Mechanisms Limiting the Northward Extent of the Northern Summer Monsoons over North America, Asia, and Africa* 

CHIA CHOU 
Institute of Earth Sciences, Academia Sinica, Taipei, Taiwan 

J. DAVID NEELIN 
Department of Atmospheric Sciences, University of California, Los Angeles, Los Angeles, California 

(Manuscript received 11 October 2001, in final form 22 July 2002) 

ABSTRACT 

Mechanisms determining the poleward extent of summer monsoon convergence zones for North America, Asia, and Africa are examined in an intermediate atmospheric model coupled with a simple land model and a mixed layer ocean. Observations show that thermodynamical factors associated with the net heat flux into the atmospheric column provide favorable conditions for the monsoon convergence zone to extend farther poleward than actually occurs. To understand the discrepancy, a series of experiments are designed to test the importance of mechanisms previously examined in the South American case by the authors, namely, soil moisture, ventilation, and the interactive Rodwell–Hoskins mechanism. The latter refers to the interaction between baroclinic Rossby wave dynamics and convective heating. In North America, experiments suggest that ventilation by both temperature and moisture advection is a leading effect. The interactive Rodwell–Hoskins mechanism tends to favor east coast rainfall and west coast dryness. In Asia, ventilation by moisture advection is particularly important and the interactive Rodwell–Hoskins mechanism tends to favor interior arid regions and east coast precipitation. Overall, these dynamical factors are crucial in setting the poleward extent of the convergence zone over North America and Asia. Africa differs from the other continents because of the high surface albedo over much of northern Africa. Because there is less positive net flux of energy into the atmospheric column, convection is less thermodynamically favored and the dynamical factors, ventilation and the interactive Rodwell–Hoskins mechanism, have a weaker impact on preventing poleward extent of the convergence zone.

1. Introduction 

Land–ocean heating contrast is considered fundamental to summer monsoon circulations (Webster 1987; Young 1987). One might conjecture that the thermodynamics of ocean–atmosphere–land interaction could be the controlling factor in the forcing of monsoons. However, examining the relation between solar heating of the continent and the associated monsoon rainfall, the rain zone does not extend as far poleward as the maximum heating would seem to indicate. This suggests that other mechanisms besides the land–sea heating contrast determine the northward extent of the summer monsoon. Many studies (Lofgren 1995; Meehl 1994; Xue and Shukla 1993; Yang and Lau 1998; Nicholson 2000) discuss the importance of land processes, such as soil moisture and surface albedo, in affecting the magnitude and position of the monsoon. Topography, such as the Tibetan Plateau, also affects the monsoon circulation (Flohn 1957; Murakami 1987; Meehl 1992; Yanai and Li 1994; Wu and Zhang 1998). This is the third in a series of papers in which we examine some large-scale dynamical mechanisms that mediate land–sea contrasts and affect the poleward extent of monsoons, and compare them with land processes.

Chou et al. (2001, hereafter CNS) discuss mechanisms that affect the seasonal movement of continental convergence zones for the summer monsoon on an idealized, rectangular continent. Because of strong summertime insolation, thermodynamic conditions for establishing a convective zone appear favorable far poleward in the continent. The poleward extent of continental convection is largely determined by dynamical effects, such as ocean heat transport, ventilation, and Rossby wave dynamics (the interactive Rodwell–Hoskins mechanism) as well as by soil moisture feedbacks. The ventilation is defined as the import into the continents of low moist static energy air from the ocean.
regions or the export of high moist static energy out of the continents. The term “interactive Rodwell–Hoskins” (IRH) mechanism was coined by CNS to describe the way Rossby wave–induced subsidence to the west of monsoon heating interacts with the convergence zone. Rodwell and Hoskins (1996, 2001) examined the importance of baroclinic Rossby wave dynamics in creating subsidence to the west of specified heating regions. The adjective “interactive” in IRH stresses that the spatial pattern of the monsoon heating itself is determined interactively with the subsidence regions.

CNS conducted a series of numerical experiments that showed impacts of all of these processes, especially ventilation and the IRH mechanism, on the idealized continent. The conclusions might be modified and the relative importance of various mechanisms might differ in application to more realistic geography because of differences between continents. For instance, the width, shape, latitude, relation to surrounding oceans, and albedo of the particular continents may play a role.

The South American case has been discussed in Chou and Neelin (2001, hereafter CN). Examining various mechanisms, ventilation has a dominant impact in limiting the poleward extent of the South American summer monsoon. Here we would like to compare the impact of the various mechanisms on the Northern Hemisphere summer monsoons in North America, Asia, and Africa. There are many studies of detailed aspects of each of the monsoon systems: North American (e.g., Douglas et al. 1993; Higgins et al. 1997; Barlow et al. 1998; Yu and Wallace 2000); African (e.g., Palmer 1986; Xue and Shukla 1993; Lare and Nicholson 1994; Rowell et al. 1995; Eltahir and Gong 1996; Cook 1997; Semazzi and Sun 1997; Janicot et al. 1998); and Asian (Ramage 1971; Riehl 1979; Luo and Yanai 1984; Murakami 1987; Ding 1992; Yasunari and Seki 1992; Lau and Yang 1996; Li and Yanai 1996; Webster et al. 1998; Annamalai et al. 1999; and references therein). What we hope to add is a particular moist dynamical perspective on the large-scale aspects of these systems. No topography is used in this study, so we can examine the dynamical mechanisms in a situation that has the realistic land–ocean contrast and other geographic factors, but is slightly simpler to analyze. The Tibetan Plateau has a significant impact on the evolution of the Asian summer monsoon (Yanai and Li 1992), so the assumption of no topography implies caveats on the discussion of the Asian summer monsoon. Nevertheless, we should be able to get some insight into mechanisms in an Asia-like continent that possibly affect the Asian monsoon.

Second, we slightly refine some of the experimental design used in CNS and CN. To examine the ventilation effect, CNS and CN suppressed the horizontal advection terms that mediate it, namely, $\mathbf{v} \cdot \nabla T$ and $\mathbf{v} \cdot \nabla q$, where $T$ is temperature, $q$ is moisture, and $\mathbf{v}$ is horizontal wind. This is useful for hypothesis testing, but the remaining transport terms, such as $T \nabla \cdot \mathbf{v}$ and $q \nabla \cdot \mathbf{v}$ do not integrate exactly to zero across the domain (although cancellation occurs to a large extent) so the energy and moisture budgets of the model are not perfectly conservative in these experiments. To avoid this problem, the irrotational (purely divergent) part of wind ($\mathbf{v}_\ast$) is used to calculate the advection instead of suppressing the entire advection $\mathbf{v} \cdot \nabla$ terms. This also cleanly separates effects of the nondivergent and purely divergent flow. Details are discussed in section 2. To test the impacts of Rossby wave dynamics, CN and CNS used experiments that suppress the beta effect in part of the domain. We modify the design of this experiment to help distinguish between the effects of conventional Rossby wave dynamics and possible effects that might be associated with a “local Hadley circulation.”

A detailed description of experiment design and review of the model is given in section 2. The northern summer climate of the simulation and observations are compared in section 3. Estimates of an important variable, the net energy input into the atmospheric column ($F^{net}$), are presented and discussed. In sections 4 and 5, the North American monsoon and the Asian monsoon are examined, respectively. The most different case, the northern African monsoon, is examined in section 6, followed by a discussion.

2. The model and experiment design

a. The model

An atmospheric model of intermediate complexity (Neelin and Zeng 2000), a simple land surface model (Zeng et al. 2000) and a mixed layer ocean model (CNS) are used here. The atmospheric model uses analytical solutions for deep convection derived by Neelin and Yu (1994) and Yu and Neelin (1994) from the Betts and Miller (1986, 1993) deep convection scheme as leading basis functions for a Galerkin expansion in the vertical. We refer to this quasi-equilibrium tropical circulation model with a single vertical structure of deep convection for temperature and humidity as QTCM1. QTCM1 mimics a general circulation model (GCM) with the Betts–Miller deep convection scheme over deep convective regions, but it is roughly equivalent to a two-layer model far from deep convective regions. A cloud-radiation scheme is used to calculate effects of deep cirrocumulus/cirrus clouds. This parameterization includes a linearized longwave radiation parameterization (Chou and Neelin 1996) based on a full general circulation model scheme (Harshvardhan et al. 1987), a simplified shortwave radiation parameterization (Zeng et al. 2000) derived from Fu and Liou (1993) with prescribed marine aerosol, and an empirical parameterization for deep and cirrocumulus/cirrus cloud fraction (Chou and Neelin 1999). A modified version of QTCDM version 2.2 was used in these experiments. The main differences between this and model version 2.1 used in CNS and CN are in the treatment of vertical advection of momentum.
and diffusion terms, and small improvements to the treatment of cloud-radiative processes. These changes are documented as a subset of the version 2.3 release.

To simulate the interaction between the atmosphere and land surface, an intermediate land surface model with a single land surface layer for both energy and water budget is used. There are three land surface types—forest, grass, and desert—differing in soil moisture, field capacity, roughness length, and evapotranspiration characteristics (see Zeng et al. 2000). Prescribed surface albedo from the Darnell et al. (1992) climatology is used. The version of the model used here does not include topography.

A mixed layer ocean model (50-m fixed depth) with prescribed divergence of ocean heat transport, known as the Q flux (Hansen et al. 1988), is used to examine interaction between the atmosphere and the ocean. The seasonal Q flux is estimated from a 10-yr model run with climatological SST. With the prescribed Q flux, SST is determined by the energy balance between latent heat, sensible heat, surface radiative flux, and Q flux. The resulting surface heat fluxes are discussed in section 3.

b. Refinement of the experiments for ventilation

In order to examine the ventilation effect, CNS suppress the advection terms \( \mathbf{v} \cdot \mathbf{\nabla} T \) and \( \mathbf{v} \cdot \mathbf{\nabla} q \) in the temperature and moisture equations. Adiabatic cooling and moisture convergence associated with the divergence are retained. This approach does indeed show the importance of these terms in the transition from a regime typical of tropical dynamics to a different dynamical balance that limits the extent of the convection zones. However, suppressing these terms causes errors in the conservation of energy and moisture, the magnitude of which is hard to estimate in advance. Furthermore, it is desirable to separate more clearly the effects of the divergent (irrotational) part of the flow from those of the nondivergent part. We present a refinement of these experiments by suppressing the part of the advection associated with the nondivergent wind. Both the energy budget and moisture budget thus retain conservation properties in certain cases. The moisture equation can be written

\[
\partial_t q + \mathbf{v}_e \cdot \nabla q + \mathbf{v}_x \cdot \nabla q + \omega \partial_z q = \cdots, \tag{1}
\]

where \( \mathbf{v}_e \) is nondivergent part of horizontal wind, \( \mathbf{v}_x \) is irrotational part of horizontal wind, \( \omega \) is vertical velocity in pressure level, and \( q \) is specific humidity. By definition,

\[
\nabla \cdot \mathbf{v}_e = 0, \tag{2}
\]

so

\[
\mathbf{v}_e \cdot \nabla q = \nabla \cdot (\mathbf{v}_e q). \tag{3}
\]

Taking the area-weighted integral of the advection by \( \mathbf{v}_e \) through the entire domain,

\[
\int_D \mathbf{v}_e \cdot \nabla q \, dA = 0, \tag{4}
\]

using periodic boundary conditions in longitude and \( \mathbf{v}_e \cdot \mathbf{n} = 0 \) at the poleward boundaries, where \( \mathbf{n} \) is the normal direction at the boundary. Thus dropping \( \mathbf{v}_e \cdot \nabla q \) for the testing purposes does not affect domain-integrated conservation. Similar results hold for temperature. Usually \( \mathbf{v}_x \) is much smaller than \( \mathbf{v}_e \) so using \( \mathbf{v}_x \) to calculate the advection terms reduces the advection very substantially. This experiment thus largely suppresses the ventilation mechanism that is mediated by \( \mathbf{v} \cdot \nabla q \) terms. The climatology of \( \mathbf{v}_e \) and \( \mathbf{v}_x \) will be discussed in section 3a.

In order to limit the impacts on the climatology through effects outside the region of interest, we define target regions over which the experiments are applied for each continent. These target regions are specified in section 2d. When applying the compression of \( \mathbf{v}_e \cdot \nabla q \) experiment to target regions, (4) becomes instead

\[
\int_{\text{Region}} \mathbf{v}_e \cdot \nabla q \, dA = \oint \mathbf{v}_e q \cdot \mathbf{n} \, dl, \tag{5}
\]

where \( \oint (\cdot) \, dl \) is around the boundary of the target region. Suppressing \( \mathbf{v}_e \cdot \nabla q \) thus creates an experiment equivalent to stopping the moisture flux by \( \mathbf{v}_e \) at the boundaries of the target region and otherwise conserves moisture (and enthalpy in the \( T \) equation). The experiment thus has a simple physical interpretation. Furthermore, we have some freedom in choosing boundaries to reduce the size of this term. As the target region area becomes large, the right-hand side of (5) divided by the domain size becomes small relative to average precipitation in the region. For instance, for the no-ventilation experiment in the North American case (section 4) the average precipitation over the target region is 6.3 mm day\(^{-1}\) whereas the omitted moisture sink due to (5) is 0.3 mm day\(^{-1}\). For the Asian case in section 5, the omitted moisture source due to (5) expressed as a domain average is 0.3 mm day\(^{-1}\), compared to an average precipitation of 8.8 mm day\(^{-1}\). Corresponding effects in the temperature equation are of similar magnitude.

In all these experiments the irrotational (purely divergent) part of the wind is consistently kept, that is, both the terms involving the divergence itself, and the \( \mathbf{v}_x \cdot \nabla \) terms. The vertically integrated moisture equation may be written equivalently as

\[
\partial_t q + \mathbf{v}_x \cdot \nabla q + \omega \partial_z q = E - P + \cdots, \tag{6}
\]

\[
\partial_t q + \mathbf{\nabla} \cdot \mathbf{\hat{v}}_x q = E - P + \cdots, \tag{7}
\]

using \( \omega = 0 \) at top and bottom of the model where \( \mathbf{\hat{v}}_x \) is a mass-weighted vertical integral, \( E \) is evaporation, and \( P \) is precipitation.

We remark that the nondivergent component of flow cannot increase the maxima of moisture and cannot increase moisture beyond values in the upstream flow. The
divergent component of the flow can increase the maxima or decrease the minima of the moisture field. Diagnostics of the moisture budget that display \( v_q \) as a moisture transport tend to be misleading in that they are dominated by \( v_q q \).

c. Testing the importance of Rossby wave dynamics by suppressing \( \beta \)

The momentum equation can be written

\[
\frac{D}{Dt} \mathbf{v} + f_0 \mathbf{k} \times \mathbf{v} + (f - f_0) \mathbf{k} \times \mathbf{v} + (f - f_0) \mathbf{k} \times \mathbf{v}' = -\nabla \phi + \ldots, \tag{8}
\]

where \( f_0 \) is a reference value of the Coriolis parameter, \( \mathbf{v} \) is zonally averaged across a target region, \( \mathbf{v}' \) is the departure from the zonal average in a target region, \( \mathbf{k} \) is a vertical unit vector, and \( \phi \) is geopotential. The \( \beta \) effect is contained in the \( f - f_0 \) terms. In previous experiments (CNS, CN), we suppressed effects associated with Rossby wave dynamics by dropping the \( f - f_0 \) terms in the momentum equation within the target region. However, one might consider that for a case where this is applied across the whole hemisphere, it also impacts the effects of variable \( f \) on conservation of angular momentum in the Hadley cell. Even for a target region of smaller longitude extent, one might worry that in addition to affecting Rossby wave dynamics, these experiments also impact the dynamics of meridional overturning circulations similar to the Hadley circulation within the continental target region.

Suppose we define local Hadley circulation as an overturning circulation involving the zonal mean across the target region. We caution that the concept of a local Hadley circulation is hazily defined from a dynamical point of view: at zonal scales much larger than the target region, a local overturning circulation is simply part of a large-scale stationary Rossby wave. Any rising or sinking motion may be partially compensated to the east or west of the region, rather than purely meridionally. Nonetheless, the terminology is used in the field and we would like to address whether making the distinction between zonal mean dynamics across the region and Rossby wave dynamics within the region is important.

More precisely, we use the following means of separating the dynamics of such a local Hadley circulation from the dynamics of Rossby waves within the target region. In (8) \( (f - f_0) \mathbf{k} \times \mathbf{v} \) represents Coriolis effects on the local Hadley circulation while \( (f - f_0) \mathbf{k} \times \mathbf{v}' \) is associated with Rossby wave dynamics within the target region. By suppressing the latter term, while keeping \( (f - f_0) \mathbf{k} \times \mathbf{v} \), an experiment with “partial-\( \beta \) effect,” that is, having the \( \beta \) effect on zonally averaged wind in the target region, is defined. This maintains the \( \beta \) effect of a local Hadley circulation while suppressing Rossby waves on scales smaller than the domain size. (For zonal scales much larger than domain, the target region can still act like a coarse-grained section of a larger Rossby wave.) In the partial-\( \beta \) experiments, there is no east–west asymmetry within the domain arising from Rossby wave dynamics. If the partial-\( \beta \) experiments and the no-\( \beta \) experiments behave similarly, then it removes the question of whether one need worry about distinguishing local Hadley cell dynamics from Rossby wave dynamics more generally.

There are edge effects associated with the change in \( f \) at the domain boundaries as noted in CN. Rossby edge waves can occur and the anticipated decay scale inside the region is less than the radius of deformation, on the order of a few hundred kilometers. Choice of target region boundaries is discussed below.

d. Experiments

A control run for seasonal climatology from a 10-yr average is used to compare to other experiments. Besides the control run, seven different experiments are designed to examine various effects that limit the northward extent of northern summer convection zones. A summary is given in Table 1. The experiments are applied, in turn, to each of the target regions described below. An additional set of experiments using a modified albedo specification is conducted for the African case, as described in section 6. Results are for 10-yr averages after a spinup period (10 yr in most cases) to ensure statistical equilibrium has been reached.

The first experiment, “saturated soil moisture,” examines the effect of land hydrology; soil moisture in
a target region is set to field capacity, so water supply to the atmosphere by evapotranspiration is not a limiting factor. The second experiment, “no ventilation,” is designed to examine the ventilation effect by suppressing the contributions of the nondivergent flow $v_n$ to the horizontal advection of moisture and temperature as described in section 2b. A pair of experiments with $v_n \cdot \nabla T$ suppressed (“no T ventilation”) and $v_n \cdot \nabla q$ suppressed (“no q ventilation”) are used to examine the importance of the energy transport and the moisture transport, respectively.

To examine the effects of Rossby wave dynamics and the local Hadley cell, three experiments are conducted. The “no-$\beta$” experiment, uses only the $f_k \times v$ term of (8), so no $\beta$ effect acts within the target region. With a combination of constant $f_0$ and no ventilation (“no ventilation and no $\beta$”), the dynamical effects on determining the position of the monsoon rainfall disappear, so only the thermodynamical factor affects the monsoon and it can be examined here. Using $f_0$ changes the circulation thus the ventilation. In order to examine the effect of the local Hadley cell, the no-$\beta$ and no-ventilation experiment is used to compare with the “no-ventilation and partial-$\beta$” experiment where $(f - f_0) \times v$ is suppressed.

The model modifications for each experiment are applied in a target region for North America, Asia, and Africa, respectively. A value of $f_0 = f(7.5^\circ N)$ is used in the North American and African cases, while $f_0 = f(12.5^\circ N)$ is used in the Asian case. The size of each target region is chosen to cover the continent or subsection of continent and large enough to minimize edge effects on the forcing region. For the North American case, the target region is between $7.5^\circ - 52.5^\circ N$ and $180^\circ - 45^\circ W$. For the Asian case, the target region is between $12.5^\circ - 52.5^\circ N$ and $162^\circ - 146^\circ E$. For the African case, the target region is between $7.5^\circ - 42.5^\circ N$ and $22.5^\circ W - 51^\circ E$. Edge effects are largest near the target region boundaries, but can have some impacts within the target region. In addition to Rossby edge waves in the no-$\beta$ and partial-$\beta$ experiments, other experiments may have potential edge effects. There is a trade-off in determining the target region size for such experiments. A larger domain for the target region places more distance from the edge, but impacts the model climate over a larger scale. We have chosen the domains with this compromise in mind, and then have conducted many sensitivity tests (not shown) with variants of these target regions to ensure the edge definition does not affect the conclusions. We have also conducted experiments (not shown) with the entire Northern Hemisphere as the target region to verify that results are qualitatively similar.

3. Northern summer climate

a. Precipitation and wind fields

Figures 1a and 2a show July climatology of precipitation from the model and the observations, respectively. In the Tropics, the model precipitation is similar to the Xie–Arkin precipitation (1996). The North American summer monsoon is a little weak and does not extend far enough northward compared to observations. It compares favorably with some GCM simulations (Yang et al. 2001). Considering that both topography and mesoscale effects are known to affect Mexican monsoon precipitation (Berbery 2001), this result is reasonable. The Asian monsoon is simulated by the model, although the model does not have distinct maximum precipitation over the Bay of Bengal. We emphasize that the absence of the Tibetan Plateau in the model certainly impacts particular aspects especially at local scales. However, the large-scale aspects of the rainfall pattern, including the northward extent along the east coast tend to be simulated. The African monsoon is well established, although the precipitation is slightly stronger than observed. The northward limit of the equatorial convergence zone occurs in the Sahel region as observed. The model precipitation associated with storm tracks in the Atlantic and Pacific is located farther east than in observations, possibly associated with weaker storm development in the summer season. This also affects the precipitation along the east coast of North America.

The model wind fields at 850 (Fig. 1b) and 500 mb (Fig. 1c) are similar to the National Centers for Environmental Prediction (NCEP) reanalysis (Figs. 2b and 2c). At 850 mb, the trade winds are slightly stronger than the NCEP data. The cross-equatorial flow over the Indian Ocean is simulated by the model, but the northward component is weak compared to the observations. This is associated with strong precipitation over the western part of the Indian Ocean, south of the corresponding maximum precipitation from the observation. This northeastward flow joins the trade winds from the western Pacific flowing into east Asia. Distinct subtropical anticyclonic circulation, such as over the northeastern Pacific and the Atlantic are generally well simulated by the model. The Atlantic subtropical high is weak. The anticyclonic circulation associated with the Pacific subtropical high in the model (Fig. 1b) is farther east than the observation in Fig. 2b. The strong northward flow near Arizona at 850 mb, which transports low-level moisture into this region is not simulated, perhaps contributing to weaker North American monsoon rainfall. At 500 mb the lower portion of the subtropical jet may be seen (the 300-mb view may be seen in Fig. 3). The location in both hemispheres is reasonable compared to NCEP data in Fig. 2, although in the Northern Hemisphere the flow extends slightly southward over the Pacific–North American sector, which may affect the North American monsoon.

Over the African continent, the model simulates low-level westerlies (Fig. 1b) between the equator and $15^\circ N$, which are a main component of the African monsoon. North of the westerlies, the Harmattan easterlies close the circle of a cyclonic flow associated with a thermal
low centered over Saharan Africa. At 500 mb, distinct easterlies, the African easterly jet (AEJ) from the tropical easterly jet, and the Harmattan easterlies are found. However, the AEJ in the model is weaker and the altitude of the AEJ is a little higher than in observations. At 600 mb this feature is weak in the model. The AEJ is caused by positive meridional temperature gradients at the surface associated with strong meridional soil moisture gradients and a reversal of the meridional temperature gradient in the lower troposphere, around 600 mb (Cook 1999; Thorncroft and Blackburn 1999). In the model, the reversal of the temperature gradient is not simulated. Thus the AEJ is not distinguished from the tropical easterlies at the upper troposphere.

The nondivergent and irrotational parts of wind at 300 mb are shown in Fig. 3. The irrotational part of the wind is plotted at a scale 1/6 that of the nondivergent part of wind. Maps of total 300-mb wind (not shown) look very similar to the nondivergent wind in Fig. 3b. Because the nondivergent wind is stronger, using only the irrotational part of wind to calculate the advection terms of $T$ and $q$ can significantly reduce the ventilation
effect, as described in section 2b. The upper-level divergent flow is closely related to low-level convergence. The strongest features are associated with the convection zones, especially in the tropical western Pacific and Indian Oceans.

b. The net energy input into the atmosphere, $F_{\text{net}}$

Here $F_{\text{net}}$ is defined as the net energy input into the atmospheric column:

$$F_{\text{net}} = S^\rightarrow_h - S^\leftarrow_h - S^\rightarrow_t + S^\leftarrow_t - R^\rightarrow - R^\leftarrow + E + H,$$

where subscripts $s$ and $t$ on the solar ($S^\rightarrow$ and $S^\leftarrow$) and longwave ($R^\rightarrow$ and $R^\leftarrow$) radiative terms denote surface and model top, and $R^\rightarrow_t \approx 0$ has been used. The $E$ is evaporation and $H$ is sensible heat flux. In general, $F_{\text{net}}$ must be balanced by divergence of the moist static energy transport in the vertically integrated moist static energy equation. This includes a term associated with the purely divergent component of the circulation, $\vec{h} \cdot \vec{\omega}$, where $h$ is moist static energy and $\omega$ the vertical velocity, and advection terms $\vec{v} \cdot \nabla T$ and $\vec{v} \cdot \nabla q$. In the Tropics, horizontal gradients of temperature and thus $\vec{v} \cdot \nabla T$ are small and $\vec{v} \cdot \nabla q$ is relatively small. Positive $F_{\text{net}}$ must be balanced by rising
motion and divergent circulation (including low-level convergence). A number of theoretical studies have made use of this balance, including evidence that the effectiveness of the $F_{\text{net}}$ forcing is determined by the gross moist stability of the troposphere (Neelin and Held 1987; Neelin and Yu 1994) which is relatively constant within convective zones (Yu et al. 1998). The rising motion is associated with low-level convergence and moisture convergence. Zeng and Neelin (1999) examine how $F_{\text{net}}$ drives moisture convergence in deep convective regions over land. Sobel and Bretherton (2000) show that a constant temperature approximation captures general features of convection zones equatorward of 15° latitude by a balance of $F_{\text{net}}$ and the divergent circulation (although $v \cdot \nabla q$ terms improve the simulation). It is thus surprising that the spatial distribution of $F_{\text{net}}$ has been discussed relatively little based on observations.

Figure 4 shows $F_{\text{net}}$ and the net heat flux into the surface ($F_s = S^\downarrow + R^\uparrow - E - H$) for July, as estimated from observations. The top-of-atmosphere radiative fluxes are from the Earth Radiation Budget Experiment (ERBE; Ramanathan 1987). The surface fluxes are from the Comprehensive Ocean–Atmosphere Data Set (COADS; Woodruff et al. 1987), except that the Darnell et al. (1992) estimate for the surface shortwave is used. Over land, a condition of zero net surface flux is used to obtain $F_{\text{net}}$. This assumption holds for timescales longer than a day. Since some of the observed components of $F_s$, $E$, $H$, $S^\downarrow$, $S^\uparrow$, $R^\uparrow$, and $R^\downarrow$, are difficult to measure, errors are larger in $F_s$ than in the top-of-atmosphere data. Because the zero $F_s$ condition is even more accurate, the estimate of $F_{\text{net}}$ is considerably more accurate over land than over the ocean.

In Fig. 4, $F_{\text{net}}$ is positive in most of the regions of oceanic convective zones in the Tropics including the Pacific intertropical convergence zone (ITCZ) and South Pacific convergence zone (SPCZ). An exception is a narrow strip near the equator in the western Pacific that has a slightly negative value of $F_{\text{net}}$ while the convection is strong. This is likely due to errors in the surface flux data over the region—the net surface flux $F_s$ implied by the COADS and Darnell datasets is over 60 W m$^{-2}$ into the oceans, whereas other estimates suggest it is smaller (Cronin and McPhaden 1997). This would imply
Fig. 4. Jul climatology from observations for (a) the net flux of energy into the atmospheric column, $F_{\text{net}}$ (net surface and top-of-atmosphere shortwave and longwave radiation, sensible heat, and latent heat fluxes); (b) the net surface flux $F_s$. Both are in W m$^{-2}$, shaded for positive values.

a substantial positive $F_{\text{net}}$ in this region. Nonconvective regions in the Tropics tend to coincide with strongly negative $F_{\text{net}}$ in this estimate. If surface solar is overestimated (e.g., Arking 1999), the pattern would remain essentially the same while the zero line would shift. The overall tendency would remain for the net flux forcing of the atmospheric column to be positive in convection zones and to be negative in dry regions.

Turning to the more accurate estimates of $F_{\text{net}}$ over continental regions, the winter hemisphere continents have strongly negative $F_{\text{net}}$ outside the Tropics, as expected. During the summer season, North America and Asia both have strong positive $F_{\text{net}}$. This positive thermodynamic forcing of the atmospheric column extends very far poleward, and is roughly of similar magnitude from the subtropics to the Poles. This is because the summer insolation has a large positive input extending all the way poleward. The part of outgoing longwave radiation (OLR) associated with temperature has large spatial scales due to dynamical effects on upper-tropospheric temperature and does not fully compensate for the large summer insolation into the atmosphere-land column. Cloud effects can modify both longwave and shortwave but do not change the basic picture.

This evaluation of summer continental $F_{\text{net}}$ raises the question regarding the summertime convection zones: what sets the poleward edge of the convection zone? Positive thermodynamic forcing of the tropospheric column exists over the whole summer continent. In a tropical dynamical regime, this would be balanced by rising motion with associated low-level convergence. Low-level convergent motions supply moisture so unless some other factor intervenes, a convection zone is established. Since the convergence zones and associated strong convection do not move nearly so far poleward as the region of positive $F_{\text{net}}$, clearly a transition to some other dynamical regime must occur in the summer continents. Comparing the summertime precipitation (Fig. 2a) to $F_{\text{net}}$ over summer hemisphere continents makes this clear. The region of strong precipitation has a sharp edge not associated with changes in the smoothly varying $F_{\text{net}}$. This implies some dynamical mechanism must set the edge of the convergence zone, not the simpler thermodynamic balance of positive net energy input with rising/divergent motions. Chou and Neelin show the Southern Hemisphere summer counterpart of Fig. 4a, and discuss similar considerations in southern summer where positive $F_{\text{net}}$ likewise extends far poleward over southern continents.

Northern Africa differs from Asia and North America. In Fig. 4, most of equatorial and northern Africa has positive $F_{\text{net}}$ but over desert regions its magnitude is
considerably smaller than typical summer continental values. This is principally due to large surface albedo, which exceeds 0.4 in some parts of the Sahara. Albedo effects in northern Africa have long been considered a contributor to regional deserts (Charney 1975), so we can anticipate differences between the African case and the other continents.

The zero surface heat flux condition makes land–atmosphere interaction very different from ocean–atmosphere interaction where heat storage and heat transport by the ocean create nonzero surface heat flux. This is quantified in Fig. 4b, which shows the net surface heat flux. Over the Northern (summer) Hemisphere, a large heat flux into the ocean exceeds 100 W m$^{-2}$ over most of the subtropical to higher-latitude oceans. The leading contributor to this would be ocean heat storage. The heat flux into the ocean in the deep Tropics tends to persist in the annual average and thus must be substantially associated with ocean heat transport. The combination of net heat flux into the ocean and zero heat flux over land (Fig. 4b) produces a strong land–ocean contrast. The consequence for the atmospheric heating is positive $F_{\text{net}}$ over land and negative $F_{\text{net}}$ over ocean (see Fig. 4a) in the summertime subtropical and higher latitudes.

These observations of $F_{\text{net}}$ and the implications outlined above appear not to have been discussed in the literature, although there have been a number of studies that estimate zonal averages of related quantities (e.g., Carissimo et al. 1985). Trenberth and Solomon (1994) use the European Centre for Medium-Range Weather Forecasts (ECMWF) analyses to present the divergence of the atmospheric energy transport, which must be equal to $F_{\text{net}}$ for a system that conserves energy. Their results tend to agree qualitatively for large regions, although they are contaminated by spatially noisy signals. Since the ECMWF analysis does not maintain energy conservation there is good reason to believe that the analysis in Fig. 4 is the more accurate, especially over land regions where the accuracy is that of the ERBE dataset.

The model simulation of $F_{\text{net}}$ and $F_s$ is shown in Fig. 5. Comparing Fig. 5a to Fig. 4a, the pattern of $F_{\text{net}}$ from the model is similar overall to the observations. Deep convection regions over the oceans in the Tropics have a positive value of $F_{\text{net}}$. Over land, positive $F_{\text{net}}$ over northern continents and negative $F_{\text{net}}$ over southern continents are found. The simulated $F_{\text{net}}$ tends to be about 15 W m$^{-2}$ stronger than the observations. Over regions that have small cloud cover, such as the Sahara desert and the cold tongue of the Pacific Ocean, a contributing factor is that the OLR from the model (not shown) is less than the observations over clear-sky regions. Nonetheless, the model results still show relatively low $F_{\text{net}}$ over the Sahara desert as observed. In Fig. 5b, $F_s$ is positive over the summer hemisphere oceans (except for
a small region in the Indian Ocean where simulated cloud fraction is too large). This creates a strong ocean–atmosphere contrast in \( F_{\text{net}} \) in subtropical to high latitudes, as observed.

4. North American case

Figure 6 shows the results for the control run and the seven experiments discussed in section 2d and listed in Table 1 applied to the North American target region. The control run in Fig. 6a as discussed in section 3 has a weaker North American monsoon than observed: the rain zone is a little too far south and the east coast does not receive enough rain. This may be due in part to the lack of topography (Broccoli and Manabe 1992) and in part to the slight shift of the storm track eastward from the coast. Figure 6b with saturated soil moisture shows some impacts. There is greater northward extension of the North American monsoon rainfall over the southwest region of the United States compared to Fig. 6a. The precipitation over the eastern part of the continent in Fig. 6b also increases. In the saturated soil moisture case, the corresponding evaporation increases and enhances local precipitation. The increase of evaporation also reduces surface temperature over North America and produces anticyclonic circulation anomalies. These anomalies transport extra moisture from the Gulf of Mexico into the south and southeast of the United States. Thus the movement of the monsoon and the increase of the rainfall is the result of both local evaporation and low-level moisture convergence. The saturated soil moisture results here may overestimate the effect of soil moisture, perhaps because of the weaker North American monsoon in the control. Other soil moisture experiments (Fennessy and Shukla 1999; Zeng et al. 2000) show smaller sensitivity of the summer rainfall though the tendency is consistent.

To test the ventilation mechanism via suppressing the horizontal advection of \( T \) and \( q \), only the irrotational component of the wind \((\mathbf{v}_q)\) is used to calculate the advection as described in section 2b. The precipitation (Fig. 6c) becomes much more intense, especially over the south and east of the continent, and the rain zone extends northward over much of the continent. Clearly, the ventilation has a very strong impact on the extent of the rain zone (much stronger than soil moisture). The intervention of turning off ventilation is rather drastic, so the resulting high precipitation is not something one expects to occur in realistic climate situations, but the experiment does show that ventilation is a leading order effect.

Some east–west asymmetry remains even where ventilation is suppressed. In Fig. 6c, the southwestern United States is the only area with small precipitation. For different versions of the model, the size of this less rainy area varies, but the east–west asymmetry is very consistent and it is consistent with the idealized monsoon case in CNS. It is thus likely associated with Rossby wave effects, suggesting these are important to establishing arid regions in the Southwest.

A test of the effect of Rossby wave dynamics is presented in Fig. 6d for the no-\( \beta \) experiment, that is, using the reference value of the Coriolis parameter, \( f_0 \) over the region. Almost the entire continent is covered by convective rain over 2 mm day\(^{-1} \). Bearing in mind the caveat that suppressing \( \beta \) is a rather heavy intervention in the dynamics, the experiment does show that dynamical factors related to \( \beta \) have a significant impact. The precipitation in Fig. 6d does not have much east–west asymmetry, suggesting that the \( \beta \) effect is important in creating the east–west asymmetry of the convection zone. Unfortunately, these results cannot be compared directly to Fig. 6c to examine whether ventilation or Rossby wave effects have a stronger impact because the modification of \( f \) greatly alters the circulation and thus changes the ventilation as well. To further examine the effects of ventilation and the \( \beta \) effect, the no-ventilation and no-\( \beta \) experiment is shown in Fig. 6e. The east–west asymmetry of the precipitation has disappeared. With the dynamical factors suppressed, such as the ventilation and the Rossby wave effect, the thermodynamical factors (shown by \( F_{\text{net}} \)) control the convection zone, so the distribution of precipitation follows \( F_{\text{net}} \) extending northward.

Figure 6f shows the results of an experiment aimed at distinguishing the effects of Rossby wave dynamics from effects that could be described as a local Hadley cell. Of the two terms in (8) associated with the \( \beta \) effect, this experiment has \((f - f_0)\mathbf{k} \times \mathbf{v}\) included and \((f - f_0)\mathbf{k} \times \mathbf{v}'\) suppressed over the North American region. Ventilation is also suppressed, so this no-ventilation and partial-\( \beta \) experiment may be compared with Fig. 6e where the \( \beta \) effect is suppressed entirely, as is ventilation. Figure 6f shows that precipitation covers the entire North American continent. The convergence zones in both experiments extend far poleward; the case in Fig. 6e in fact does not extend as far as Fig. 6f although we do not attribute much significance to the differences in the far northern continent. Including the \( \beta \) effect on the local Hadley cell clearly does not prevent the convection zone from extending northward. Thus descent in a local Hadley cell can be eliminated as a significant mechanism limiting the extent of the convection zone. The northward edge is affected by the Rossby wave dynamics, which concurs with the IRH mechanism as described in CNS.

A pair of experiments designed to examine the effects of \( \mathbf{v}_\phi \cdot \nabla T \) and \( \mathbf{v}_\phi \cdot \nabla q \), respectively, on the ventilation mechanism are shown in Figs. 6g and 6h. From the energy budget of the control run, (not shown) \( \mathbf{v}_\phi \cdot \nabla T \) has a cooling tendency exceeding 60 W m\(^{-2} \) over most of the continent north of the monsoon region, whereas \( \mathbf{v}_\phi \cdot \nabla q \) has magnitude less than 30 W m\(^{-2} \) except in small regions of moistening or drying. Thus it would appear that \( \mathbf{v}_\phi \cdot \nabla T \) is more important for balancing positive \( F_{\text{net}} \) over the continent than \( \mathbf{v}_\phi \cdot \nabla q \), as is discussed.
Fig. 6. Jul precipitation for experiments over the North American region: (a) control run, (b) saturated soil moisture, (c) no ventilation, (d) no $\beta$, (e) no ventilation and no $\beta$, (f) no ventilation and partial $\beta$, (g) no $T$ ventilation, and (h) no $q$ ventilation. Contour interval is 2 mm day$^{-1}$. See Table 1 and section 2d for details of the experiments.
in CNS. When suppressing the effect of \( \mathbf{v}_a \cdot \nabla T \), a strong northward extension of the rain zone might be expected. However, the result in Fig. 6g does not show any northward movement of the rain zone. Without \( \mathbf{v}_a \cdot \nabla T \), the continent cannot cool down efficiently, so the surface and air temperature increase significantly. The value of low-level moisture required to create convective available potential energy (CAPE) and initiate convection thus increases. The oceanic moisture values imported by \( \mathbf{v}_a \cdot \nabla q \) are lower than what is required to initiate convection, so most of the continent remains without a sustained convergence zone. When suppressing the effect of \( \mathbf{v}_a \cdot \nabla q \), the moisture of the atmosphere increases relative to the control in the southern continent while the temperature remains about the same. Figure 6h shows a very slight northward extension of the rain zone. However, the positive \( F_{\text{net}} \) is balanced by \( \mathbf{v}_a \cdot \nabla T \), so rising and convergent motions are not generated over most of the continent and the change in the convection zones remains small.

Comparing Figs. 6g and 6h to Fig. 6a, either advection term is enough to prevent the convergence zone extending northward. A strong nonlinear effect is thus involved in the ventilation mechanism. When suppressing one advection term, the effect of the other term on cooling and drying is enhanced. The results of the experiments are thus very different from attributions based on the budget analysis in which both advection terms appear to contribute to preventing northward extension of the convergence zone. One should thus use caution when employing budget analysis to understand mechanisms.

A similar nonlinearity is noted in the impact of suppressing the ventilation mechanism and that of suppressing the \( \beta \) effect. Suppressing either creates large changes in the continental precipitation so ranking their importance must be done with caution.

We have repeated the experiments shown in Fig. 6 with different model versions and find the general conclusions to be robust. There is considerable sensitivity in the quantitative distribution of precipitation because of the drastic changes that the experiments produce. The results here are used only to qualitatively estimate the impact of these mechanisms on determining the northward extent of the convergence zone.

5. Asia-like continent case

We reiterate that the Tibetan Plateau is important to producing a realistic Asian monsoon. Using a continent that has the geography of Asia without topographic effects can, however, give us a first glimpse of mechanisms that can affect the Asian monsoon. GCM simulations that remove this plateau (Hahn and Manabe 1975; Zhisheng et al. 2001) do retain many large-scale aspects of the monsoon convection zones. Caveats on the climatology in Fig. 7a include the following: it does not have a distinct feature of maximum precipitation over the Bay of Bengal, the precipitation in south and Southeast Asia is confined in a smaller region, and the precipitation along the northeast coast is a little higher than observed. Any of these might be affected by the absence of the Tibetan Plateau, though GCMs that have full topography have similar problems (Sperber and Palmer 1996).

Experiments in the Asia-like continent case, overall, have many similarities to the North American case. Saturated soil moisture (Fig. 7b) has a similar impact on extending the convergence zone modestly northward and enhancing the precipitation over the east coast. When ventilation associated with \( \mathbf{v}_a \) is suppressed, the poleward extension of the convergence zone is substantial (Fig. 7c). The convergence zone has east–west asymmetry favoring convection over the eastern part of the continent and descent to the west. In the interior of the continent, reduced precipitation may be partially due to the albedo effect, which reduces \( F_{\text{net}} \) by reflecting solar radiation. The no-\( \beta \) experiment (Fig. 7d) shows significant poleward extension and no obvious east–west asymmetry. With both the ventilation associated with \( \mathbf{v}_a \) and the \( \beta \) effect suppressed, again poleward extension and little east–west asymmetry are seen, similar to the no-ventilation and partial-\( \beta \) experiment.

A difference between Figs. 7e and 7f, which is also found in the North American case, is that the convergence zone in Fig. 7e does not extend as far north as in Fig. 7f. The whole region becomes much warmer and higher OLR allows \( F_{\text{net}} \) to approximately balance. The convergence zone is terminated at high latitudes as it extends toward regions of lower moisture. For the no-ventilation and partial-\( \beta \) experiment, the target region might feel the effect of Rossby wave dynamics for scales at the domain size and larger, which allows descent to be communicated westward. However, in all these experiments, the small differences are less important than the result that including \( \beta \) effects on a domain-averaged “local Hadley cell” is clearly not a barrier to poleward extension of the convergence zone. Thus earlier conclusions regarding the impact of the IRH mechanism seem to hold.

A difference between the Asian case and the North American case comes when turning to the no-\( T \) and no \( q \) ventilation experiments seen in Figs. 7g and 7h, respectively. When \( \mathbf{v}_a \cdot \nabla T \) is suppressed in the no-\( T \) ventilation case (Fig. 7g), the \( \mathbf{v}_a \cdot \nabla q \) term is able to prevent the convection zone from extending poleward and the simulation remains similar to the control. Suppressed \( q \) ventilation alone has an effect comparable to the no-ventilation case, at least for a region in and somewhat poleward of the control convection zone in central Asia. Still, the convection zone in Fig. 7h does not extend as far poleward or into the continental interior as in the no-ventilation experiment (Fig. 7c). This may be contrasted to the North American case where keeping either \( \mathbf{v}_a \cdot \nabla T \) or \( \mathbf{v}_a \cdot \nabla q \) could prevent poleward extension of the convection zone. In the Asian case, \( \mathbf{v}_a \cdot \nabla q \) alone...
Fig. 7. As in Fig. 6 but over Asia.
can maintain a climate similar to the control, but $\nabla \cdot \mathbf{v}$, $\nabla T$ alone is less effective at preventing rainfall increase. Conjectures are the following: (i) $T$ ventilation is less efficient because the wind has traveled across greater distances of continental region in the Asian case; (ii) there is a considerable northerly flow component in the wind in parts of the region north of the Asian monsoon, and this may be particularly effective at bringing low $q$ air to disfavor convection over the continent.

Understanding the circulation involved in ventilation admittedly is more complicated in Asia than in other continents because of the wider continent (even without the Tibetan Plateau). The path of air masses responsible for ventilation and the relation to oceanic effects are less clear. The effect of the Tibetan Plateau obviously needs to be addressed to understand the actual impact of these mechanisms on the extent of the Asia monsoon. The effects on east Asia in these results are more realistic than those in the immediate vicinity of the plateau. However, the results here suggest that ventilation does tend to be a cooling and drying effect, akin to the North American and South American (CN) cases. Overall, high moist static energy air is exported from the continental region and by one path or another is replaced by lower moist static energy air.

6. African case

In experiments for the African case (Fig. 8), one encounters major differences from North America and the
Asia-like continent. Saturated soil moisture (Fig. 8b) does create a region of low but nonzero precipitation over much of the desert region. Saturated soil moisture is a more drastic assumption for Africa than for other continents because of climatological low soil moisture over northern Africa, so this result might be somewhat unrealistic. The sensitivity is higher than found in Douville et al. (2001). The implied evaporation (not shown) over the artificially saturated desert is very large, roughly similar to neighboring ocean regions. Ventilation is only able to remove about half this moisture input, the remainder of which precipitates. Further examination of the results in Fig. 8b shows that the region of intense rainfall in the equatorial continental convection zone extends farther northward compared to the control run (Fig. 8a). Increases in evaporation tend to increase boundary layer moisture and the resulting increase in the boundary layer moist static energy (see, e.g., Pal and Eltahir 2001) tends to favor convection (when it is sufficient to meet convective instability criterion). The increase in low-level moisture is partly due to local evaporation and partly to increased evaporation upward affecting moisture advection.

The no-ventilation experiment (Fig. 8c) has much less effect than for other continents. Convection is still very tropically confined over most of Africa. Suppressing ventilation does allow poleward extension up the east side of the region, which creates rainfall over Arabia. Rossby wave descent due to the Asian monsoon is not very effective at preventing this increase of precipitation over Arabia—which suggests caveats on the original Rodwell and Hoskins (1996) hypothesis for this region. Experiments with an enhanced Asian summer monsoon, such as the no-ventilation experiment in Fig. 7c, also show only small southward movement of the African summer monsoon. This suggests that the specific regional aspects of the Rodwell and Hoskins (1996) hypothesis may be the leading effects, even though the general mechanism is found to be important in other regions. The obvious limitation on poleward migration of convection in the African region experiments is the albedo, as already noted in the discussion of $F_{\text{net}}$ in section 3b, especially over the Sahara. Precipitation contours actually tend to follow albedo contours, as may be seen from the reference contour of albedo added to Fig. 8c. The high surface albedo over the Sahara prevents the convergence zone from extending northward. Suppressing ventilation does allow poleward extension up the east side of the region, which creates rainfall over Arabia.

The no-ventilation and no-$\beta$ (Fig. 8e), and no-ventilation and partial-$\beta$ (Fig. 8f) experiments encounter a similar effect to the no-ventilation experiment (Fig. 8c). The convergence zone cannot move very far poleward over northern Africa due to the albedo effect. These experiments do affect the regions to the immediate north of the control-run convection zone, so dynamical mechanisms have some effect in Sahel. But they are not nearly as important as on other continents.

To confirm the importance of the albedo effect in northern Africa, a set of constant surface albedo experiments is conducted. Albedo is set to a constant value of 0.26 over the African continent in the entire set of experiments including the control. The results are shown in Fig. 9. We note that the difference between the constant albedo control run (Fig. 9a) and the observed albedo control run (Fig. 8a) has a dipole pattern in precipitation (not shown), which is consistent with Xue and Shukla (1993).

Qualitatively, the results for the dynamical-factor experiments in the constant albedo case (Fig. 9) are similar to other continents. Suppressing the ventilation mechanism (Fig. 9c) permits the convection zone to expand over almost the whole of northern Africa including the Sahara. This is also the case for experiments with no ventilation and no $\beta$ or partial $\beta$ (Figs. 9e and 9f). Suppressing $\beta$ alone also permits a substantial, though lesser, expansion. Thus, one spatial variations in albedo are removed as a factor, the dynamical factors become the controlling mechanisms as on other continents. This may have implications for paleoclimate modeling of the African monsoon, in which albedo feedback is believed to be a major—but not sole—factor (e.g., Claussen et al. 1999; Joussaume et al. 1999). In a situation where albedo is reduced over the Sahara, such as the constant albedo control in Fig. 9a, the dynamical factors set the poleward limit of the convection zone. It is reasonable to hypothesize that they would enter similarly in an albedo feedback case.

Some differences between the dynamical factor experiments over Africa and over other continents may also be noted, in particular for the no-ventilation experiment (Fig. 9c). There is little east–west asymmetry caused by Rossby wave interaction with the heating within the African continent. The Rossby wave effect associated with the convection over southern Asia (i.e., the Rodwell–Hoskins hypothesis) may play a role in this. The Asian influence would tend to disfavor the eastern part of the region, while regional Rossby wave effects would disfavor the western side and the two may tend to balance out. Finally, we note that albedo does enter as a factor on other continents as well but the effects are modest compared to the dynamical mechanisms, as may be anticipated from $F_{\text{net}}$ (Fig. 4).

7. Discussion

Motivated by observational estimates of the net energy flux into the atmosphere over summer continents, a series of numerical experiments are conducted for Northern Hemisphere summer monsoons for each of the North American, Asian, and African regions. The experiments are aimed at testing mechanisms that potentially limit the poleward extent of the summer season convergence zones. The mechanisms, outlined in CNS
and CN, are soil moisture feedback, the ventilation mechanism, and the IRH Rossby wave mechanism. From a methodological point of view, we have verified conclusions of CNS and CN for a variant of the no-ventilation experiments that maintains conservation and more clearly separates purely divergent and nondivergent components of wind. The experiments designed to test Rossby wave effects do so by suppressing $\beta$, that is, by setting the Coriolis parameter to a constant value in a target region. However, one might worry that the original experiment could also impact effects of a “local Hadley circulation.” By including the $\beta$ effect on the zonal mean across the target region, we can examine effects of a local Hadley cell and Rossby wave dynamics separately. Qualitatively similar results are found for the “no-ventilation and no-$\beta$” experiment and the “no-ventilation and partial-$\beta$” experiment: the convection zone extends much farther poleward and less east–west asymmetry is found in either case. Thus it supports the previous focus on Rossby wave dynamics as responsible for the east–west asymmetry of the precipitation within the continent.

The net flux of energy into the atmosphere, $F_{\text{net}}$, is similar for South America (CN), North America, and Asia for their respective summer seasons: positive $F_{\text{net}}$ extends far poleward in the continents in observations and simulation. For Africa, positive $F_{\text{net}}$ extends less far poleward because of the strong surface albedo. In the previous studies for an idealized continent and for the South American monsoon, the thermodynamical factor associated with positive $F_{\text{net}}$ is a necessary condition for establishment of a convergence zone over conti-
ments, but the dynamical factors determine how far poleward the convergence zones can extend. Overall, the importance of dynamical factors in setting poleward extent of the convergence zone seems to apply in North America and Asia as in South America.

Although it is difficult to assign relative importance to ventilation versus the IRH mechanism, experiments here suggest that ventilation is very important for each case except for Africa with its strong surface albedo. The Rossby wave effect is important in favoring east coast rainfall and west coast dryness in North America. Likewise in Asia, Rossby wave dynamics favor increased rainfall toward the east and dryness in the interior. Because of the drastic changes in each experiment and the nonlinearity of the response, care must be taken when interpreting the results. For instance, in North America suppressing either ventilation associated with $v \cdot \nabla T$ alone or ventilation associated with $v \cdot \nabla q$ alone has little effect on extending northward of the convergence zone. Either advection term alone can prevent the poleward extension of the convergence zone. Over Asia, ventilation associated with moisture advection appears to be more important than the cooling associated with $v \cdot \nabla T$.

Africa is clearly different than other continents because of the high surface albedo found in large regions. This reduces the positive value of $F_{\text{net}}$ over much of northern Africa. Thus, even very modest values of moist static energy transport by the nondivergent component of the flow can balance this energy input and prevent convergence. Even when the dynamical factors, such as ventilation and Rossby wave dynamics, are suppressed, little northward extension of the convergence zone is found. Over the high albedo regions, even a small warming of the land–atmosphere column can cause $F_{\text{net}}$ to become negative and thus limit poleward extension of the convergence zones. Contrasting with the other continents, a first requirement for the summertime poleward movement of the convection zones is that the simple thermodynamical factor associated with $F_{\text{net}}$ should favor convection. In Africa this requirement ceases south of the Sahara. If this first requirement is met, which is the case even far poleward in Asia and North America, then the dynamical factors determine the poleward extent of the convergence zone.

These experiments suggest a modification to the view of monsoon systems. Moisture availability is typically not the leading order effect in limiting monsoon extent, if other factors favor establishment of a low-level convergent circulation. The irrational (purely divergent) part of the flow is very effective at increasing moisture sufficiently to meet convective criteria, if the moist static energy budget terms are such as to sustain the convergence. Except for northern Africa, the extent of monsoon systems is thus set by the transition from a tropical dynamical regime, in which positive net energy input into the column is balanced by rising motion and low-level convergence, to a midlatitude dynamical regime, in which advection terms $v \cdot \nabla T$ and $v \cdot \nabla q$ are of leading importance in energy transport (where $v$ is the nondivergent component of the wind). Aspects of this transition have been previously noted in analysis datasets by Barlow et al. (1998) who discuss the transition from a balance of diabatic heating with adiabatic cooling in the monsoon region, and with $v \cdot \nabla T$ in the region poleward of the North American monsoon. The results here also make clear that $v \cdot \nabla q$ tends to act to oppose poleward extension of the convection zone. This is because the nondivergent component of the flow cannot increase moisture above values that occur in the upstream region, and values over cooler oceans tend to be smaller than the values that would be required to initiate convection over the warm continent.

A summary of this view of summer monsoon convection zones is as follows. Unless the albedo is very high, there is a positive net energy flux into the atmospheric column extending far poleward in the continent. This favors low-level converging motions that supply sufficient moisture for a convergence zone up until such latitude as ventilation by the nondivergent component of the flow is able to balance the energy input. The divergent circulation is subject to baroclinic Rossby wave dynamics, so the spatial pattern of the convergence zone and descent region are linked. This favors greater poleward extension of the convergence zone on the eastern side of the continent and arid regions west of that. The geometry of a particular continent can have a considerable quantitative impact on the interplay of these mechanisms, but for the large scale this scenario seems to hold for each of the North American, Asian, and South American summer convection zones.

Acknowledgments. This work was supported under National Science Foundation Grant ATM-0082529, National Oceanic and Atmospheric Administration Grants NA86GP2003 and NA86GP2004, National Aeronautics and Space Administration Grant NAGS-9358, and National Science Council Grant NSC89-2811-M-034-0001.

REFERENCES


Arking, A., 1999: Bringing climate models into agreement with observations of atmospheric absorption. J. Climate, 12, 1589–1600.


—, and —, 1993: The Betts–Miller scheme. The Representation


1 FEBRUARY 2003

CHOU AND NEELIN
