The Origin of Monsoons

Winston C. Chao
Laboratory for Atmospheres, NASA Goddard Space Flight Center, Greenbelt, Maryland

Baode Chen
GEST Center, University of Maryland, Baltimore County, Baltimore, Maryland

(Manuscript received 12 September 2000, in final form 11 May 2001)

ABSTRACT

The monsoon is interpreted as an intertropical convergence zone (ITCZ) substantially away (more than 10°) from the equator and the existence of the ITCZ does not have to rely on land–sea contrast. Land–sea contrast can provide a favorable longitudinal location for the ITCZ but this role can be replaced by sea surface temperature contrast in the longitudinal direction. Thus, the interpretation of the monsoon presented herein differs from the long-held fundamental belief that its basic cause is land–sea thermal contrast on the continental scale in the sense that the existence of landmass is not considered a necessary condition for monsoons. Through general circulation model experiments, support has been found for this interpretation. The Asian and Australian summer monsoon circulations are largely intact in an experiment in which Asia, maritime continent, and Australia are replaced by ocean with sea surface temperature (SST) taken from that of the surrounding oceans. Thus, in these areas land–sea contrast is not a necessary condition for monsoon. This also happens to the Central American summer monsoon. The same thing can also be said about the African and South American summer monsoons, if these continents are replaced by ocean of sufficiently high SST. It is also shown that in the Asian monsoon the change resulting from such replacement is due more to the removal of topography than to the removal of land–sea contrast. In the Asian and Australian winter monsoons land–sea contrast also plays only a minor role.

The origin of the ITCZs and their latitudinal locations have been previously interpreted by Chao. The circulation associated with an off-equator ITCZ, previously interpreted by Chao and Chen through a modified Gill solution and briefly described in this paper, explains the monsoon circulation. The longitudinal location of the ITCZ is determined by the distribution of surface conditions. ITCZs favor locations of high SST as in the western Pacific and Indian oceans, or tropical landmass, due to land–sea contrast, as in tropical Africa and South America. Thus, the role of landmass, when it is important, in the origin of monsoons can be replaced by ocean of sufficiently high SST. Furthermore, the ITCZ circulation extends into the tropics in the other hemisphere to give rise to the winter monsoon circulation there. Also through the equivalence of land–sea contrast and high SST, it is argued that the basic monsoon onset mechanism proposed by Chao is valid for all monsoons.

1. Introduction

The notion that continental-scale land–sea contrast is the main reason that monsoon circulation exists has been a long-held fundamental belief since the time of Halley (1686; e.g., Wallace and Hobbs 1977, which was adopted by Holton 1992; Webster et al. 1998). The purpose of this paper is to scrutinize this notion. The central idea of this notion states that in summer, radiative heating of the continent (e.g., Asia) gives rise to a continental-scale thermal low, and surrounding this thermal low in its southeast direction the low-level wind flows in from southwest. This low-level inflow creates a convergence of moisture, which maintains cumulus convection. In winter, radiative cooling of continent gives rise to a thermal high and to its southeast the low-level wind is from the northeast. Continental-scale land–sea contrast does undoubtedly exist. However, whether it really acts as the main driving force of the monsoon has not been tested in numerical experiments as will be done in this work. The results of our test demonstrate that land–sea contrast is not a necessary condition for the existence of the monsoon. There has been an increasing recognition in the recent years that the monsoon is inextricably tied to the heating in the intertropical convergence zone (ITCZ; e.g., Chao 2000; Hoskins and Rodwell 1995). The origin of the ITCZ has been interpreted by Chao (2000). We interpret the monsoon as an ITCZ substantially away (more than 10°) from the equator and the existence of the ITCZ does not have to rely on land–sea contrast. Land–sea contrast can provide a favorable longitudinal location for the ITCZ but this role can be replaced by sea surface temperature contrast in the longitudinal direction. A brief qualitative explanation of
why the off-equator ITCZ is the source of monsoon circulation can be offered based on the circulation field associated with the ITCZ heating (Chao and Chen 1999) and will be described in section 3. The existence of the ITCZ does not always have to rely on land–sea contrast on the continental scale. This is hinted in the fact that in February the ITCZ close to Australia (and its associated monsoon circulation) covers a longitudinal range several times as long as that of Australia and thus cannot possibly be caused mainly by the land–sea contrast associated with Australia. Also the aqua-planet monsoons obtained in Chao (2000) and in Yano and McBride (1998) exhibit characteristics of observed monsoons, such as low-level southwesterly winds in the monsoon precipitation region and its neighborhood to the south, wind reversal at upper levels, and cross equatorial flows. Yet, these may not be sufficient to argue that the ITCZ in the Asian summer monsoon is not mainly due to land–sea contrast. A purpose of this work is to provide a convincing argument. The role of land–sea contrast in other monsoons will also be examined.

In this work the role of continental-scale land–sea contrast in the origin of monsoons is examined through numerical simulation with the Goddard general circulation model (GCM; Takacs et al. 1994). The Asian and Australian monsoon circulations (and other monsoons) are obtained in a 4-yr integration, and then the integration is repeated with Asia, the maritime continent, and Australia replaced by ocean. The sea surface temperature (SST) at each affected grid is specified as the SST at the first grid to its east that is an ocean grid in the first experiment. The latter integration shows that the monsoon circulation pattern over the would-be locations of south Asia and Australia, and the surrounding region, has largely remained. The results discount land–sea contrast as the main cause of the Asian monsoon. A third experiment is the same as the first (the control) except that the topography of Asia, the maritime continent, and Australia is reduced to zero. This experiment reveals that in the Asian summer monsoon, the difference between the first two experiments is generally due more to the removal of topography than to the removal of land–sea contrast. They show that the difference between the second and the third experiment is mainly in the longitudinal location of the maximum precipitation. Additionally, in the Asian and Australian winter monsoons land–sea contrast also plays only a modifying role. Although land–sea contrast plays only a modifying role in the Asian and Australian (and Central American) monsoons, additional experiments show that it is the main reason that ITCZ (and thus monsoon) exists in Africa and South America. Thus, monsoons can be classified into two groups depending on whether land–sea contrast plays a major role. It is also argued that the role played by land–sea contrast in the monsoon is basically equivalent to that played by SST contrast. Where land–sea contrast is important for the monsoon, the monsoon can still exist if the land is replaced by ocean of sufficiently high SST.

The model used, the Goddard GCM, is described and the experimental results are presented in section 2. Section 3 gives a brief qualitative explanation of why, according to our interpretation, the monsoon can be interpreted as an ITCZ located away from the equator in combination with its circulation field. As a result of this study, after minor modification to take into account land–sea contrast, the monsoon onset mechanism proposed by Chao (2000), which is based on ITCZs in an aqua-planet model with zonally uniform SST, can be applied to all places, not just western Pacific. Further discussions and a summary are offered in the last section.

2. Model and experiments

The latest version of the Goddard Earth Observing System general circulation model (version 2) is used. A 4° lat × 5° long grid size and 20 levels are used with 4 levels below 850 hPa. This model uses the discrete dynamics of Suarez and Takacs (1995). The relaxed Arakawa–Schubert scheme (RAS; Moorthi and Suarez 1992) is a main feature of the model. This scheme gives almost identical time-mean results as the original Arakawa–Schubert scheme at much reduced computational cost. The RAS is used in conjunction with a rain–evaporation scheme (Sud and Molod 1988). The large-scale moist and dry convection remain the same as documented in Kalnay et al. (1983). The boundary layer and turbulence parameterization, a level 2.5 second-order closure model, is that of Helfand and Labraga (1988). The long wave radiation package is that of Chou and Suarez (1994). The short wave radiation package is that of Chou (1992) and Chou and Lee (1996). The prognostic cloud water parameterization of Del Genio et al. (1996) is used. Sea surface temperature, sea ice, and snow are from observations (for details, see Takacs et al. 1994). Other features include land surface process parameterization (Koster and Suarez 1996) and gravity wave drag parameterization (Zhou et al. 1996).

The initial condition for the experiments is that of a January 1987 European Centre for Medium-Range Weather Forecasting analysis. Excluding clearly stated modifications, in all experiments the model was run for four years with observed boundary conditions, including the observed SST. Only the last 3 yr of the output were used for analysis. After the control run (Fig. 1), an experiment, E1, was run with all land grids between 60°E and 180° (except Antarctic) replaced by ocean and the SST of each affected grid is specified as that of the first grid on its east side that is an ocean grid in the control. Figures 2a and 2c show the August rainfall and 850-hPa wind arrows of E1 averaged over the last 3 yr. Clearly, in E1 the ITCZ, or the rainy region, associated with the Asian summer monsoon still exists. However, the precipitation region does not extend as far to the north (to cover regions where Tibet and southern China
FIG. 1. (a) Aug and (b) Feb precipitation and (c) Aug and (d) Feb 850-hPa winds averaged over last 3 yr of the control run.

were) as in the control run (Fig. 1a); this is very similar to the results obtained by Hahn and Manabe (1975) in a GCM experiment without mountains. The high precipitation in the Arabian Sea and the Bay of Bengal lessens in E1 and the precipitation pattern in the Indian Ocean and western Pacific becomes more zonally uniform. The circulation field associated with this ITCZ in E1 shows southwesterlies in the ITCZ region, cross equatorial flow in the Indian Ocean, and Somali jet in the low-levels (Fig. 2c), monsoon features observed in control (Fig. 1c) and observation. In the regions where Tibet and southern China would normally be located there is an easterly region at low-levels. This feature is hardly detectable in the control run. Figures 2b and 2d show the same plots for February. The ITCZ and the associated circulations in the western Pacific and Indian Ocean is shifted eastward. This shift is also apparent in the 850-hPa wind field (Fig. 3c) which shows that the westerlies associated with the ITCZ is shifted eastward changing from E1 to E2. The August precipitation in the southern Indian Ocean is somewhat reduced in E1 (comparing with the control) and is further reduced in E2. Also the easterly region at low level over Tibet and China is similar to that in E1 but extends further eastward into western Pacific. Moreover, the reduction of precipitation over Tibet and China remains. Thus the effect of removing Asian landmass is due much more to the removal of topography than to the removal of land–sea contrast, although in some small areas the opposite is true. However, one must be reminded of the fact that the land–sea contrast in E2 is not exactly the same as that in the control, since the topography of the two experiments is different.
Results of E2 in February show the similar results (Figs. 3b,d). They show that the northeasterlies in southern China, Indochina, and India, the central feature of Asian winter monsoon, remain in both E1 and E2. This reveals the relatively minor role of land–sea contrast in its impact on the Asian and Australian monsoons (in both summer and winter).

It has been observed that in the Indian summer monsoon region, the upper tropospheric meridional temperature gradient south of the Tibetan Plateau is reversed (Flohn 1957; He et al. 1987; Yanai et al. 1992). Our results also show such a feature. Figure 4a shows the mass-weighted average temperature between 200 hPa and 500 hPa in August averaged over the last three years of the control run. It shows a maximum over the Tibetan Plateau. This maximum still exists in E1 (Fig. 4b) but it is moved eastward and equatorward. In E2 (Fig. 4c) it is moved further eastward. In all three experiments the reversal of meridional temperature gradient in the upper tropospheric levels in the Asian summer monsoon region are clearly simulated. Thus it can be concluded that the heating due to the Tibetan Plateau, though contributing to the meridional temperature gradient reversal at these levels, is not a necessary condition for such reversal. In February the upper tropospheric meridional temperature gradient reversal in the Australian monsoon region also exists (not shown) but its magnitude is much smaller than that in the Asian summer monsoon.

Experiments similar to E1 and E2 are done for the cases of removing American and African landmass and topography. In the E1 type experiments when a land grid is replaced by ocean the SST is specified by linearly interpolating the SST’s of first ocean grids on east and west sides that are ocean grids in the control run. When interpolation cannot be done, SST is copied from one side as in E1. Figure 5 shows precipitation in the case of removal of the Americas. In August the precipitation region in Central America remains; but that in Mexico is reduced considerably. Figure 5b shows the February results. The precipitation region in South America disappears. The precipitation region in eastern Pacific just west of Central America is much enhanced. In both August and February results in the case of removal of topography in the Americas (Figs. 6a,b) are similar to that in the control. The enhancement in the eastern Pacific west of Central America remains. Thus it can be concluded the land–sea contrast plays a major role in the South American monsoon but not in the Central American monsoon.
For the Mexican monsoon both topography and land–sea contrast appear to be important, although a more definitive determination can not be made due to the coarse grid size used in the model. Figure 7 shows the resolution of the topography in the model. It is quite obvious that although the topography is better resolved in the Asian continent, it is not well resolved in Central America and Mexico. Thus the impact of topography on monsoon in the latter regions should be studied with a mesoscale model. The same thing can probably be said about the monsoons in Africa and South America.

Similar experiments of removal of the African (and Arabian) landmass and topography show that the land–sea contrast plays a major role in African monsoon. Figures 8a,b show the corresponding precipitation in August. In August without the African landmass, precipitation in the location previously occupied by Africa disappears and the precipitation in the Asian and Central American monsoons increases. Also, the Somali jet is barely discernible (not shown). This is consistent with the prevailing interpretation that the Somali jet depends on the topography in east Africa for its existence. In February, the results show (not shown) that the SST in the former region of southern Africa is high enough to keep some precipitation. In an additional experiment, the experiment of removal of Africa is repeated but the assigned SST of the affected grids is increased by 3°C. The results (Fig. 9) show that the August African monsoon reappears. The high strength of this monsoon indicates that the 3°C increase is excessive. This reveals that the role of land–sea contrast, when important, can be replaced by sufficiently high SST.

To summarize, our experimental results show that land–sea contrast plays a major role in monsoons in South America, Africa (excluding southern Africa), and Mexico; and a minor role in monsoons in southern Asia (including India), Australia, and Central America. Thus for land–sea contrast to play a major role in monsoon the landmass has to be sizeable and covers the latitude of ITCZ. Also, the role of land–sea contrast, when important, can be replaced by ocean of sufficiently high SST.

As the main conclusion of these experiments, the conventional interpretation for the origin of monsoons, which depicts land–sea contrast on the continental-scale as the main cause (implying landmass is indispensable), is problematic. In the following section our interpretation for the origin of monsoons will be presented.
3. Monsoon as off-equator ITCZ and its associated circulation

We need to reflect on what monsoons are and what causes them. Although the precise definition of a monsoon varies according to the personal preference of individual investigators, our definition is not far from the consensus. In a time (say, monthly) mean sense, a monsoon is a continental-size convective system in the tropics. When in the Northern Hemisphere it has, again in a time mean sense, a predominate southwesterly flow at the low levels converging toward the continental-scale precipitation region. At the upper levels the wind direction reverses; that is, the predominate flow is northeasterly. Thus there is a large vertical wind shear that is not found elsewhere in the Tropics. The low-level southeasterly flow can be traced back to a southeasterly flow in the Southern Hemisphere. In a Southern Hemisphere monsoon, everything is the same except that the meridional component of the winds reverses its direction. Thus a monsoon contains a sizable precipitation area and its associated circulation field. Convection (which results in precipitation) and the circulation field interact with each other. Thus it is not correct to say which one causes the other. It has been known for some time that a planet with constant sea surface temperature and constant solar zenith angle in longitude, latitude, and time can generate ITCZs (Suni 1992; Chao 2000; Kirtman and Schneider 2000). Chao’s experiments with such settings show that when RAS is used, a double ITCZ is obtained in the beginning and shortly after only one ITCZ at around 15° away from the equator remains. The meridional cross section of zonally averaged zonal and meridional winds associated with the ITCZ (Fig. 10) show the characteristics of monsoon circulation, that is, southwesterly flow at the low levels converging toward the precipitation area and large ver-
tical wind shear in the monsoon area. Therefore it can be concluded that the basic mechanism responsible for the existence of monsoons is earth’s rotation.

The monsoon circulation field in Fig. 10 is quite weak comparing with that of Fig. 3 of Chao (2000). The difference is that in the latter there is an SST peak at 30°N. The cumulus schemes used in the two experiments are different, but this difference has much a lesser impact on the strength of the monsoon circulation [an experiment similar to Fig. 3 of Chao (2000), but using RAS also shows a strong monsoon circulation.

To summarize, the existence of a landmass is not a necessary condition for the existence of monsoons. Under the settings of uniform SST and solar angle, earth’s rotation alone is sufficient to generate monsoon circulation. But such circulation is much weaker than the observed monsoon circulation. To get realistic strength, realistic pole-to-equator SST gradient, and an SST peak substantially (about more than 10°) away from the equator are necessary. Another factor that helps to give more realistic monsoon circulation strength is zonal asymmetry, which allows ITCZ to concentrate in a few longitudinal ranges. Zonal asymmetry is achieved by SST gradient in the zonal direction and by existence of landmass at or near the latitude of the ITCZ.

To a large extent the Gill (1980) solution of linear response of circulation field to an imposed stationary heating field on a β-plane gives a fairly accurate depiction of tropical circulation surrounding a heating source. However his heating field has a cosine function in its longitudinal distribution and thus does not depict the heating in the ITCZ well. This can be modified by taking a running mean of the entire Gill solution in longitudinal direction to get a flat top zonal profile in the heating field (Chao and Chen 1999):
where $f(x)$ is the Gill solution of $L = 2$, and $\varepsilon = 0.1$. Coordinate $x$ is the nondimensional length in the zonal direction with unity equal to about 1000 km. The zonal running mean turns the Gill heating field closer to an ITCZ heating field; that is, the zonal distribution of the heating field is now a cosine function with its top flattened and covers a much wider longitudinal domain. Figure 11 shows the solution for the asymmetric heating case (i.e., the zonal running average of Gill’s Fig. 3b). As explained by Chao and Chen (1999), the bulk part of the ITCZ heating area and the area to its immediate west are occupied by southwesterlies. This corresponds very well to the low-level southwesterly circulation field of the Indian summer monsoon. Quantitatively, the modified Gill solution, when given heating field of strength equivalent to what is observed, gives southwesterly of a magnitude comparable to the observed monsoon southwesterly. There is a cross-equatorial flow converging toward the ITCZ. Also there is an easterly region to the north of the ITCZ whose maximum wind speed is not quite as large as that of the westerly component of wind in the ITCZ. Moreover, the Gill solution assumes a wind direction reversal at upper levels. When Fig. 11 is flipped upside-down (such that the $x$-axis is still pointing to the east), the circulation field corresponds to that of an ITCZ to the south of the equator and there are northeasterlies north of the equator, a situation resembling the northeasterlies in the winter monsoon over India and Indochina when the ITCZ is just north of Australia. In summary the basic characteristics of monsoons such as a wide precipitation region, south-
westerlies (for Northern Hemisphere monsoons) covering the precipitation region and its neighborhood (northwesterlies in the southern hemisphere) at low-levels, cross-equatorial flow at low-level, and circulation reversal at high-levels are all found in the off-equator ITCZ. Consequently, in our interpretation of the origin of monsoons we equate monsoon with the off-equator ITCZ and its associated circulations. The origin of the ITCZ has been discussed by Chao (2000).

4. Discussions and summary

Although we have used model data of only 3 yr long, since the gross monsoon features are consistent in all 3 yr, we can consider our results more than enough to prove our points. In fact, the claim that land–sea contrast is the main cause for monsoons can be discounted, if the data of just 1 yr model run without land–sea contrast show monsoons.

The easterly zonal wind components in Tibet, southern China and the neighboring western Pacific in Figs. 2c and 3c, which are located north of the monsoon rainy region, correspond to the easterly zonal wind component north of the ITCZ in Fig. 11. Thus these easterlies can be considered part of the monsoon circulation. They appear to be interrupted by topography and cover only a very narrow range in Fig. 1c.

Our interpretation of the origin of monsoons is considerably different from the textbook interpretation. To detail our interpretation further, let us consider again the aqua-planet setting with zonally uniform SST. In summer the ITCZ is zonally uniform (in a time-averaged sense over at least 10 days, for example) and at low levels of the ITCZ region there are southwesterlies (for Northern Hemisphere monsoons) and winds flow across equator to converge to the ITCZ (see Fig. 3 of Chao 2000). Thus this combination of precipitation and circulation has all the characteristics of monsoon (a zonally uniform aqua-planet monsoon) and thus can be called monsoon. The circulation pattern on the ITCZ side of the equator is the summer monsoon, and that on the other side of equator is the winter monsoon. When zonally nonuniform SST is introduced into the setting, the ITCZ is no longer zonally uniform and prefers longitudinal ranges where the SST is high. Also when introduced, land, if covering the latitude of the ITCZ, is also a favorable location, if it is surrounded by oceans of relatively low SST. Thus in our interpretation land–sea contrast, similar to zonally nonuniform SST, merely makes a zonally uniform ITCZ (in an aqua-planet with zonally uniform SST) zonally nonuniform. Moreover, the circulation pattern associated with a zonally nonuniform ITCZ is well explained by the modified Gill solution as described in the preceding section.

Chao (2000) proposed a theory for monsoon onset based on numerical experiments on the ITCZ over an aqua-planet with zonally uniform SST. His theory identifies monsoon onset with the sudden jump of the latitudinal location of the ITCZ. He hypothesized that his theory is valid not only over the western Pacific, where the situation resembles aqua-planet, but also elsewhere in the tropics. Our present study, by the arguments that the monsoon is the off-equator ITCZ and that the ex-
istence of the ITCZ does not have to rely on land–sea contrast, and that land–sea contrast merely offers a favorable location for the ITCZ, gives support to his hypothesis. The two types of attraction on the ITCZ, due to earth’s rotation and latitudinal peak of zonally uniform SST (Chao 2000), still exist (with some modifications) when the presence of zonally nonuniform SST and land–sea contrast is taken into account. And his basic subcritical instability interpretation of monsoon onset does not change.

In summary, the monsoon is interpreted as an inter-tropical convergence zone (ITCZ) substantially away (more than 10°) from the equator and the existence of the ITCZ does not have to rely on land–sea contrast. Land–sea contrast can provide a favorable longitudinal location for the ITCZ but this role can be replaced by sea surface temperature contrast in the longitudinal direction. Our experiments with a general circulation model show that continental-scale land–sea contrast is not the main cause of the Asian (including Indian) and Australian summer monsoons by demonstrating that the summer monsoons in these regions still exist, when Asia, the maritime continent, and Australia are replaced by oceans. Also in the Asian monsoon the effects of the removal of continents are in general due much more to the removal of topography than to land–sea contrast. Nevertheless, further experiments show that land–sea contrast is crucial to the monsoons in northern Africa and South America. Also the role of landmass in the origin of monsoons, when important, can be replaced by ocean of sufficiently high SST. Our interpretation for the origin of monsoons is that the summer monsoon is the ITCZ and its associated circulation when the latitudinal location of the peak of the ITCZ is substantially (more than 10°) away from the equator. The origin of the ITCZ has been interpreted by Chao (2000). Also, the circulation associated with the ITCZ has been interpreted by Chao and Chen (1999) and a brief description of it was offered in the preceding section. This circulation has all the characteristics of summer monsoon circulation. The longitudinal location of the ITCZ is determined by the distribution of surface conditions. ITCZ favors longitudinal locations of high SST as in western Pacific, Indian Ocean, and oceans neighboring Central America or a tropical landmass, due to land–sea contrast, as in tropical Africa and South America. The ITCZ circulation extends to the tropics in the other hemisphere to give rise to the winter monsoon circulation there. Also through the equivalence of land–sea contrast and high SST, it is argued that the basic monsoon onset mechanism proposed by Chao (2000) is valid for all monsoons.

Note added in Proof: A repeat of experiment E1 with the SST of the converted grids reduced by 3° did not result in substantial changes in the Asian monsoon.

Acknowledgments. This work was supported by NASA Office of Earth Sciences.

REFERENCES

Takacs, L. L., A. Molod, and T. Wang, 1994: Documentation of the


