Gregarious Tropical Convection

BRIAN E. MAPES

Department of Atmospheric Sciences, University of Washington, Seattle, Washington

(Manuscript received 24 March 1992, in final form 8 October 1992)

ABSTRACT

A heat source with a vertical profile like that of observed tropical mesoscale convective systems (MCSs) is shown to cause, through inviscid gravity wave dynamics, upward displacement at low levels in a mesoscale region surrounding the heating. Typical values are ~10%–30% area contraction at the surface everywhere within 270 km of the heating 6 h after it starts. As a result, conditions near an existing MCS (but beyond the area of MCS outflow) become more favorable for the development of additional convection. This theory predicts that cloud clusters should be gregarious. Infrared satellite imagery confirms that almost half of the cold cloudiness observed in a month over the oceanic warm pool region was contributed by just 14 objectively defined multiday "superclusters."

1. Introduction

Precipitating deep convection in the tropical troposphere depends crucially on the supply of humid air in the lowest few hundred meters. Any circulation that deepens the humid boundary layer by horizontal area contraction thus tends to favor deep convection, while subsidence (low-level divergence) tends to suppress convection. The purpose of this paper is to demonstrate how observed tropical convective heating causes low-level area contraction in its near vicinity, through inviscid gravity wave dynamics. As a result, this theory predicts that tropical convection should be gregarious. Satellite data confirm that cloud clusters over the oceanic warm pool do indeed occur in inseparably tangled "superclusters."

In this paper, the response of a stratified fluid to a single spatially and temporally localized heating event, with vertical structure like that observed in tropical mesoscale convective systems (MCSs), is considered. Upward motion occurs in the heated region of the fluid, such that adiabatic cooling nearly compensates the heating. The heating is realized as warming elsewhere in the stratified atmosphere, wherever the adiabatic compensating subsidence occurs. In recent years, a clear dynamical picture has emerged of when, where, and how the compensating subsidence associated with convective ascent takes place (Bretherton and Smolarkiewicz 1989, hereafter BS89; Nicholls et al. 1991, hereafter NPC).

The subsidence takes place in gravity wave–like pulses that travel away from the heat source. These pulses undergo dispersion with respect to vertical wavelength, the deeper vertical structures traveling faster. The notion that the response of a stratified fluid to imposed heating can be resolved into a spectrum of vertical modes is more than just a mathematical expedient: the various vertical modes really do separate themselves out with distance from the heating (as shown below in a laboratory fluid). One consequence of this dispersion is that adiabatic upward displacements at low levels are generated in the fluid by a positive-only heat source that resembles tropical MCS heating.

The idea of a positive feedback between deep convection and adiabatic vertical motions induced in its stratified environment is not new. Most studies of this question, however, have focused on the propagation and longevity of large, intense midlatitude MCSs (e.g., Hoxit et al. 1976; Raymond 1987; Cram et al. 1992). These intense midlatitude events are generally observed near frontal zones, in areas of significant quasigeostrophic lifting, and in strong vertical wind shears. As a result, the role of convective heating in the intensification and propagation of these systems has remained somewhat obscure. Nevertheless, the feasibility of a "wave-CISK" or "forced gravity wave" (Raymond 1987) model of intense convective systems has been reasonably well established by analytical and numerical studies (reviewed in Cram et al. 1992). The present results indicate that even when convective heating does not produce localized gravity wave lifting intense enough to *trigger* additional convection, it can still act to *favor* additional convection in its mesoscale vicinity.

The dispersion of vertical modes excited by heating in a stratified fluid is reviewed and demonstrated in section 3. The response of the atmosphere to an idealized but realistic positive-only MCS heating profile,

Corresponding author address: Dr. Brian E. Mapes, Department of Atmospheric Sciences, AK-40, University of Washington, Seattle, WA 98195.

consisting of just two vertical modes, is considered in section 4. Section 5 contains satellite observations showing that the deep convection over the oceanic warm pool, in which the heating profile measurements were made, is indeed gregarious: it occurs in inseparably tangled multiday "superclusters." A summary and discussion follow in section 6.

2. ICAPE and inhibition: The importance of low-level area contraction

A useful energetic measure of the viability of convection is convective available potential energy, or CAPE ("positive area" on a thermodynamic chart). CAPE is a parcel property (units: J kg^{-1}) that depends on the parcel's equivalent potential temperature θ_e and on the ambient density profile through which the parcel is assumed to rise. CAPE is very sensitive to parcel θ_e: the CAPE of a parcel with θ_e = 360 K, rising in an atmosphere characterized by a mean sounding from the tropical warm pool region, is 2370 J kg^{-1}, while for a 355 K parcel it is 1180 J kg^{-1} and for 350 K a mere 260 J kg^{-1}. Observed variations in upper-tropospheric temperature soundings alter these numbers by only about ±150 J kg^{-1} (Figs. 11 and 13 of Mapes and Houze 1992).

In the limited range between the lowest θ_e capable of positive buoyancy and the highest observed, the linear approximation CAPE = 260 + 211(θ_e − 350) is quite adequate. With this formula, profiles of $\theta_e(p)$ from soundings through the boundary layer can be converted to profiles of CAPE(p). In general, θ_e and therefore parcel CAPE decrease monotonically with height above the surface (Fig. 12 of Mapes and Houze 1992), from CAPE values of 1000–2000 J kg^{-1} at the surface to zero for parcels of the ambient air above about 900 mb.

Consider now an integrated CAPE ("ICAPE") or, in discrete form, the mass-weighted sum of CAPE for all low-level parcels that have positive CAPE:

$$\text{ICAPE} = \sum_{\text{CAPE}>0} \text{CAPE}(p)\frac{\delta p}{g}. \qquad (1)$$

The units of ICAPE are joule per square meter, so it is a property of an atmospheric column rather than of an arbitrarily chosen parcel. Boundary-layer area contraction increases ICAPE by a factor of $(1-\Delta)^{-1} \approx 1+\Delta$ for small Δ, where $\Delta = \delta A/A$ is the area contraction factor. Note that vertical resolution better than the "standard levels" of 1000 mb, 925 mb, etc., may be needed to evaluate ICAPE accurately.

In addition to the ICAPE increase, boundary-layer convergence also causes adiabatic cooling of the ambient air at low levels. Parcels rising through this cooler environment reach their level of free convection (LFC) at a lower altitude. In other words, boundary-layer convergence decreases the convective inhibition ("negative area" on a thermodynamic chart), so that a lesser amount of lifting by intense, localized triggering mechanisms is required to set off deep convection. This effect is probably more relevant to the development of new convection than the increase of ICAPE, but is not as easily quantified. In any case, there can be little doubt about the net result of these two effects: *boundary-layer contraction favors additional deep convection.*

3. Dispersion of vertical structures in a heated stratified fluid

The dispersion relation for internal gravity waves is particularly simple in the hydrostatic limit (Raymond 1983; BS89; NPC). In that case, the horizontal phase and group velocities are both given by

$$c = Nm^{-1}, \qquad (2)$$

where N is the buoyancy frequency and m is the vertical wavenumber. This low frequency or hydrostatic limit is adequate for the quasi-horizontal, nonoscillatory flows into and out of a heated region, though not for the highly tilted, upward-propagating stratospheric gravity waves that convection also excites. Equation (2) predicts that deeper flow structures will propagate out into a fluid from a heat source more rapidly than will shallower flow structures. This dispersion effect can be demonstrated in a laboratory fluid.

Figure 1 shows the response of a continuously stratified fluid (salt water) to a localized buoyancy forcing (a salty convective plume). The photographs have been inverted, as is traditional in plume literature, so the plume appears to go upward like atmospheric convection. The plume in Fig. 1 is not meant to realistically represent atmospheric convection. It is simply one example of a strong, rapid buoyancy forcing with complex vertical structure. The initially vertical dye lines, illuminated by a strobe light at time intervals of half a buoyancy period, register the gradual arrival in the far-field fluid of flow structures of ever smaller vertical scale.

The initial dye-line displacements show only the deepest component of the heating: the gravest mode of the layer, half a wave spanning the depth. As time goes on, shallower flow structures gradually propagate out from the forced region. In the long-time limit, all the dye lines attain the same shape (in the absence of rotation effects). The irreversible horizontal flows, associated with the vertical motions in the plume, are confined behind boundaries that propagate out from the plumelike gravity waves—dispersively—into the ambient fluid.

Emanuel (1983) and NPC have described the transient response to convective heating as consisting of "gravity waves." While this characterization is not incorrect, it obscures important aspects of the response. Linear gravity waves propagate vertically and, averaged over a wavelength, do not transport heat or modify the mean state of the fluid through which they propagate

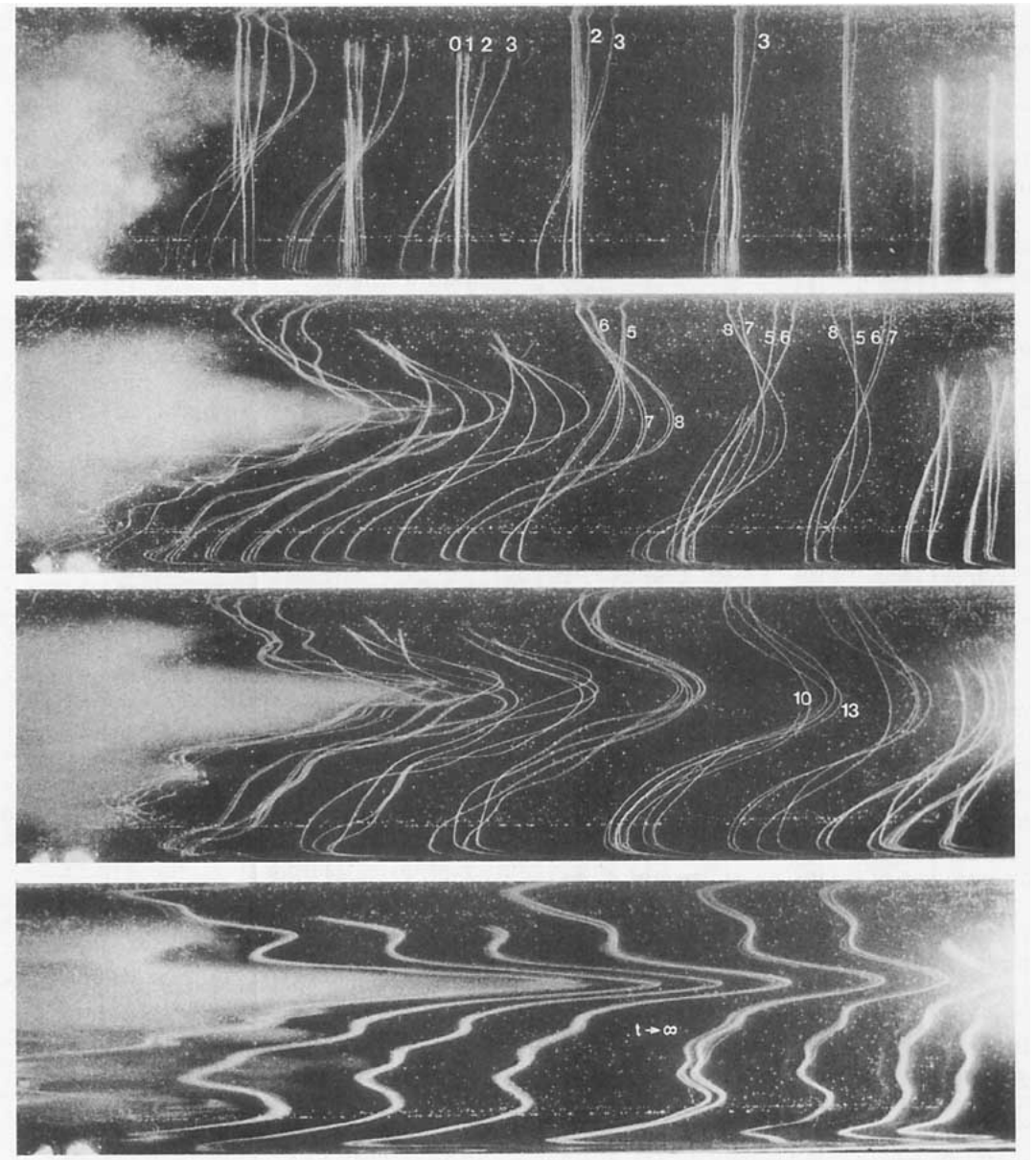

FIG. 1. Strobe photographs of the successive positions of initially vertical dye lines in continuously stratified salt water. Numbers indicate the times, in units of π/N (half buoyancy periods), at which the line positions were recorded, after the pulse of "heating" (salt buoyancy forcing). Photographs are inverted.

(Emanuel 1983). But in addition to the extremely complex, oscillatory gravity wave field excited by convection, there are important irreversible (nonoscillatory) flows that reflect the net heating. The fluid shown in Fig. 1 contained numerous oscillatory motions, almost beyond description in their complexity. But more importantly, the dye lines moved in response to the *net* buoyancy forcing, and never returned to their original positions.

The pulse launched by the start-up of a heat source of a single vertical wavelength (see Fig. 6 of BS89; Figs. 3, 6 of NPC) propagates by the same mechanism as a gravity wave: at its leading edge, the pressure gradient force causes some subsidence, leading to positive buoyancy, which in turn gives rise to a downward pressure force in the adjacent fluid, and so on. This mechanism is encapsulated in the gravity wave dispersion relation, and therefore gravity wave theory predicts quite well the propagation speed of the pulse. But the pulse is not an ordinary wave, in the sense that it does not have periodic structure in time or space. It is somewhat analogous to a tidal bore in water, which propagates at the speed of an ordinary surface water wave, but which irrevocably changes the depth of the water

as it passes. In reference to this analogy, the heating-induced horizontally propagating gravity wave pulses whose passage causes net vertical displacements in a stratified fluid will here be referred to as "buoyancy bores." Note that the use of the word "bore" is not meant to imply that high-amplitude, nonlinear, or turbulent effects are necessarily important to the propagation of the pulses.

The horizontal structure of buoyancy bores comes from a convolution of the horizontal structure of the heat source and the temporal structure of its initiation. Bretherton and Smolarkiewicz derived the bore shapes, in both slab and radial symmetry, for heating with sinusoidal vertical structure and δ-function structure in the horizontal, starting instantly. The response to more complex heating distributions can be constructed by superposition. In slab symmetry, the bore is simply a step function (Fig. 6 of BS89), while in radial geometry the structure is more complex (plotted in Fig. 4a). For a more realistic heating field, distributed in the horizontal and starting gradually, the buoyancy bores are broader in structure, but the vertical displacements affected by their passage are the same.

More important for the present paper, solutions for heating with complex *vertical* structure can be built up by superposition of buoyancy bore solutions for the various vertical modes excited by the heating (see e.g., Fig. 5 of NPC). For simplicity, just two vertical modes, quantized by a rigid lid, are considered below. The adiabatic upward displacements, demonstrated in theoretical models in sections 4a and 4b, are confirmed in section 4c with the use of a fully compressible nonhydrostatic model with no rigid lid.

4. Low-level area contraction forced by heating

Consider for simplicity a heat source that excites just two vertical modes of a stratified fluid: the gravest mode, with a wavelength twice the depth of the heating, and a shallower mode. The simplest choice for this second mode is the one with wavelength equal to the depth of the fluid. Hereafter, the two modes are numbered with indices $l = 1, 2$ (twice the vertical wavenumbers of $1/2$ and 1), respectively. The effect of altering the choice $l = 2$ for the second mode is explored in Eq. (9). The heating profile considered in this section is

$$Q = Q_0[\sin(\pi z/H) - \sin(2\pi z/H)/2], \quad (3)$$

where H is the depth of the fluid and Q_0 is in buoyancy source units. This heating profile (plotted in Fig. 3a) is positive at all altitudes, and is a decent approximation to observed tropical convective heating profiles.

Mature mesoscale convective systems consist of deep convective cells and stratiform precipitation areas that evolved from earlier convective cells. The low-level condensation heating in convective updrafts is partially canceled by evaporative cooling in convective and mesoscale downdrafts, while in the upper troposphere, condensation heating in the mesoscale updraft of stratiform precipitation areas adds to heating in convective updrafts. Radiative heating also tends to be concentrated aloft. The resulting *net* heating profile for the whole system looks like (3): positive at all levels, but with its maximum value in the upper troposphere (Houze 1982, 1989).

Alternatively, and more conveniently, the thermal forcing can be specified as horizontal mass divergence D, under the assumption that adiabatic cooling (N^2w) balances the heating Q in the heated region. Utilizing the mass continuity equation, and substituting Q from (3), it follows that

$$D = \partial(\rho w)/\partial z \approx -\partial/\partial z(\rho Q/N^2) \quad (4a)$$

$$= \pi\rho Q_0/HN^2[-\cos(\pi z/H) + \cos(2\pi z/H)] \quad (4b)$$

$$\equiv D_1 + D_2.$$

The assumption (4a) is well met in the models used in this paper, and apparently in the tropical troposphere as well.

The curves D_1, D_2, and D are plotted in Fig. 2a. Note that the two modes have equal amplitudes in terms of horizontal mass divergence, even though the second-mode heating is only half as large, because of the differentiation in (4a).

Figure 2b shows two measures of mean horizontal mass divergence profiles in Australian tropical MCSs. The solid line comes from composite rawinsonde measurements around the Gulf of Carpentaria, while the dashed line is from airborne Doppler radar measurements within MCSs (see Mapes and Houze 1993a for measurement details). Also shown is the gravest mode of the troposphere, plotted with an amplitude equal to its projection onto the rawinsonde-observed profile. The Doppler measurements represent larger, more vigorous MCSs than the rawinsonde composite measurements. The two datasets differ in several ways, but they agree on one essential point, which they share with the idealized profile D in Fig. 2a: *the lower-tropospheric convergence into these tropical MCSs over the oceanic warm pool peaks significantly above the surface.* It is precisely this feature that leads to the low-level adiabatic upward displacements, which are the subject of this paper.

The measurements in Fig. 2b are averages over entire MCSs, and in the case of the rawinsonde measurements, over the entire life cycle of MCSs. But it should be noted that this elevated peak in the low-level convergence is observed even in the convective subportions of MCSs alone (Mapes and Houze 1993a). Note that the actual difference $D - D_1$ between the observed and gravest-mode profiles in Fig. 2b is dominated by a higher wavenumber that $l = 2$, perhaps $l = 4$ or more.

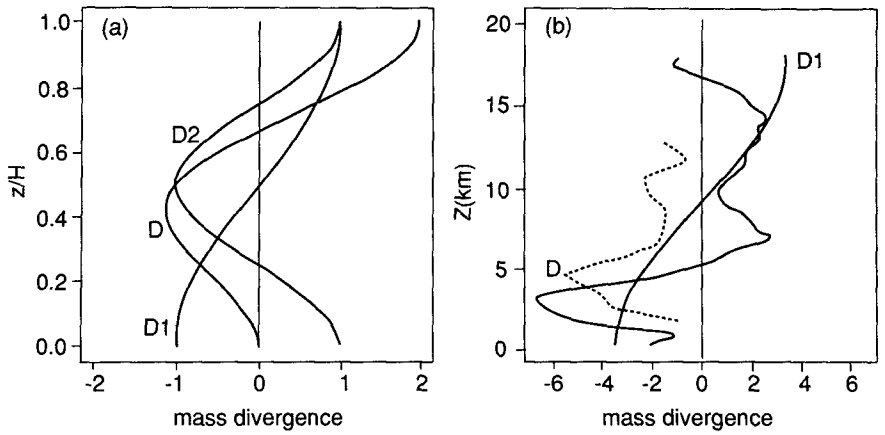

FIG. 2. Profiles of mass divergence (kg m^{-3} s^{-1}): (a) D, D_1, and D_2 for the idealized calculations of section 3a. (b) Observed MCS mass divergences (D) and the gravest-mode component D_1. The solid D line shows mean mass divergence from the Gulf of Carpentaria rawinsonde array, averaged over the whole life cycle of the four MCSs in the Frank and McBride (1989) composite study, multiplied by 10^6. The dashed D line is the mean of all 93 Doppler radar–measured profiles in nine mature MCSs in northern Australia (26 convective area profiles, 37 stratiform area profiles, and 30 intermediate profiles; see MH93a), multiplied by 10^5.

As a result, the actual induced vertical displacements should be more intense and localized [see (9)] than the results of this idealized calculation indicate.

The analytic solutions of BS89 and NPC for the buoyancy field in space and time following the onset of heating with the vertical profile (3) can be evaluated directly to yield vertical displacement δz. Those solutions, for heat sources with both slab and radial symmetry, are given in section 4b. But first, it is worth considering a geometrical line of reasoning about horizontal displacements. This horizontal displacement viewpoint has intuitive appeal, for two reasons. First, as discussed above, tropical MCS heating is expressed in the atmosphere, and measured by humans, as horizontal wind divergence. Profiles of horizontal displacement (as expressed, for example, in the dye lines of Fig. 1) directly reflect features of the horizontal divergence profile. Second, horizontal displacements induce geostrophic currents in a rotating fluid and hence have a special physical importance, as discussed in section 4d.

a. Horizontal displacements of dye lines near a heat source

This argument relies only on the fact that, for a slab-symmetric heat source, all the adiabatic vertical motion occurs in the buoyancy bores, narrow compared to the horizontal distances they traverse, which propagate at speeds given by (2).

Consider the situation, depicted in Fig. 3a, some time τ after the initiation of the heat source Q near $x = 0$, in an otherwise motionless, nonrotating, stratified Boussinesq fluid with a rigid lid. Figure 3a shows horizontal wind profiles, dye lines and their initial positions, and schematic buoyancy bores at time τ. The two bores have propagated away from the heat source, at gravity wave speeds [from (2)] $c_l = NH/\pi l$, out to distances L and $2L$, where $L = c_2\tau$. Behind (to the left of) the bores, horizontal winds associated with each mode (U_1, U_2) blow, satisfying mass continuity. This figure is derived from the results shown in Fig. 6 of NPC.

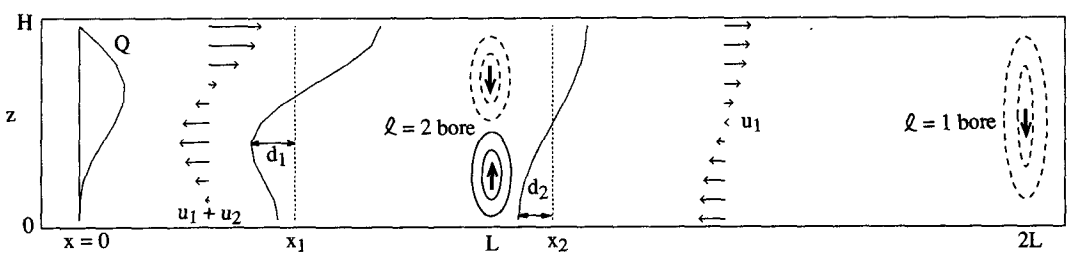

FIG. 3. Schematic of the buoyancy bores, horizontal winds, and integrated displacements of material lines (solid) relative to their initial positions (vertical dashed lines) a time τ after the initiation of the slab-symmetric heat source $Q(z)$ near $x = 0$. The two buoyancy bores have reached L and $2L$, where $L = NH\tau/2\pi$. Note that the material lines have moved closer together at low levels, indicating area contraction remote from the heated region. Adapted from Fig. 5 of NPC.

Note that the dye lines at time τ (solid) are closer together at the surface than they were initially (dashed), indicating area contraction in the unheated fluid between them. Dye line 2 has moved toward the heated region at low levels, driven solely by the gravest-mode component of the heating, while dye line 1, closer to the heating, has moved in a way that more closely matches the actual horizontal divergence profile in the heated region.

It is easy to estimate the displacements, d_1 and d_2, of two initially vertical dye lines initially at positions x_1 and x_2 (see Fig. 3a). Dye line 2 has moved in response to the $l = 1$ wind field, U_1, acting on it for the amount of elapsed time $(2L - x_2)/c_1$ after the $l = 1$ bore passed it; therefore,

$$d_2 = U_1(2L - x_2)/c_1. \quad (5)$$

Similarly, dye line 1 has also moved in response to U_1 since the $l = 1$ bore passed it and, additionally, in response to U_2 since the $l = 2$ bore passed it:

$$d_1 = U_1(2L - x_1)/c_1 + U_2(L - x_1)/c_2. \quad (6)$$

To simplify matters, take $x_2 \sim L$. Then the mean surface contraction $\Delta = \delta A/A$ in the area between the dye lines can be evaluated:

$$\begin{aligned}\Delta_{\text{mean}} &= (d_2 - d_1)/(x_2 - x_1) \\ &= -(U_1/c_1 + U_2/c_2) \\ &= -U_{\max}/c_1. \end{aligned} \quad (7)$$

The second equality follows from the substitution $U_2(z = 0) = -U_1(z = 0) \equiv U_{\max} > 0$, since $c_1 = 2c_2$. At the surface Δ_{mean} is negative, indicating net contraction, and does not depend on x_1, indicating that the low-level area contraction is uniform in the area behind the $l = 2$ bore (all the convergence happened during the passage of the bore).

A numerical evaluation of (7) goes as follows. The amplitude of mass divergence in MCSs (convective plus stratiform regions) is typically 0.5×10^{-4} kg m^{-3} s^{-1} (Fig. 2b, Figs. 6–8 of Mapes and Houze 1993a). Hence, for a slab heating region 200 km wide, $U_{\max} = 5$ m s^{-1} (symmetry about $x = 0$ accounts for a factor of 2). For $c_1 = 50$ m s^{-1}, $\Delta = -U_{\max}/c_1 = -0.1$ in a region that at $\tau = 6$ h extends 540 km from the origin. As a result, ICAPE increases by about 10% in that region (neglecting the effects of upper-tropospheric subsidence—see section 4c).

The choice of $l = 2$ for the second bore was arbitrary, and observations indicate that a higher wavenumber dominates the difference between D and D_1 at low levels (Fig. 2b). Retracing the arguments above, replacing the factors of 2 that resulted from the choice $l = 2$ with l, shows that (7) is a special case of the more general formula:

$$\begin{aligned}\Delta_{\text{mean}} &= U_{\max}[1 - l \max(D_l)/\max(D_1)]/c_1 \\ &= U_{\max}[1 - l^2 \max(Q_l)/\max(Q_1)]/c_1. \end{aligned} \quad (8)$$

The *greater* the vertical wavenumber of the *second bore,* the *greater the low-level area contraction,* though it takes place in a smaller area. For example, for the more realistic case of $l = 4$ rather than $l = 2$ (8-km rather than 16-km wavelength of $D - D_1$, see Fig. 2b), the numerical evaluation above would instead yield $\Delta = -0.3$ in a region 270 km in half-width at $\tau = 6$ h.

The factor of l^2 in (8) indicates that high-vertical-wavenumber components are far more important than their amplitudes in terms of *heating* (or vertical velocity) might lead one to believe. In other words, the traditional presentation of measurements of horizontal divergence in an integrated form, as vertical velocity or apparent heat source, deemphasizes important higher-wavenumber features. For example, the elevated peak in low-level convergence that is central to the present results is inherent, but far from apparent, in the shape of the Australian MCS heating profiles diagnosed by Frank and McBride (1989).

b. *Analytic solutions: Slab versus radial symmetry*

The low-level area contraction factor Δ can also be straightforwardly evaluated from the analytical solutions of BS89 and NPC. In particular, buoyancy b is, outside the heated region, related to vertical displacement by $b = -N^2\delta z$, so that the low-level area contraction factor is

$$\Delta = \delta A/A = -\delta z/z = [b/N^2 z]_{\text{small } z}. \quad (9)$$

For the slab symmetry case, Eq. (13) of NPC can be evaluated, superimposing the two vertical modes in (3), to give the buoyancy b in the region behind the second buoyancy bore (where their bracketed quantity asymptotes to π):

$$b = U_{\max}N[\sin(\pi z/H) - \sin(2\pi z/H)]. \quad (10)$$

Note that b is negative at low levels, indicating that net upward displacement has occurred there. Evaluating (9) with the small z limit ($\sin\theta \sim \theta$) of (10) yields $\Delta = -U_{\max}/c_1$, in agreement with (7) and its generalization (8).

The buoyancy bore solution for radially symmetric heating is given in Eq. (10) of BS89. Figure 4a shows a plot of $\Delta(r)$ at $\tau = 6$ h, calculated by evaluating (9) above, using the BS89 solution for the buoyancy field. The bores appear as unrealistic sharp discontinuities, because all of the heating was assumed to occur at $r = 0$. The response to a more realistic horizontal and temporal heating distribution can be easily built up by superposition. The response is scaled for a heat source characterized by $U_{\max} = 5$ m s^{-1} at 100-km radius, comparable to the scaling below (7).

Note that $\Delta < 0$ everywhere inside the second bore, as found above for the slab-symmetric heat source. But the magnitude is less, as a consequence of the convex radial geometry. Presumably, for concave geometries, in which heating is arranged surrounding an enclosed area, this geometric spreading effect could reverse, and

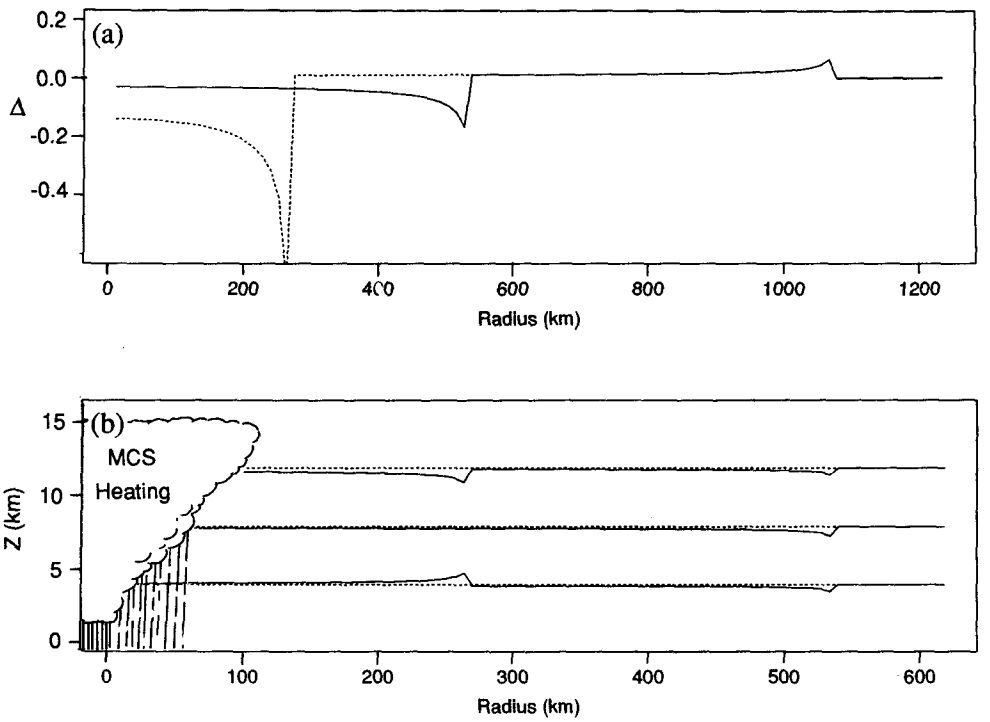

FIG. 4. Linear response of a stratified fluid to a radially symmetric heat source (delta function of radius), scaled so that the divergent wind $U_{max} = 5$ m s^{-1} at a radius of 100 km, comparable to the scaling below (8) for a line source. Calculated from Eqs. (10) and (11) of BS89. Discontinuities result from the artifice of delta-function heating. (a) The area contraction factor Δ at the surface at time $\tau = 6$ h. The solid curve results from the choice $l = 2$ for the second vertical mode, while the dashed line results from the choice $l = 4$. (b) The positions of three material surfaces (solid) initially horizontal (dashed), at time $\tau = 3$ h.

instead focus the buoyancy bore response. This mechanism could explain the frequent observation in satellite pictures that areas that are surrounded by convective cloudiness erupt into activity soon thereafter, too soon for outflows from the initial convection to be responsible.

A schematic view of the vertical displacements of initially horizontal material lines (e.g., isentropes), induced by a radially symmetric two-vertical-mode heat source, is shown in Fig. 4b. Low-level isentropes have been lifted, and upper-level isentropes depressed, near the heat source, while all the isentropes are depressed in the far-field fluid, out to the $l = 1$ bore.

c. *Confirmation by a numerical model with open upper boundary*

The theoretical framework utilized above, while pleasingly simple, has a number of unrealistic aspects. The most serious deficiencies are the hydrostatic and Boussinesq approximations, and the rigid-lid upper boundary condition. These concerns are especially critical in light of new evidence that NPC may contain errors in its plots of the no-lid case (D. Durran and C. Bretherton, personal communication). The Klemp–Wilhelmson compressible model, described in Miller and Durran (1991), was used to simulate the response of an unbounded compressible fluid like the atmosphere to a slab-symmetric heat source with the two-mode vertical structure (3). A smooth horizontal distribution of heating, with half-width 20 km, was used. The stratification was specified as $N = 10^{-2}$ s^{-1} up to 16 km, and $N = 2.6 \times 10^{-2}$ s^{-1} in the stratosphere above that. A radiation upper boundary condition was used. The buoyancy field, after two hours of the positive-only heating, is shown in Fig. 5.

The two separate "buoyancy bores" can be clearly seen near $x = 180$ km and 360 km. Near the heat source at low levels, upward displacements (dashed) have taken place everywhere within $\sim$150 km of the heating, just as in the analytic results above. The $\sim$100-m upward displacement at altitude 1 km indicates a boundary-layer area contraction of $\Delta \sim 10\%$, similar to the analytic results. Elsewhere in the troposphere, buoyancy perturbations are positive. In addition, gravity waves propagate upward into the stratosphere (not shown).

While the low-level upward displacements tend to increase ICAPE, the associated downward displacements in the upper troposphere tend to decrease ICAPE, as any parcels rising through that warmed layer will experience less buoyancy. Two points are important to note. First, the model buoyancy perturbations are overestimates, owing to the fact that the real tropical

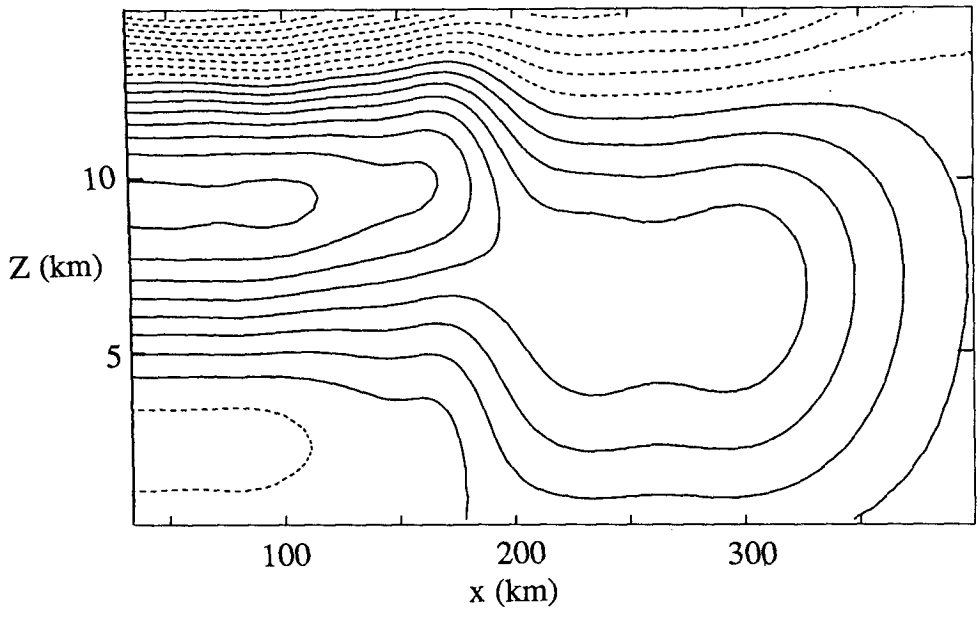

FIG. 5. The buoyancy field (proportional to vertical displacement) computed by a fully compressible numerical model with a radiation upper boundary condition. Positive-only slab-symmetric heating with the vertical profile (3) was applied for 2 h to an initially resting fluid. Negative values (dashed) at low levels indicate adiabatic upward displacements. For a heat source scaled to have midlevel inflow $U_{max} = 5$ m s^{-1}, the contour interval is ~125-m vertical displacement. Zero contour omitted.

upper troposphere has much lower static stability than the lower troposphere (e.g., Fig. 13 of Mapes and Houze 1993a), and to the simplified vertical structure of heating (especially the 16-km wavelength of the second mode). Australian monsoon soundings show that the standard deviation of temperature in the upper troposphere is <1°C (Mapes 1992). Second, any such upper-tropospheric temperature changes can only affect deep convection once it is well under way, and are more likely to modulate the depth to which convection penetrates than its frequency of development or its vigor at low levels. As mentioned in section 2, increases in ICAPE, while easy to quantify, are not the only way in which low-level upward displacements favor the development of deep convection.

d. Discussion and extension to long time scales

It may seem counterintuitive that a heat source, which is positive at all levels and causes mean ascent at all levels in the heated region, also causes a net adiabatic *cooling* in the surrounding fluid. Is not the environmental flow supposed to consist of "compensating subsidence?" One way of looking at this apparent paradox is to say that the subsidence in the deep $l = 1$ bore is in fact "overcompensating subsidence" at low levels, because the heating is small there. The $l = 2$ bore then comes along with low-level ascent that overcompensates in the opposite sense, but in a smaller area. The net effect is that the mean low-level temperature perturbation, averaged over the entire area af-

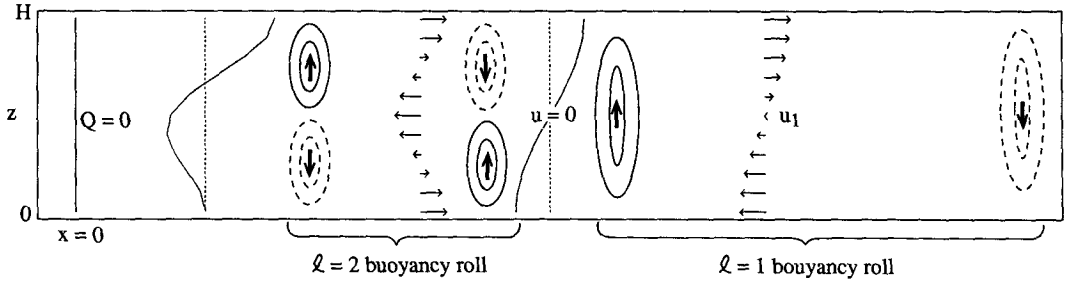

FIG. 6. Two "buoyancy rolls." Schematic of the buoyancy bores, horizontal winds, and horizontal displacements of material lines at a time τ, after the two-mode heat source (3) acted near $x = 0$ from time 0 until time $\tau_2 < \tau/2$. The $l = 1$ "buoyancy roll" (bounded on each end by $l = 1$ buoyancy bores) has completely outrun the $l = 2$ buoyancy roll. Adapted from Figs. 5 and 10 of NPC.

fected by the heat source (i.e., all the way out to the $l = 1$ bore) is consistent with the net heating. The vertical profile of subsidence does match the profile of ascent in the heated region, but the subsidence is not uniformly distributed in the horizontal.

The key feature of the heating profile (3) that leads to inviscid lifting is that the vertical gradient of heating has its maximum value elevated. More specifically, if $D > D_1$ at low levels, then the far-field winds, with a gravest-mode profile like D_1, may converge with the nearer-field winds, which match more closely the actual divergence profile D.

Consider now the (slab symmetric) case in which the heating terminates at some time τ_2. As the heating terminates, new bores are launched, of opposite polarity to those launched when heating began. An increment of heat equal to $Q\tau_2$ has been irreversibly added to the atmosphere, and persists in the form of temperature perturbations within two propagating "buoyancy rolls," one for each of the two vertical modes. Buoyancy rolls are simply roll circulations, with buoyancy bores at each end (see Fig. 10 of NPC). Figure 6 shows the buoyancy bores, winds, and dye-line displacements long after the heating has ended (at a time $\tau > 2\tau_2$), when the two buoyancy rolls have become completely separated. At this time, the area with low-level area contraction $\Delta < 0$ is confined to the inside of the $l = 2$ buoyancy roll.

Buoyancy rolls propagate forever, leaving no trace of warming or vertical displacement near the heated region, in the absence of rotation and dissipation. However, on the earth, these assumptions break down. Rotation and dissipation have a similar effect: they both introduce a time scale, and thereby a distance scale within which the response to a heat source is trapped (e.g., Bretherton 1987).

Consider first the case of rotation. As indicated in Fig. 6, the passage of the $l = 1$ buoyancy roll causes irreversible net *horizontal* displacement of a dye line. In a rotating fluid, geostrophic currents in and out of the page are induced by such displacement. The resulting wind shear requires, in order for thermal wind balance to hold, a horizontal temperature gradient, so the buoyancy rolls must leave behind some of their buoyancy as they propagate. In a sense, then, buoyancy rolls are gradually absorbed by a rotating fluid. The *e*-folding distance for this absorption (analogous to an optical depth for the absorption of light) is the familiar Rossby deformation radius from geostrophic adjustment theory. Rossby deformation radii are given by $\lambda_R = c_l/f$ [Eq. (7.5.4) of Gill 1982]. Note that λ_R is smaller for shallower vertical modes. Numerically, λ_R for the gravest mode $l = 1$ is $\sim$500 km at 40° latitude. The result of heating a stratified fluid on the equatorial β plane is more complex, and is the subject of an upcoming manuscript.

In the related problem with linear damping rather than rotation, the corresponding deformation radii are $c_l\tau_d$, where τ_d is the dissipation time scale. More sophisticated analysis suggests that the appropriate dissipation parameter involves the geometric mean of viscosity and thermal diffusivity (Emanuel 1983). In any case, in either a rotating or dissipative fluid, *deeper structure travels farther, just as it traveled farther in a given amount of time* in the transient dynamics considered above.

An example of the rotational trapping of shallow structure can be seen in the heat-forced Hadley circulation calculations of Schubert et al. (1991). Conveniently, the authors happened to choose a heating profile with its maximum gradient elevated above the surface. Their Figs. 3b and 5b show the displacements of potential vorticity lines, which act just like dye lines outside the heated latitudes. It is perhaps clearer from their zonal wind fields (Figs. 3c, 5c) that "gravest vertical mode" structure dominates the solution at latitudes remote from the heating, while shallower components of the heating are trapped nearer the heated region. Trajectories in the meridional plane (their Fig. 4b) indicate low-level area contraction at latitudes −5° to 0°, remote from the heating, because of the elevated gradient of their specified heat source.

The dissipative trapping of shallow components of a heat source is well illustrated by Geisler and Stevens (1983). The authors reconsidered the vertical structure of the Gill (1980, described in Gill 1982) model of the equatorial Walker circulation as a steady (forced and damped) Kelvin wave. They, too, happened to choose a vertical heating profile with an elevated gradient. Figure 3 of Geisler and Stevens (1983) shows their results: zonal wind as a function of distance east of the heating. Again, the gravest vertical structure dominates the steady wind response at large distance, while shallower features of their sharp heating profile are expressed only locally. As a result, their model predicts a steady ascending low-level easterly flow across the Pacific, rising toward the elevated base of heat source in the western Pacific. In a fully dynamic model, such ascending flow might tend to cause convection east of the heat source, resulting in an eastward propagation of the equatorial heat source, as in the Chang and Lim (1988) two-vertical-mode Kelvin wave–CISK theory of the 40–50-day oscillation.

In summary, the wind field induced by a heat source exhibits ever-simpler vertical structure as a function of distance from the heat source, both in the transient problem and in models with rotation or dissipation. Thus, if the deep, simple far-field inflow has a larger value at the surface than does the more detailed near-field flow, low-level contraction occurs in the intervening environment.

5. Gregarious tropical convection

The observation that tropical convection is highly clustered has long been puzzling from a theoretical point of view, as many simple theories predict that conditional instability should favor convective ele-

ments at the smallest possible scale: individual clouds separated as widely as possible from each other (e.g., Lilly 1960; Bretherton 1987, 1988). One particularly evident form of clustering is that convective cells occur in multicellular MCSs (frequently termed "cloud clusters" in satellite literature). The main organizing mechanism behind MCSs is apparently that the pool of cold downdraft outflow air at the surface triggers new convective cells (Houze and Betts 1981; Mapes and Houze 1992). But the detrainment of moisture by initial convective cells can also favor multicellular MCS development (Dudhia and Moncrieff 1987). Randall and Huffman (1980) have shown that gregarious behavior or "clumping" develops naturally when randomly triggered clouds can create—by whatever mechanism—a zone around themselves that is more favorable for new cloud development than the more distant environment. By the same token, if MCSs act to favor their neighbors, then they should be observed to occur in clumps.

Recent studies indicate that MCSs are indeed gregarious. Subjective observations of "super cloud clusters" (Nakazawa 1988) have been given an objective meaning by Mapes and Houze (1993b), who defined a *supercluster* as follows. A cloud cluster within an infrared satellite image is a spatially connected cold cloud area, of at least moderate size (>5000 km^2), with blackbody temperature lower than 208 K (−65°C). This temperature threshold is among the lowest in use in infrared satellite studies of convection, and corresponds roughly to the radar echo (precipitation) area within the MCSs investigated by airborne radar during EMEX (Mapes and Houze 1992). A *time cluster* is a set of temporally connected cloud clusters, where two cloud clusters are said to be temporally connected if their spatial overlap from one satellite image to the next (3-h intervals) exceeds half the size of either cloud, or 10 000 km^2. A supercluster may then be defined as any time cluster with a duration exceeding 48 hours. The arbitrary parameters listed here (except for the present colder temperature threshold) were selected by Williams and Houze (1987), who used these definitions in a computer algorithm to track cloud clusters observed near Borneo during the winter monsoon experiment (WMONEX).

The algorithm described above has been applied objectively to GMS infrared satellite data covering the 1987 Australian monsoon (14 January to 14 February), during which the horizontal divergence measurements presented in Fig. 2b were obtained. The resulting time clusters are plotted in Fig. 7. The plotted values are the durations of time clusters, in hours, in characters whose size (area on the page) is proportional to the total amount of cold cloud coverage (area × time) accounted for by each time cluster. It is evident that a great deal (43%) of the total cold cloud coverage was associated with superclusters, several of which exceeded 100 h in duration. The Australian monsoon appears as a series of superclusters, some of which included one or more tropical cyclones. But equatorial superclusters, with no tropical cyclones, are also evident. Experimenting with the parameters of the algorithm does not make the superclusters go away.

The superclusters generally consist of several distinct

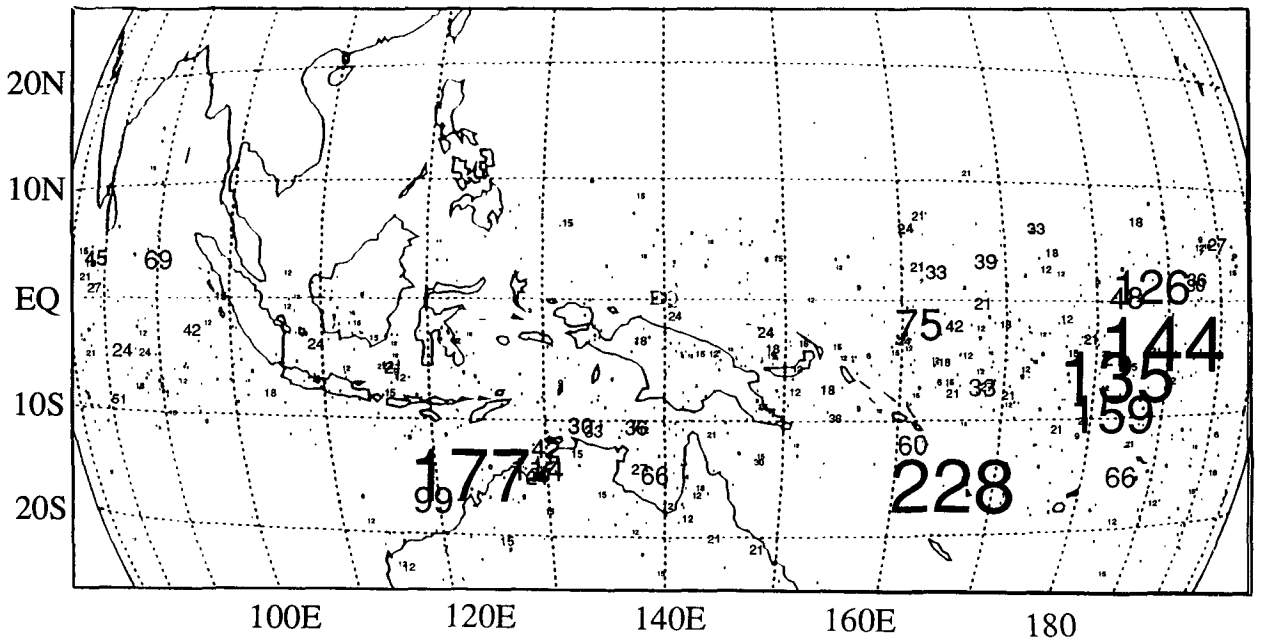

FIG. 7. "Time clusters" (see text) 14 January to 15 February 1987. The plotted values are the durations of objectively tracked cold cloud entities, in hours. Numbers are centered at the location of the centroid of each time cluster at its time of maximum size. The area of the plotting characters on the page is proportional to the total amount (area × time) of cold cloudiness (proxy for convection) subsumed within each time cluster. See Mapes and Houze (1993b) for details.

cloud clusters at any given instant, but the *cloud clusters merge, split, and mingle so gregariously* that it becomes impossible—and perhaps meaningless—*to track them as individual entities.*

6. Summary and discussion

The compensating subsidence that balances the upward mass flux in a heated region of the atmosphere takes place in propagating gravity wave pulses, here called buoyancy bores, which emanate from the heat source when it starts. These buoyancy bores propagate like gravity waves, so their speed is well predicted by gravity wave theory: $c = \lambda/\tau_b$, where λ is vertical wavelength and τ_b is the buoyancy period $2\pi/N$ (about 10 min in the tropical lower troposphere). A complex heat source has multiple vertical wavelengths embedded within it, and the buoyancy bores corresponding to these wavelengths separate out from each other with distance as they propagate away from the heat source. As a result, a heat source with characteristics like a tropical MCS (an elevated peak in the low-level convergence) causes upward displacements at low levels in the nearby atmosphere, thus favoring the development of additional convection nearby.

These results conflict with a traditional assumption, valid only for infinite uniform fields of cumulus clouds, that the compensating subsidence associated with convective ascent occurs between the clouds (e.g., Bjerknes 1938; Asai and Kasahara 1967; the "slice method," Holton 1979, p. 335). That assumption leads to the oft-cited conclusion that convection acts to suppress nearby convection, such that "conditional instability produces maximum growth rates for motions on the scale of ordinary cumulus clouds" (Holton 1979, p. 355). In light of that conclusion, the observation that tropical deep convection is highly clustered has generally been attributed to the organizing influences of separate, preexisting, large-scale disturbances.

This historical presumption underlies and colors much of the literature concerning the interaction of tropical convection with larger-scale circulations. One result has been a predilection on the part of some scientists to interpret observations of *mean* upward motion within sounding arrays that contain deep convection as *large-scale* upward motion, and even as a large-scale *forcing,* to which the convection is merely a response. For example, Arakawa and Schubert (1974) show in their Fig. 13 observations that the cloud work function A (a quantity like CAPE) does not dramatically *increase* during periods with deep convection and mean upward motion, as "expected." Their interpretation was that convection responds quickly to the deep mean ascent, assumed to represent slow, large-scale "forcing" and represented by the suggestive symbol F_c. Another equally logical interpretation is that the large-scale tropical atmosphere responds rapidly and passively to convection.

The intermediate interpretation, advocated in this paper, is that tight feedbacks exist between tropical deep convection and the environmental flows with which it satisfies mass continuity. If the convection has a horizontal divergence profile like that observed in MCSs over the world's warmest waters, the feedbacks can be positive, as shown herein, leading to self-organizing superclusters (Mapes and Houze 1993b). Persistent, clustered convection will tend on the rotating earth to spin up coherent large-scale circulations or, in the language of linear wave dynamics, to excite planetary waves. Advection, and possibly quasi-balanced (with quasigeostrophic as a special case) dynamical vertical motions, associated with these large-scale circulations can become important contributors to the evolution of the whole system. But the relationship between tropical deep convection and the motions of the fluid in which it is embedded is not as separate, or as one-sided, as the language of "large-scale forcing" and "convective response" implies.

Acknowledgments. For the intellectual seeds of this work, I am indebted to Professor C. Bretherton. The photograph in Fig. 1 was taken in the GFD lab of Professor P. Rhines. Mr. R. Pandya performed the numerical simulation shown in Fig. 6. G. C. Gudmundson edited the manuscript. Mr. R. Zehr of CIRA/CSU supplied the satellite data from which Fig. 7 was prepared. This work was supported by the U.S. National Oceanic and Atmospheric Administration's Hurricane Research Division and the National Science Foundation, under NSF Grant ATM9008406.

REFERENCES

Arakawa, A., and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment. Part I. *J. Atmos. Sci.,* **31,** 674–701.

Asai, T., and A. Kasahara, 1967: A theoretical study of the compensating downward motions associated with cumulus clouds. *J. Atmos. Sci.,* **24,** 487–496.

Bjerknes, J., 1938: Saturated adiabatic ascent of air through a dry-adiabatically descending environment. *Quart. J. Roy. Meteor. Soc.,* **108,** 325–330.

Bretherton, C. S., 1987: A mathematical model of nonprecipitating convection between two parallel plates. Part I: "Linear" theory and cloud structure. *J. Atmos. Sci.,* **44,** 1809–1827.

——, 1988: A mathematical model of nonprecipitating convection between two parallel plates. Part II: Nonlinear theory and cloud field organization. *J. Atmos. Sci.,* **45,** 2391–2415.

——, and P. K. Smolarkiewicz, 1989: Gravity waves, compensating subsidence, and detrainment around cumulus clouds. *J. Atmos. Sci.,* **46,** 740–759.

Chang, C.-P., and H. Lim, 1988: Kelvin wave-CISK: A possible mechanism for the 30–50 day oscillations. *J. Atmos. Sci.,* **45,** 1709–1720.

Cram, J. M., R. A. Pielke, and W. R. Cotton, 1992: Numerical simulation and analysis of a prefrontal squall line. Part II: Propagation of the squall line as an internal gravity wave. *J. Atmos. Sci.,* **49,** 209–225.

Dudhia, J., and M. W. Moncrieff, 1987: A numerical simulation of quasi-stationary tropical convective bands. *Quart. J. Roy. Meteor. Soc.,* **113,** 929–967.

Emanuel, K. A., 1983: Elementary aspects of the interaction between cumulus convection and the large-scale environment. *Mesoscale Meteorology: Theories, Observations, and Models,* D. K. Lilly and T. Gal-Chen, Eds., D. Reidel, 551–575.

Frank, W. M., and J. L. McBride, 1989: The vertical distribution of heating in AMEX and GATE cloud clusters. *J. Atmos. Sci.*, **46,** 3464–3478.

Geisler, J. E., and D. E. Stevens, 1982: On the vertical structure of damped steady circulation in the tropics. *Quart. J. Roy. Meteor. Soc.*, **108,** 87–93.

Gill, A. E., 1982: *Atmosphere–Ocean Dynamics.* Academic Press, 662pp.

Holton, J. R., 1979: *An Introduction to Dynamic Meteorology.* Academic Press, 391 pp.

Houze, R. A., Jr., 1982: Cloud clusters and large-scale vertical motions in the tropics. *J. Meteor. Soc. Japan,* **60,** 396–410.

——, 1989: Observed structure of mesoscale convective systems and implications for large-scale heating. *Quart. J. Roy. Meteor. Soc.,* **115,** 425–461.

——, and A. K. Betts, 1981: Convection in GATE. *Rev. Geophys. Space Phys.,* **19,** 541–576.

Hoxit, L. R., C. F. Chappell, and J. M. Fritsch, 1976: Formation of mesolows or pressure troughs in advance of cumulonimbus clouds. *Mon. Wea. Rev.,* **104,** 1419–1428; **105,** 1063.

Lilly, D. K., 1960: On the theory of disturbances in a conditionally unstable atmosphere. *Mon. Wea. Rev.,* **88,** 1–17.

Mapes, B. E., 1992: The Australian monsoon and its mesoscale convective systems. Ph.D. dissertation, University of Washington, 184 pp.

——, and R. A. Houze, Jr., 1992: An integrated view of the 1987 Australian monsoon and its mesoscale convective systems. Part I: Horizontal structure. *Quart. J. Roy. Meteor. Soc.,* **118,** 927–963.

——, and ——, 1993a: An integrated view of the 1987 Australian monsoon and its mesoscale convective systems. Part II: Vertical structure. *Quart. J. Roy. Meteor. Soc.,* in press.

——, and ——, 1993b: Cloud clusters and superclusters over the oceanic warm pool. *Mon. Wea. Rev.,* **121,** 1398–1415.

Miller, P. P., and D. R. Durran, 1991: On the sensitivity of downslope windstorms to the asymmetry of the mountain profile. *J. Atmos. Sci.,* **48,** 1457–1473.

Nakazawa, T., 1988: Tropical super clusters within intraseasonal variations over the western Pacific. *J. Meteor. Soc. Japan,* **66,** 823–839.

Nicholls, M. E., R. A. Pielke, and W. R. Cotton, 1991: Thermally forced gravity waves in an atmosphere at rest. *J. Atmos. Sci.,* **48,** 1869–1884.

Randall, D. A., and G. J. Huffman, 1980: A stochastic model of cumulus clumping. *J. Atmos. Sci.,* **37,** 2068–2078.

Raymond, D. J., 1983: Wave-CISK in mass flux form. *J. Atmos. Sci.,* **40,** 2561–2572.

——, 1987: A forced gravity wave model of self-organizing convection. *J. Atmos. Sci.,* **44,** 3528–3543.

Schubert, W. H., P. E. Ciesielski, D. E. Stevens, and H. Kuo, 1991: Potential vorticity modeling of the ITCZ and the Hadley circulation. *J. Atmos. Sci.,* **48,** 1493–1509.

Williams, M., and R. A. Houze, Jr., 1987: Satellite-observed characteristics of winter monsoon cloud clusters. *Mon. Wea. Rev.,* **115,** 505–519.