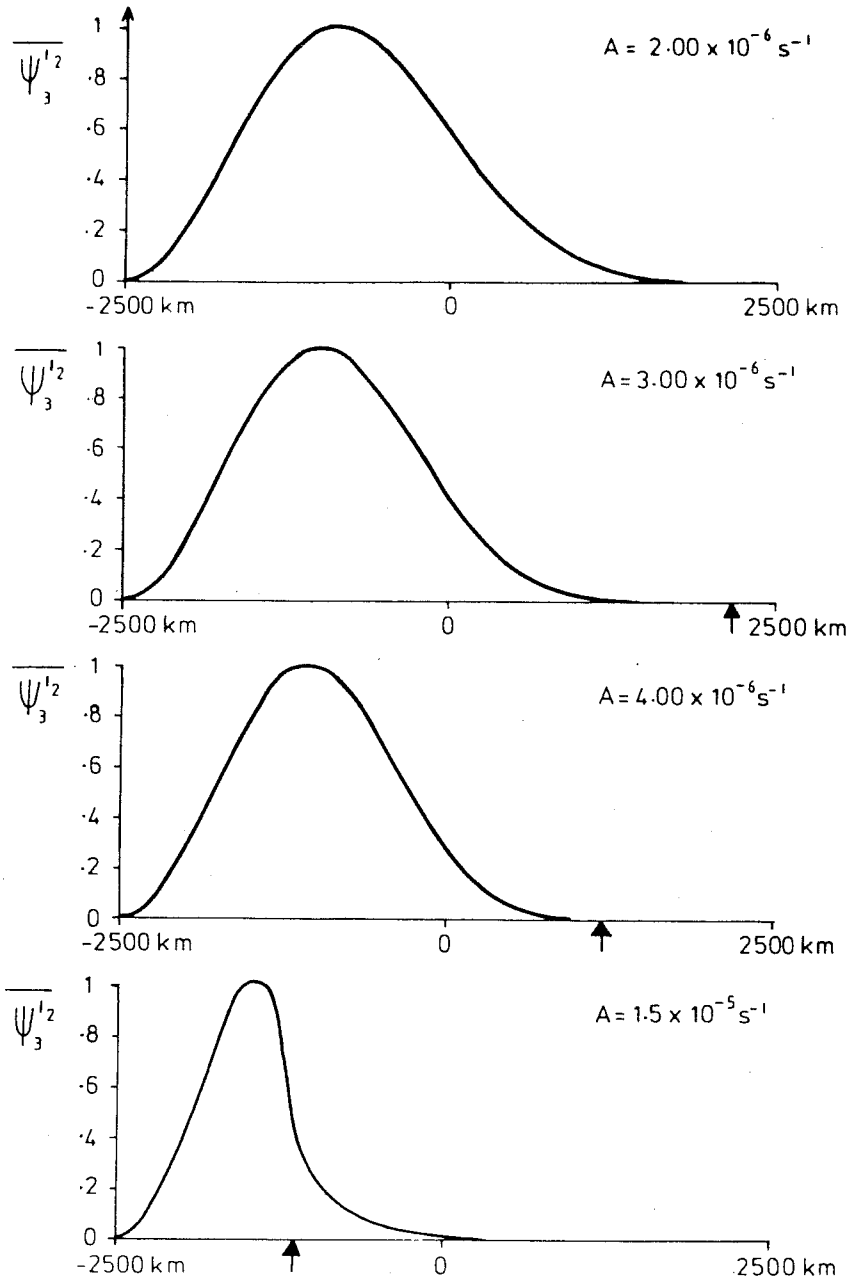

Fig. 10. Showing the streamfunction amplitude of the most unstable wavenumber 7 normal mode for various values of A.

indicated by the arrow; note that the mode has significant amplitude only at those latitudes where the steering level exists within the fluid.

It must be emphasised that the shears employed in the previous runs are not unrealistically large, yet growth rates are reduced by a considerable factor. Even larger reductions, or indeed complete stabilization, are possible if dissipative (Ekman pumping) terms are included. Furthermore, similar growth rate reductions are found when more realistic background states, including jets in which the artifical sidewall boundary conditions play no role, are used in place of the basic flow defined by Eq. (6). Details are given in James (1987).

Let us now return to the simulations carried out with different values of drag described in the previous section. The initial state in both the high and low drag cases was identical and had zero surface wind. As the eddies developed and the flow settled down towards its climatological state, the temperature gradients were weakened and a barotropic component of the wind was introduced. In the high drag runs, this surface wind was small while the reduction of the temperature gradients was substantial. In the low drag case, the climatological temperature field was rather close to the radiative equilibrium state, while the surface wind was very large. Following the ideas introduced by these calculations for the two layer system, let us see whether the strong horizontal shear in the low drag climate plays a significant role in reducing the level of eddy activity.

The linear normal modes which can grow on these zonal flows can be calculated using the spectral model to generate the elements of a large matrix of complex elements which describe the tendencies due to changes in each independent degree of freedom of the system. The eigenvectors of this matrix are the normal modes of the discretised model system and can be determined using standard methods. The solid curves in Fig. 11 show the growth of the most unstable normal mode as a function of zonal wavenumber for the initial states. These are identical, but of course the higher drag in the one case reduces the growth rates of the normal modes slightly. The dashed curves show the growth rates for the climatological state. In the low drag run, they are over an order of magnitude smaller than in the initial state. It is worth remarking that the most unstable normal modes shown in that curve had very noisy and small scale meridional structure;

it seems unlikely that they could grow to very large amplitude. The reduction of growth rates for the high drag case is less, consistent with the larger eddy activity in that run.

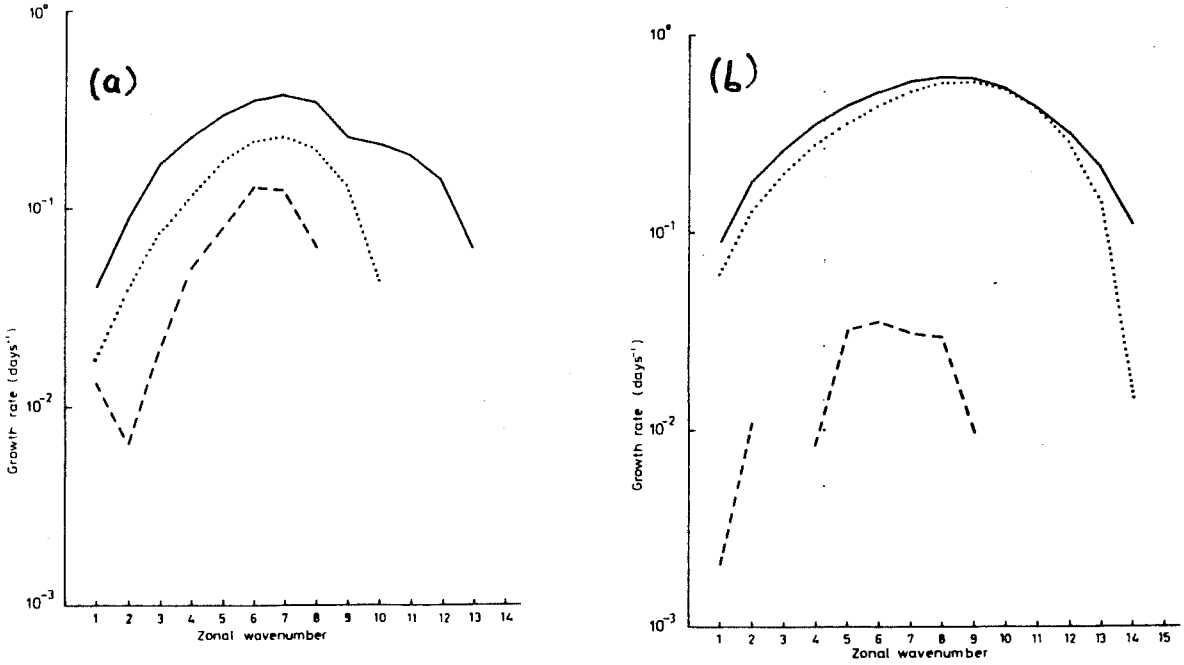

Fig. 11. Growth rate versus zonal wavenumber in (a) the high drag, (b) the low drag runs. The solid curve is for the initial (radiative equation) basic state and the dashed curve the climatological mean. The dotted curves are for the climatological state but with a barotropic correction setting the surface winds to zero.

Finally, the dotted curves show the importance of the horizontal shears in controlling the level of eddy instability in the two runs. The climatological state for both runs was modified by subtracting the surface wind from every level. This barotropic correction leaves the temperature field unaltered and has very little effect on the potential vorticity gradients (which, for large Richardson number, are dominated by vertical, not horizontal, curvature of the zonal wind field). But clearly the effect on the growth rates is dramatic. In the low drag run, growth rates return essentially to the values for the initial state, showing that horizontal shear was the dominant mechanism in stabilizing the flow. Even in the high drag case about half the stabilization must be attributed to the rather weak horizontal barotropic shears developed in the final climatology. It turns out that this stabilization by horizontal shear is a fundamental effect in the life cycles of baroclinic waves (described by, for example, Simmons and Hoskins, 1978); James (1987) showed that the collapse of the instability in the latter part of the life cycle is due principally to the development

of horizontal shear, not to the source of available potential energy becoming exhausted.

5. THE BAROTROPIC GOVERNOR

The momentum fluxes associated with unstable baroclinic waves are rather difficult to discuss in general terms, and require a full solution of the (generally inseparable) eigenvector problem for their determination. But certain general principles appear to hold, at least in a wide range of examples; these are conveniently summarized in terms of the familiar Lorenz energy cycle discussed in the first section.

The rate of change of total eddy energy can be written:

$$\frac{dE}{dt} = - \left\langle [u^*v^*][u]_y - \frac{R^*f}{\Delta p\ \sigma^2} [v^*T^*][T]_y \right\rangle \qquad (7)$$

where the quasi-geostrophic framework has been used. In balanced flow, horizontal temperature fluxes invariably imply vertical temperature fluxes, so that I do not distinguish between the eddy available potential and kinetic energies; E refers simply to the total eddy energy. The angle brackets denote integration over the entire atmosphere and other notation has its usual meaning. Baroclinic instability is characterised by a negative correlation between the poleward temperature flux and the temperature gradient, that is, by down gradient temperature fluxes. In contrast, the momentum fluxes tend to be positively correlated with the horizontal shears and so lead to a reduction in the rate of generation of eddy energy. This direct effect of momentum fluxes is not especially large, and is not the dominant process in accounting for the reduced growth rate in a horizontally sheared environment. Rather, the structural changes of the normal mode restrict the correlation between $[v^*T^*]$ and $[T]_y$ and so the growth rates are reduced.

However, the direction of the momentum fluxes is almost always such as to increase the horizontal shears in the vicinity of the developing disturbances. In the classical lifecycle experiments of Simmons and Hoskins (1978) a barotropic jet was added to the zonal flow as a result of the momentum fluxes associated with the evolving baroclinic disturbance. Similarly, zonal mean circulation statistics for the observed atmospheric circulation show that the transient momentum fluxes act to generate westerlies somewhat poleward of the latitudes of maximum eddy activity (see, for

example, Oort (1983)). Thus the depression belt tends to lie in a region of horizontal shear generated by the eddies themselves. James and Gray (1986) suggested that this effect is sufficiently ubiquitous in realistic situations to be idealised in the modified picture of the Lorenz energy cycle, shown in Fig. 12. We called this picture the "barotropic governor"; the eddies generate momentum fluxes which increase the horizontal shear in regions where baroclinic activity is concentrated. This shear tends to reduce the baroclinic activity until some climatological balance between the generation of energy by differential heating and its dissipation, essentially by friction, is achieved. The feedback between the horizontal shear and the generation of eddy energy governs the level of eddy energy and eventually the zonal kinetic energy, and plays a central role in ensuring that the sources and sinks remain in balance.

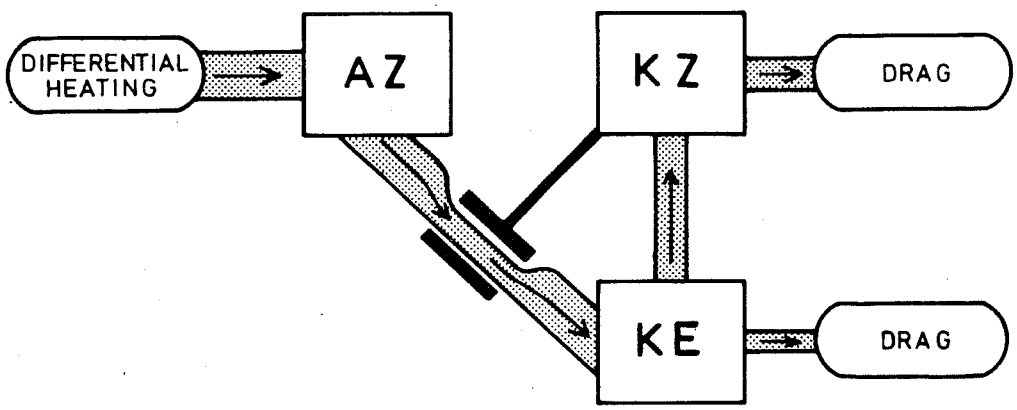

Fig. 12. Illustrating the "barotropic governor".

5. IMPLICATIONS

Fig. 13 shows the current ECMWF wind error for 10-day forecasts during the Northern Hemisphere winter. The error is broadly barotropic in structure, with westerly errors at 50°N and easterly errors around 30°N. The typical horizontal shear is as much as 2×10^{-6} s^{-1}, quite enough to cause significant reductions in the baroclinic growth rates and hence to the poleward temperature fluxes, according to the results discussed in the previous section.

It can easily be imagined that as a consequence of the error in the westerly

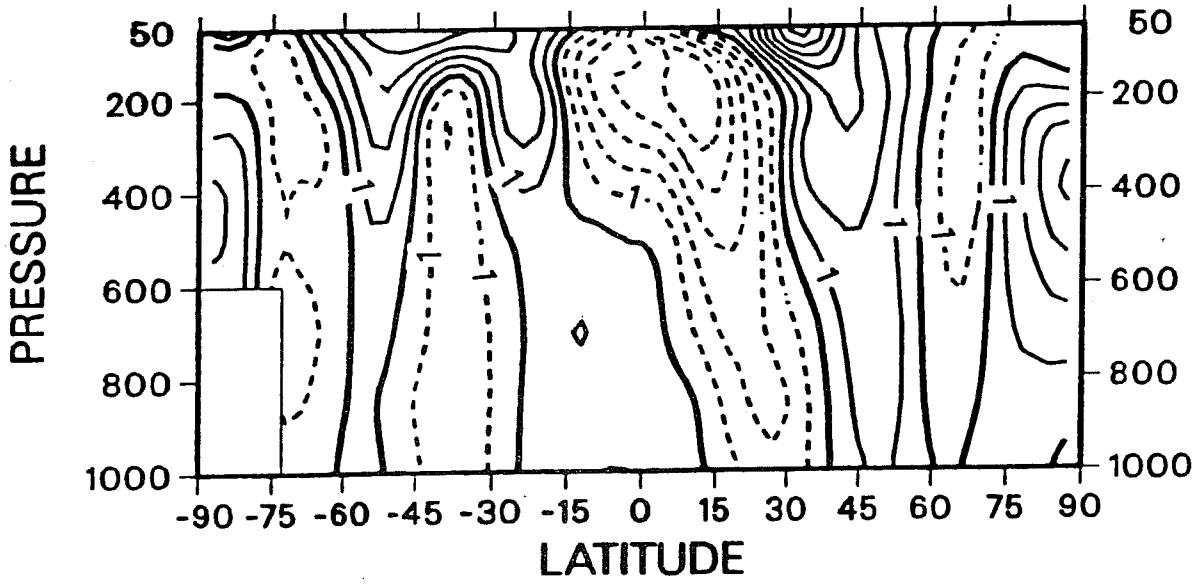

Fig. 13. The error in the 10-day forecast zonal wind at ECMWF for the period December 1986 to February 1987. Contour interval 1 m s^{-1}.

wind, there will be a tendency by day 10 of the forecast for the eddy temperature flux to be too small and consequently for the warming of the tropics to be exacerbated. The point is a fairly straightforward one which nevertheless needs emphasising periodically. An error in, for example, the tropic temperature field need not be the direct consequence of an error in the tropical diabatic heating. It could be the consequence of a global deficiency in the atmospheric dynamics which might have its origin in the parametrization of mechanical dissipation in midlatitudes. The barotropic governor mechanism suggests a particular route whereby this type of interaction might take place. It seems to have received rather little attention in the literature previously, but in view of the apparently ubiquitous nature of the governor mechanism (which is related to older ideas of "negative viscosity"), I suggest it could well be important when considering systematic errors in forecasting and general circulation models.

Most of the discussion here has been in the context of the zonal mean flow. The barotropic governor also apparently plays a role in the zonal variations of the flow, and in determining the length of the major storm tracks. The storm tracks of the northern hemisphere start over the western side of the major ocean basins. The change of surface drag from land to ocean may well play a crucial role in the extra generation of eddy kinetic energy in such regions (Paul Valdes, personal communication). As depressions go through

their lifecycle, a developing surface wind is generated which in turn helps to account for the greater stability of the flow over the eastern part of the ocean basin. A failure to dissipate these westerlies adequately could distort the length and shape of the storm track and lead to important systematic errors in the distribution of transient eddy activity by the end of a 10-day forecast.

The ideas discussed in this section are of course speculative. But they could easily be investigated using the kinds of simplified models and techniques mentioned in the previous sections. It is hoped to extend the work described in James and Gray (1986) and James (1987) in these directions.

References

Charney, J.G. and Stern, M., 1962: On the stability of internal baroclinic jets in a rotating atmosphere. J. Atmos. Sci., 19, 159-172.

Held, I.M. and Hou, A.Y., 1980: Nonlinear axially symmetric circulations in a nearly inviscid atmosphere. J. Atmos. Sci., 37, 515-533.

James, I.N., 1987: Suppression of baroclinic instability in horizontally sheared flows. J. Atmos. Sci., 44, 3710-3720

James, I.N. and Gray, L.J., 1986: Concerning the effect of surface drag on the circulation of a baroclinic planetary atmosphere. Q.J. Roy. Met. Soc., 112, 1231-1250

James, I.N. and Hoskins, B.J., 1985: Some comparisons of atmospheric and internal baroclinic instability. J. Atmos. Sci., 42, 2142-2155.

Kung, E.C. and Tanaka, H., 1983: Energetic analysis of the global circulation during the special observation periods of FGGE. J. Atmos. Sci., 40, 2575-2592.

Oort, A.H., 1983: Global atmospheric circulation statistics, 1958-1973. NOAA Professional Paper 14 (U.S. Department of Commerce).

Simmons, A.J. and Hoskins, B.J., 1978: The life cycles of some nonlinear baroclinic waves. J. Atmos. Sci., 35, 414-432.