Chapter 1

THE ELEMENTARY HADLEY CIRCULATION

Peter J. Webster
School of Earth and Atmospheric Sciences and School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0340, U.S.A.

Abstract

In the most basic terms, the Hadley circulation can be thought of as a large-scale overturning of the atmosphere driven by latitudinal heating gradients, extending roughly between the Tropics of Cancer and Capricorn covering roughly half the surface area of the planet. Rising air occurs near the equator with subsidence in the subtropics. The circulation has a strong seasonal variability. It is manifested during the equinoxes as a pair of relatively weak cells with a common rising zone near the equator termed the Intertropical Convergence Zone (ITCZ). A much stronger cross-equatorial cell marks the soistitial seasons with rising motion in the summer hemisphere and widespread descending air in the winter hemisphere. The meridional circulations are instrumental in determining where tropical rainfall occurs and where the great deserts are located. Variability of the location and intensity of the Hadley circulation (or its regional manifestation such as the monsoons), through the ages has helped shape the history of mankind, either spawning regions of civilization by providing an abundance of rainfall for agriculture or destroying them by periods of drought. The variability of the Hadley circulation is also manifested on interannual times scales as an important component of the waxing and waning of El Niño in the Pacific Ocean, perturbing seasonal climates worldwide.

The Hadley circulation was the first phenomenon to be described by using the physical insight of the natural system emerging out of the Renaissance. Both Halley (1686) and Hadley (1735) provided basic accounts of the physical processes that drive the meridional cells. However, a detailed examination of the phenomenon, using data sets that are now available, shows that many questions cannot be answered in the confines of the Halley-Hadley model. For example, what limits the latitudinal extent of the cells? What is the role of the Hadley system in balancing the planetary heat budget? What factors determine the vertical scale of the Hadley circulation? Why is there considerable longitudinal variability in the strength of the circulation? How does the ocean interact with the atmospheric Hadley circulation and is there an oceanic counterpart?

An attempt is made to answer these questions from a fundamental physical perspective. It is found, for example, that the vertical transport of heat and the heat balance of the tropics in the ascending branch of the Hadley circulation are difficult to understand without considering "undiluted hot convective towers," first considered by Riehl and Malkus (1958). An explanation of the depth of tropical convection follows by consideration of the magnitude of the sea surface temperature (SST) and the stability of the tropical atmosphere. Furthermore, both the atmosphere and the ocean meridional cells contribute to the poleward transport of heat. In the atmosphere, it is the instabilities of the Hadley cell (the middle latitude eddies or waves in the westerlies) that complete the transport of heat towards the poles. It is shown that the atmospheric Hadley circulation drives an oceanic circulation that acts as a negative climate feedback. Finally, a simple model of the combined ocean-atmosphere system is presented that underlines the importance of both the oceanic and the atmospheric Hadley circulations in balancing the heat budget of the planet.

1. INTRODUCTION

In both hemispheres, steady and brisk winds flow across the subtropical oceans in a general westward equatorial direction, converging gradually towards the equator. These fair weather winds, referred to as the trade winds,[1] merge together in a more stormy and rainy environment in a region of low pressure called the equatorial trough. This area was termed the doldrums by early mariners because of the generally sluggish light winds, which provide only slow and uncertain progress or "tread" in addition to frequent storms and squalls. In the vicinity of the equator, the converging warm tropical air is forced to rise, and on reaching the upper troposphere it flows poleward until it reaches the subtropics, where it descends and flows towards the equator in the surface layers. The high humidity of the air rising over the equator results in the largest rainfall rate on the planet. Furthermore, the descending branch of the tropical cell is responsible for the vast dry regions of the subtropical oceans and continental areas, forming great deserts over land and the subtropical high-pressure belt over the oceans. In addition, the strongly seasonal monsoon circulations, which bring copious amounts of summer rainfall to over 60% of the global population, are regional manifestations of these meridional circulations. Collectively, this

[1] It is generally thought that the name "trade winds" has a commercial origin because of their use by earlier traders crossing the Atlantic Ocean for the Americas. However, Philander (1998) suggests that the name has a nautical origin, reflecting the steadiness of the winds coming from the word "tread" which refers to the steady path of a ship's progress.

broad-scale circulation pattern, which dominates the general circulation and covers nearly half of the surface area of the globe, is called the Hadley circulation.

To a significant degree, the Hadley circulation has determined where human civilization has grown and flourished. At the same time, temporal variations in the location and strength of the Hadley circulation may have been responsible for the demise of early civilizations in North Africa, Mesopotamia, and North America (e.g., Weiss and Bradley 2001). On a much shorter time scale, the circulation is intimately linked with the aperiodic El Niño/Southern Oscillation (ENSO) phenomenon, which brings drought or flood to different parts of the tropics on interannual time scales (Bjerknes 1966, 1969). Thus, in terms of its vast scale and its influence on the habitability of planet, the Hadley circulation is arguably the most important of all climate features. Yet, despite the fact that most scientists are familiar with the phenomenon, and while it is the climate feature one first identified as an entity, the physical processes that drive the Hadley circulation and determine its variability are not well understood at all.

With the importance of the Hadley circulation in mind, it is clear that a physical understanding of the system is essential. An absence of this knowledge makes it very difficult both to advance knowledge of the present climatology, thus limiting our ability to predict the variability of the present climate and weather, and to interpret proxy data laid down by past climates. It will be equally difficult in the absence of a thorough physical understanding to determine future configurations of planetary climate in the presence of changing constituents of the atmosphere and through the influences of longer-term solar variability.

1.1 Climatology of the Hadley Circulation

Despite a lack of a thorough physical understanding of the system, the fundamental forces that drive the Hadley circulation are well known. In fact, they are the same factors that drive all motions on the planet: pressure gradient forces arising from differential radiative heating between the equator and the poles and the rotation rate of the planet. Figure 1-1a shows the zonally averaged solar radiation absorbed at the surface of the planet (S, top panel), the long-wave radiation emitted to space by the planet (I_E, middle panel), and the net columnar radiation (R_{TOT}, defined as $S - I_E$, bottom panel) plotted as a function of latitude. Three averages (annual, ANN; boreal summer, June through August, JJA; and boreal winter, December through February, DJF) are plotted. The absorbed solar radiation possesses an extremely strong seasonal variability, with maximum values in the summer subtropics.

The variability of S is in sharp contrast to the distributions of emitted long-wave radiation (I_E), which is relatively constant in the tropics and subtropics but shows substantial seasonal changes at high latitudes. Overall, at all times

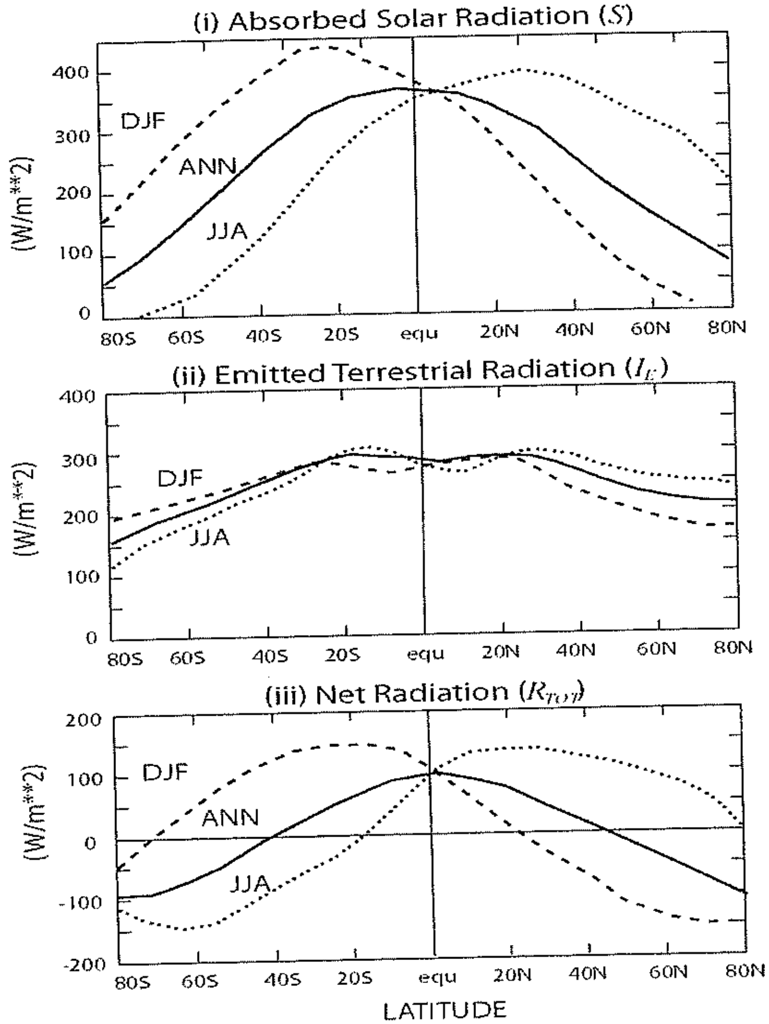

Figure 1-1a. Absorbed solar radiation (S, upper panel), emitted long-wave radiation at the top of the atmosphere (I_E, middle panel), and net radiation (R_{TOT}, bottom panel) as a function of latitude for the annual average (heavy solid line), boreal summer (June–August, JJA, dotted line) and boreal winter (December–February, DJF, dashed line). Note that in the equatorial regions the emitted terrestrial radiation stays essentially constant. The strong seasonal variability in the net radiation comes from solar radiation variability. Units W m^{-2}.

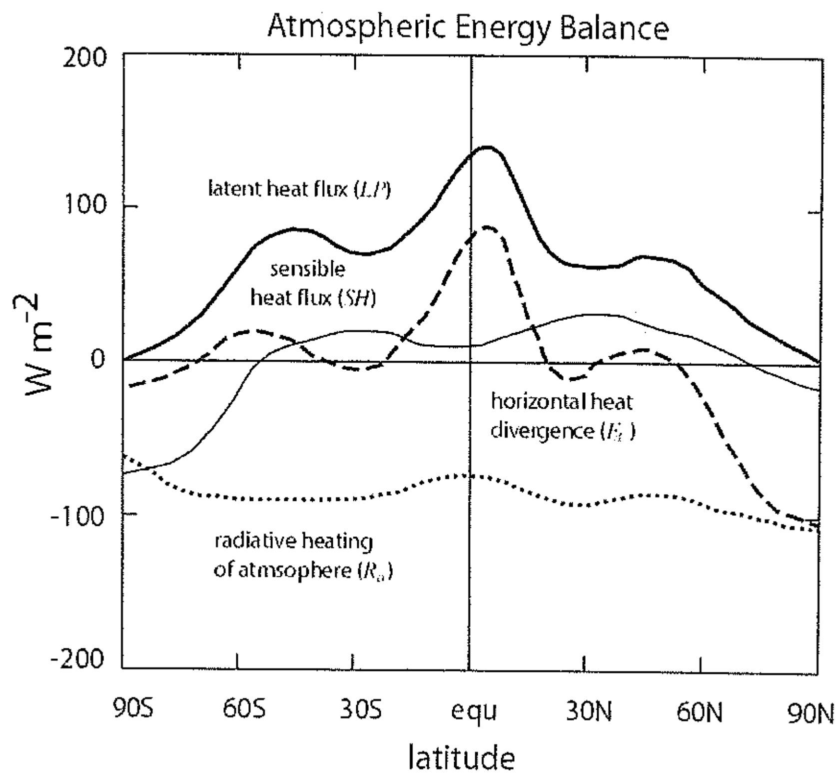

Figure 1-1b. Atmospheric energy budget showing heating due to condensation (LP, heavy solid line, where L is the latent heat of condensation and P is the precipitation rate), sensible heating (SH, light solid line), net radiative heating of the column (R_a, dotted line), and the advection of heat by atmospheric motions (F_d, dashed line). Units W m^{-2}.

of the year, there is net heating in the tropics and substantial cooling at higher latitudes, especially in the winter hemispheres. In general, the summer hemispheres show net heating between the equator and the poles and possess a much smaller equator-to-pole radiational heating gradient than in the winter hemispheres. However, it is important to note that while there exists net heating in the tropics and net cooling at higher latitudes, the long-term seasonal average temperature distribution remains relatively constant. This being the case, it is clear that there must be a net transport of heat between the tropics and higher latitudes and that this transport can only be accomplished by fluid motion in the atmosphere and the ocean forced by pressure gradient forces resulting from the radiational heating imbalances. Furthermore, as the winter and summer temperatures of both hemispheres remain much the same from year to year (ignoring for the moment possible global warming effects), the heat transports and radiational cooling to space that provide this long-term equilibrium must occur on subseasonal time scales.

To learn more about the processes that produce the long-term thermal equilibrium of the planet, it is useful to look closely at the columnar atmospheric heat budget. Figure 1-1b shows the latitudinal distributions of the components of the columnar atmospheric energy balance. The energy balance is given by $\Delta F = LP + SH + R_A$, where LP represents the release of latent heat through precipitation P, where L is the latent heating of condensation coefficient, S represents the sensible heating of the atmosphere, R_A is the net heating of the column (the difference between the net radiation at the top of the atmosphere and the surface), and ΔF is the net heating of the column. The units of all of the terms are W m^{-2}. The figure shows substantial heating from the condensation of evaporation at the surface and, to a lesser degree, from the heating of the atmosphere at the planetary surface. The heating occurring at any latitude is balanced by a combination of radiative cooling of the atmosphere to space and latitudinal transports of the residual energy by the fluid motions of the atmosphere and the oceans.

Figures 1-2a–d show latitude-height cross sections of the long-term average meridional stream function (ψ, lower panels) and the zonal wind component ($[\overline{u}]$, upper panels) for the boreal spring (March–May, MAM), JJA, the boreal fall (September–November, SON), and DJF, respectively. Following Peixoto and Oort (1992), the mass stream function ψ is defined by:

$$[\overline{v}] = \frac{g}{2\pi a\cos\varphi}\frac{\partial\psi}{\partial p} \quad \text{and} \quad [\overline{\omega}] = -\frac{g}{2\pi a^2\cos\varphi}\frac{\partial\psi}{\partial\varphi} \qquad (1)$$

where $[\overline{v}]$ and $[\overline{\omega}]$ are the climatological values of the meridional and vertical velocity components averaged around latitude circles. The bars in Equation (1) denote time averages and the square brackets zonal averages around latitude circles. The vertical velocity in Equation (1) is the time rate of change of pressure (dp/dt), and as pressure decreases with height a negative value of $[\overline{\omega}]$ denotes upward motion. The terms a and g represent the radius of the planet and acceleration due to gravity and φ is latitude, respectively.

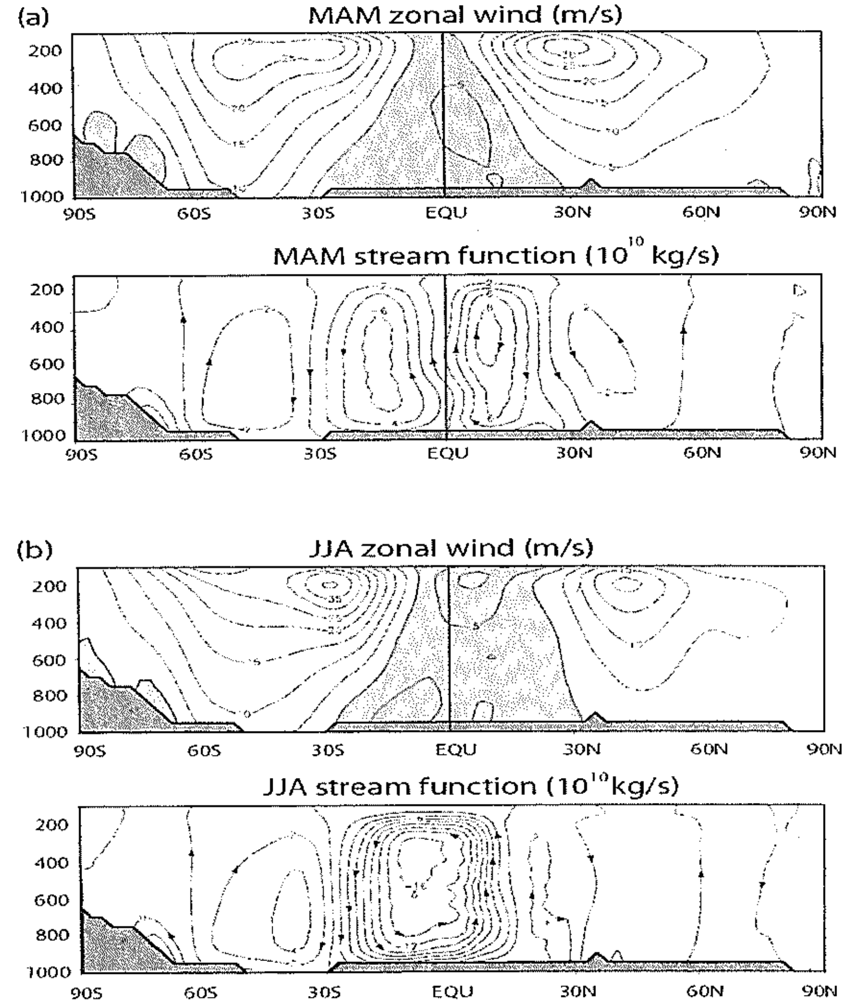

Figure 1-2. The zonally averaged circulation of the atmosphere showing zonally averaged zonal wind component ($[\overline{u}]$, upper panel, m s^{-1}) and mass stream function (ψ, lower panel, 10^{10} kg s^{-1}) for (a) March–May (MAM), (b) June–August (JJA), (c) September–November (SON), and (d) December–February (DJF), plotted against latitude and pressure (hPa). Zonally averaged elevation of the planet and negative (easterly) zonal winds are shaded. The stream function denotes the average flow in the latitude-height plane. Weaker Ferrel cells are evident in all seasons at higher latitudes. NCEP/NCAR reanalysis data were used to compute 50-year seasonal averages.

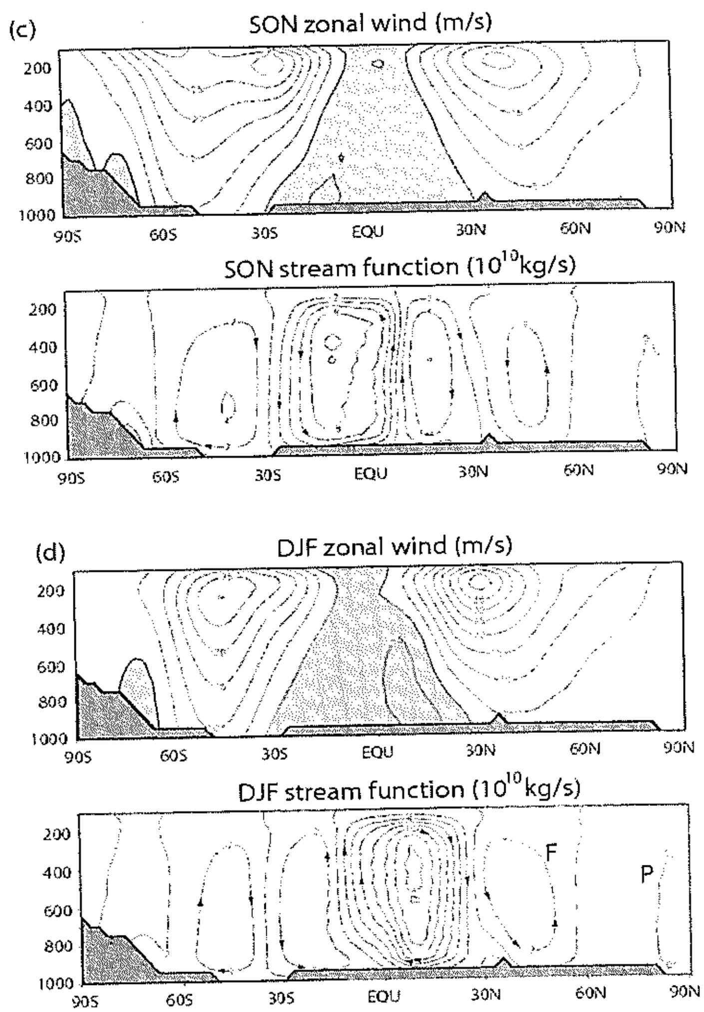

Figure 1-2. Continued.

The meridional mass stream function, as defined in Equation (1), shows the average trajectories of air parcels in the latitude-height plane. Note from Equation (1) that the vertical velocity is proportional to the latitudinal gradient of the stream function. Therefore, a tighter packing of the streamlines in latitude is indicative of a stronger vertical velocity. Similarly, a greater vertical stacking of the streamlines implies a stronger meridional

wind component. In total, the zonal velocity component $[\bar{u}]$ and the fields of ψ completely represent the time- and zonally averaged flow in the latitude-height plane. The equinoctial meridional circulations (lower panels of Figs. 1-2a and c) each possess a pair of meridional overturnings between the equatorial regions and the subtropics on both sides of the equator that combine near the equator to produce a relatively narrow region of ascent defining the mean Intertropical Convergence Zone (ITCZ). At all times of the year the vertical extent of the ascent, and thus the vertical scale of the Hadley circulation, is between 12 and 15 km. In the upper troposphere, air diverges away from the ITCZ region towards the poles. The character of the solstitial circulations is very different (lower panels of Figs. 1-2b and d) to that found at the equinoxes. The relatively weak dual cells are replaced by much stronger single cells that straddle the equator, with ascent in the summer hemisphere and descent in the winter subtropics. The vertical velocities of these solstitial cells are roughly twice the magnitude of equinoctial values. In all seasons, weak reverse cells occur in both hemispheres at higher latitudes, but these are faint replicas of the tropical cells.

The zonal wind distributions (top panels, Figs. 1-2a–d) are divided into broad regimes: lower tropospheric easterly regimes in the tropics (the trades) and broad westerlies with maximum strength in the upper troposphere in the mid-latitudes near 300–200 hPa levels. In the solstitial seasons the stronger of the two westerly jet streams is located in the winter hemisphere poleward of the strong upper tropospheric cross-equatorial flow and strong winter hemisphere trades. In contrast, during the equinoctial seasons the mid-latitude westerly jets and the low-level trades are fairly symmetrical about the equator.

Figure 1-3 plots the meridional distribution of rainfall over the oceans and the continental regions separately for DJF and JJA. The two distributions are remarkably different, particularly in the relative positions of the maxima. Precipitation over land occurs much more poleward than over the ocean, suggesting the strong influence of the heated continents in the summer hemispheres. In contrast, the maximum precipitation over the oceans appears to follow the annual variation of the sea surface temperature (SST)[2] (bottom panel). Noting that the precipitation maxima are generally synonymous with the ascending part of the Hadley cell, the large differences in the locations of the land and ocean precipitation maxima suggest that there is substantial variation of the circulation in different longitude bands.

[2] There is an exception to this rule. Except in times of El Niño, the ITCZ, and the warmest SST in the East Pacific, resides in the Northern Hemisphere during both the boreal summer and winter.

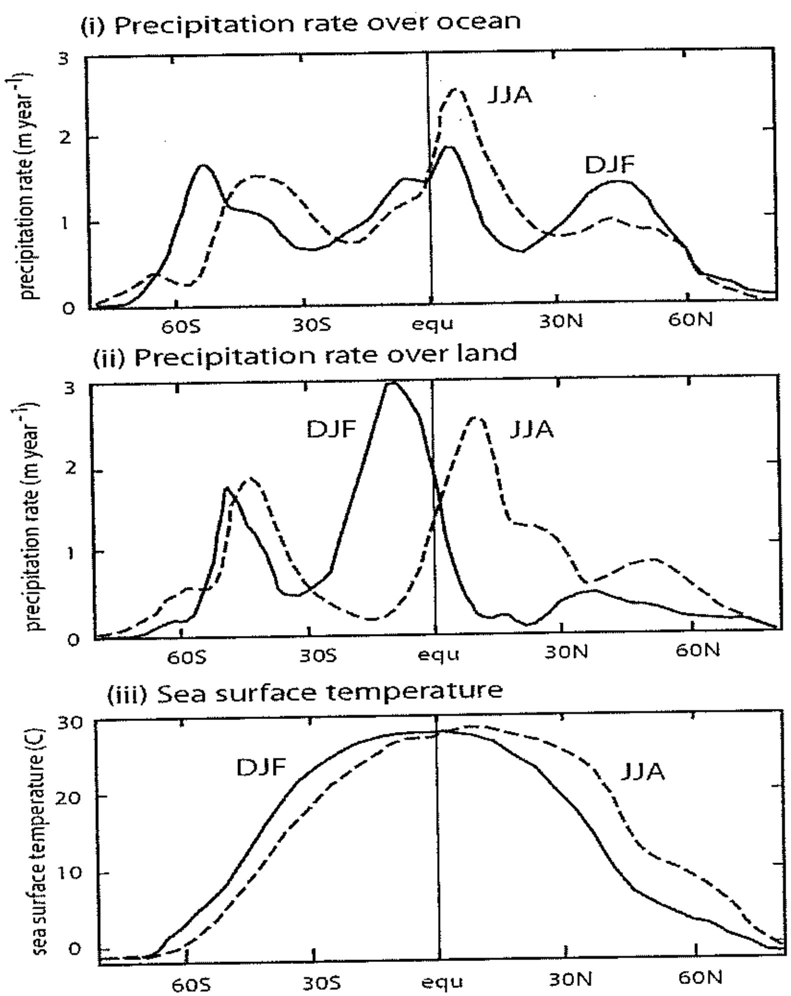

Figure 1-3. Precipitation rate (P, units m year^{-1}) over the global oceans (i) for JJA and DJF and land areas (ii). Ocean precipitation possesses a maximum in the Northern Hemisphere in both seasons due to the much colder water in the Southeast Pacific, which does not support strong convection in either season. Warmer SSTs (panel iii) in the Northern Hemisphere are consistent with increases in precipitation rate during JJA. Precipitation over land occurs at higher latitudes, reflecting the impact of heated continents and the monsoon circulations. Rainfall plots adapted from Jaeger (1976).

1.2. A List of Questions

The climatological features of the Hadley circulation described above are well known and appear in many textbooks (e.g., Peixoto and Oort 1992), with seasonal variability of the Hadley circulations being described in terms of radiative heating and angular momentum arguments. Yet, there are a number of characteristics of the Hadley circulation that are rarely addressed and which raise some interesting questions. For example:

(1) Why is the Hadley circulation limited in latitudinal extent? Why doesn't it extend all the way to the poles rather than being constrained in its location to the tropics and the subtropics?

(2) What determines the vertical scale of the Hadley circulation? Why isn't it 2 or 50 km rather than the observed 12–15 km?

(3) What factors determine the location of the near-equatorial precipitation or, more fundamentally, the region of ascending motion in the Hadley circulation?

(4) Why is the Hadley circulation so much stronger in the solstitial seasons compared to the equinoctial seasons?

(5) Why is there longitudinal variability in the strength of the Hadley circulation?

(6) Is there an oceanic counterpart of the Hadley regime? If so, is it a component of the overall system such that the Hadley cell may be thought of as a coupled ocean-atmosphere phenomenon?

These are essential questions that require consideration if we are to understand the nature of the Hadley circulation and, by extension, the vital components of the ENSO phenomenon and the monsoons.

We begin the investigation of these six questions by discussing the primary physical processes that create the Hadley cell. A historical perspective is developed starting with discovery of the circulation over 300 years ago. The successes and shortcomings of these early explanations are discussed. The dynamical constraints that determine the vertical and horizontal extent of the circulation are explored and the factors that determine where precipitation occurs are elucidated. It is argued that the monsoon circulation may be considered as a modified Hadley circulation. Finally, we develop the concept of an "oceanic Hadley-like circulation" and discuss its cooperative e, teamed with the atmospheric Hadley circulation, in balancing the heat dget of the planet.

PHYSICAL NATURE OF THE HADLEY CIRCULATION

1. Early Explanations

The first attempts to explain the physical nature of the tropical cli- ate came from two scientists: Sir Edmund Halley (1656–1742) and Sir eorge Hadley (1685–1786), both of whom were intrigued with the weather servations accumulating from around the world. For over 200 years, ariners had kept careful records of their shipping and trading routes and corded in some detail the wind, weather, and state of the ocean. These ob- rvations were of such widespread interest and economical value that com- lations were published and made available to mariners setting out on mmercial ventures around the world (Royal Society 1699; Kutzbach 987). When these observations were viewed collectively, scientists such as alley and Hadley found persistent features of regional weather that showed oth strong seasonality and year-to-year repeatability. First, Halley (1686) nd then Hadley (1735) attempted to find physical rules that would explain e patterns of wind and weather they had uncovered.

Utilization of the concept of seasonality of the global wind systems ertainly predated the expansion of European commerce in the fifteenth and ixteenth centuries. For example, the peoples of the Pacific Ocean traded ver vast distances for centuries prior to the arrival of the Europeans, mak- ng use of the trade winds (e.g., Kirch 2000). There is also much historical vidence for the use of the seasonally reversing monsoon winds as the cen- erpiece of flourishing trade between South Asia and East Africa and hroughout most of the Indian Ocean basin as early as 4,000 BP (Warren 987). These same trade routes, anchored in timing to the annual cycle, are till utilized to the present day. What made the efforts of Halley and Hadley o special, though, was that they were the first to apply the emerging under- standing of physical concepts, developed during the Renaissance, to the problem of explaining the direction and seasonal variation of the tropical wind systems.

Sir Edmund Halley was a mathematician, physicist, and astronomer of considerable renown, remembered most for his work on the orbits of

comets.[3] However, it is arguable that his astronomical work pales in comparison to some of his other scientific contributions. He provided an astronomical basis for the determination of longitude at sea, a problem of enormous consequence at the time, even though his method was difficult to implement. He solved the problem of representing three-dimensional quantities in two-dimensional form by introducing the concept of contours or isogenic lines. The first chart provided a global view of magnetic variation from true north, providing a method of correction for compass readings made by mariners. He used the same construct to produce the first surface wind climatology. Figure 1-4 shows climatologies of the Atlantic and the Indian Oceans from Halley (1686).

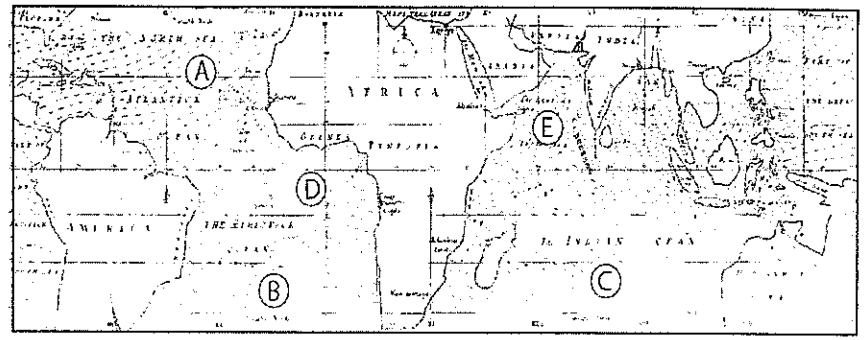

Figure 1-4. Halley's original map of the surface winds in the Indian and Atlantic Oceans. The map accurately represents the constancy of the trade winds throughout the year in the North and South Atlantic Ocean (A, northeast trades; B, southeast trades) and the South Indian Ocean (C, southeast trades). It also depicts where the winds reverse seasonally off the coast of West Africa (D, African monsoon) and in the North Indian Ocean (E, Indian monsoon). Adapted from Halley (1686).

From a meteorological point of view, Halley's most enduring work is an identification of the basic forcing mechanisms that drive the trade winds and the monsoon circulations shown in Figure 1-4. It was the ". . .

[3] Halley did forecast the return of the comet that bears his name. In fact, it has reappeared on time four times since his initial forecast, although the first appearance occurred after his death. In addition to the scientific works described briefly here, he was also a diplomat, a professor of geometry at Oxford University, and a captain in the Royal Navy. In addition to all of these accomplishments, he performed one task of enormous importance. He convinced Isaac Newton to consolidate his scientific research and publish *Philosophiae Naturalis Principia Mathematica*. And, in fact, he even paid for the publication costs (Bryson 2003).

action of the Suns Beams upon the air and water . . ." which produced a dynamic effect on the property of the overlying surface such that ". . . according to the Laws of Staticks, air which is less rarified or expanded by heat [i.e., over the colder regions] . . . must have a Motion towards those parts . . . which are more rarefied [i.e., air over warmer regions] . . . to bring it to an Equilibrium . . . ." (Halley 1686). Essentially, Halley had described the dynamic forcing of air by differential radiative heating. Applying this general principle to the observed surface winds, Halley went on to hypothesize that ". . . as the cold and dense Air, by reason of its greater Gravity, presses on the hot and rarified [i.e., low-level convergence of air resulting from differential heating] . . . [it] must ascend . . . and being ascended it must disperse to preserve the Equilibrium [i.e., upper air divergence to conserve mass] by a contrary current which must move from those parts where the greatest heat is: So that by a kind of Circulation, the North-East Trade below, will be attended by a South-Westerly above, and the South-Easterly with a North West Wind above . . . ." .Halley had managed to describe the manner in which differential heating would result in fluid motion. Although he appeared to omit an explanation of why there was an easterly component of the surface trade wind regime, the anticipation of an upper-level return flow was a remarkable achievement predating upper-air observations by almost 300 years.

A more complete explanation for the orientation of the tropical wind fields was left to Sir George Hadley, a lawyer and meteorologist, after whom the tropical meridional circulation was finally named. Noting that ". . . the causes of the general trade-winds have not been fully explained by any of those who wrote on that subject . . . " Hadley (1735) invoked the impact of the ". . . diurnal motion of the Earth . . ." as the factor that determined the directions of both the surface and upper tropospheric winds. He noted that ". . . setting aside the diurnal motion of the Earth, the tendency of the air would be from every side towards that part [of the planet] where the Sun's action is most intense at the time, and so a N.W. [northwesterly] wind would be produced in the morning and a N.E. [northeasterly] in the afternoon, by turns, on this side of the parallel of the Sun's declination, and a S.W [southwesterly] and S.E. [southeasterly] on the other . . . ." Hadley continued his deduction by supposing the atmosphere was at rest relative to the rotating planet and letting the air be forced to move from higher latitudes to the equator by the same forces proposed by Halley. This exercise allowed him to conclude that ". . . as the surface of the Earth at the equator moves so much faster [i.e., greater tangential velocity] than at the Tropicks [from which it follows that] . . . the air as it moves from the Tropicks towards the equator, having less velocity than parts of the Earth it arrives at, will have a relative motion contrary to the diurnal motion of the Earth, which combined

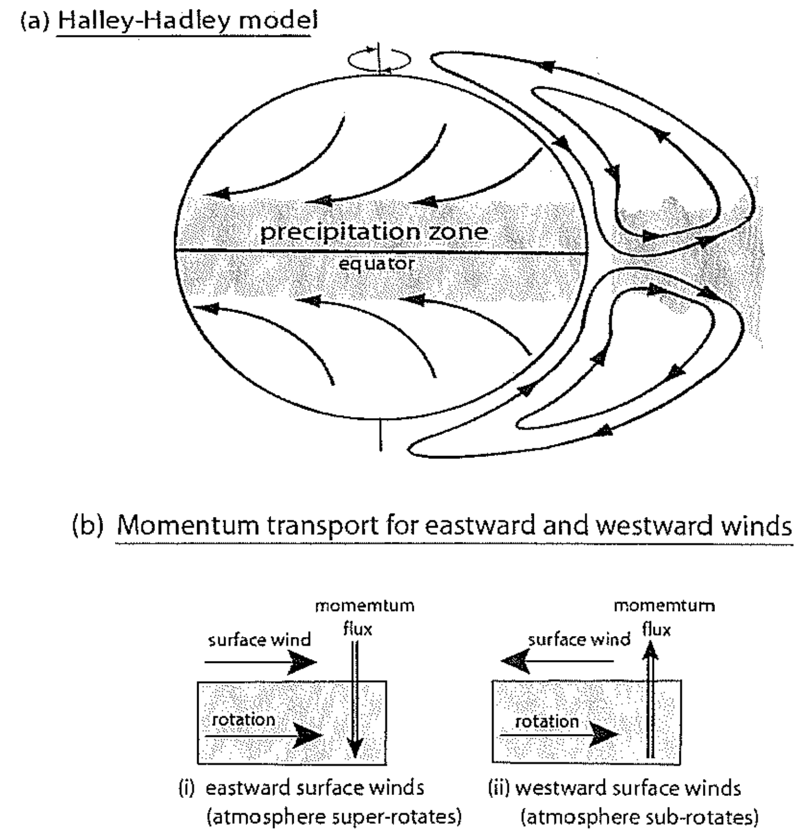

Figure 1-5. (a) The Halley-Hadley model showing one direct cell in each hemisphere. Onl one precipitation zone would be produced by this system, located in the vicinity o the equator. The impact of rotation is to turn the equatorial surface flow towards th west over the entire globe. The resultant easterly winds would mean that the at mosphere was rotating at a rate less than the rotation of the planet (i.e., sub rotating). (b) Schematic diagram of the transfer of momentum between the atmos phere and the solid earth for westerly (eastward) and easterly (westward) winds. I the winds over the entire planet were one-signed (as in panel a), the transfer o momentum would also be of one sign. Conservation of angular momentum woul insist that the atmosphere would accelerate at the expense of the solid earth (or vic versa). To account for the observed quasi-steady state of the kinetic energy of th atmosphere, and to explain the existence of extratropical westerlies, a series of ad ditional cells were proposed.

with the motion towards the equator, a N.E. wind will be produced on thi side of the equator, and a S.E. on the other . . . becoming stronger and mor and more easterly and be due east at the equator . . . ." Noting that the theoretical winds, so produced, were much stronger than observed, he invoked frictional dissipation at the surface to reduce its speed. Hadley also extended the argument to show the upper tropospheric return flow, noting that the raised air at the equator had a greater tangential speed than at higher latitudes and then, under the action of Halley's forces, the air would have an ever increasing westerly component as it moved towards the poles. Noting that the ascending air at the equator will move poleward, he used the same argument to explain the generation of upper atmospheric westerlies. A schematic of the Halley-Hadley model is shown in Figure 1-5a.

Together, Halley and Hadley described the physical essence of the mean meridional circulations of the tropics.[4] Halley (1686), by extension, also provided an explanation for the regional monsoons and the role of differential heating between the land and the ocean in producing seasonal reversibility of the winds. In effect, Halley's explanation also extends to zonally asymmetric circulations along the equator as being forced by longitudinal gradients of thermal forcing. An example of such a circulation is the surface flow along the equator, which converges in the western Pacific Ocean and rises in the vicinity of the Indonesian Archipelago, forced primarily by heating gradients associated with the ocean warm pool in the western Pacific Ocean and the colder upwelling waters of the eastern Pacific Ocean. These are the circulations responsible for the macro-scale correlations found in atmospheric variables by Walker (1924, 1928) such as the Southern Oscillation. The zonal circulations were later called the "Walker circulations" (Bjerknes 1969).

2.2. Higher-Latitude Issues of the Halley-Hadley Theory

While the Halley-Hadley theory explains a significant portion of the variance of the mean seasonal state of the tropical atmosphere, it does not explain the entire climate state of the higher latitudes. The theory calls for two cells, each occupying an entire hemisphere, with rising motion occurring near the equator and descent over the rest of the planet, in essence providing one maximum in precipitation near the equator. In contrast, Figure 1-2 shows that the Hadley cell is confined to the tropics. Figure 1-3 shows, in

[4] Perhaps a fairer name for the circulation would have been the Halley-Hadley circulation. After all, Hadley's application of the conservation of angular momentum rested completely on Halley's explanation of the thermal forces that produce an equatorward motion of surface air parcels and his understanding of the conservation of mass. It is not known to this author who was first to call the circulation the "Hadley circulation."

addition to precipitation aligned with the rising branches of the Hadley cell, maxima in the extratropics as well. Furthermore, if a Hadley cell is driven just by surface pressure gradient forces moving air towards the equator, then all surface winds would have an easterly component as they converge towards the equator, while the return flow would be completely westerly as air moves poleward. Three questions remain: (1) How can we explain the surface westerlies at higher latitudes within the confines of the Halley-Hadley theory? (2) Why don't the observed cells extend all the way to the poles? and (3) What causes the secondary precipitation maxima in the extratropics? The first question was answered to a large degree by Hadley (1735). Answers to the second and third questions depend on the instability of the Hadley circulation itself.

The law of conservation of angular momentum requires that the total angular momentum of the earth system remains constant. That is, the combined angular momentum of the atmosphere, solid earth, and oceans is fixed[5]. Because of frictional effects it is possible for momentum to be transferred between the atmosphere and the ocean/land interface. If the surface winds were to have an easterly component everywhere, the atmosphere would gain westerly momentum from the solid earth/ocean at all points on the globe. That is, the atmosphere would gain angular momentum while the solid earth/ocean would lose angular momentum as shown in Figure 1-5b. However, as the length of day (a measure of the angular momentum of the solid component of the system) remains constant within narrow bounds, as does the integrated kinetic energy of the winds, then the time-average net transfer of angular momentum between the atmosphere-ocean-land and the atmosphere must equate to zero. This balance can occur only if there are surface westerlies as well as easterlies.

The necessity for the existence of both easterlies and westerlies at the surface of the planet can be proved by using a very simple argument. Consider a featureless planet without topography. Assume that the average eastward force per unit area by the atmosphere on the surface of earth at some latitude ϕ is given by $F(\varphi)$ that may be written as $F(\varphi)=-\alpha[\overline{u}]$ where α is some frictional coefficient. Then the total torque on the atmosphere about earth's axis is:

$$aF(\varphi)\cos\varphi = -a\alpha\left[\overline{u}\right]\cos\varphi \qquad (2)$$

5 Strictly, the moon is also part of the earth system and should be included in our argument. However, for simplicity, we ignore the moon, noting only that through tidal friction effects that the radius of the orbit of the moon and rotation rate of the planet are slowly changing but on time scales that are not of concern to us here.

Noting that the area of a zonal strip between latitudes φ and $\varphi + d\varphi$ is $2\pi a^2 \cos\varphi \, d\varphi$, we can write an expression for the total torque in a strip of width $d\varphi$ as:

$$2\pi a^3 F(\varphi)\cos^2\varphi \, d\varphi \tag{3}$$

The total torque of the surface winds on the solid earth can be calculated as the integral of expression (3) across all latitudes. For long-term equilibrium the total torque must equal zero—i.e.,

$$\int_{-\pi/2}^{\pi/2} F(\varphi)\cos^2\varphi \, d\varphi = -\int_{-\pi/2}^{\pi/2} \alpha\left[\bar{u}\right]\cos^2\varphi \, d\varphi = 0 \tag{4}$$

Between Equations (2) and (4) and noting that $\cos^2\varphi$ and α are positive, it is clear that for the net torque to be zero, not only must $F(\varphi)$ change sign but the net torque exerted by westerly winds must equal the net torque by the easterly winds. So there must be both surface easterlies and westerlies. Therefore, a pair of Hadley cells with surface easterlies everywhere and westerlies aloft (Fig. 1-5a) cannot produce a wind system that remains constant in time.[6]

It is clear from Figure 1-2 that the Hadley circulation does not extend beyond the subtropics, in contrast to the Halley-Hadley model shown in Figure 1-5a. To explain the surface westerlies, it was argued that a second cell was required of the opposite sign. This is the mid-latitude Ferrel cell, marked as "F" in Figure 1-2d. The Ferrel cell has a region of common descent with the Hadley cell in the subtropics and rising motion at higher latitudes close to the extratropical precipitation maxima shown in Figure 1-3. It was conjectured that the turning of the poleward lower tropospheric wind was the reason for the prevalent surface westerlies. Weak polar cells ("P" in Fig. 1-2d) can also be identified. These polar cells were used to explain the weak polar easterlies. In total, six cells (two Hadley, two Ferrel, and two polar cells) were needed to describe the mean structure of the atmosphere.

Hadley (1735) had a simpler idea. He noted that, "The same principle as necessarily extends to the production of West Trade-Winds without

[6] In this argument we have assumed a featureless planet. If there are mountains, then the torque that the mountains impart on the atmosphere must be taken into account. Thus, in Equation (4), there would be an extra term on the left-hand side of the equation. In the Northern Hemisphere the frictional torque imparted by mountains is about 25% of the total frictional torque, but comparatively much smaller in the Southern Hemisphere (Peixoto and Oort 1992).

the Tropicks. . . ." [i.e., the extratropical westerlies]. Noting that the heated air at the equator will rise ". . . to make room for the air from cooler parts . . ." it will ". . . spread itself abroad over the other air and so its motion must in the upper regions must be to the north and south of the equator. Being got up at a distance from the surface of the Earth, it will soon lose great part of its heat and therefore acquire density and gravity to make it approach the surface again . . . ." As was explained earlier, this subsiding air will have a strong westerly component so that the surface ". . . thereby become a westerly wind . . . ." Hadley's arguments are seductive. They are consistent with the need for there to be surface easterlies and westerlies on the surface of the globe in order to balance the eastward and westward torques. Furthermore, they remove the necessity of having to explain the existence of surface westerlies by the Coriolis turning of the weak Ferrel cell poleward flow in the mid-latitudes.

However, neither the Ferrel cell nor Hadley's concept of descending westerly momentum can explain the latitudinal limitation of the Hadley cell or the existence of eastward propagating transient systems that dominate the weather of higher latitudes. Clearly, other arguments are necessary to explain the observed nature of the general circulation.

2.3. The Hadley Circulation and the Poleward Transport of Heat

During the last few decades it has become clear that the indirect Ferrel cell is merely a by-product of the very strong poleward transport of energy by the waves in the westerlies. These waves result from the process of baroclinic instability that is, in essence, the instability of the Hadley circulation itself. Charney (1947) and Eady (1949) independently developed the concept of baroclinic instability. Baroclinic instability arises from the rapid growth of perturbations in an environment where the basic flow varies in a particular manner in the environment. Of particular importance is the variation of the zonal flow $[\bar{u}]$. It is found that if positive shear exists (e.g., surface westerlies or easterlies surmounted by stronger westerlies) and if it is located sufficiently far from the equator, wave-like disturbances will grow rapidly that receive their energy from the basic flow itself. Using basic flows similar to what is observed in the subtropics, Eady and Charney showed that the fastest growing modes were the same scale as those observed in the extratropics: the familiar transient waves associated with extratropical weather. If the shear is strong enough, small motions within the flow will grow rapidly, transferring potential energy to kinetic energy. As they grow at the expense of the background flow, they reduce the temperature gradient and, as all instabilities do, tend to stabilize the system. During this process of stabilization, the waves transfer heat and momentum polewards, allowing the energy balance of the planet to be achieved.

Considered in tandem, the Hadley circulation and the extratropical baroclinic waves are synergetic. The Hadley circulation arises as a means of transferring excess tropical heat towards the poles in an attempt to achieve planetary thermal equilibrium, as was discussed earlier. As the planet is rotating, the upper tropospheric poleward flow becomes increasingly westerly as latitude increases. For the Hadley circulation to accomplish the necessary poleward heat transfer, the meridional temperature gradient must increase to values much larger than are observed. The well-known thermal wind equation:

$$\frac{\partial[\bar{u}]}{\partial z} \approx -\frac{g}{f[\bar{T}]a}\frac{\partial[\bar{T}]}{\partial \varphi} \qquad (5)$$

shows the consequences of an increasing latitudinal temperature gradient. Here $[\bar{u}(z,\varphi)]$ is the zonally averaged zonal wind component; $[\bar{T}(z,\varphi)]$ is the temperature of the atmosphere; and g, f, and a are the gravitational constant, the Coriolis parameter ($2\Omega\sin\varphi$ where φ is latitude), and planetary radius, respectively. The equation shows that the vertical shear ($\partial[\bar{u}]/\partial z$) will increase as the latitudinal temperature gradient (i.e., as $\left|\partial[\bar{T}]/\partial\varphi\right|$ increases). This is the vertical shear in the westerly winds to which Eady and Charney referred and which is unstable to small perturbations. Viewed broadly, one can think of the baroclinic waves as more efficient transporters of heat than the Hadley circulation, tending to stabilize the flow and reduce the meridional temperature gradient and the vertical shear. Also, rising motion in the low-pressure sector of the baroclinic waves produces the mid-latitude precipitation maxima seen in Figure 1-3.

Two factors could alter the shear and possibly the location of baroclinic waves: the rotation rate ($\Omega\ s^{-1}$ in the thermal wind equation 5) and magnitude and form of the factors that might alter the latitudinal temperature gradient. Hunt (1979) used a general circulation model (GCM) to test the sensitivity of Hadley circulation to rotation rate, varying it upwards and downwards by a factor of 5. The Hadley circulations produced in these experiments were found to be tightly bound about the equator for higher rotation rates, but more expansive in latitude than they are at present for slower rotation rates. In addition, the band of extratropical waves also moved their location correspondingly equatorward. These relocations are consistent with larger values of vertical shear being closer to the equator, which is evident from Equation (5). Although changes in rotation rate on the scale considered by Hunt have not occurred in the more recent history of the planet, the experiments illustrate the role rotation plays in determining the structure of the tropical climate.

The second possible factor, the alteration of the latitudinal temperature gradient, allows for more plausible scenarios. Changes in the latitudinal distribution of albedo were probably associated with the oscillation of high-latitude ice margins and variations of vegetation between the ice ages and the interglacials. Such variations may influence the magnitude and location of $\left|\partial[\bar{T}]/\partial\varphi\right|$ and with it the location of the zone of extratropical disturbances. For example, Shin et al. (2003) used a coupled general circulation model to study the climate at the last glacial maximum (LGM) and found a stronger Hadley cell, which was attributed to the stronger latitudinal SST gradient, and enhanced precipitation in the extratropical storm tracks. However, the stronger Hadley circulation appears at odds with interpretations of proxy data for the last glacial maximum that suggested much drier tropics, at least in the western Pacific Ocean/Indonesian region (e.g., Webster and Streten 1978). Probably more thought and more numerical experiments are warranted to explore this important topic.

2.4. Disposition of Energy in the Hadley Circulation

In addition to the important role the Hadley circulation plays in transporting heat poleward, it is a critical player in the vertical transport of heat in the tropics. The heat balances displayed in Figure 1-1 might suggest a rather simple energy balance picture. An excess of heating occurs in the tropics (because of greater solar heating compared to long-wave cooling) and a deficit occurs at higher latitudes (greater long-wave cooling than solar heating). Balance is accomplished by oceanic and atmospheric meridional heat transports. Figure 1-6a shows a schematic of the processes listed above, including estimates of their respective magnitudes. Balancing the heat budget of the tropics is complicated because heating and cooling of the tropical atmosphere take place at different vertical levels. The three main heating and cooling processes are:

(1) *Radiative heating*: As the atmosphere is largely transparent to incoming solar radiation, most of the energy that arrives at the top of the atmosphere is reflected by clouds or the planetary surface, absorbed in the upper layers of the ocean or close to the surface of the

land areas, and, at the long-wave end of the solar spectrum, absorbed by atmospheric water in ice, liquid, or vapor form. About 50% of the incident radiation at the top of the atmosphere is absorbed at the planetary surface. Solar heating of the atmosphere occurs indirectly by the transfer of heat at the surface through the turbulent transfer of sensible heat (generally large over land and smaller over the oceans) and latent heat through evaporative processes (generally smaller over land than the oceans). The atmospheric column is also heated by long-wave radiation from the surface that is offset to a large degree by the heating of the surface by down-welling atmospheric long-wave radiation. The latent heating is not realized immediately except as an increase in the specific humidity of the boundary layer.

(2) *Condensational heating:* As the moist air converges towards the equator under the action of pressure gradient forces, it is forced to rise and as it does so cools, condensing water vapor and producing deep convective clouds. Overall, the trade wind regime acts as a vast solar collector, with most of the energy being released in the ITCZ. In this manner, the ITCZ acts as the boiler box of the tropical heat engine. Through these processes, solar heating of the surface has relocated, in effect, to the upper troposphere where it can be radiated efficiently to space, be advected to higher latitudes, or some combination of both processes.

(3) *Radiative cooling:* With little water vapor or other greenhouse gases in the atmosphere above the tropopause, the planetary system can cool rather effectively to outer space. From Figure 1-1, we have noted that the cooling to space is fairly constant seasonally and decreases from the equator to the poles by about a factor of 2. It is interesting to observe that the heat loss to space by radiative processes in the upper troposphere is about two and a half times the net heat transferred poleward by atmospheric motions. The factors that determine the ratio of the two forms of heat loss by the tropics turn out to be determined by how quickly the atmosphere of a planet cools compared to how quickly heat can be moved from one location to another by atmospheric and oceanic motions.

Consider the situation where the tropical atmosphere transfers heat upwards in clouds as shown in Figure 1-6a but where this heat is lost at a rapid rate to space through radiative cooling. In this situation, none is available for transfer to higher latitudes. On the other hand, consider the same situation but now assume that the dynamical transports are so strong that an atmospheric parcel does not have time to cool radiatively. In the first instance, the atmospheric column would be said to be in radiative equilibrium with outer space. In the second instance, dynamics completely overshadow radiative effects. These effects may be termed radiative and dynamic limits, respectively. One might ask, Why aren't all of the atmospheric columns between the atmosphere and the pole in radiative balance, thus negating the need for the horizontal transfer of heat? Clearly on earth this radiative "solution" does not occur, although there are examples in the solar system where the conditions are nearly met. To determine the situation on earth, we need expressions for the radiative and dynamical time scales.

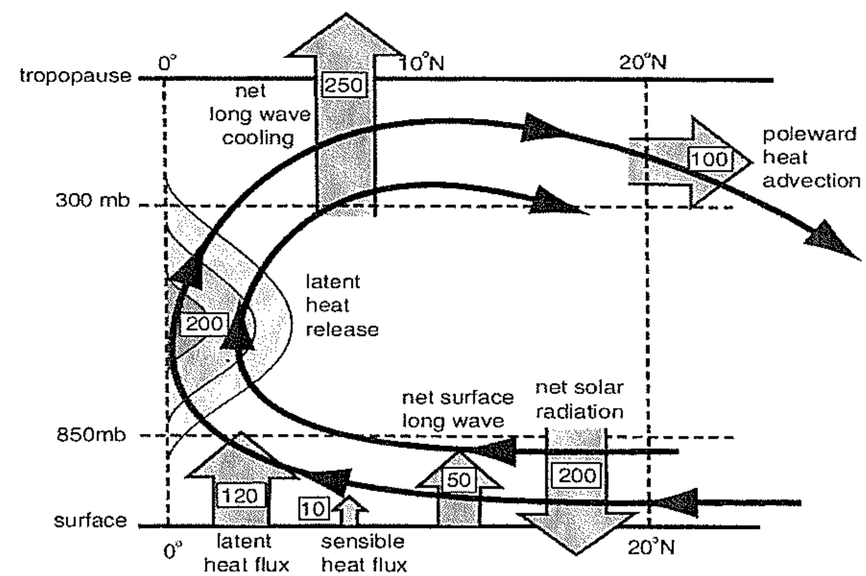

Figure 1-6a. Heat balance in the near-equatorial section of the Hadley circulation (solid dark arrows). Heating of the surface occurs from either direct solar heating (200 W m^{-2}) or by down-welling atmospheric long-wave radiation. Heating of the atmosphere occurs at the surface by direct sensible heating (10 W m^{-2} over oceans, higher over land), by a net long-wave radiative flux (50 W m^{-2}), and in the mid-troposphere by the condensation of water vapor that initially evaporated at the surface (200 W m^{-2}). At the top of the atmosphere, radiative cooling to space with a magnitude of about 250 W m^{-2} takes place. The residual (100 W m^{-2}), is transported poleward initially by the Hadley circulation. It is important to note that the regions of heating and cooling in the tropical atmosphere are separated distinctly in the vertical. That is, the tropics are heated at the surface and middle atmosphere and are cooled in the upper troposphere.

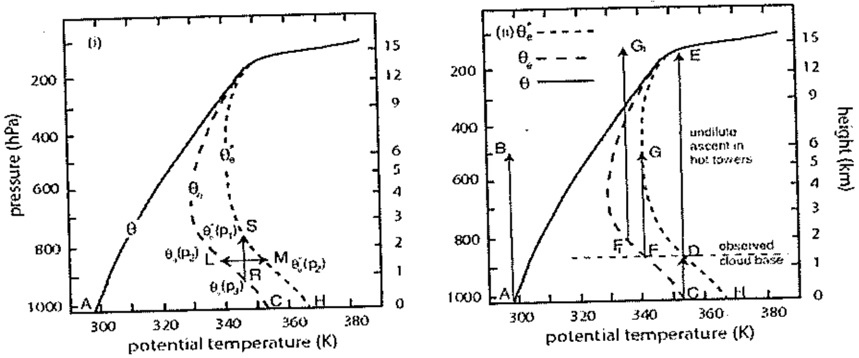

Figure 1-6b. The profiles of potential temperature θ, equivalent potential temperature θ_e, and equivalent potential temperature saturated at the local vertical temperature θ_e^* for the climatological tropical atmosphere. Panel (i) shows the three profiles where it can be seen that from about 600 mb and above, the atmosphere is stable ($\partial\theta_e^*/\partial z \geq 0$). Panel (ii) shows the consequences of ascent of moist parcels vertically lifted from different locations in the tropical column. It is clear that a parcel lifted from the surface (point C) requires the least work to lift it to its condensation point (D). Furthermore, this parcel, having the highest equivalent potential temperature in the lower troposphere, will release the greatest amount of latent heat if it is lifted through the atmosphere. The trajectory D–E shows the path of vertical ascent through the undilute cores of Riehl and Malkus's (1958) hot towers.

The temperature change of an atmosphere is given by (e.g., Curry and Webster 1998):

$$\frac{dT}{dt} = \frac{g}{C_p}\frac{\partial F}{\partial p} \tag{6}$$

where C_p is the specific heat of the atmosphere and F is the net radiative flux divergence between the atmosphere and outer space where:

$$F = 0 - \sigma T_e^4 \tag{7}$$

and it is assumed that there is no flux of long-wave radiation from outer space and the planet radiates to outer space at the equivalent temperature of

the planet defined as $T_e = [S(1-\alpha)/4\sigma]^{1/4}$ where S is the incident solar radiation, α the planetary albedo, and σ the Stefan-Boltzman constant. Between Equations (6) and (7), we can find an expression for the cooling rate of the planet:

$$\frac{dT}{dt} = -\frac{g}{C_p}\frac{\sigma T_e^4}{np_0} \tag{8}$$

The factor n has been introduced to provide a measure of the long-wave optical depth. For a strongly absorbing atmosphere, we set $n = 1$ (see Curry and Webster 1998, chapter 14). For a less-absorbing atmosphere such as earth or Mars, which have optically thinner atmospheres, $n < 1$. From Equation (8) it is clear that the cooling rate of the planet is determined by a combination of the mass of the planet (given by surface pressure p_0), as well as its heat capacity (C_p), radiative equilibrium temperature (T_e), and its long-wave optical properties (n). Thus, a hot atmosphere (large T_e) with little mass (small p_0) will cool quicker than a cooler and more massive atmosphere.

We define a radiative time scale (τ_{rad}) as the time it takes for the planet to cool by a certain fraction from its initial temperature.[7] This is referred to as the *e*-folding rate and provides a time scale for the cooling of the planet by 0.37 of its initial value. τ_{rad} can be calculated by integrating Equation (6) in between the times when the temperature is T_e and when it cools to T_e/e. That is:

$$\tau_{rad} = \int_{t(T_e/e)}^{t(T_e)} dt = \int_{T/e_e}^{T_e} \frac{dt}{dT}\, dT \tag{9}$$

and using Equation (8) gives:

$$\tau_{rad} = \frac{np_0 C_p}{g\sigma}\int_{T_e/e}^{T_e}\frac{1}{T^4}dT = \frac{np_0 C_p}{4g\sigma(T_e^3 - T_e^3/e^3)} \approx \frac{np_0 C_p}{4g\sigma T_e^3} \tag{10}$$

---

[7] It might seem that a more logical time scale might be the halving time of temperature; that is, the time it would take to halve the initial temperature. However, it is mathematically convenient to choose the base of the natural logarithm, *e*, and consider when the temperature drops to $1/e$ of its initial value. "*e*" has a value of approximately 2.71828. Such a time scale is referred to as the *e*-folding time scale.

Finally, we require a dynamical time scale. This is defined simply as the time for the wind of characteristic speed U to move across over some characteristic planetary distance taken here as the radius of the planet, a. Then:

$$\tau_{dyn} = \frac{a}{U} \tag{11}$$

The ratio of the dynamical and radiative time scales determines whether radiative or dynamical effects will dominate a climate. Between Equations (10) and (11) we find:

$$\varepsilon = \frac{\tau_{dyn}}{\tau_{rad}} = \frac{4ag\sigma T_e^3}{Unp_0 C_p} \tag{12}$$

We can understand the relative allocation of energy between radiative cooling at the top of the atmosphere and lateral poleward advection by considering the extremes of the ε:

(1) $\varepsilon >> 1$: For such situations, the dynamical time scale must be very much larger than the radiative time scale, meaning radiative effects will dominate. On Mars, for example, $\varepsilon \approx 200$ because of the smallness of the atmospheric mass (6 hPa surface pressure compared to 1,000 hPa on earth) and a very weak greenhouse effect. With such a large value of ε, the horizontal dynamical fluxes of heat are small compared to the local radiative cooling to space, meaning that a parcel of air advected along by the winds will cool rapidly to space so that the initial temperature signature of the parcel would be lost relatively quickly. The dominance of radiative effects illustrates why the observed equator-to-pole temperature gradient and the dayside-nightside temperature gradient along a line of latitude on Mars is so much larger than that observed on earth.

(2) $\varepsilon << 1$: In this case, the dynamic transports of heat dominate over radiative effects. In such a system, a parcel advected by the motion of the atmosphere will cool very slowly and maintain its initial signature as it moves poleward. Venus has a value of ε of about 0.003 because of the massiveness of the atmosphere (the surface pressure on Venus is about 10,000 times and 100 times that of Mars and earth, respectively). Thus, during the time it takes a parcel to be advected from the equator to the pole, the atmospheric parcel would have cooled imperceptibly. Consequently, the equator-to-pole temperature gradient is almost zero and there is little difference between daytime and nighttime temperatures.

(3) $\varepsilon \approx 1$: A value of unity suggests parity between dynamical and radiative effects on an atmospheric parcel. Earth, for example, has a value of ε of about 0.4. On its poleward motion, an advected parcel will cool radiatively but not sufficiently rapidly that its initial thermodynamic signature will be lost by the time the parcel reaches its destination. The comparatively similar values of the two time scales is the reason why in the upper troposphere, there are both an advection of heat towards the poles by the Hadley circulation and significant local cooling to space of roughly equal magnitudes.

To a large degree, we have accounted for the disposition of heat once it reaches the upper troposphere. However, the manner in which the vertical transport of heat takes place in the upward branch of the Hadley cell is more difficult to explain for two reasons: (1) Air is not transported upward by large-scale ascent as is suggested by Figure 1-6b, and (2) there exists a distinct stable layer in the middle troposphere that makes vertical transport difficult. These two issues are strongly linked.

Satellite pictures suggest a very different view of the tropics to that suggested by both Figures 1-2 and 1-6a. By and large, the tropics are relatively cloud free and the predominant feature is weak subsiding air over broad cloudless regions. Interspersed within this subsidence are deep cloudy regions where very strong ascent is associated with convection that links the boundary layer with the upper troposphere. The net ascent shown in Figure 1-6a is really a composite of the regions of vigorous ascent and slow descending air outside the disturbed regions. We can understand the problem by considering some simple energy concepts. Figure 1-6b shows profiles of potential temperature θ, equivalent potential temperature θ_e, and equivalent potential temperature saturated at the local vertical temperature θ_e^*, which are defined in a convenient energy form as, respectively:

$$\left.\begin{aligned} C_p\theta &= C_pT + gz & \text{(a)} \\ C_p\theta_e &= C_pT + gz + Lq & \text{(b)} \\ C_p\theta_e^* &= C_pT + gz + Lq^* & \text{(c)} \end{aligned}\right\} \tag{13}$$

These three equations represent the dry static energy, the moist static energy, and the saturated moist static energy, respectively. Here, L is the latent heating coefficient of condensation, q is the observed or ambient vapor pressure, $q^*(T)$ is the saturated vapor pressure at temperature T of the height of the parcel, and C_p is the specific heat of air at constant pressure. The θ and θ_e profiles shown in Figure 1-6b represent average conditions in the tropical atmosphere. The θ_e^* profile shows the equivalent potential temperature profile if the air were saturated at the ambient temperature at some height z. The letters A, C, and H represent the three temperatures at the surface, values of which can be found by setting $z = 0$ in (Eqs. 13a–c). From Equations (13) it can be seen that the temperature would rise by 55 K if all of the water vapor in the surface parcel were condensed (cf. Eqs. 13a and b) or by 65 K if the parcel were saturated (cf. Eqs. 13a and c). This 10 K difference at the surface between the ambient equivalent potential temperature and the saturated value changes with height, as the temperature of the atmosphere (and hence the saturated vapor pressure) is a function of height. The greatest difference between θ_e and θ_e^* occurs in the mid-troposphere. Saturation of a parcel and the release of latent heat occur if the parcel is lifted and cooled adiabatically (constant θ) and if the parcel encounters the θ_e^* curve, as occurs in panel (i) of Figure 1-6b, when a parcel is lifted and cooled along the line R–S. Finally, for the atmosphere to be unstable at some height, there is the condition that $\partial\theta_e / \partial z < 0$ (e.g., Curry and Webster 1998). Clearly, this is not the case above 600 mb (4 km), suggesting that the middle and upper troposphere are very stable.

Panel (ii) of Figure 1-6b shows the outcome of lifting air parcels in the tropical atmosphere. If the air is dry ($q = 0$), then a parcel will follow a trajectory A–B. As the potential temperature θ is constant for adiabatic ascent and as the potential energy gz increases as the parcel is lifted, then T must decrease with ascent. Thus the parcel will always be cooler than its environment and can never generate buoyancy, indicating an extremely stable state. Consider now a moist parcel located at the surface (point C). If the parcel is lifted to D, it becomes saturated and latent heat is released, which increases the buoyancy of the parcel. As it turns out, this is roughly the observed cloud base of deep convective clouds in the tropics. Consider a parcel located at higher elevation (F) in the atmosphere. Upon lifting, the parcel will eventually reach saturation but not until point G near 500 mb. Nearly 5 times the amount of work[8] is necessary to raise a parcel from F to G than to lift a parcel from C to D in order to achieve saturation and the release of latent heat. Thus, there is great efficiency in raising lower tropospheric parcels of moist air compared to moist parcels located higher in the troposphere. It should also be noted that the fact that the parcel is moist does not ensure that latent heat will be released if it is lifted and cooled. Consider a parcel at height F_1. No matter how high the parcel of moist air is lifted (e.g., to G_1), it never encounters the saturated equivalent potential temperature (θ_e^*) curve and it will always be cooler than its environment. Such a parcel will never generate buoyancy, and by the time that it reaches the upper troposphere (if somehow it were lifted to that height), it will be cooler than the environment.

The difficulties in raising parcels through the tropical mid-tropospheric stable layer were first identified in the landmark study by Riehl and Markus (1958). Consider once again the lifting of a parcel from level C to level D where saturation occurs (panel ii, Figure 1-6b). The parcel will continue to rise using the buoyancy generated by the release of latent heat of condensation. However, dry, low θ_e air from air surrounding the cloud will be entrained into the cloud, potentially reducing the buoyancy and curtailing the vertical ascent of the parcel and the height of the cloud. Riehl and Malkus proposed the concept of the "hot tower," which is a cumulonimbus cloud that possesses an "undilute" saturated core that would funnel moist and warm boundary layer air through the stable mid-troposphere, protecting it from dilution with the low θ_e air. To provide such protection, the cloud must be of significant horizontal dimensions. In addition, the ascending air must have the highest specific humidity, originating near the surface (position C). Also, the amount of work needed to raise a surface parcel is less than anywhere in the tropical atmospheric column. Large convective elements are common in the tropics, especially in the region of the ITCZ. Riehl and Malkus calculated that about 1,500–5,000 hot towers would be needed each day to perform the required vertical heat transport. Viewed from a tropics-wide perspective, this is not a large number.

Whereas the hot tower hypothesis answers many of the problems that we have encountered in transporting surface heat into the upper troposphere, one problem remains. Some form of mechanical lifting is necessary to saturate the parcel as it is raised from C to D. A number of mechanisms have been proposed, all of which depend on the generation of a low-level

[8] Work is defined as the product of the force needed to lift the parcel ($m\,g$ where m is the mass of the parcel) times the distance moved in the vertical Δz; that is, the work required to lift a parcel from point C to D is $m\,g(z_D - z_C)$.

cyclonic circulation and frictional convergence. In the conditional instability of the second kind (CISK) mechanisms, the converging moist air rises and the release of latent heat lowers the surface pressure, which increases the inflow and so on. This mechanism was introduced by Ooyama (1964) and Charney and Eliassen (1964) to explain tropical cyclone development, and later by Bates (1970) to explain instabilities of the ITCZ. Unfortunately, there is little observational evidence for CISK in nature. We will propose a second mechanism for lifting parcels to saturation in Section 4.4.

2.5. Driving Forces of the Hadley Circulation

Condensational heating of the equatorial atmosphere does two things: It raises the temperature of the atmospheric column, and it decreases the pressure at the surface relative to that in the subtropics. These two impacts are used to support the common argument that the Hadley circulation is driven by an equator-directed surface pressure gradient between the subtropical high-pressure zone and the equatorial trough. While this argument is not incorrect, it is incomplete. The Hadley circulation is also driven by a reversed pressure gradient in the upper troposphere. In fact, if there were no surface pressure gradients at all, there would still be meridional cells, as long as the equatorial regions were warmer than the subtropics.

Consider the following argument. If the equation of state ($p = \rho RT$) is combined with the hydrostatic equation ($\partial p/\partial z = -\rho g$), and it is assumed that the temperature of an atmospheric column is constant (i.e., T = constant), we obtain:

$$\partial \ln p/\partial z = -g/RT\,, \tag{14}$$

where p, ρ, and R are the atmospheric pressure, density, and gas constant, respectively. Equation (14) shows that that the warmer the atmospheric column, the slower the pressure decreases with height. Now consider two columns with mean temperatures T_{EQ} and T_{ST} representing an equatorial and a subtropical column, respectively, where $T_{EQ} > T_{ST}$. From Equation (14) it is clear that the pressure decreases less rapidly with height in the warm column compared to that in the cold column. Thus, if the equatorial surface pressures in the tropics and subtropics were the same, then at some height above the surface the pressure in the warm column will be greater than in the cold column, producing a pressure gradient aloft that would drive air poleward. We may see this formally by integrating Equation (14) in the vertical between the surface ($z = 0$) and some height in the column ($z = z_1$).

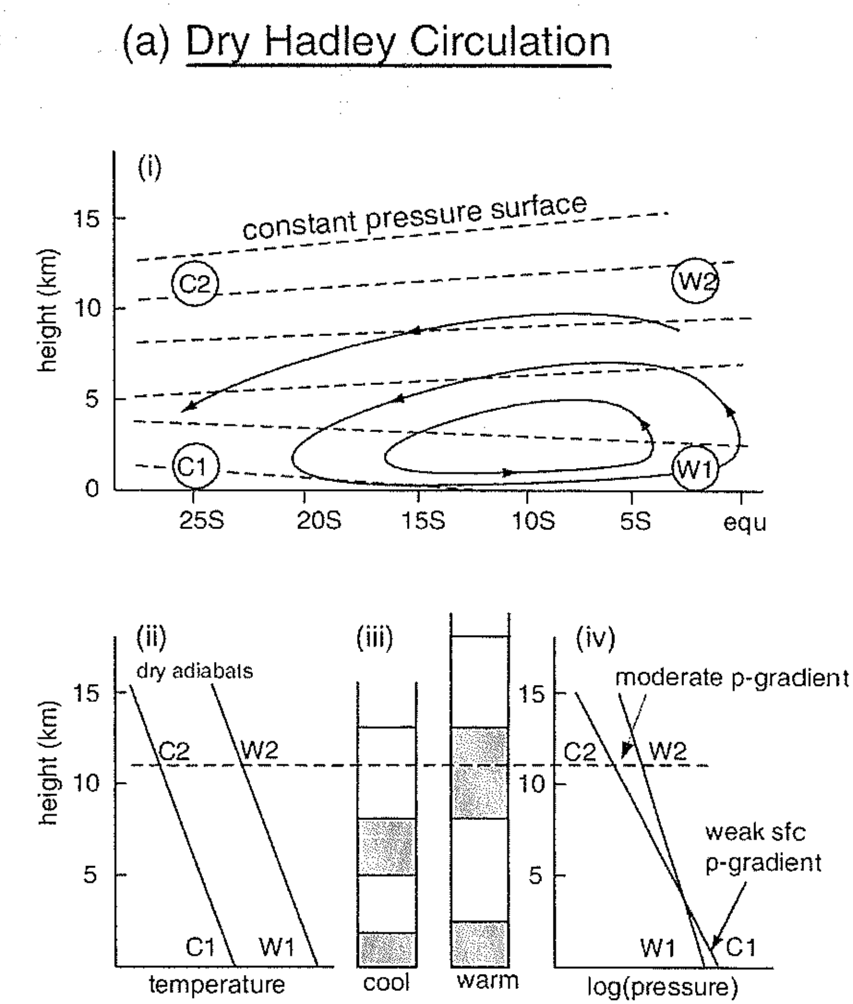

Figure 1-7a. Schematic of the Hadley circulation between the subtropics and the equator in an environment where there is no latent heat of condensation. The dashed lines in panel (i) show the slope of the pressure surfaces as indicated by Equations (7) and (8). Panel (ii) shows the distribution of temperature with height for the warm column over the heated continent (W) and the cold column in the subtropics (C) at two heights (1 and 2). In the dry atmosphere, the temperature decreases upward dry adiabatically. Panel (iii) shows the expansion of layers of different temperatures. Panel (iv) shows the variation in pressure with height. At the surface there is a weak pressure gradient where (p(C1) > p(W1)) while at higher levels the pressure gradient is reversed (p(W2) > p(C2)). The pressure gradient reversal occurs in the lower troposphere and a weak shallow meridional cell ensues.

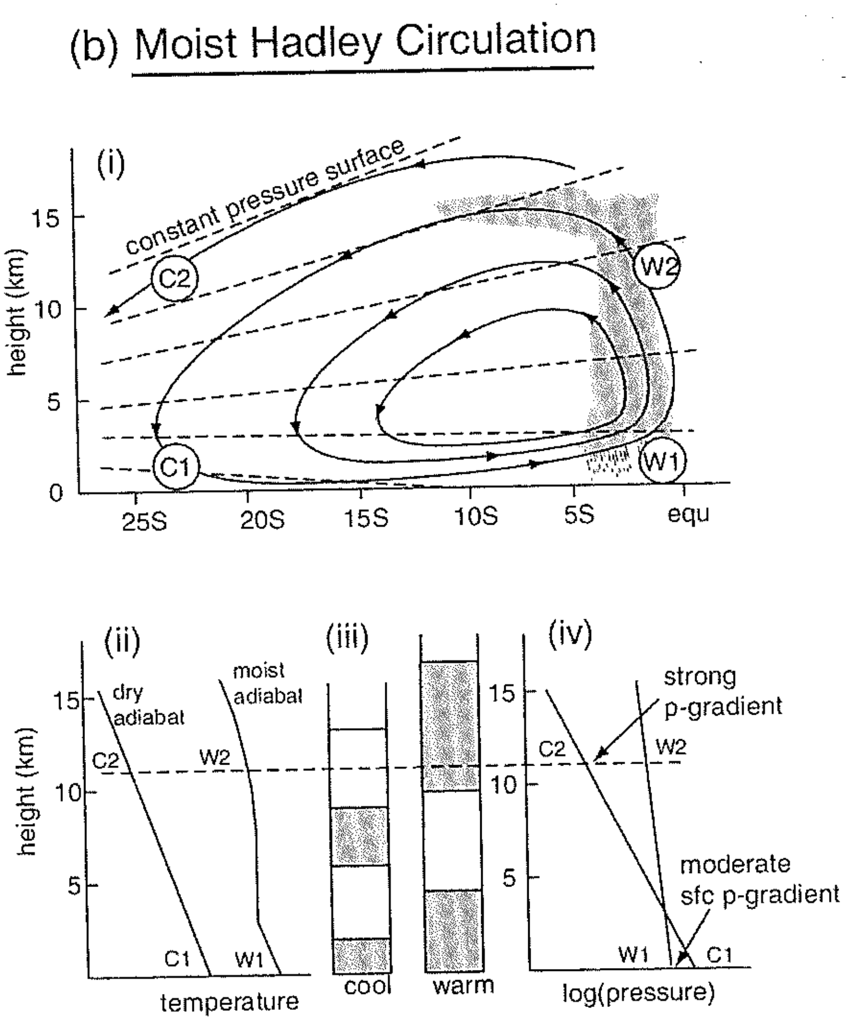

Figure 1-7b. As in Fig. 1-7a except where latent heating occurs in the ascending region of the Hadley cell. In the region of ascent, the vertical temperature lapse rate is close to moist adiabatic (panel ii), and expansion of the layers is more extreme in the heated region so that the upper-level pressure gradient between the warm and cold column is extremely large (i.e., (p(W2) >> p(C2)). The result is a system with more tilted isobars (dashed lines, panel i) and a deeper and more vigorous Hadley circulation (p(W2) > p(C2)). Like the monsoons and sea breezes, the Hadley circulation is driven largely by a reversed upper-level pressure gradient.

Performing this operation for the equatorial and subtropical columns, we obtain the following vertically dependent difference in pressure between the columns:

$$\Delta(\ln p_{z=z_1}) - \Delta(\ln p_{z=0}) = -\frac{gz_1}{R}\left(\frac{1}{T_{EQ}} - \frac{1}{T_{ST}}\right) \tag{15}$$

where Δ refers to the difference between the equatorial column and the subtropical column. If the differences in the surface pressure were zero, then:

$$\Delta(\ln p_{z=z_1}) = -\frac{gz_1}{R}\left(\frac{1}{T_{EQ}} - \frac{1}{T_{ST}}\right) \tag{16}$$

where $\Delta(\ln p_{z=z_1}) > 0$ if $T_{EQ} > T_{ST}$. Also, it should be noted that the pressure difference increases linearly with height. Furthermore, from observations, we know that $\Delta(\ln p_{z=z_1}) > \Delta(\ln p_{z=0})$ so that the Hadley circulation is driven by two opposing pressure gradients, one located at the surface and the second, the stronger, at higher levels in the atmosphere. The same basic physics described above drive thermal circulation ranging from small-scale sea breezes to very large-scale monsoons. These factors are shown schematically in Figure 1-7a for a dry Hadley circulation (i.e., assuming there are no moist processes) and in Figure 1-7b for a Hadley circulation with moist processes.

3. THE VERTICAL SCALE OF THE HADLEY CIRCULATION

The Hadley circulation occupies the entire depth of the troposphere extending up to heights greater than 15 km, consistent with the vertical extent of deep convective clouds. However, this is not the same as heights expected from simple fluid dynamical arguments. A series of studies (e.g., Matsuno 1966; Webster 1972; Gill 1980) examined the fundamental structure of waves in the tropical atmosphere that were, in essence, a subset of global modes identified and catalogued by Longuet-Higgins (1968). These studies showed that modes in a fully stratified atmosphere could be represented by modes in a shallow fluid of a particular depth. The depth of the fluid is referred to as the "equivalent depth" of the mode.[9] For slow large-scale modes, the vertical scale is between 2 and 3 km and much smaller than the observed vertical scale of the Hadley circulation.

The difference between the predicted and observed scales of tropical motion can be reconciled by noting that the theoretical models mentioned above did not consider moist processes or, more particularly, dissipative processes associated with convection. Chang (1977) found that the observed scales of the modes would be the same as those predicted with both heating and dissipation in the formulations. To a large degree, the vertical scale of the circulation is set by the heating profile associated with the latent heat release. As the maximum vertical velocity occurs where the maximum heating occurs, we can see from Figure 1-2 that it should be located in the middle troposphere centered near 7–8 km as is shown in Figure 1-6a. But why should tropical heating be located in the mid-troposphere?

Webster (1994) and Emanuel et al. (1994) argue that the magnitude of the tropical SST determines the vertical structure of the tropical atmosphere. From Clausius-Clapeyron considerations (e.g., Curry and Webster 1998), we know that the saturated vapor pressure of water increases exponentially with temperature so that an air parcel over the warm equatorial ocean will have a larger water vapor pressure than a parcel over a cooler ocean. Thus, surface air flowing from higher latitudes towards the equator in the trade winds (forced by a pressure gradient formed in response to the SST gradient) will possess increasingly saturated vapor pressure as the parcel warms. The warmer the SST at the end of the parcel's journey, the greater the moisture content of the air parcel. As the air masses converge, ascent of the moist air occurs, producing a nearly moist adiabatic temperature profile. In essence, the warmer the SST, the greater the release of latent heat in the atmospheric column and the higher a parcel will ascend. A schematic of the relationship between the magnitude of the SST, the vertical temperature profile, and the latent heat profile if ascent were to occur is shown in Figure 1-8. This figure is supported by observational evidence that the SST and the tropopause height appear to be closely related.

[9] Formally, the equivalent depth is contained within the separation coefficient when the vertical and horizontal parts of the equations are formally separated. Separation requires that only relatively simple thermal structures can be considered. But even with this simplicity, considerable insight can be gained about the fundamental nature of real atmospheric waves.

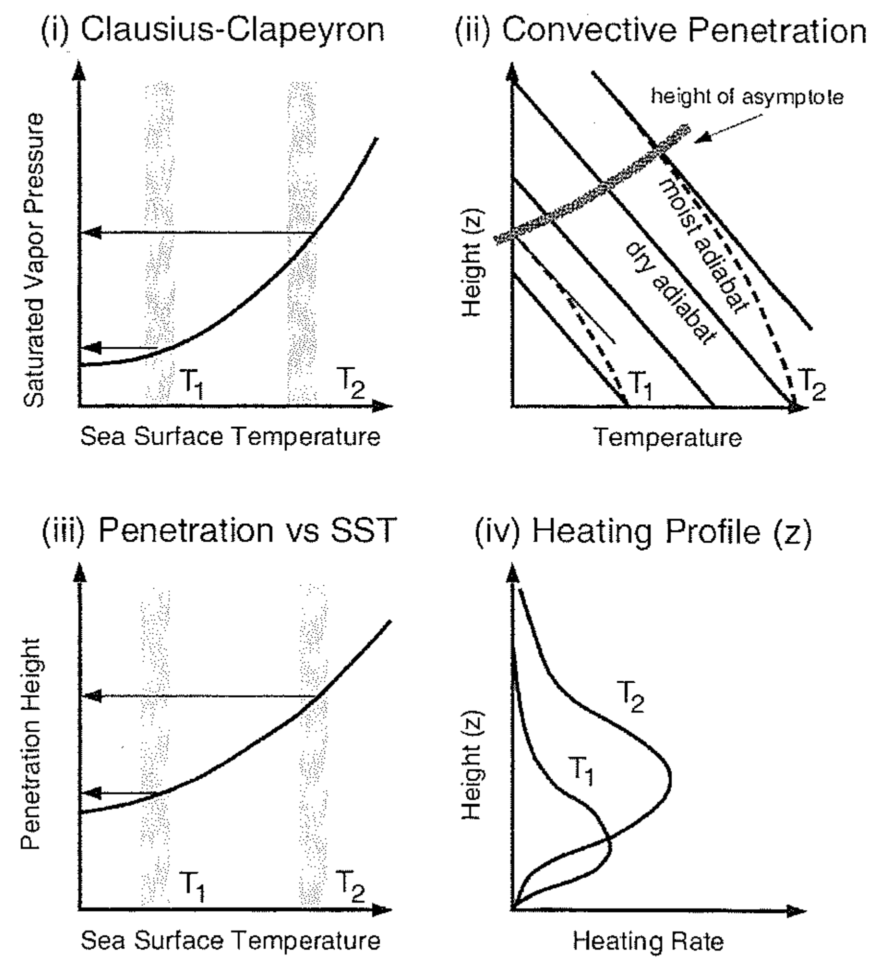

Figure 1-8. Schematic diagram relating the distribution of organized convection over the oceans with SST. Two columns with temperatures T_1 and T_2 (where $T_2 > T_1$) represent, for example, the subtropics and the equatorial trough. Panel (i) shows the Clausius–Clapeyron relationship (i.e., $e_s = e_s(T)$) and its relationship to the convective penetration height (panels ii and iii) over the ocean. Assuming the same relative humidity distribution in each column and that convection has occurred, the magnitude of the heating in the warm column will be greater than that in the cold column and will occur exponentially higher in the atmosphere (panel iv). Adapted from Webster (1994).

4. LOCATION OF THE ITCZ

Observationally, the location of the ITCZ can be identified by satellite as a minimum in outgoing long-wave radiation (OLR)[10]. Seasonal means of OLR show the ITCZ as a meandering narrow band of cloudiness close to the equator. In general, the ITCZ resides near the equator or in the summer hemisphere. An exception is in the eastern Pacific Ocean, where it resides to the north of the equator for the entire year, except occasionally in spring, when it forms as a double ITCZ on either side of the equator. Two factors appear to be responsible for its location: the magnitude of the equatorial SST and strength of the cross-equatorial pressure gradient. We will briefly consider each of these factors.

4.1. Influence of SST on the Location of the ITCZ

From the discussion in Section 3, it would seem that the deepest and strongest equatorial convection (and hence the location of the ITCZ) should coincide with the locations of the warmest SSTs. However, this is a common misconception. Figure 1-9 shows OLR plotted a function of SST for regions in the Indian Ocean and the equatorial Pacific. If SST were the sole determining factor of the strength of convection, then the OLR-SST relationship would be a straight line. However, in the Pacific Ocean we note that as SST increases, the OLR-SST relationship becomes increasingly less linear. In fact, the OLR takes on a broad range of values indicating both cloudy and cloudless regions in the regions of highest SST. In the Indian Ocean box, where the SST is generally warmer than in the Pacific Ocean, there is little relationship between SST and OLR. Thus, although there is a general increase in convection as SST increases, especially in the Pacific Ocean, it is clear that where the SST is warmest, other factors must come into play in determining the location of convection.

The latitudinal distributions of SST, the mean atmospheric sea level pressure (MSLP), and convection (inverse of OLR) in the Indian Ocean, the eastern Pacific Ocean, and the central Atlantic Ocean are shown in Figure 1-10. In the boreal winter, the SST maximum is poleward of the convection by about 5° of latitude. The minimum MSLP coincides with the SST maximum near 15°S. Similar displacements occur in the Atlantic and the eastern Pacific during summer. In the Atlantic Ocean during the boreal winter,

[10] OLR is a satellite product and is the integral of radiation emitted from the surface of the planet, the atmosphere, and clouds. The tops of deep convective clouds are cold and thus minima in the OLR represent deep clouds.

maxima in SST and convection occur in the same locations as the MSLP minima. An exception occurs in the summer Indian Ocean. We will return to this case presently. There, the mean monthly convection occurs over a broad latitudinal range.

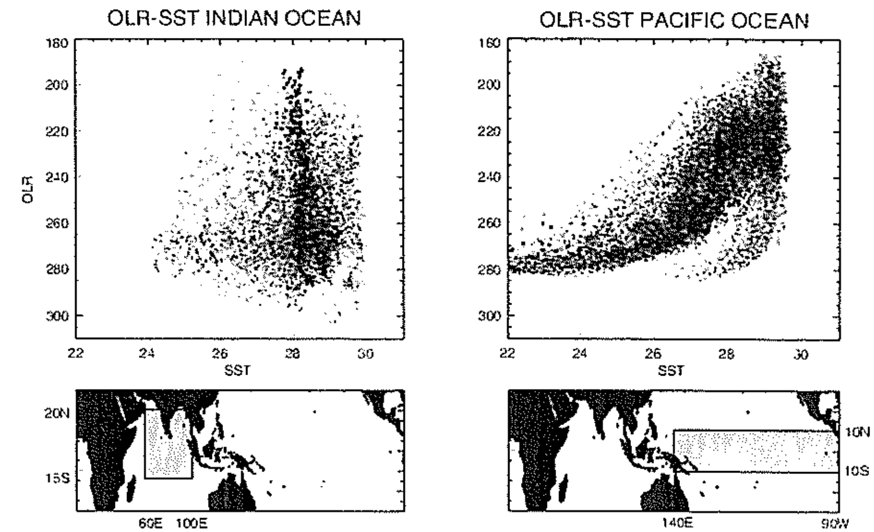

Figure 1-9. OLR (W m^{-2}, inverted scale) as a function of SST (°C) for the Indian Ocean (shaded region, left map, 20°N–15°S, 60°E–100°E) and the equatorial Pacific Ocean (shaded region, right map, 10°N–10°S, 140°E–90°W). OLR-SST pairs were computed from mean monthly data over the 11 y period 1980–1990 in 2° x 2° longitude-latitude squares within the shaded regions. In general, small values of OLR indicate deep clouds and high values indicate shallow convection or its absence. In the Pacific region, there is a general increase in the height of convection with increasing SST as suggested by the heating-SST relationships shown in Fig. 1-8. However, at the warmest temperatures (>28°C) there is a breakdown of this relationship. In the Indian Ocean, there is an absence of a relationship between OLR and SST altogether, although it is noted that the SST is in general >26°C. The lack of correlation at high SSTs and in the Indian Ocean suggests that other processes besides SST, such as atmospheric dynamics, are important in determining where convection occurs and also preclude the appearance of deep convection in the warmest regions of the tropical oceans.

In summary, Figures 1-9 and 1-10 suggest that maximum values in SST do not necessarily set the location of the ITCZ. We note from Figure 1-

10 that the greatest deviations between maximum SST and minimum OLR occur where there is a strong cross-equatorial pressure gradient. However, when there is a co-location of maximum SST and minimum OLR, there is a minimal cross-equatorial pressure gradient.

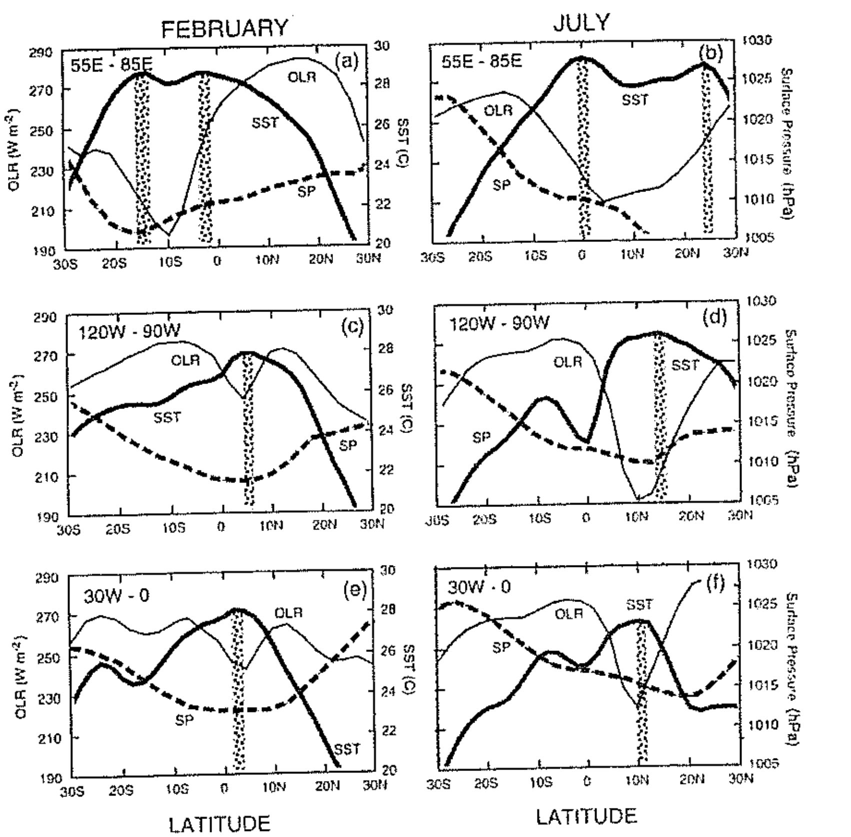

Figure 1-10. Distributions of OLR (thin solid lines), SST (heavy solid lines), and mean sea level pressure (SP: dashed lines, hPa) for three longitudinal bands: (a) and (b) Indian Ocean, 55°E to 85°E; (c) and (d) eastern Pacific Ocean, 120°W to 90°W; (e) and (f) eastern Atlantic Ocean, 30°W to 0°. (a), (c), and (e) are for February 1992, and (b), (d), and (f) are for July 1992. Stippled areas denote areas of maximum SST. Adapted from Tomas and Webster (1997).

4.2. Influence of Cross-Equatorial Pressure Gradients

Tomas and Webster (1997) noted that, in addition to the SST maximum occurring on the poleward side of strong convection, the location of the zero absolute vorticity line[11] ($\eta = 0$) lay just equatorward of the maximum convection. They also found that where the $\eta = 0$ contour was located was directly related to the magnitude of the cross-equatorial pressure gradient. In Figure 1-11a, two situations are shown that depict, respectively, weak and strong cross-equatorial pressure gradients. Figure 1-11b plots the location of the zero absolute vorticity gradients as a function of cross-equatorial pressure gradient. Irrespective of the domain over which the pressure gradient is calculated, the correlations are very high, indicating an extremely strong relationship between the location of convection and the pressure gradient. Noting that the regression curves in Figure 1-11b pass though zero, it is clear that if there is no pressure gradient that the zero absolute vorticity line (and convection) resides at the equator. When the cross-equatorial pressure gradient was positive (high pressure north of the equator, low pressure to the south), convection lay south of the equator, and vice versa. In general, the stronger the pressure gradient, the more the $\eta = 0$ contour was displaced away from the equator.

It is clear that the location of convection is associated with the cross-equatorial pressure gradient. But what produces this gradient? A first thought might be that the convection itself may be responsible. But if this were the case, convective maxima and MSPL minima would coincide, and this does not occur unless there is no cross-equatorial pressure gradient (see Figure 1-10).

The question thus becomes: Why is a cross-equatorial pressure gradient important in determining the location of convection? Without explanation, Lindzen and Hou (1988) made the interesting observation from a modeling study that a heating function of a given magnitude produced a much larger response if it were placed off the equator rather than on the equator. It turns out that the cross-equatorial pressure gradient effect and the sensitivity to where heating is placed in the Lindzen and Hou model are related. The relationship resides in the manner in which the dynamics of the tropics produce circulations that may enhance tropical convection under certain circumstances.

[11] Absolute vorticity, η, is given as the sum of the earth's vorticity plus the relative vorticity, i.e., $\eta = f + \zeta = 2\Omega \sin\varphi + (u_y - v_x)$.

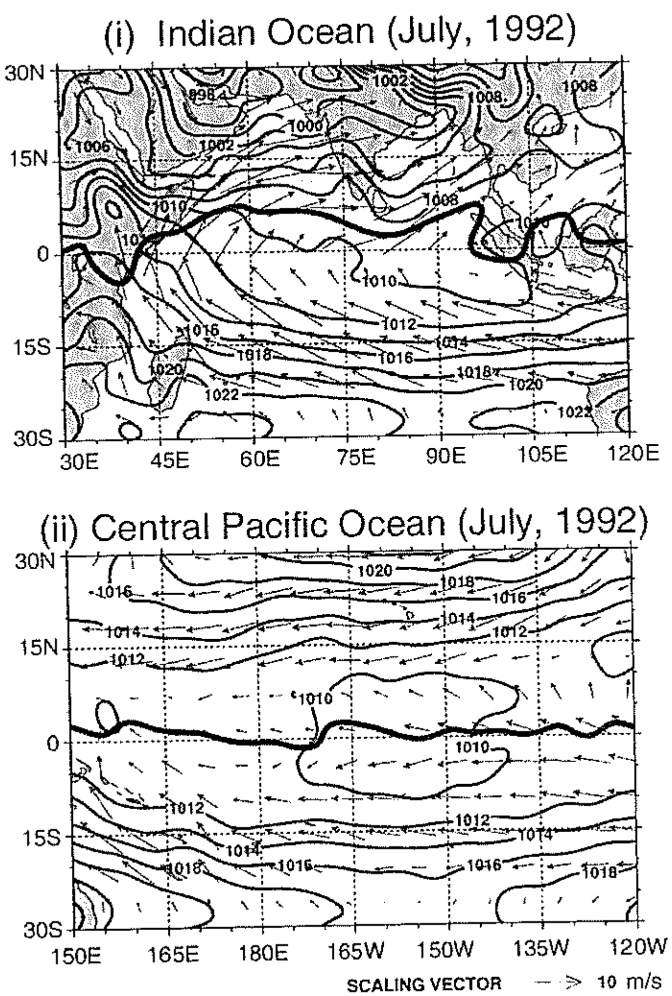

Figure 1-11a. Distributions of sea level pressure (hPa) and horizontal wind (m/s) at 925 hPa commonly encountered about the equator. (i) Indian Ocean, 30°E to 120°E, and (ii) Pacific Ocean, 150°E to 120°W. Bold lines denotes the zero absolute vorticity contours. In the western Pacific warm pool (ii), where there is an absence of a cross-equatorial pressure gradient, the zero absolute vorticity gradient lies close to the equator. In the monsoon region (i), where there is a strong cross-equatorial pressure gradient, there is a substantial displacement of the zero absolute vorticity line from the equator.

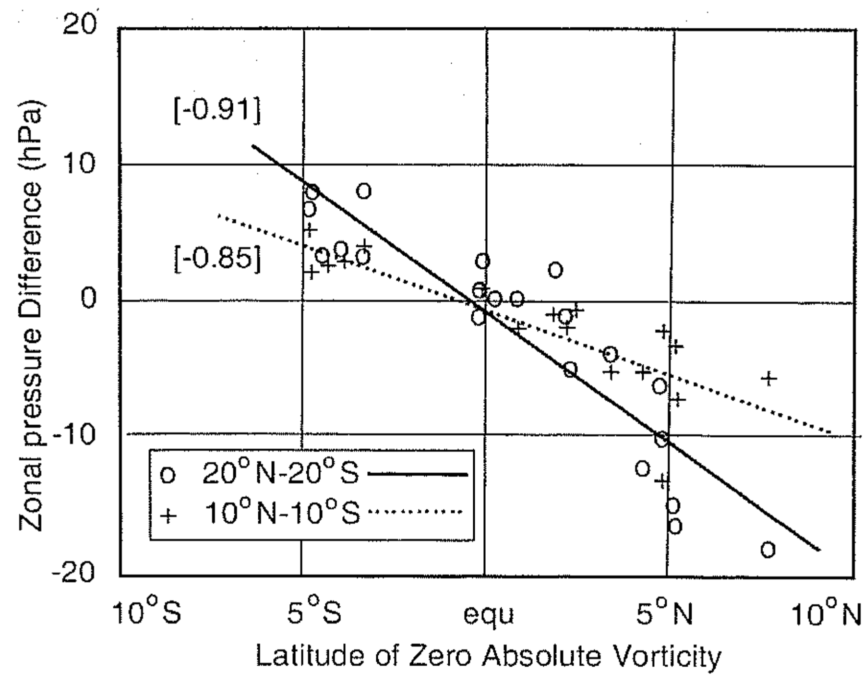

Figure 1-11b. Latitude of zero absolute vorticity at 850 hPa and the magnitude of the cross-equatorial pressure gradient calculated between 10°N and 10°S and between 20°N and 20°S. The dotted and solid lines are regression lines. Regression line and correlation are shown for calculations across two latitude zones. Correlations are exceptionally high (–0.85 and –0.91, respectively). European Center for Medium Range Weather Forecasts data were used. Adapted from Tomas and Webster (1997).

In the tropics, absolute vorticity is dominated by earth's vorticity (f) so that η is negative south of the equator and positive to the north. However, if there is advection from one hemisphere to the other under the action of a cross-equatorial pressure gradient, then it is possible for the $\eta = 0$ contour to reside away from the equator so that regions of negative η can reside north of the equator and vice versa. In this situation, the cross-equatorial pressure gradient advects absolute vorticity from one hemisphere to the other. Such a situation is dynamically unstable (Tomas and Webster 1997). The form of the instability is called "inertial instability," which is a subset of a larger class of instabilities called "symmetric instability" (Emanuel 1979). In the case of a high pressure in the Southern Hemisphere and a low pressure to the north of the equator, the advection of absolute vorticity will cause an "invasion" of anticyclonic vorticity into the Northern Hemisphere, resulting in an acceleration of the parcel away from its original position. The only way of stabilizing the system is to produce cyclonic vorticity to neutralize the anticyclonic vorticity. Creation of cyclonic vorticity is accomplished by "vortex tube stretching" through the establishment of a vertical circulation to the north of the equator (Tomas and Webster 1997). As the atmosphere is moist, the forced vertical ascent will produce a strong release of latent heat and an enhanced circulation. In fact, a close examination of the Lindzen-Hou results shows the existence of the $\eta = 0$ contour just equatorward of the enhanced convection.

4.3 Sequence of Processes for the Formation of Organized Off-Equator Convection

The simple instability theory, and some corroboration from observations, suggests a sequence of processes leading to convection away from the equator but not necessarily over the warmest SSTs. The sequence is as follows:

(1) *A cross-equatorial advection of absolute vorticity:* Organized, zonally orientated convection occurs off the equator when there is a north-south cross-equatorial pressure gradient. The strongest gradients occur over the Indian Ocean in both the boreal summer and winter, and over the Atlantic and eastern Pacific Oceans during the boreal summer. These latitudinal SST distributions produce large-scale pressure gradients spanning 15°–20° on either side of the equator. The macro-scale cross-equatorial pressure gradients, produced by the background SST distribution, must be sufficiently strong to produce a divergent wind field that advects absolute vorticity across the equator. The advection of absolute vorticity across the equator renders air parcels in the near-equatorial summer hemisphere to be inertially unstable.

(2) *The forcing of a stabilizing secondary circulation:* In order to relax the instability, a secondary circulation forms, which produces cyclonic vorticity to compensate for the importation of anticyclonic vorticity from the winter hemisphere. This secondary circulation occurs as a convergence-divergence doublet about the zero absolute vorticity line. Together they form a vertical circulation.

(3) *Convective conditional instability:* The existence of a cross-equatorial pressure gradient is necessary but not sufficient for the formation of off-equator convection. In addition, the SST must be sufficiently high for there to be a conditionally unstable atmosphere and for latent heat release to accompany the secondary circulation. Without the release of latent heat, the ensuing circulation will remain shallow and be governed by the equivalent depth of the mode.

4.4. Vertical Transport of Heat at the Equator

In Section 2.4, we discussed the problem of moving heat from the boundary layer to the upper troposphere where the system can cool to space or advect heat polewards. Riehl and Malkus (1958) noted that broad-scale ascent could not accomplish the transport and that heat could only be transported vertically by very deep cumulus clouds within their "protected" cores. The problem is describing a mechanism that allows convection to be initiated. Specifically, what mechanism takes care of the work necessary to lift a parcel from the surface (point C) to saturation height (point D) in Figure 1-6b?

We suggest that the mechanisms responsible for the lifting are the secondary circulations resulting from the inertial instability associated with the advection of absolute vorticity across the equator under the action of cross-equatorial zonal pressure gradients. The upward vertical motion in the ascending part of the stabilizing cell essentially reduces the stability of the atmospheric column, allowing deep penetrative convection to develop that would help accomplish the vertical transport of heat. Tomas and Webster (1997) describe the generation of deep convection as spasmodic with roughly the inertial frequency of the latitude of the zero absolute vorticity line (roughly 5° from the equator and thus 4–5 days). They speculate that the aperiodic convection may be the source of easterly waves that propagate westward with deep convective elements. In this manner, the generation of deep convection and the consequent balancing of the equatorial heat budget is accomplished by instabilities of the ITCZ or, in other words, the instabilities of the ascending branch of the Hadley circulation.

5. MONSOONS AND THE HADLEY CIRCULATION

The largest cross-equatorial pressure gradients exist in the vicinity of the great continents of Asia, Africa, and Australia. During the summer, these continental regions receive copious precipitation at the same latitudes where other locations receive very little precipitation. In addition, the strength of these regional monsoon circulations has a very strong influence on the zonally averaged circulation distributions seen in Figure 1-2. For example, if the South Asian monsoon region were removed from the JJA mean stream function, the zonally averaged distribution would be much weaker and the region of ascent would be found to be much closer to the equator. Consideration of Figure 1-3 supports this contention. Figure 1-12 shows the zonal velocity component and the stream function averaged zonally across the narrow band 70°E–90°E for JJA and DJF. This particular band was chosen to give a cross section through the region of strongest precipitation in the South Asian monsoon.[12] The most apparent difference between the monsoon cross sections (Fig. 1-12) and the global charts is that monsoon fields are very much stronger. For comparison, consider the region of descent in the Southern Hemisphere in Figures 1-2b and 1-12a. Noting that the vertical velocity is proportional to the latitudinal gradient of the stream function, we have gradients between the equator and 30°S of 16 x 10^{10} kg s^{-1}/30°latitude for the zonally averaged case and 68 x 10^{10} kg s^{-1}/30°latitude for the summer monsoon. That is, vertical velocities associated with the monsoons are about a factor of 4 larger. Similar differences can be seen in the boreal winter (cf. Figs. 1-2d and 1-12b).

The zonal velocity structures for the summer and winter monsoons are very different. During winter, an extremely strong westerly jet stream resides equatorward of the Himalayas. During summer, this jet weakens and moves poleward of the mountains and is replaced by an upper tropospheric easterly jet stream. The easterly jet results from the strong heating of the atmosphere over South Asia, which reverses the temperature gradient between the equator and 30°N. Equation (5) shows that under these circumstances, the winds must become increasingly westward with height (i.e., if $\partial T/\partial\varphi > 0$ then $\partial u/\partial z < 0$). The resultant easterly jet extends westward across North Africa and determines to a large degree the demarcation of wet

[12] In Figure 1-2, the stream function was computed by using zonally averaged values of the vertical and meridional components with the constraint that at any latitude the vertically averaged meridional velocity component must average to zero in the long-term mean because of mass continuity. However, averages over finite meridional bands, like those shown in Figure 1-12, are not constrained, simply because flow can return at other longitudes. That is, within a narrow longitude band, conservation of mass would also need to take into account the zonal velocity.

and dry regions in the Sahel (Nicholson and Flohn 1980; Webster et a 1998). The reversal of the temperature gradient is unique to all of the mor soon regions and probably is associated with the elevated sensible and later heating associated with the Himalayas (Webster et al. 1998).

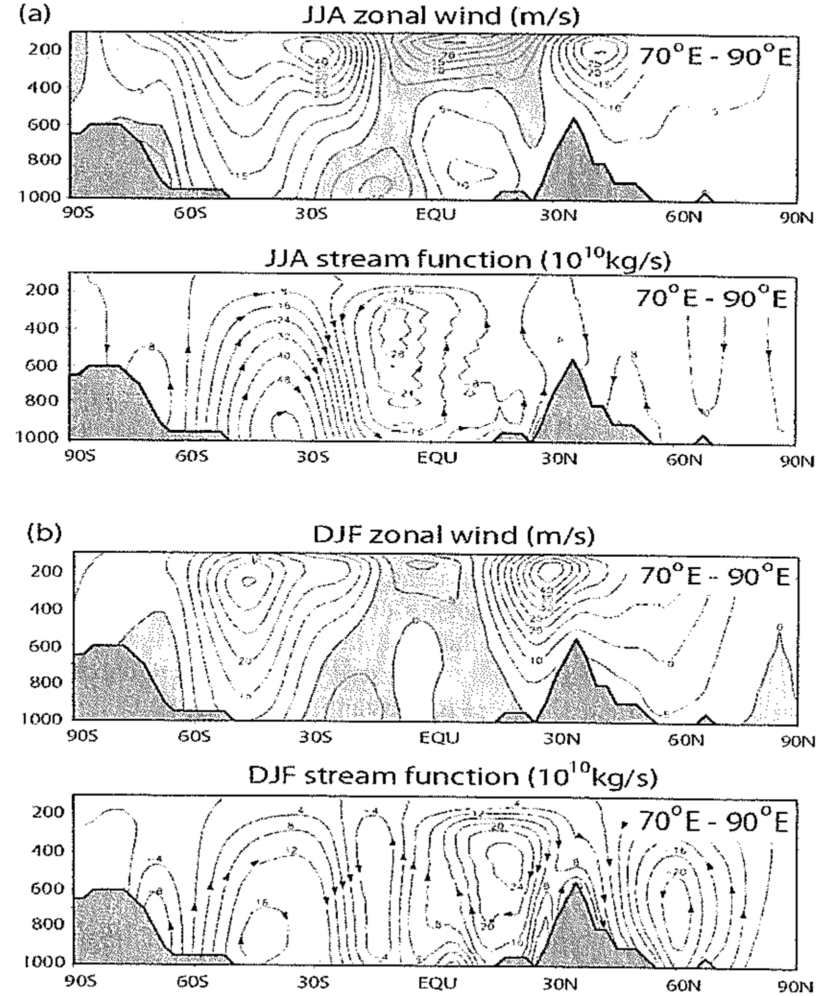

Figure 1-12. Same as Fig. 1-2 but for the monsoon regions. Zonally averaged zonal wind (upper panel, m s^{-1}) and mass stream function (lower panel, 10^{10} kg s^{-1}) for (a June–August (JJA) and (b) December–February (DJF), plotted against latitude an pressure (hPa). Contours show averages between 70°E and 90°E. Zonally average elevation of the planet is shaded, with the Himalayas and the Tibetan Plateau pre dominating near 30°N and the Antarctic continent in the south. It should be note that, unlike the zonally averaged mass stream function shown in Fig. 1-2, th stream function might not necessarily conserve mass, as zonal variability needs t be taken into account.

In essence, the same physics that control the Hadley circulation control the dynamics of the monsoon circulation, with a large number of other factors coming into play. First, in the Indian Ocean sector during the boreal summer, the cross-equatorial pressure gradient is stronger than elsewhere on the globe (Tomas and Webster 1997). Second, during the boreal spring the SST of the Indian Ocean is warmer than in any other tropical region. Third, the land area is elevated, so the impact of heating is emphasized because it occurs higher in the troposphere (Webster et al. 1998). Air forced across the equator by the strong cross-equatorial pressure gradient acquires a very high specific humidity. During the boreal winter, the cross-equatorial pressure gradient reverses due principally to the intensity of the Siberian High. However, Australia is a relatively flat continent in comparison to Asia and so there is no counterpart to the Northern Hemisphere elevated heat source and maximum heating lies close to the equator. In combination, these factors lead to a weaker winter monsoon than its summer counterpart.

6. OCEAN HADLEY CIRCULATION

So far, we have considered the impact of solar heating of the surface of the planet from a purely atmospheric perspective, noting that the atmosphere is heated from below through a combination of turbulent latent and sensible heat, and long-wave radiative fluxes. Over land areas, most of the surface heating is returned to the atmosphere. However, in the ocean a large amount of solar energy penetrates to substantial depths with *e*-folding penetrations of about 15 m (e.g., Webster 1994). As turbulent wind mixing allows heat to be stored within the ocean mixed layer, not all of the radiative energy arriving at the ocean surface is returned immediately to the atmosphere. Thus, the equatorial oceans have the potential of accumulating heat and increasing substantially in temperature.

Does the ocean have a role in maintaining the heat balance of the tropical regions? At first glance, it might appear that the oceanic transport of heat would have the wrong sign and lead to a positive feedback between the strength of the Hadley circulation and the monsoons. Winds flow towards the ITCZ in both the Hadley and monsoon regimes. If the heat transport were to be in the same direction as the winds, then the oceans would accumulate heat near regions of convection. This increased SST gradient would drive an even stronger Hadley circulation or monsoon and produce even stronger convection. That is, there would be a strong positive feedback.

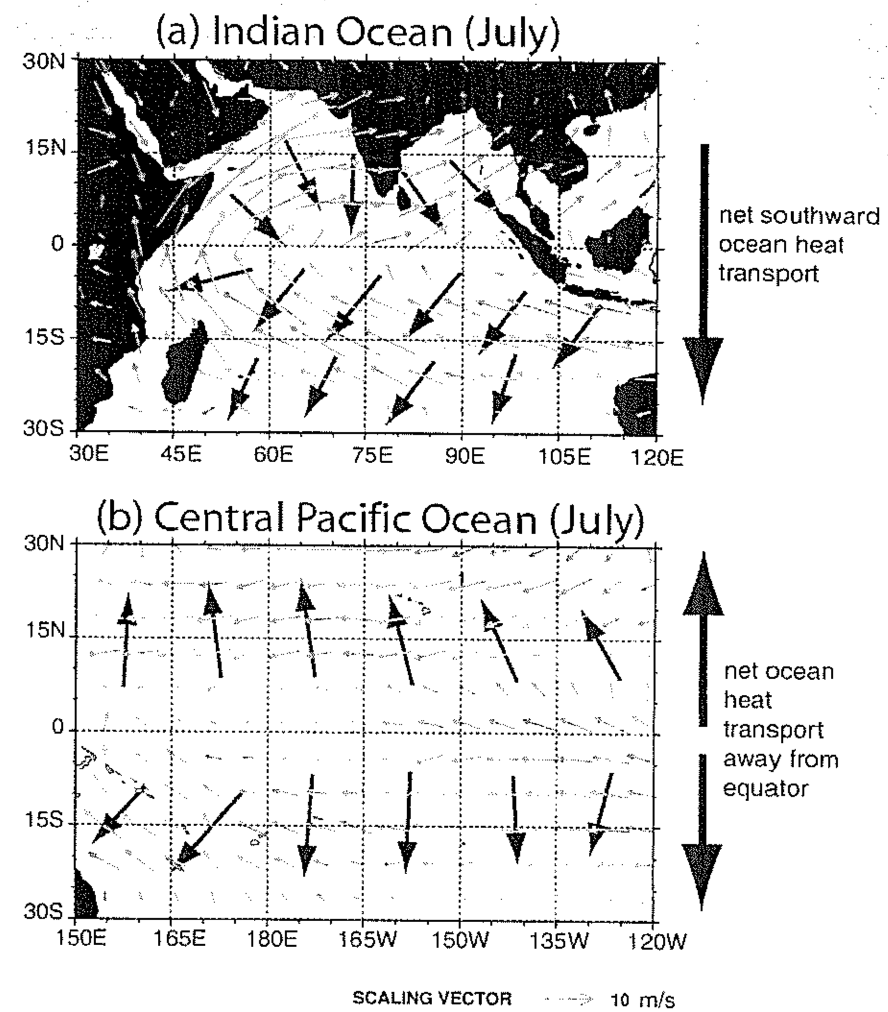

Figure 1-13. Surface wind distributions (gray vectors) for (a) the boreal summer monsoon in the Indian Ocean, and (b) the trade wind regime of the Pacific, and Ekman transport by the ocean (black arrows). The total Ekman transport mass transport is to the right of the surface wind in the Northern Hemisphere and to the left in the Southern Hemisphere. The overall effect of this wind-driven oceanic transport is to advect heat in the opposite direction to that advected by the divergent part of the wind field.

However, in reality, the ocean response to the Hadley and monsoon lower tropospheric winds act in the opposite sense and provide an exceedingly efficient negative feedback that regulates the strength of the circulation. This is because the effect of the rotation of the earth manifests itself by producing mass transports (totaled through the upper layers of the ocean) that are orthogonal to the direction of the surface wind—to the right of the surface wind in the Northern Hemisphere and to the left of the surface wind in the Southern Hemisphere. This effect is referred to as Ekman transport (see e.g., Loschnigg and Webster 2000). Thus, on a rotating planet, ocean mass and heat transport by the trade winds is away from the equator in both hemispheres. In the monsoon regions, where the flow is across the equator, the mass and heat transport are towards the winter hemisphere at all latitudes. Thus, although the atmospheric heat transport is towards the ITCZ or monsoon continental regions, the ocean transport is in the opposite direction (Fig. 1-13).

Noting that the Ekman transport is a wind-driven circulation and that it produces a heat transport that is opposite to that occurring in the atmosphere, it is easy to see why there is a negative feedback between the atmosphere and the ocean. Consider the situation where, for some reason, there is a stronger than average Hadley or monsoon cell, perhaps because the SST in the equatorial regions is warmer than normal. The lower-level winds will be stronger and, as a result, the ocean heat transport will be increased but in the opposite direction. The Hadley or monsoon cell will then be reduced in intensity because the ocean heat content near the ITCZ will be less, reducing the SST, and consequently, the intensity of convection. The opposite effect is also true. If the Hadley or monsoon is weak for some reason, wind-driven ocean transport will be also weak, heat will accumulate in the equatorial near-surface layers, and the atmospheric cell will return to average strength. Webster et al. (2002) describe this regulatory process in great detail for the monsoon circulation and use it as an argument to explain why the monsoon precipitation over South Asia does not deviate greatly from year to year and that there are rarely successive years of above-average or below-average rainfall.

Finally, it is interesting to contrast the manner in which the ocean and the atmosphere are heated and cooled. Figure 1-14 shows a comparison of the two systems. As was discussed at length in Sections 2.5 and 2.6, the atmosphere is heated at the planetary surface and also in the mid-troposphere, the latter through latent heat release in the ITCZ. Furthermore, it is cooled at the top of the atmosphere and by upper-tropospheric poleward transports of heat by the winds. The ocean, on the other hand, is heated and cooled at its surface through a net positive flux of heat at low latitudes and a net negative flux at high latitudes, respectively. Near the equator, the atmos-

phere is conditionally unstable, at least in the lower troposphere. The uppe ocean immediately below, however, is statically stable with warm, fresl water overlying colder, more saline, deeper water. The surface cooling a higher latitudes, on the other hand, renders the upper ocean convectivel unstable. Deep plumes of subsiding water result from this cooling, espe cially in the northern Atlantic Ocean and to some extent in regions aroun Antarctica. The heating and cooling of the atmospheric and oceanic Hadle cells and the regions where each is either convectively stable or unstabl appear as reversed mirror images.

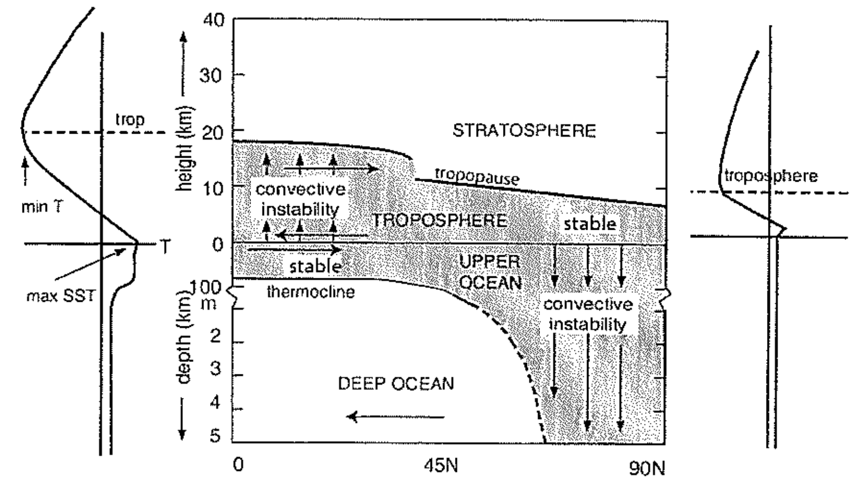

Figure 1-14. Schematic cross section of the atmosphere and the ocean showing the manner ir which the ocean and the atmosphere interact. Arrows indicate the direction of hea flux in the atmosphere and the ocean. Vertical temperature sections through the atmosphere and the ocean are shown, left (tropics) and right (polar regions). In the tropics, the warm SST renders the troposphere convectively unstable and heat is transferred to the upper troposphere, where it is either radiated away or advected tc higher latitudes. The ocean below, heated by the sun and freshened by rainfall, is extremely stable, on the other hand. The stability of the upper ocean minimizes vertical mixing of heat downwards, thus maintaining the high SSTs. At highei latitudes, on the other hand, cooling at the surface renders the ocean convectively unstable, in sharp contrast to the atmosphere above. Overall, the ocean and the atmosphere are reversed mirror images of each other. The shaded area may be referred to as the joint interactive zone of the ocean and the atmosphere, where interactions can take place on relatively short time scales. Adapted from Webster (1994).

7. SOME CONCLUDING REMARKS

The Hadley circulation and its local manifestations, such as the Asian-Australian monsoon system, impact most of the world's population both in the long term, by defining regions of habitability, and also in the short term by being involved in events such as ENSO. These circulations result from radiative imbalances between the equator and higher latitudes. However, because of the role of rotation, the importance of moist thermodynamics, and the involvement of the oceans, physical explanations of the Hadley circulation are rather complicated. Viewed collectively, the ocean and the atmosphere appear as a self-regulating system that is probably of importance in stabilizing the climate of the planet. It is important that the physics of the coupled ocean-atmosphere Hadley circulation be understood. For example, it is essential that data laid down by past climates be interpreted in a proper physical context. It is clear that future climate states will depend a great deal on how the Hadley circulation and the monsoons react to increases in greenhouse gases. Only with well-configured models will it be possible to determine the degree to which the negative feedbacks between the atmosphere and the oceans might mitigate changes imposed by anthropogenic effects or whether or not the feedbacks will be positive.

8. ACKNOWLEDGMENTS

This chapter resulted from a research grant from the program in Climate Dynamics of the National Science Foundation (NSF Grant: ATM 0328842). Thanks to the Royal Society and JSTOR at http://www.jstor.edu for permission to reproduce Figure 1-4 from Halley (1687) and for the availability of the papers by Edmund Halley and George Hadley. We are grateful for the use of the NCEP/NCAR reanalysis data set.

9. REFERENCES

Bates, J.R. 1970. Dynamics of disturbances on the intertropical convergence zone. *Quarterly Journal of the Royal Meteorological Society* 96: 677–700.

Bjerknes, J. 1966. A possible response of the atmospheric Hadley Circulation to equatorial anomalies of ocean temperature. *Tellus* 18: 820–829.

Bjerknes, J. 1969. Atmospheric teleconnections from the equatorial Pacific. *Monthly Weather Review* 97: 163–172.

Bryson, B. 2003. *A Short History of Nearly Everything*. New York, New York: Broadway, 560 pp.

Chang, C.P. 1977. Viscous internal gravity-waves and low-frequency oscillations in the trop ics. *Journal of the Atmospheric Sciences* 34: 901–910.

Charney, J.C. 1947. The dynamics of long waves in a baroclinic westerly current. *Journal o Meteorology* 4: 135–163.

Charney, J.C., and A. Elaissen. 1964. On the growth of the hurricane depression. *Journal o the Atmospheric Sciences* 21: 68–75.

Curry, J.A., and P.J. Webster. 1998. *Thermodynamics of Atmospheres and Oceans*. International Geophysics Series, Vol. 65, San Diego, California, Academic Press, 471 pp.

Eady, E.T. 1949. Long waves and cyclone waves. *Tellus* 1(3): 33–52.

Emanuel, K.A. 1979. Inertial instability and mesoscale convective systems. I. Linear theor of inertial instability in rotating viscous systems. *Journal of the Atmospheric Sciences* 36: 2425–2449.

Emanuel, K.A., J.D. Neelin, and C.S. Bretherton. 1994. On large-scale circulations in convecting atmospheres. *Quarterly Journal of the Royal Meteorological Society* 120 1111–1143.

Gill, A.E. 1980. Some simple solutions for heat-induced tropical circulations. *Quarterl Journal of the Royal Meteorological Society* 106: 447–462.

Hadley, G. 1735. Concerning the cause of the general trade-winds. *Philosophical Transactions of the Royal Society of London* 39: 58–62.

Halley, E. 1686. An historical account of the Trade Winds, and Monsoons, observable in th seas between the Tropicks, with an attempt to assign the physical cause of the sai Winds. *Philosophical Transactions of the Royal Society of London*, 16: 153–168.

Hunt, B.G. 1979. Influence of the earth's rotation on the general circulation of the atmosphere. *Journal of the Atmospheric Sciences* 36: 1392–1408.

Kirch, P.V. 2000. *On the Road of the Winds: An Archaeological History of the Pacific Islands before European Conquest*. Berkeley, California: University of Californi Press, 424 pp.

Kutzbach, G. 1987. Concepts of monsoon physics in historical perspective: The Indian monsoon (seventeenth to early twentieth century). In, Fein, J.S., and P.L. Stephen (eds.). *Monsoons*. Chichester, U.K.: Wiley Interscience., pp. 159–210.

Lindzen, R.S., and A.Y. Hou. 1988. Hadley circulations for zonally symmetric heating centered off the equator. *Journal of the Atmospheric Sciences* 45: 2416–2427.

Longuet-Higgins, M.S. 1968. The eigenfunctions of Laplace's tidal equations over a sphere. *Philosophical Transactions of the Royal Society of London A*. 262: 511–607.

Loschnigg, J., and P.J. Webster. 2000. A coupled ocean-atmosphere system of SST regulation for the Indian Ocean. *Journal of Climate* 13: 3342–3360.

Matsuno, T. 1966. Quasi-geostrophic motions in the equatorial area. *Journal of the Meteorological Society of Japan*, Series II, 44: 25–43.

Nicholson, S.E., and H. Flohn. 1980. African environment and climate changes and the general atmospheric circulation in late Pleistocene and Holocene. *Climate Change* 2: 313–348.

Ooyama, K. 1964. A dynamical model for the study of tropical cyclone development. *Geofisica International* 4: 187–198.

Peixoto, J.P., and A.H. Oort. 1992. *Physics of Climate*. Melville, New York: American Institute of Physics, 520 pp.

Philander, S.G. 1998. Is the temperature rising? The uncertain science of global warming. Princeton, New Jersey: Princeton University Press, 262 pp.

Riehl, H., and J.S. Malkus. 1958. On the heat balance in the equatorial trough zone. *Geophysica* 6: 503–538.

Royal Society. 1699. *Directions for Sea-men, Bound for Far Voyages. Philosophical Transactions*, Royal Society (London).

Shin S.I., Z. Liu, B. Otto-Bliesner, E.C. Brady, J.E. Kutzbach, and S.P. Harrison. 2003. A simulation of the last glacial maximum climate using the NCAR-CCSM. *Climate Dynamics* 20(2–3): 127–151.

Tomas, R., and P.J. Webster. 1997. On the location of the intertropical convergence zone and near-equatorial convection: The role of inertial instability. *Quarterly Journal of the Royal Meteorological Society* 123: 1445–1482.

Walker, G.T. 1924. World Weather I . *Memoirs of the India Meteorological Department* 24: 75–131.

Walker, G.T. 1928. World Weather III. *Memoirs of the India Meteorological Department* 2: 97–106.

Warren, B.A. 1987. Ancient and medieval records of the monsoon winds and currents of the Indian Ocean. In, Fein, J.S., and P.L. Stephens (eds.). *Monsoons*. Chichester, U.K.: Wiley Interscience, pp. 137–158.

Webster, P.J. 1972. Response of the tropical atmosphere to local steady forcing. *Monthly Weather Review* 100: 518–541.

Webster, P.J. 1994. The role of hydrological processes in ocean-atmosphere interaction. *Reviews of Geophysics* 32: 427–476.

Webster, P.J., and N.A. Streten. 1978. Late Quaternary ice-age climates of tropical Australasia: Interpretations and reconstructions. *Quaternary Research* 10: 279–309.

Webster, P.J., C. Clark, G. Chirikova, J. Fasullo, W. Han, J. Loschnigg, and K. Sahami. 2002. The monsoon as a self-regulating coupled ocean-atmosphere system. *Meteorology at the Millennium*. San Diego, California: Academic Press, pp. 198–219.

Webster, P.J., T. Palmer, M. Yanai, R. Tomas, V. Magana, J. Shukla, and A. Yasunari. 1998. Monsoons: Processes, predictability and the prospects for prediction. *Journal of Geophysical Research* 103 (TOGA special issue): 14451–14510.

Weiss, H., and R.S. Bradley. 2001. What drives societal collapse? *Science* 291(5504): 609–610.

Chapter 2

HADLEY CIRCULATION DYNAMICS

Seasonality and the Role of Continents

Kerry H. Cook
Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, New York14853-1504, U.S.A.

Abstract The equations that govern the Hadley circulation are reviewed, and the observed circulation is described. Atmospheric general circulation model (AGCM) simulations are used to evaluate the dominant zonally averaged momentum and thermodynamic balances within the Hadley regime.

A diagnostic application of the governing equations is used to identify the mechanisms of the Hadley circulation's seasonal evolution between equinox and solstice states. A "vertical driving" mechanism acts through the thermodynamic balance, and is important for regulating the circulation's strength when heating differences between seasons are close (within ~5°) to the equator. A "horizontal driving" mechanism acts through the horizontal momentum equations and is more effective off the equator. Unlike the results from axisymmetric models in which the prescribed heating is always close to the equator, the horizontal forcing mechanism is responsible for most of the Hadley circulation seasonality in the reanalysis and GCM simulations.

The presence of continental surfaces introduces longitudinal structure into tropical diabatic heating fields, and pulls them farther from the equator. The winter Hadley cells in a simulation with continents are much stronger than in a simulation with no continents, and the summer cell is half the intensity of that when continents are included. The strengthening of the winter cell occurs through an increase in low-level wind speeds, which enhances the zonal momentum flux from the surface into the atmosphere. The development of strong monsoon circulations in the Northern Hemisphere summer and the convergence zones of the Southern Hemisphere (South Pacific [SPCZ], South Atlantic [SACZ], and South Indian Ocean [SICZ] convergence zones) shifts mass out of the subtropics, lowers the zonal mean subtropical highs, and weakens the summer cell.