Mechanisms of Monsoon Low-Frequency Variability: Surface Hydrological Effects

PETER J. WEBSTER
Division of Atmospheric Physics, CSIRO, Aspendale, Victoria, Australia
(Manuscript received 3 November 1982, in final form 28 March 1983)

ABSTRACT

Observations indicate that monsoon systems are characterized by orderly large-scale and low-frequency variations. With a time scale of two weeks and sometimes longer, regions of ascending motion are observed to form to the north of the equator and propagate slowly northward across southeast Asia. The propagation appears to be associated with the "active-break sequence" of the summer monsoon which acts as a local modulator on the activity of the synoptic-scale disturbances.

A zonally symmetric non-linear two-layer model containing an interactive ocean and a "continental" poleward of 18°N is used to investigate the mechanisms which produce the observed low-frequency variability. Only when a full hydrology cycle is considered does the model produce variations which resemble the observed structures. Mechanisms are traced to include the interaction of the components of the total heating function. The sensible heat input in the boundary layer, although considerably smaller than the other heating components, destabilizes the atmosphere ahead of the ascending zone allowing the moist convective heating component to move northward slightly ahead of the band of precipitation. The poleward encroachment of these components of the heating forces the vertical velocity, which is proportional to the total heating, to move poleward also. The poleward movement is aided by the evaporative cooling of the precipitation moistened ground on the equatorial side of the rising motion which reduces the sensible heat input and effectively stabilizes the troposphere and thus reduces the convective heating in that sector while at the same time reducing the latent heat flux. The time scale of the event is determined by the rate of evaporative drying behind the ascent and the formation of a new zone of ascent in the vicinity of the coastal margin. A schematic representation of heating intercomponent interaction and dynamic feedback is given and the generality of the mechanism to other observed situations is considered. The hypotheses developed and tested in this study underline the importance of the role of ground hydrology related processes in large-scale atmospheric dynamics.

1. Introduction

Traditionally, the monsoon is characterized as a relatively reproducible and smooth seasonal phenomenon with a variable amplitude which differentiates between a "good" and a "bad" monsoon year. In actual fact the subsynoptic temporal structure of the monsoon is much richer. Besides a very variable synoptic behavior, the monsoon regions of Southeast Asia, Indonesia-Australia and Africa are characterized by low frequency (sub-seasonal) modulations of the synoptic scale monsoon weather with each region exhibiting its own special features. A period during which a region receives considerable precipitation, usually a manifestation of a sequence of disturbances, is often referred to as an active monsoon period. On the other hand, the absence of precipitation for a prolonged period is usually referred to as the break monsoon (i.e., a "break" in the active monsoon). The "active-break" sequence appears to apply a multi-week modulation to the regional precipitation patterns. The predominant time scale of the variation of rainfall in the monsoon areas is summarized by Ramage (1971).

A number of studies have attempted to isolate the character of the low frequency variability of the monsoon from observations. Krishnamurti and Bhalme (1976), Murakami (1977), Sikka and Gadgil (1980), among others have found two principal energy peaks in the 10–15 day and 30–40 day period ranges. Krishnamurti and Bhalme related the variation of the precipitation over India to an oscillation of the entire interhemispheric monsoon system and suggested that variations in the Mascarene High in the South Indian Ocean were transmitted through the low level monsoon flow over the Western Indian Ocean to the Indian sub-continent region with roughly a 10–15 day periodicity.

Using a 5-year set of satellite visible brightness data Sikka and Gadgil (1980)² have related the low frequency variations in precipitation over India to the

---

1 Present affiliation: Department of Meteorology, Pennsylvania State University, State College, PA 16802.

2 The domain has been enlarged subsequently to include the Western Pacific Ocean and Indonesian region for successive northern hemisphere summers and is reported in Gadgil (1982).
successive northward migration of a maximum cloudiness zone (the MCZ), a phenomena which they show to be fairly coincident with the 700 mb trough. Two favorable locations for the MCZ were found. These are a more usual location over the monsoon zone north of 15°N and between the equator and 10°N. The southern location was only occupied intermittently when the northerly MCZ was absent or weak. Reestablishment of the MCZ to the north occurred by a northward migration of a new zone forming in the south.

Figure 1 presents a simplified interpretation of the principal quasi-periodic migration of the MCZ as determined by Sikka and Gadgil (1980). The numbers indicate the approximate period of the existence of an MCZ event. Although the cloud band occasionally remained in a northward location for an extended period (~40 days), the average period is about 15 days. The variability is enclosed within a well-defined seasonal envelope. Propagations early in the summer are confined close to the coastal margin, a path returned to in the early fall. During summer the propagations extend much further poleward. Thus, according to Sikka and Gadgil the character of the monsoon during each summer season is one of orderly low-frequency transience which tended to support the observations of Krishnamurti and Bhalme (1976) and Murakami (1977). Within their scheme a “break-monsoon” (i.e., the absence of organized precipitation over a large segment of the Indian subcontinent) occurs when the monsoon MCZ in the north weakens prior to a reestablishment by the northward migration of the southern MCZ.

A recent study by McBride (1983) considered the variation of cloudiness in the Australasian–Indonesian region during the Southern Hemisphere summer (i.e., in the precipitating region of the winter monsoon). During the period of the Winter Monsoon Experiment (December 1978 through March 1979) McBride found that large scale variations in cloudiness were apparent and were associated with active and break periods in monsoon precipitation. However, unlike the variation found for the Indian monsoon by Sikka and Gadgil (1980), the variations were distinctly longitudinal in character and varied only weakly in latitude. There appears to be no similar study for the African monsoon region.

A number of mechanisms have been proposed for the low-frequency variability of the planetary scale monsoon. Most refer to some type of circulation modulation induced by feedbacks between the dynamics of the system and some aspect of the heating. For example, Sikka and Gadgil (1980) attribute the MCZ migration to variations of surface heating produced by the cloudiness gradient which is produced by the existence of the MCZ itself. Similar feedbacks emerge from Krishnamurti and Bhalme (1976) and undoubtedly are of importance, at least over the continental regions. (The initial surface temperature tendencies over continents introduced by cloudiness factors have been estimated theoretically by Stephens and Webster, 1979, 1981).

Most theoretical and modeling studies of the monsoon have concentrated on the structure of the mean monsoon (e.g., Washington and Daggapaty, 1975; Hahn and Manabe, 1975 and many others). Despite a constant insolation assumption, even these studies realize low frequency variability. In the latter paper, low frequency modulations are observable in diagrams portraying time-sections and the transients appear to possess similar time scales to those observed. Beyond the general circulation models investigation of monsoon variability with variable insolation has generally been restricted to the simpler phenomenological models such as those considered by Webster et al. (1977), Webster and Lau (1977) and Webster and Chou (1980a,b, hereafter referred to as WC1 and WC2) in

---

3 The original diagram of Sikka and Gadgil (1980) (their Fig. 4) showed time sections of the 700 mb trough as well as the northern and southern limits of the MCZ. A central location is plotted in Fig. 1.

4 Stephens and Webster (1983) discuss in detail the dual time scales for which cloud–radiation interaction may impact climate. Over land areas, time scales are relatively short. Over oceans, because of the differential absorption of the long wave and visible streams and the manner in which clouds can alter the ratio of the magnitude of long wave and visible, the time scales of interaction introduced by the cloud may be rather long.
which an "aggregate" philosophy was adopted allowing the physical complexity of the system to be successively incremented from experiment to experiment. For example, Webster and Lau (1977) considered a dry monsoon model (i.e., no hydrologic cycle) allowing only the effect of a heating differential between an interactive and evolving ocean and the land. In a series of experiments, WC1 and WC2 extended the model scheme to include a full hydrology cycle and compared differences in both the mean seasonal model monsoon and the subseasonal variability of the model monsoon WC2 via a series of controlled experiments. The latter paper produced evidence of a subseasonal variability of the time scale suggested by Krishnamurti and Bhalme (1976) and of the character observed by Sikka and Gadgil (1980).

In summary, data studies, general circulation models and phenomenological studies alike appear to be showing similar low frequency, orderly sub-seasonal variability. The aim of this study is to identify the basic physical mechanisms governing its structure. As a vehicle towards this aim we will use the zonally symmetric version of the Webster and Lau (1977) paper used in WC1. It is noted that a major result of Webster and Lau was the importance of the east-west ocean–land contrast. However, in order to simplify the system we choose, in the first instance, the zonally symmetric version assuming that the principal mechanism of poleward-equatorial propagation will exist in the zonally symmetric model. In this, we aim to provide insight into the more complicated structures of the real atmosphere and the large scale general circulation models.

2. Model simulations of low-frequency variability

a. Model description

The WC1 version of the Webster and Lau (1977) multi-domain model is a zonally symmetric two-layer non-linear primitive equation model with an interactive advective mixed-layer ocean which predicts, in stable ocean regimes, the mixed-layer depth and temperature and a meridional heat transport by the ocean. In unstable regions where the mixed layer is undefined, a convective adjustment scheme is used for the calculation of the surface temperature. The transition point between the stable and unstable zones being computed at each iteration. Evolving sea- and land-ice extents are also calculated. The temperature of the continental surface is determined from energy balance considerations.

In all experiments discussed as follows the continent–ocean boundary is placed at 18°N, thus crudely approximating the mean southern coastline of Asia. With this stipulation the model is assumed to be representative of the longitude band 60°E to 100°E. As discussed above it is appreciated that lateral effects may be of considerable importance to the structure of the monsoon, although for simplicity we ignore lateral effects in this study.

The governing zonally symmetric set may be written as

\[ \dot{u}_t + G(\bar{u}) - \bar{u}\left(f + \frac{1}{a}\tan\phi\bar{u}\right) = \bar{F}_1, \]  
\[ \dot{v}_t + G(\bar{v}) + \bar{u}\left(f + \frac{1}{a}\tan\phi\bar{u}\right) = -\frac{1}{a}\bar{\Phi} + \bar{F}_2, \]  
\[ \dot{T}_t + G(\bar{T}) - k\bar{T}_w/p = Q_{TOT} + F_3, \]  
\[ \dot{\bar{q}}_t + G(\bar{q}) = \bar{S} + \bar{F}_4, \]  
\[ \bar{\Phi}_p = \frac{-RT}{p}, \]  
\[ \bar{\omega}_p + \frac{(\bar{v}\cos\phi)\bar{\omega}}{a\cos\phi} = 0, \]

where \( G \) is the operator,

\[ G(X) = -\frac{(\bar{X}\bar{v}\cos\phi)\bar{\omega}}{a\cos\phi} - (\bar{X}\bar{\omega})_p. \]

The equations are written for the 750 and 250 mb levels which have the subscript signatures of 1 and 2 respectively on the dependent variables,

\[ Q_{TOT} = Q_{LH} + Q_{SEN} + Q_{RAD}, \]

where \( Q_{SEN} \) and \( Q_{RAD} \) are the sensible and radiative heating rates. In (8) \( Q_{LH} \) is the latent heating rate and is made up of two components; the heating due to convective effects \( Q_{CON} \) and the heating due to broad scale convection \( Q_{LS} \). Thus the latent heating of a column is given by

\[ H = \frac{L}{2C_p} \left( |q - \bar{q}| + |\bar{q} - \bar{\bar{q}}| \right) + \Delta \left( \int_0^{\infty} \eta(1 - b) \frac{L}{C_p} \omega'q'dp \right), \]

where the first term represents the heating due to broad scale convection and the second the convective heating in moist unstable regions. In (9) \( \Delta \) is the Heaviside function which is unity for \( \bar{\omega} < 0 \) but 0 for \( \bar{\omega} > 0 \), and operates on a convective heating scheme which is a combination of Ooyama (1969) and Anthes (1977). Here \( \eta \) and \( b \) are effective controls on the intensity and moisture availability and are given by

\[ \eta = 1 + \frac{\theta_{\bar{e}_b} - \theta_{\bar{e}_2}}{\theta_{\bar{e}_2} - \theta_{\bar{e}_1}}, \]

\[ \bar{\bar{q}} \] See, for example, the discussion by Stephens and Webster (1979) regarding longitudinal variations in heating and also the theoretical results of Webster and Lau (1977) and Webster (1980).
\[ b = \begin{cases} \frac{1 - R}{1 - R_c}, & R > R_c \\ 1, & R \leq R_c \end{cases} \]  

(11)

In (10) \( \theta_1 \) and \( \theta_c \) are the equivalent potential temperatures of levels 1 and 2 and \( \theta_e \) the equivalent potential temperature of the boundary layer. In (11) \( R \) is the layer relative humidity and \( R_c \) a critical relative humidity.

The two forms of latent heating have entirely different characteristics. Through (10), the convective latent heating is strongly linked to lower boundary heating and the resulting dry convective instability in the lower troposphere. Thus, if moisture is available in a convective unstable region, deep rapid heating of the entire column will ensue. On the other hand, the large scale latent heating requires condensation achieved by broad scale ascent of moist air. The latter process will occur over a larger region depending only on the moisture distribution. It will tend to be smaller than \( Q_{CV} \). The more vigorous convective heating will be restricted to the heated land areas or very warm ocean regions.

Evaporation is calculated over ocean and land in an identical manner using the bulk aerodynamic formulation

\[ E_0 = \rho_s V_s |V_s| C_v (q_s - q_h). \]  

(12)

Over the ocean an infinite moisture supply is assumed whereas over land a water budget is maintained such that

\[ E = \begin{cases} E_0, & W \geq W_c \\ E_0 \frac{W}{W_c}, & W < W_c \end{cases} \]  

(13)

where \( W \) is the moisture holding capacity of the soil and \( W_c \) some maximum value, assumed to be 0.1 m. If the ground water exceeds \( W_c \), it is assumed to be lost as runoff. In (13) \( W \) is calculated as the difference of the precipitation and the evaporation \( E \). The ground hydrology enters into the surface heat budget in another manner by altering the ground albedo and the evaporation, as described in (12). A relationship is assumed which linearly decreases the ground albedo from a maximum of 0.2 when \( W = 0 \) to 0.1 (the same as the assumed ocean albedo) when \( W > W_c \).

The following paragraph defines the simple constant flux boundary layer of the model. The vertical fluxes of sensible and latent heat at the lower boundary are defined by:

\[ \begin{align*}
H_S &= \rho_s C_D |V_s| C_p (T_s - T_h) \\
H_L &= \rho_s C_D |V_s| L (q_s - q_h)
\end{align*} \]  

(14)

where the subscripts \( s \) and \( h \) denote the quantities at the surface and the reference level which is defined at the top of the constant flux layer. As used in (14) and (12) \( |V_s| \) is the velocity at the lower boundary and is assumed to be one half the prognostic velocity calculated at 750 mb. The constant flux layer, assumed to be 80 m in height, is taken to represent the lower boundary of the atmosphere. To calculate quantities at level \( h \) we assume that the flux from the lower surface can be described in the lower layer in terms of an eddy diffusion. Thus we can write

\[ -u \frac{\partial T}{\partial Z} = - \frac{u}{\Delta Z} (T_1 - T_h) = \rho_s C_D |V_s| (T_s - T_h) \]

so that

\[ T_h = \frac{\eta T_s + T_1}{\eta + 1}, \]  

(15)

where \( \eta = \rho_s \Delta Z C_D |V_s| / \mu \) and \( \Delta Z \) is the thickness between the 750 mb level and the top of the constant flux layer. Here \( C_D \) is a drag coefficient defined by

\[ C_D = \left[ \frac{k_0}{\ln \left( \frac{h + Z_0}{Z_0} \right)} \right]^2 \]

where \( k_0 = 0.4 \) is the von Karman constant and \( Z_0 \) is the roughness parameter of the underlying surface.

In this study we will test the hypothesis that the monsoon variability is produced by the intercomponent interaction of the total heating function (8). As modulations have occurred in model results with assumed constant cloudiness distribution (e.g., Hahn and Manabe, 1975, and WC2), we will test the hypothesis that the variability occurs by interaction between the two components of \( Q_{LH} \) and \( Q_{SEN} \). In order to isolate the interaction of these two heating functions, the cloudiness is set constant and 30% coverage is assumed. That is, we will assume that the radiative heating is non-interactive, except, of course, that the ground albedo and surface temperature is a function of soil moisture which is a function of precipitation. Thus, even with a fixed cloudiness a weak \( Q_{RAD} \) feedback does exist. Actually, the neglect of interactive clouds does not appear to be a major factor in that a change in sign of the feedback would not be involved. In fact, the feedback taking place between the sensible and latent heating is consistent with that which would be attributable to cloud-radiational heating feedback.

In WC1 the two layer formulation was chosen because it was considered adequate to represent the basic modal form the monsoons which appear as deep divergent structures occupying most of the troposphere in the vertical. An increase in vertical reso-

\footnote{We have become aware of similar variability occurring in a zonally symmetric version of the NASA-GLAS general circulation model in a study by Goswami and Shukla (1983).}
lution would not necessarily alter the form of the response. However, the price one pays for the two level formulation is acrudeness of the representation of the reference level in terms of interpolations of variables measured at the surface and 750 mb.

b. Experiments

Webster and Chou (1980a) performed experiments with the model discussed above in which the evolving monsoon driven by a variable insolation was studied under a number of situations. The experiments considered the monsoon circulation of an ocean covered earth with and without a hydrology cycle (the “dry” and “moist” ocean experiments of WC1 and an ocean covered earth with a polar continental cap north of 18°N. The latter geography was also considered with and without a hydrology cycle (the “dry” and “moist” ocean-continent experiments). Except for the “moist ocean-continent” experiment, all other experiments produced a smooth seasonal variation lagging behind the insolation maxima but with no subseasonal structure. With respect to the sensitivities of the various parameterizations used in the model, we refer the reader to Webster and Lau (1977) and WC1. We will consider only the results of the moist ocean-continent experiment in this paper.

The following procedure was adopted. The model was initialized as a horizontally isothermal atmosphere with an initially prescribed surface temperature. With the sun held in the equinoctial position the model was integrated for one year during which time an equilibrium state was approached. The model was run for a further three years with the insolation undergoing the seasonal cycle. Results are shown for years 2 and 3 after the equinoctial equilibrium state. Day 365 on all diagrams refers to the spring equinox one year following the commencement of variable insolation.

c. Results

Figures 2a–d show the surface temperature ($T_s$), the lower (750 mb) and upper (250 mb) tropospheric zonal velocity components ($\bar{u}$ and $\bar{u}_z$) and the vertical velocity ($\bar{w}$; actually $dp/dt$) as functions of time and latitude.

Within a large-period seasonal modulation two main subseasonal variations may be observed. The first is a monsoon onset which occurs about 40 days after the spring equinox. The onset may be seen in the vertical velocity fields as a substantial latitudinal jump from south of the equator to a position just poleward of the continental demarkation (the solid

---

**FIG. 2a.** Time–latitude section of the surface temperature ($T_s$, K) for the 1½ year period between the spring equinox (day 365) to the autumnal equinox (day 910). The continuous horizontal line represents the model coast located at 18°N. The three boxes labelled A, B and C show the location of the latitude–time regions detailed in Figs. 3a, b and c, respectively. Contour interval; 4 K.

**FIG. 2b.** As in Fig. 2a except for the 750 mb zonal wind component $\bar{u}$ (m s$^{-1}$). Dashed contours denote easterly winds. Contour interval; 2.5 m s$^{-1}$.
horizontal line at $18^\circ$N). The abruptness of the onset is in sharp contrast to the autumn withdrawal of the monsoon which appears as a more orderly progression of the ascending band southward. The second variation exists after the onset and remains throughout the Northern Hemisphere summer. The $W_c$ fields are characterized by maxima which appear to migrate northward from south of the coast to well north of the coast with about a bi-weekly periodicity. The other fields also possess similar features. For example, the continental surface temperature (Fig. 2a) shows regions of cooling which appear to coincide in location and timing with accelerations of the upper and lower tropospheric zonal flow (Figs. 2b and c). In late summer and early fall, the position of the ascending centers moves equatorward towards the land–sea boundary. It should be noted that the variability ceases during the Northern Hemisphere winter when the maxima vertical velocity is south of the equator and over the ocean where it assumes a fairly constant location and magnitude. This cessation of variability follows a gradual diminution of activity during the fall.

Three boxes, marked A, B and C are inscribed on Figs. 2a–d and define three time–latitude sections between $2^\circ$S and $62^\circ$N for the 30-day time periods between days 450–480, 770–800 and 810–840. The boxes enclose three specific cases which are enlarged in Figs. 3a–c to facilitate detailed study. Each figure has four panels which show the surface temperature, the vertical velocity, a measure of the vertical shear ($\bar{u}_2 - \bar{u}_1$) and the lower tropospheric flow upon which is superimposed a measure of the latitudinal shear ($\bar{u}_1$), which is proportional to the low level relative vorticity. Regions of intense low level cyclonic shear are stippled. The sections of Fig. 3 will be used as the principal means of comparison of the model results with the real atmosphere.

The detailed characteristics of Fig. 3 may be summarized as follows. An event (or epoch as termed by Sikka and Gadgil, 1980) may be traced in the vertical velocity fields as rising motion commencing over the ocean to the south of the land–sea demarkation. Maximum intensification occurs over land 3 to 4 days later in the initial phase and some $12^\circ$ to $16^\circ$ farther inland some 7 days later with a slight reduction of the intensity of the event between the two maxima. During mid-summer (sections A and C) the maxima furthest away from the coast during each epoch is the strongest whereas early in summer the more southerly maxima is strongest. An interesting feature of each figure is that an event, while reproducible in their basic characteristics (i.e., they successively move northward from south of the coast with a period in the 13–15 day range and each possesses a double maximum) possesses its own character.

Accompanying each northward excursion of the vertical velocity maxima is a surface temperature
minimum lagging behind by a day or so. Furthermore the vertical shear undergoes considerable change in keeping with the increasing easterlies aloft and the increasing westerlies below. In the lower troposphere intensification of the relative vorticity field occurs and local intense cyclonic maxima are produced near the inland maximum. The cyclonic vorticity maxima tend to lag the vertical velocity maxima by a couple of days. In sympathy with the evolving fields noted above, substantial changes are also occurring among all the components of the total heating field. We will discuss these subsequently.

3. Comparison of model results with observations

The basic periodicity of the summer oscillations found in the model tends to be similar in many respects to the variations inherent in the analyses of Krishnamurti and Bhalme (1976), Murakami (1977) and Sikka and Gadgil (1980). Murakami found two basic periodicities in his spectrum analyses of a number of observed fields. These were related to a 5-day peak, located in northern India and associated with transients in the monsoon trough and a larger scale 15-day variability. The latter peak was likened by Murakami to a “...large scale standing oscillation...of the Hadley-type which affects both the summer and winter hemispheres...” With the zonal symmetry model constraint it is impossible to establish a relationship between the model results and the observed 5-day spectral peak which most likely pertains to individual disturbances in the monsoon region. However, it should be noted that the relative vorticity fields of Figs. 3a-c do possess cyclonic maximum at the northern extension of each of the cell excursions which may be interpreted as an approach to instability of the shear zone when the rising motion is at its most intense. That is, the vorticity maxima may indicate regions in
which disturbances would form if the model system were more general.

Another feature which is quite similar to observations is the seasonal variation of the locus of the maximum vertical velocity events as found by Sikka and Gadgil (1980). The northward penetration of the events tends to increase following the monsoon onset to reach a maximum in late summer before retreating southward in the early fall. Such model behavior may be seen by comparing Figs. 1 and 2b.

Of two principal modes of monsoon transience noted by Sikka and Gadgil (i.e., the 15 day and 40 day periodicities), the model seems more capable of reproducing the shorter period biweekly variability along with the principal zones of activity. The longer period does not seem to be a property of the model.

4. Variability in surface variables and heating fields

Large scale variations in the surface variables and in the components of the total heating field occur simultaneously during the latitudinal event propagations. These will be discussed in detail through two event cycles during the 449–468 day period which corresponds to part of section A (Fig. 3a).

Figure 4 shows the variability of the surface albedo (%) and the ground moisture (cm) which resemble in character (although with a slight lag) the vertical velocity fields. Through each cycle the albedo at 26°N varies from 5 to 8%. The 3% reduction in albedo following the precipitation results from the accumulation of ground water which varies by as much as 1.5 cm through each cycle. With the ground at saturation (i.e., 10 cm of absorbed water), the land surface albedo is reduced to values similar to that of the adjacent ocean.

The relationships between the vertical velocity, the total heating $Q_{TOT}$ and boundary condition terms are shown from two different perspectives in Figs. 5 and 6. Fig. 5 plots time sections of the variables at three latitudes (18°N, 20°N and 34°N) for the period 449–468 days, whereas Fig. 6 shows latitude sections of the fields at days 449, 453, 457 and 461. Both sections
pass through two evolving events which are cross-referenced between Figs. 5 and 6 by the use of the identifiers A, B and C. In these figures A is a remnant of a previous excursion prior to day 449.

The variation of the vertical velocity and the total heating field is shown in Fig. 5a. Near the coastal margin, $W_z$ and $Q_{TOT}$ are almost exactly in phase and only slightly less so at 34°N. Such phase correspondence may be expected from simple scaling considerations as discussed in e.g., WCI. However considerable phase differences occur between the various components of $Q_{TOT}$ and between these components and the vertical velocity field. Furthermore a significant latitudinal dependence appears to exist in the phase relationships. These properties are apparent in Fig. 5b which plots $Q_{CV}$, $Q_L$ and $Q_{SEN}$ as functions of time at the three latitudes. Near the coast variations in $Q_{SEN}$ (this component being substantially smaller than the other two heating forms) and $Q_{CV}$ appear to coincide in time. However farther north $Q_{SEN}$ tends to lead $Q_{CV}$ which, in turn, precedes $Q_L$ by about two days. These are the three terms (each out of phase) which sum to $Q_{TOT}$ that is almost exactly in phase with $W_z$. 

In Fig. 6a the vertical velocity is plotted with $T^*_i$ and $W^*$ which represent, respectively, the deviations of $T_i$ and $W$ from their mean monthly values at a particular latitude. Two points are apparent. First, the $T^*_i$ is almost exactly out of phase with $W^*$. The negative correlation arises from the increased availability of moisture for evaporation by ground water accumulation following precipitation and the subsequent lowering of temperature. Decreases in $T^*_i$ occurs despite the increase of absorption of radiation at the surface due to the lower albedo caused by the higher ground water. The second point refers to the relative position of the maximum upward velocity component and the extrema in the $T^*_i$ fields. In each latitude section a maximum and a minimum in $T^*_i$ lies poleward and equatorward of the $W_z$ maximum, respectively. Presumably this is due to the ground water accumulation due to precipitation associated with $W_z$. The effect, though, is to produce a maxi-

FIG. 3c. As in Fig. 3a except for region C of Fig. 2.
mum in $Q_{\text{SEN}}$ ahead of the rising motion and a minimum behind it.

Figure 6b shows the latitudinal sections of $Q_L$ and $Q_{CV}$ plotted with $W_z$. Here $Q_{CV}$ tends to lead the $W_z$ maximum in latitude whereas $Q_L$ tends to lag. From Fig. 5b we can see that $Q_L$ lags $W_z$ both spatially (latitudinally) and temporally whereas $Q_{CV}$ tends to lead $W_z$ in both dimensions. Furthermore by comparing Figs. 6a and 6b it seems that the $Q_{CV}$ distribution is nearly in phase with the $T^*_z$ field, and, consequently, the $Q_{\text{SEN}}$ field.

5. Mechanisms

In the following paragraphs the experimental results will be condensed into a consistent theory describing the low frequency variability of the monsoon flow. The consistency of these arguments with theories which may have arisen had we used a system with greater degrees of freedom or more physical processes are also discussed.

a. Heating intercomponent interaction and dynamic feedback

The perturbation of the $Q_{\text{SEN}}$ field by the precipitation distribution (noted in Fig. 5 and inferred in Fig. 6) may have two distinct effects.

First, it can continually displace the total diabatic heating profile ($Q_{\text{TOT}}$) slightly poleward, and, consequently, as can be inferred from Fig. 5a, cause a corresponding poleward movement of the vertical velocity maximum. With the movement of the vertical velocity maximum, the precipitation maximum will also move poleward. The effect of this feedback will be slow as $Q_{\text{SEN}}$ is an order of magnitude smaller than either $Q_L$ or $Q_{CV}$. However, the incremental erosion of the $Q_{\text{TOT}}$

---

Fig. 4. Variation of the surface albedo (%) and the ground water (cm) between days 449 and 468.

Fig. 5. Variation of (a) the vertical velocity field $W_z$ (solid curve; units $10^{-3}$ mb s$^{-1}$) and the total heating rate $Q_{\text{TOT}}$ (dashed curve; units $10^{-5}$ K s$^{-1}$) and (b) the sensible heating rate $Q_{\text{SEN}}$ (heavy dashed curve; right hand scale), the large scale convective heating rate $Q_L$ (dashed curve) and the convective heating rate $Q_{CV}$ (solid curve). Fields are plotted at latitudes 18°N, 20°N and 34°N between days 449 and 468. Individual events are marked with A, B, C and D for recognition. Note that the vertical velocity $W_z$ is plotted as a positive quantity to facilitate comparison.
field behind the $W_s$ maximum and the small growth ahead is consistent with the slow creep northwards of the $W_s$ field. It should be noted that the effect is consistent with the variation of latent heat flux expressed in (14). As the sensible heat flux decreases with the lowering of $T_s$, the latent heat flux also diminishes with the lowering of $q_s$, which at saturation is purely a function of $T_s$. The reduction of $q_s$, of course, occurs simultaneously with an increase in $q_h$, which is largest in the regions of precipitation. That is $(q_s - q_h)$ and, therefore, $H_t$ decreases.

The second effect deals with the destabilization of the lower troposphere where $Q_{SEN}$ is a maximum (i.e., ahead of the $W_s$ maximum). The consequences of the destabilization and the effect that it has on the components of the heating field may be seen from Fig. 6b and also in the functional form of $Q_{CV}$. In each panel of Fig. 6b (except in the region very close to the coast which will be discussed separately) the $Q_{CV}$ maximum is slightly poleward of the $W_s$ maximum and close to the $T_s^*$ extremum. From (10) the effect of the destabilization of $Q_s$ is quite clear. As $T_s^*$ increases, $q_s$ increases and the intensity factor $q_t$ rapidly grows. With the increase of $q_t$ and the subsequent accelerating inflow of moisture from the south, the moisture availability effect ($b$) defined in (11) also increases. On the other hand, the $Q_L$ maximum, which depends upon the broad scale lifting processes, lags equatorward behind the vertical velocity field.

In the above paragraphs we have established two principal effects of the precipitation-perturbed $Q_{SEN}$ field. Important both effects act in the same sense and move the $Q_{TOT}$ field poleward but for slightly different reasons. In the first case the $Q_{TOT}$ field is modified by the addition of the perturbed $Q_{SEN}$ field. In the second case, the $Q_{SEN}$ input destabilizes the layer ahead of the rising motion and allows another component of $Q_{TOT}$, $Q_{CV}$, to move ahead of the $W_s$ maximum.

Slightly different circumstances exist in the vicinity of the coastline during the formation of an event. Between Figs. 6a and 6b it can be seen that after event B leaves the coastline (day 452) and proceeds inland, the maximum component of the total diabatic heating near the coast is $Q_s$ where $Q_{SEN}$ possesses a local minimum (see Fig. 5b). As $T_s^*$ increases $Q_{SEN}$ increases (Fig. 5b) which effectively destabilizes the lower layers and allows $Q_{CV}$ to increase. The effect of this process may be seen in the "birth" of event C in Fig. 6b on day 456.

As event C forms, event B has moved inland leaving in its wake a maximum in ground water and a

The vertical solid line at $18^\circ$N denotes the coastal location. Note that the vertical velocity $W_s$ is plotted as a positive quantity to facilitate comparison.
minimum in $T^*$. However, following evaporational drying at the surface behind B and with the precipitation occurring near the coast within C, a gradient in $Q_{\text{SEN}}$ develops and destabilization occurs ahead of C followed by the subsequent poleward motion of the $W_s$ maximum. With time, event C grows at the expense of B as the rapidly increasing release of latent heat within C reduces the moisture flux from the ocean region to B.

Finally, we should examine the location of the initial development of an event and ask the question of why it does not move initially equatorward instead of poleward. Reexamining Figs. 3a–c it can be noted that on the days prior to the establishment of the maximum over land (e.g., on days 826 and 837 in Fig. 3c) there existed a region of rising motion to the south of the coast and that it coincides with a maximum in $Q_{l}$, which is the only component of $Q_{\text{TOT}}$ of any magnitude which remains following the northward excursion of the previous event. The other components ($Q_{\text{CV}}$ and $Q_{\text{SEN}}$) have been reduced and moved inland by precipitation events. However, over the ocean the $Q_{l}$ field remains fairly constant. Over the land near the coast, $Q_{\text{SEN}}$ (followed by $Q_{\text{CV}}$) grows rapidly following the evaporation of ground water. Adding the $Q_{\text{CV}}$ and the $Q_{\text{SEN}}$ fields to the $Q_{l}$ field causes a shift in $Q_{\text{TOT}}$ (and hence $W_s$) towards the north.

The large-scale latent heat release remains large over the ocean region because of the abundance of moisture near the coast (cf. 9). As the sea surface warms during late summer the relative magnitude of $Q_{l}$ relative to the other components of $Q_{\text{TOT}}$ increases. The effect of this may be seen in the more equatorial location of the start of each event during late summer.

The interactive mechanisms described above which produce the variability in the monsoon flow are illustrated schematically in Fig. 7. In separate panels and at three times, the interactive components of the total diabatic heating and the subsequent dynamic responses are characterized.

b. Extensions

1) Seasonal Variation

Within the confines of the theory developed in (a), it is possible to include an explanation of the seasonal variability. In Fig. 2d it can be seen that the poleward extent of the event excursions closely follow the distribution of surface temperature. For example, the most northward propagations occur in the middle and late summer when the land area is heated well inland. During early summer and again in the fall when the temperatures away from the coast are considerably cooler, the extent of the poleward excursions is less. The reason for this seasonal behaviour rests in the dependence of the lower troposphere destabilization by sensible heating being basically surface temperature dependent, at least to first order. That is, the temperature of the relatively dry surface ahead of the precipitation eventually determines the extent of the propagation of the zone.

Whereas the extent of the penetration of the precipitation is very seasonal, the frequency of the model events is fairly set at about 12–15 days. Neither the model nor the interpretive theory provides a suggestion as to why observed monsoon breaks are often of longer duration during late summer.

2) Variability during the Northern Winter

Figure 2 shows the absence of any appreciable variability when the convective region is in the Southern Hemisphere and over the interactive ocean. Rather than an organized latitudinal propagation of events, the region of ascent exists in a rather narrow band with a fairly constant amplitude. In fact the region of ascent appears confined to lie within the 300 K sea surface temperature contour which is intriguing because in the observed tropics organized convection appears to be similarly constrained.

The absence of variability over the oceans is a test of the applicability of the theory. In the first place the basic dynamic structure of the ocean (here modeled as an advective mixed layer) provides temperature variations which are much longer than those over land. Ocean temperature variations are essentially decoupled from precipitation events as the sea surface temperature is a function of the response of a deep
advective mixed layer rather than that of a surface energy budget. Alternatively, one may say that the surface temperature is calculated relative to infinite ground moisture so that both the latent heat flux and the sensible heat flux are essentially constant. The effect of this is to decouple the two components of heating $Q_{sen}$ and $Q_{LH}$.

3) LAGS BETWEEN THE HEATING FIELD AND THE DYNAMIC RESPONSE

The vertical velocity field has been shown to correlate well with the heating field (see Fig. 5a) although through a rather complicated superposition of the components of heating. On the other hand, the dynamic response as manifested, for example, by the relative vorticity fields of Fig. 3 lag both the heating fields and the vertical velocity fields by about two days. This lag, an understanding of which is important in order to relate low frequency events with the formation of higher frequency modes, in more a general system, is consistent with vorticity generation arguments.

Upward vertical motion and consequent vortex tube stretching creates cyclonic vorticity which will accumulate as long as the sense of the vertical motion remains the same. The maximum rate of production will occur when $W_0$ is a maximum which is in agreement with Fig. 3. The offset between the maxima of $W_0$ and $u_0$ is due to the decreased rate of cyclonic vorticity generation, as $W_0$ decreases, being overtaken by the destruction of vorticity by dissipative processes.

4) TIME SCALES

In a multivariate structure as we have considered here, the reason for a particular time-scale of motion is difficult to ascertain. If we were advocating a dynamic instability mechanism the problem would be simpler. However, our basic mechanism relates to the interaction and feedback between various heating functions and the dynamics. The rate of the northward creep of the zone of rising motion and precipitation is determined by the erosion of the total heating function in the region of precipitation by reduced fluxes of sensible and latent heat and by the sensible heat destabilization ahead of the zone. However, the life cycle of the event is determined by the evaporation time scales behind (equatorward of) the precipitating zone. The drying and warming changes this region from being a source of latent energy to a source of sensible energy with the ability to generate rising motion in situ. The birth of the next event, determined by the time-scale of evaporation, heralds the decay of the precipitating zone to the north, by reducing the horizontal flux of latent energy.

As the heating functions are parameter choice dependent to a certain extent, the two-week period found in this study should only be considered suggestive of the fact that a low-frequency modulation of the monsoon may be expected through inter-component heating function interaction via ground hydrological processes.

c. Alternative hypothesis: Radiation–cloud interaction and dynamic feedback

As mentioned earlier the form of the model used in the experiments does not allow for variable cloudiness. However, in attempts to describe the observed variability, Sikkha and Gadgil (1980) and Krishnamurti and Bhalme (1976) have suggested that the modification of the radiation stream by the clouds associated with forced ascent may be responsible for the propagation of the zones or the reason that particular periodicities appear to exist in the data. Thus for completeness it is important to decide if the radiation–cloud interaction and dynamic feedback works in the same sense as the interaction between the diabatic heating components.

Radiation–cloud interaction alone cannot be the complete answer to the event propagation. Over the ocean the changes in the net radiation at the surface probably has little effect on the time scales considered here. Efficient mixing by either wind stress or by convective overturn will cause the sea surface temperature to remain relatively fixed.

If the ascending zone reaches the land area, the radiative effect will be identical on either side of the cloudiness zone unless ground hydrology is taken into account. With this further degree of freedom and assuming that the zone is over land, the radiation–cloud interaction modifies the total heating profile in the same sense as the mechanism offered in Section 5a! In the dry area ahead of the clouds enhanced radiative heating will cause $T_1$ to increase and the lower troposphere to be destabilized. In the wet area behind the cloud zone a larger flux will be absorbed by the low albedo ground enhancing the drying and making the area more conducive for the formation of a further event. In this manner, cloud–radiation interaction in a system with full hydrology will work in the same sense as the interactive heating component feedback. However, only detailed and careful experimentation will allow the time scales of the two related mechanisms to be evaluated.

6. Concluding remarks

An effort has been made to establish the physical mechanisms which may be responsible for the existence of time scales of the monsoon which, on the one hand, are considerably shorter than the seasonal insolation time scale, and, on the other hand, much longer than the time scale of instabilities which may be responsible
for synoptic disturbances. The variability in question has been well documented by Sikka and Gadgil (1980).

Using the simple zonally symmetric model developed by Webster and Chou (1980a), a number of experiments were performed in which the physical complexity of the system was successively increased. Only with a hydrologic cycle did the low frequency modulations occur. Comparison provided evidence of some similarity between model results and observational records.

The central mechanism which allowed for the successive zones of ascending motion to propagate northwards was the modulation of the total heating (which is proportional to the vertical velocity distribution) by precipitation (in turn a strong function of the vertical velocity distribution) and the subsequent changes in the sensible heating and latent heat flux. The total heating gradient is altered by either subtracting out sensible heat behind the $W_x$ maximum and adding it ahead and by creating a region of maximum convective heating ahead of the $W_x$ maximum by destabilizing the lower troposphere. Similar arguments exist for the variation of the latent heat flux.

As with all models, and in particular phenomenological models, the results must be considered as suggestive. In this model a number of key effects are ignored. For example, we ignore interactive cloudiness changes although it can be argued that radiative modulation will not change the sign of the feedback and one cannot be too sure of the magnitude or the principal time scales which would be introduced. Furthermore orographic effects are ignored and the continent is considered to be zonally symmetric even though Webster (1980) showed that the east-west differences in heating were probably as important as those in the north-south direction.

Perhaps one of the most important points revealed by this study is the importance of a process which, while relatively small in magnitude, is instrumental in causing significant modification to processes much larger than itself. In this we refer to the changes in $Q_{CV}$ induced by the modified distribution of $Q_{S}$. If we had scaled the thermodynamic equation (3) and the total diabatic heating function (8) prior to the examination of the variation (i.e., if we had assumed a priori that $Q_{TOT}, Q_{CV} \gg Q_{S}$) and had disregarded $Q_{S}$ in favour of the other components, the catalytic role of $Q_{S}$ and the basic mechanism of zone propagation would not have emerged.

Distinct differences to the observation and model results discussed above were noted by McBride (1982) who observed that during the winter monsoon over northern Australia the MCZ cloud shifts were distinctly longitudinal. However, the Southern Hemisphere summer component of the Asian–Australian monsoon system does possess some major geographic differences with the northern summer component. Indonesia spans the equator and the Australian land mass just to the south starting near 10°S, compared to an average 18°N for the South Asian coastline. Oceans embrace Australia and Indonesia to the east and west. Furthermore, in the Indonesia and the western Pacific Ocean area resides the climatological Walker circulation, a feature which itself appears to undergo a 40–50 day variability (Madden and Julian, 1971). In that sense the Indonesian–Australian region stands as a region of confluence with ascent tied to different physical mechanisms; the zonally oriented Walker circulation and the meridionally oriented monsoon cell.

The observations of McBride (1983) need not reduce the generality of the mechanisms of monsoon transitions developed here. That is, the interactive nature of the $Q_{TOT}$ components should be valid for either longitudinally or latitudinally oriented land masses. If differences were to occur it would probably be in the dynamic form of the response. Both Webster (1973) and Gill (1981) have shown that the dynamic response to prescribed forcing is strongly latitudinally dependent. Thus, the model form for heating located over a land mass in the sub-tropics of the northern summer hemisphere will be completely different to that for heating located on the equator. Depending on latitude the modal response may be separately a quasi-stationary Rossby wave, a Kelvin wave or a combination of both. It is interesting to speculate, however, on the relationship between the quasi-stationary Kelvin wave response at low latitudes (Webster, 1973, Chang, 1977, Gill, 1981) and the total heating component feedback on the 40–50 day variability of near equatorial large scale flow. Possibly the form of the response is given by the quasi-stationary modes, the diabatic heating component feedback, evolving from heating interactions over Indonesia and Australia, may provide the temporal modulation.

It is difficult at this stage to envisage the role the adjacent ocean (i.e. the Northern Pacific Ocean) may have on the structure of the variability, on either its period or its strength. To study this effect it is planned to resurrect the Webster and Lau (1977) domain average model in the geographical format used by Webster (1980).

Finally, it should be reiterated that the monsoon system is not an isolated system. A number of theoretical studies (Webster, 1981, 1982) have suggested ways in which low latitude heating events may influence higher latitudes. Such influence is likely to occur into the winter hemisphere when the westerlies encroach close to the equator (Webster, 1982). However the active-break variation with their observed 15 day and 30–40 day variability (Sikka and Gadgil, 1980) occur in the sub-tropics and one may speculate on the effect they may have both interhemispherically in the winter hemisphere and towards higher latitudes in the summer hemisphere. A study in the spirit of those listed above is currently underway.
Acknowledgments. A major part of this research was carried out while visiting the Department of Meteorology of the Naval Postgraduate School in Monterey as the G. J. Haltiner Research Chair Professor. I thank Professor R. Renard, Head of the Department, for his hospitality. I also appreciate the benefits of consultations with my colleagues Dr. A. Plumb, Prof. C.-P. Chang, and Mr. Lang Chou. This research was supported in part by the National Science Foundation, Atmospheric Research Section, under Grant ATM 80-13153.

REFERENCES


