The North American Monsoon

David K. Adams and Andrew C. Comrie
Department of Geography and Regional Development, The University of Arizona, Tucson, Arizona

ABSTRACT

The North American monsoon is an important feature of the atmospheric circulation over the continent, with a research literature that dates back almost 100 years. The authors review the wide range of past and current research dealing with the meteorological and climatological aspects of the North American monsoon, highlighting historical development and major research themes. The domain of the North American monsoon is large, extending over much of the western United States from its region of greatest influence in northwestern Mexico. Regarding the debate over moisture source regions and water vapor advection into southwestern North America, there is general agreement that the bulk of monsoon moisture is advected at low levels from the eastern tropical Pacific Ocean and the Gulf of California, while the Gulf of Mexico may contribute some upper-level moisture (although mixing occurs over the Sierra Madre Occidental). Surges of low-level moisture from the Gulf of California are a significant part of intraseasonal monsoon variability, and they are associated with the configuration of upper-level midlatitude troughs and tropical easterly waves at the synoptic scale, as well as the presence of low-level jets, a thermal low, and associated dynamics (including the important effects of local topography) at the mesoscale. Seasonally, the gulf surges and the latitudinal position of the midtropospheric subtropical ridge over southwestern North America appear to be responsible for much spatial and temporal variability in precipitation. Interannual variability of the North American monsoon system is high, but it is not strongly linked to El Niño or other common sources of interannual circulation variability. Recent mesoscale field measurements gathered during the South-West Area Monsoon Project have highlighted the complex nature of the monsoon-related severe storm environment and associated difficulties in modeling and forecasting.

1. Introduction

a. Aims and rationale

The North American monsoon, variously known as the Southwest United States monsoon, the Mexican monsoon, or the Arizona monsoon, is experienced as a pronounced increase in rainfall from an extremely dry June to a rainy July over large areas of the southwestern United States and northwestern Mexico. These summer rains typically last until mid-September when a drier regime is reestablished over the region. The North American monsoon (NA monsoon) has been studied since the early part of this century, and a fairly extensive research literature has developed that deals with a range of meteorological and climatological aspects of the phenomenon. Surprisingly, no comprehensive review of the NA monsoon literature has been presented. Several papers have briefly summarized the basic climatology of the NA monsoon (e.g., Maddox et al. 1995), but to date, no single contribution has sought to integrate the wide range of research relating to this notable feature of the atmospheric circulation over North America.

In this paper, we provide a synthesis of major climatological and meteorological research themes regarding the NA monsoon: we outline the past debate and emerging consensus on mechanisms and moisture sources; we assess the state of knowledge pertaining to the extent and variability of the NA monsoon, both spatially and temporally; and we include an overview of research on the severe storm environment and modeling efforts associated with the NA monsoon. We also address two additional issues. First, considering the
importance of the NA monsoon within the atmospheric circulation over North America, and despite several recent research papers, misunderstandings of NA monsoon dynamics and influence are still surprisingly widespread within the broader meteorological and climatological community. Second, this is an area of research that has become quite active, particularly during the last few years with the South-West Area Monsoon Project (SWAMP), and it offers many theoretical and applied avenues for further investigation. We intend this paper to be a timely contribution that will critically summarize the state of knowledge and stimulate further inquiry on the NA monsoon.

Pulling together the diverse array of NA monsoon studies into a cohesive body of literature is a challenging task. We have chosen to divide the literature into several categories, which by nature are somewhat arbitrary and partially overlapping. Many of the articles reviewed cover a number of these categories, and thus a particular study may appear in each of several appropriate sections. We review papers found primarily in the major meteorological and climatological journals, including those that focus principally on the NA monsoon as well as those less directly connected but that nonetheless provide useful insights. Also included are a small number of books and, where necessary, less easily available scientific reports. Some Spanish-language references have been included for completeness, but the reader is referred to Reyes et al. (1994) for a more complete Spanish-language reference list.

b. Climatological setting

We use the term “North American monsoon” in order to be geographically all-encompassing. It is important to note that most of the NA monsoon literature has a strong geographical bias toward the southwestern United States, especially the state of Arizona, necessitating frequent mention of these areas in this review (and in our suggestions for future research directions). We wish to emphasize that this phenomenon influences larger areas of the southwestern United States and that, in fact, it is centered over northwestern Mexico. A discussion of areal extent is presented later in the paper, but we introduce the broader monsoon region here to provide background on important physiographic characteristics.

Figure 1 highlights the major physical and topographic features of southwestern North America that relate to the NA monsoon. The region is bounded to the west by the Pacific Ocean, including the Gulf of California. Summertime sea surface temperatures (SSTs) at middle and subtropical latitudes along the Pacific Coast are mainly cool on average (< 25°C), but SSTs in the Gulf of California and tropical eastern Pacific (farther south) are warm (> 28°C). To the east, the region is bounded by the Gulf of Mexico, with warm SSTs (> 26°C), and by the central plains of the United States. The interior of the region is characterized by several large upland areas: the Colorado Plateau extends north and east to the Rocky Mountains from Arizona’s Mogollon Rim; the north–south aligned mountain ranges of the basin and range province dominate through Nevada, southwestern Arizona, and northwestern Sonora; the Mexican Altiplano (plateau) is defined by the Sierra Madre Occidental and Oriental to the west and east, respectively. In addition, the peninsular ranges of southern California and Baja California along with the Sierra Nevada play important climatological roles for the interior deserts, limiting penetration of marine moisture from the Pacific Ocean. There are also two lowland areas associated with the NA monsoon: the lower Colorado River valley and neighboring low desert areas, which play a critical role in the formation of the thermal low; and the coastal lowlands of Sonora and Sinaloa, between the Gulf of California and the western flanks of the Sierra Madre Occidental. Most of the lower elevation zones of the monsoon region, particularly the Sonoran Desert, receive up to 80% or 90% of possible summer
insolation, and summer surface temperatures can be extremely hot, often exceeding 40°C.

The combination of seasonally warm land surfaces in lowlands and elevated areas together with atmospheric moisture supplied by nearby maritime sources is conducive to the formation of a monsoonlike system. With the exception of Ramage (1971), most authors regard this circulation as a “true” monsoon, based on seasonal reversal of pressure and wind patterns, energy and mass transfers, and characteristic regimes of rainfall and temperature (e.g., Bryson and Lowry 1955a; Krishnamurti 1971; Tang and Reiter 1984; Douglas et al. 1993). Two points are worth noting with respect to our coverage of the NA monsoon: first, seasonal wind direction changes over the North American Arctic might technically be considered a monsoon, but we exclude this feature; second, the term “North American monsoon” has been applied to a more extensive phenomenon that we do not explicitly cover, which includes the wet season over Central America and even the seasonal displacement of the intertropical convergence zone (ITCZ) over the eastern Pacific Ocean, for example, in the Pan-American Climate Studies program (PACS 1996).

2. Historical perspective

Historically, research interest in the NA monsoon has its roots in the peculiar distribution of summer precipitation in this arid and semiarid region of North America. Figure 2 illustrates monthly rainfall distributions for various sites across the region, many of which display the characteristic NA monsoon mid-to late summer rainfall maximum following extremely dry conditions in May and June. Early researchers noted that July and August thunderstorms were abundant in the uplands of Arizona and western New Mexico, and relatively infrequent to the west in southern California (Campbell 1906; Beals 1922; Blake 1923). However, the origin, frequency, and distribution of this warm season convective activity were not seen as necessarily linked to some larger-scale regional control. In fact, it was this curious distribution of warm season precipitation (nearly absent in southern California and quite frequent in elevated zones of central Arizona) that spurred much of the original research that would attempt to clarify the general character of the summer circulation of the desert Southwest (Reed 1933, 1937, 1939). For example, the infrequent summer thunderstorms of the California deserts, sometimes referred to as “Sonora storms” (Campbell 1906), were first thought to have their origin in the Mexican state of Sonora, from where they would travel northward to southern California. The proximity of Sonora to Arizona made this northward movement plausible for convective storms in Arizona, but the distance to California and the localized nature of the storms once in California seemed to negate this possibility (Blake 1923). Another explanation suggested that the intensification and movement of the semipermanent thermal low pressure over the Colorado River valley were responsible for widespread rain in Arizona and California (Beals 1922). Others, too, believed that a relationship existed between the intense thermal low and summer thunderstorms in the region (Ward 1917; Blake 1922; Willett 1940). It was suspected that the resulting pressure gradient would draw Gulf of California moisture northward, enhancing instability and increasing convective activity in the mountains and deserts of the southwestern United States. Remarkably, the notion of the Gulf of California as a moisture source for NA monsoon precipitation would not appear prominently in the literature again until several decades later.

Fig. 2. Seasonal distribution of precipitation across southwestern North America. Note that northwestern Mexico shows the strongest monsoon signal, which diminishes through Arizona, New Mexico, and Nevada. Northeastern Mexico and Texas display early summer–late fall precipitation peaks, while the West Coast has a dry summer Mediterranean distribution (vertical axis of all graphs represents 180 mm with 20 mm increments). Areas south of the broken line receive greater than 50% of their annual rainfall in July, August, and September (after Douglas et al. 1993).
Following the advent of a network of upper-air soundings over the western United States, Reed (1933) recognized that the distribution of summer rains in this region was, in fact, part of a larger-scale phenomenon. The position of the “high-level anticyclone” east or west of the Continental Divide, he proposed, dictated the occurrence of rainfall events by stabilizing (westward position) or destabilizing (easterly position) the lapse rate. Moreover, he believed advection of low-level moisture along the anticyclonic trajectory was not necessary for destabilization because sufficient water vapor was presumably present throughout the region. Widespread convective activity was dependent, supposedly, on upper-level instability alone.

The first comprehensive climatology of the NA monsoon phenomenon was presented by Bryson and Lowry (1955a, b). Although these authors were not the first to recognize the monsoonal nature of the circulation regime (e.g., Ives 1949), the prominence of their work ensured that the NA monsoon would be recognized primarily as an Arizona summer feature, hence the later predominance of Arizona in the literature. Bryson and Lowry (1955a, b) provided the dynamic explanation for the sudden appearance of the summer thunderstorms that was lacking in Reed (1933). The rapid onset of rainfall in the Arizona summer monsoon was defined as a “singularity” (the recurrence of an atmospheric phenomenon near a particular calendar date), and it was found to be related to a sudden shift and readjustment in the midtroposphere hemispheric circulation. This shift in circulation, resulting from the westward and northward expansion of the Bermuda high sometime around the end of June, causes dry southwesterly midlevel winds from the Pacific Ocean to be replaced by moist southeasterly winds (Fig. 3). This sequence of synoptic patterns places Arizona under a light southeasterly flow, which seemingly implied water vapor advection solely from the Gulf of Mexico. Consequently, the Gulf of Mexico was viewed as the unique moisture source for the NA monsoon—a view that continued to appear in the literature for many years (Reitan 1957; Jurwitz 1953; Green 1963; Green and Sellers 1964; Hastings and Turner 1965) and that became a central theoretical issue in much subsequent research.

3. Moisture source debate

Reitan (1957) and, later, Rasmusson (1967) raised several important points that seemed to contradict Bryson and Lowry’s (1955a, b) conclusions. Although not directly investigating moisture sources, Reitan (1957) concluded that the greatest amount of precipitable water vapor during the Arizona monsoon is found below 800 mb. Considering the fact that the topography between Arizona and the Gulf of Mexico rises above this level, the Gulf of Mexico appeared to be an unlikely source for such low-level moisture. Similarly, Rasmusson (1967), in examining annual vapor

Fig. 3. Pentads of 500-mb geopotential height (dam) from 1979–88 showing the transition from southwest flow over the NA monsoon region in early June to east or southeast flow in July. Note rapid northward shift in height contours during last week of June and initiation of light easterly flow over northern Mexico (after Bryson and Lowry 1955b).

1 Several citations of this source have been misleading. Reitan (1957) did not suggest the Gulf of California as an important source of low-level moisture during the summer in Arizona. He, in fact, reiterates the findings of Bryson and Lowry (1955a) that the moisture comes from the Gulf of Mexico. After examining his conclusions, however, other authors raised the question of how Gulf of Mexico moisture could be advected over the Continental Divide and still provide low-level moisture in Arizona.
flux in North America for a 2-yr period, discovered a strong northward vapor flux into the southwestern United States from the Gulf of California and a much weaker inflow from east of the Rocky Mountains. Taking these observations into account, the Gulf of California and eastern tropical Pacific arose again as possible contributors to increases in moisture and concomitant rains during the NA monsoon.

a. The gulf surge phenomenon

Hales (1972) and Brenner (1974) argued that the eastern tropical Pacific off the Mexican coast, and not the Gulf of Mexico, provides the greatest amount of water vapor for the NA monsoon west of the Continental Divide. To substantiate this hypothesis, they offered a mechanism by which atmospheric moisture is transported northward through the Gulf of California, occasionally as far inland as the Great Basin Desert. The transport mechanism was described as a “surge” of moist tropical Pacific air that is channeled up the Gulf of California as a result of a low-level atmospheric pressure gradient. This pressure gradient is created as the thermal equilibrium between the tropical Pacific and the Gulf of California becomes disrupted, primarily due to the development of a cloudy, showery air mass at the mouth of the Gulf. The thermal and, hence, pressure contrasts between the cooler, showery air mass to the south and the intensely heated surface in the deserts of Arizona and eastern California drive the moisture surge. A gulf surge is often manifest as a strong shift to south–southeast winds (particularly at Yuma, Arizona) accompanied by lower temperatures and higher dewpoints, which are detectable at both upper-air stations (Empalme and Mazatlán) adjacent to the Gulf of California. For the most part, the gulf surge is a lower tropospheric feature, with all but the strongest surges typically confined below the 700-mb level. The northward influx of low-level maritime tropical air into the Sonoran Desert and the uplands of Sonora and Arizona is believed to augment convective activity. Many other investigations have subsequently supported this moisture surge hypothesis (Hales 1974; Houghton 1979; Jáuregui and Cruz 1981; Tang and Reiter 1984; Carleton 1985, 1986; Hasimoto and Reyes 1988; Reyes and Cadet 1988; Carleton et al. 1990; Badan-Dangon et al. 1991; Rowson and Colucci 1992; Douglas et al. 1993; Douglas 1995; Maddox et al. 1995; McCollum et al. 1995; Stensrud et al. 1997).

Due to their apparent role in the convective environment of northwestern Mexico and the desert Southwest, gulf surges and more generally the lower atmosphere over the Gulf of California have been a focus of recent research (Douglas 1995; Douglas and Li 1996; Stensrud et al. 1997). The actual mechanisms for surge initiation and their relationship to large-scale circulation patterns have been unclear (Hales 1972; Brenner 1974). Recent evidence from Stensrud et al. (1997) again points to the role of westward-moving tropical disturbances south of the mouth of the Gulf of California (approximately 20°N) in the development of major gulf surges. However, their findings from mesoscale model simulations for the gulf region indicate that large-scale features can indeed be linked to major surges. They contend that major surges occur when a midlatitude trough passes through western North America (around 40°N) preceding, by 1 or 2 days, the passage of an easterly wave across the same longitudinal belt (Fig. 4). The increased subsidence

**Fig. 4.** Conceptualization of the gulf surge phenomenon. Major surges are associated with easterly waves crossing south of the Gulf of California, out of phase with an upper-level midlatitude trough. Minor surges are associated with mesoscale convective systems drifting over the northern half of the Gulf (see text for details).
behind the trough over the northern end of the gulf apparently creates conditions in surface pressure and low-level temperature fields conducive to surges. The actual initiation of major surges may result from cold outflows caused by deep convection over the Gulf of California, or it may simply result from the low-level convergence associated with the easterly wave.

Minor gulf surges, those typically confined to the northern half of the Gulf (Hales 1972), can be triggered by the passage of a mesoscale convective complex passing from the mainland over the Gulf, by backdoor cold fronts approaching from the north to northeast as opposed to the typical west to northwest direction (R. Maddox 1996, personal communication; Hales 1972) or by the passage of an easterly wave without the preceding upper-level trough associated with major surges (Stensrud et al. 1997), as conceptualized in Fig. 4. The weather effects for Sonora and Arizona of these minor surges are similar to major surges; however, the former go undetected in the Empalme and Mazatlán soundings, making forecasting these surges more difficult (R. Maddox 1996, personal communication).

The gulf surge does not, in itself, necessarily account for the total Gulf of California–eastern Pacific moisture transport into the desert Southwest. Field studies (Badan-Dangon et al. 1991; Douglas 1995) describe mean southerly flow at the northern end of the Gulf of California during July and August, with a low-level jet as a notable characteristic (see Fig. 5 and related discussion in section 4a, also). Mesoscale observations from SWAMP (35 days of pilot balloon soundings and several research aircraft measurements) document the nature of the low-level jet. It is present under various synoptic configurations, exhibiting an early morning maximum over the northern Gulf of California, with the jet appearing to be strongest under surge conditions (Douglas 1995). The persistence of this low-level flow suggests significant northerly transport of water vapor from the Gulf of California. For the Sonoran Desert region, a recent analysis (Schmitz and Mullen 1996) of atmospheric water vapor flux (calculated for the time-mean wind during July and August) for the Gulf of California shows that time-mean moisture transport is dominant over that of transient flux (deviations from the mean), although the role of transient flux may be more critical for actual precipitation episodes over the region.

b. A dual moisture source

Despite the fact that the more recent literature has tended to view the tropical eastern Pacific–Gulf of California as the dominant moisture source for the NA monsoon, the Gulf of Mexico is generally still viewed as an important contributor. Most investigations of the NA monsoon subscribe to the idea of moisture advection in some combination from both the eastern tropical Pacific–Gulf of California and the Gulf of Mexico (Hales 1974; Brenner 1974; Mitchell 1976; Carleton 1985, 1986; Carleton et al. 1990; Harrington et al. 1992; Watson et al. 1994a; Stensrud et al. 1995; Schmitz and Mullen 1996). Much recent work states that low-level moisture is largely attributable to the Gulf of California and that upper-level water vapor is transported from the Gulf of Mexico. The debate is more than academic because source regions have implications for seasonal forecasts (Schmitz and Mullen 1996) and for the development of the convective environment (Stensrud et al. 1995).

Carleton (1986) identifies contributions of both sources by examining a set of thermodynamic parameters for several Mexican and United States upper-air stations. Although the importance of the sources relative to one another is not evaluated, he argues that rapid

---

Fig. 5. Schematic of selected dynamic features during the NA monsoon. The typical 850-mb late-afternoon height configuration (solid dark lines) includes heat lows over the elevated interior, which act to draw in moisture by way of low-level jets (LLJ) originating over the Gulf of California and the Gulf of Mexico (large arrow). The early morning Gulf of California LLJ core is shown via shaded isotachs (3, 5, and 7 m s$^{-1}$) and streamlines at 450 m (1500 ft) above ground level (adapted in part from Douglas 1995).
lower-level moisture surges moving up the Gulf of California and slower high-level moisture from the Gulf of Mexico work in concert during monsoon “bursts” (his term “burst” has the same connotation as that for the Southwest Asian monsoon—i.e., large surges of moisture and concomitant precipitation). Carleton et al. (1990) reaffirm this conclusion in their examination of interannual variability of Arizona rainfall during the NA monsoon. This recognition of two distinct moisture sources for the NA monsoon raises important questions about the relative contributions of the Gulf of Mexico and eastern tropical Pacific–Gulf of California during monsoon bursts, as well as the possible spatial and temporal variations in the relative contribution from each source at different heights in the atmosphere. Maddox et al. (1995) claim that this reduced role of the Gulf of Mexico as moisture source is still too great, but the results of Schmitz and Mullen (1996), focusing on the Sonoran Desert, suggest that the time-mean upper-level flow (700 mb to 200 mb) has its origin over the Gulf of Mexico, while low-level flow is attributable to the Gulf of California. Clearly, this question has not been completely resolved.

One further complicating factor in the moisture source debate is the role of the Sierra Madre Occidental in moistening the midlevels of the atmosphere. Recent empirical and modeling studies (Maddox et al. 1995; Stensrud et al. 1995; Schmitz and Mullen 1996) highlight the importance of low-level Gulf of California and eastern Pacific moisture being drawn inland over the Sonoran and Sinaloan lowlands and up the western flanks of the Sierra Madre Occidental due to topographic heating. The resulting convection redistributes moisture to higher levels. This moistened air may then mix with elevated Gulf of Mexico moisture from the east and can subsequently be transported into the southwestern United States via flow around the anticyclone. As a result, the simple Gulf of California–Gulf of Mexico moisture source dichotomy becomes much more difficult to unravel.

4. Geographic extent

a. Spatial domain

It is quite apparent that Arizona has been the principal geographic study area for NA monsoon research, particularly in relation to summer precipitation (Bryson and Lowry 1955a,b; Reitan 1957; Bryson 1957a,b; Green 1963; Carleton 1985, 1986, 1987; Moore et al. 1989; Smith and Gall 1989; Adang and Gall 1989; Carleton et al. 1990; Rowson and Colucci 1992; Dunn and Horel 1994a,b; McCollum et al. 1995; Maddox et al. 1995). This may be a result of the fact that the monsoon phenomenon was first recognized in the literature as the “Arizona summer monsoon” (Bryson and Lowry 1955a). Furthermore, it has been stated that, in the United States, the NA monsoon is most visible in Arizona (Trewartha 1981; Adang and Gall 1989; Carleton et al. 1990). Nevertheless, several investigators have expanded or shifted the geographic focus of NA monsoon influence. Reyes and Cadet (1988) and Douglas et al. (1993) have shown that the “Arizona monsoon” is merely an extension of a more pronounced Mexican phenomenon centered over the western foothills of the Sierra Madre Occidental (Fig. 6). New Mexico has also been recognized as greatly influenced by the NA monsoon (Tuan et al. 1973; Mitchell 1976; Trewartha 1981; Harrington et al. 1992; Douglas et al. 1993; Bowen 1996), and one index of monsoon influence has shown that New Mexico is the state most affected by the NA monsoon (Douglas et al. 1993). For example, note the similarities in precipitation regime between Tucson and Alamogordo in Fig. 2. The NA monsoon moisture and attendant thunderstorms even extend into the lower Colorado River valley and the Mojave Desert, and farther west to the transverse and peninsular ranges of southern California, with some regularity (Tubbs 1972).

On a much broader scale, Tang and Reiter (1984) and Reiter and Tang (1984) demonstrate that the western United States and northern Mexican plateau act to create a true monsoonal circulation pattern over the North American continent, comparable to that of the Tibetan Plateau. This expanded region of summer monsoon influence, which they designate as the “western plateau,” ranges longitudinally from the Sierra Nevada in the west to the Colorado and Wyoming Rockies in the east. The northernmost limit runs from Oregon along the Idaho–Utah border into Wyoming. This limit, according to Tang and Reiter (1984), coincides with an anticyclonic shear line that marks the northernmost extension of summer convective activity. The Mexican Plateau along the international border serves as their southern boundary. Tang and Reiter (1984) postulate a low-level jet containing tropical moisture drawn up the Gulf of California into the Great
Basin desert of Nevada and Utah as a result of monsoonal pressure gradients. This corresponds closely to ideas put forth by Brenner (1974), Hales (1974), and Houghton (1979). Likewise, moisture is transported east of the Continental Divide by a low-level jet from the Gulf of Mexico. The moisture is moved up the Continental Divide as a result of interior plateau heat-lows, thereby increasing convective activity (Fig. 5). These ideas are in agreement with the postulated dual moisture source for the NA monsoon, and they indicate that the Continental Divide may be important in determining low-level moisture trajectories.

b. Distribution of air masses

Several of the aforementioned studies have also described the NA monsoon as a contrast between dry and moist air masses in the southwestern United States and northwestern Mexico. The moist air mass, characterized as a moist tongue moving north from Mexico into the southwestern United States, is bounded to the north and west by drier air from the Pacific subtropical high (Bryson and Lowry 1955a,b). The zone of interaction between the air masses has been viewed as a definite, but shifting, boundary that exists between the moist air mass brought in by the westward extension of the Bermuda high and the drier eastern-Pacific air (Bryson and Lowry 1955a,b; Mitchell 1976; Moore et al. 1989; Adang and Gall 1989; Watson et al. 1994a). For example, while the transition from dry to moist air masses at the onset of the monsoon season is not always rapid, it has been shown to occur in as short a period as a few hours (Nastrom and Eaton 1993). In addition, gradients in temperature, water vapor content, and winds similar to those of a quasi-stationary, midlatitude front between the moist tongue and dry air have also been observed (Moore et al. 1989; Adang and Gall 1989). The boundary of this moist tongue has been described as shifting spatially in agreement with the westward-expanding Bermuda high (Bryson 1957a; Trewartha 1981; Tang and Reiter 1984; Harrington et al. 1992), from eastern New Mexico in June to western New Mexico and central and eastern Arizona by July. Eastern California and western Arizona are under its influence by August. Climatologically, this boundary

**Fig. 6.** Frequency (in percent of total hours) of cloud-top temperatures less than $-38^\circ C$ for June, July, and August of 1985–92, derived from averaged hourly infrared GOES imagery (adapted from Douglas et al. 1993). Shaded regions delineate areas of systematic deep convection.
has been defined in terms of contrasting equivalent potential temperature (Mitchell 1976). These studies of the NA monsoon boundary do not, however, examine the specific contributions provided by the Gulf of Mexico or the Gulf of California for the development of the moist tongue.

5. Spatial and temporal variability

The NA monsoon region, like other arid and semiarid locales, experiences great climatic variability on a range of spatial and temporal scales. Difficulties in understanding the variability of summertime convective activity in the southwestern United States and northwestern Mexico result from the complex interactions between synoptic and mesoscale circulation features and the extremely varied topography. Although larger-scale atmospheric motions control the broad distribution of water vapor and the general stability or instability in the atmosphere, local topographic effects in the western United States and northwestern Mexico are critical to the geographic and even temporal distribution of convective activity. In fact, dealing with these factors has been at the heart of many modeling and forecasting efforts in the region, as we discuss in section 6b.

a. Regional-scale spatial variability

Spatial variability of the NA monsoon has been examined almost exclusively in terms of precipitation. The geographic influence of NA monsoon rains is primarily centered in the western foothills of the Sierra Madre Occidental in northwestern Mexico, specifically in the Mexican states of Nayarit, Sinaloa, and Sonora (Hales 1972; Douglas et al. 1993). This region receives up to 70% of its annual rainfall in the months of July, August, and September (Mosiño and Garcia 1974; Douglas et al. 1993; Reyes et al. 1994). The intensity of convection over the Sierra Madre is apparent from satellite imagery (Douglas et al. 1993; Negri et al. 1993; Stensrud et al. 1995), as illustrated in Fig. 6. The distinctness of the July through September maximum decreases rapidly as one moves northward through Sonora toward the international border. The NA monsoon precipitation regime also becomes less notable toward the eastern flank of the Sierra Madre Occidental and onto the Mexican Altiplano. Much of northeastern Mexico actually experiences a June and September rainfall maximum with a minimum in July and August. Similarly, NA monsoon rains diminish rapidly to the west, across the Gulf of California along the Baja Peninsula and northward into California. The greater influence of subsidence under the Pacific subtropical high in these latter areas has been suggested as a reason for the steep decrease in rainfall (Hastings and Turner 1965). Yet the drop-off takes place over a relatively short distance (as compared to the scale of subsidence); a concise, supported explanation for this westward decrease is not found in the literature. North of about 26°N on the Pacific side of the peninsula, the cool waters of the California current and strong subsidence stifle summertime convection, and the winter precipitation regime of the West Coast dominates (Hastings and Turner 1965). These patterns of precipitation seasonality are illustrated in Fig. 2.

Within the United States, southern and eastern Arizona, much of New Mexico, and south-central Colorado receive precipitation related to the NA monsoon, but rainfall amounts are much more variable than in northwestern Mexico. For example, the Mogollon Rim of central Arizona and the White Mountains of eastern Arizona receive a much greater percentage of rainy days during the monsoon than the lower desert to the southwest, despite generally higher dewpoints in this latter region (although relative humidities are lower). This is not unexpected considering that convection is also greater over the elevated regions of Arizona. Across the NA monsoon area, regional-scale spatial variability in precipitation is closely linked to the penetration of moisture into the interior regions (e.g., Hales 1972; Brenner 1974), but synoptic-scale controls are believed to be equally important for the regional variability of NA monsoon precipitation in Arizona (Carleton 1986). Although not directly part of the NA monsoon, decaying tropical cyclones also contribute to the regional variation of total summer precipitation (Reyes 1988; Hereford and Webb 1992), particularly over Sinaloa, Sonora, and the southern tip of Baja California.

b. Local-scale variability

Much of the early NA monsoon research focused on the spatial and temporal variability of Arizona summer rainfall (McDonald 1956; Bryson 1957a; Green 1960). Considering the nature of air mass thunderstorms, it is not surprising that there is a great spatial variability in precipitation and a variable geographic distribution of summer thunderstorms. In Arizona, only small-to-moderate interstation correlations exist between rainfall events (McDonald 1956). Yet, diurnal variability in summer thunderstorms does show a
marked geographic dependence, particularly in Arizona (McDonald 1956; Ackerman 1959; Sellers 1958; Hales 1972, 1977; Brenner 1974; Reiter and Tang 1984; Balling and Brazel 1987; Maddox et al. 1991; Watson et al. 1994b; Maddox et al. 1995). Diurnal convective activity is generally at a maximum over the Colorado Plateau and the higher elevations of the Continental Divide during early afternoon. Over the highlands of southern Arizona and northern Sonora, maximum activity is characteristic of the early evening hours. By late evening and early morning, the coastal lowlands of northwest Mexico and the lower desert regions of southern and central Arizona experience a nighttime convective maximum (in contrast to the timing of convective storms elsewhere), as illustrated in Fig. 7. Similar diurnal precipitation variation due to topographic effects has also been noted in central New Mexico (Bowen 1996). The specific causes of the nocturnal maximum have been explored but are still not clearly understood. Hales (1977) states that midtropospheric cooling results from localized evaporation of cumulonimbus towers that develop over the higher terrain of central Arizona in the afternoon. This elevated cool layer is supposedly advected west and southwest over the low desert toward evening, thereby destabilizing the local environment (although no further support of this idea has been found). Ascending air in the late evening and early morning over the Sonoran Desert low may also support the nocturnal convection (Douglas and Li 1996), as a result of the larger lapse rate. With regard to the development of severe nocturnal mesoscale convective systems (MCS) over southern and central Arizona, McCollum et al. (1995) state that subsynoptic scale features are also of great import, particularly low-level moisture advection from the Gulf of California and localized convergence in favored areas.

c. Seasonal variability

Apart from the spatial variability that is directly related to topographic features, there is much intraseasonal variability of NA monsoon rainfall (in terms of intensity and areal extent), particularly in the southwestern United States. The highlands of eastern Sinaloa and Sonora in northwestern Mexico experience shower activity almost every day during the entire NA monsoon season. Nevertheless, there are apparent periodicities in rainfall episodes in different areas of northwestern Mexico, which may be attributable to variations in source region influence (Reyes et al. 1994). Intraseasonal variation of precipitation totals in this region may be related to the frequency of MCS and tropical perturbations that account for a great portion of the yearly precipitation (Reyes et al. 1994). Farther north, much of the intraseasonal precipitation variability may be related to the gulf surge phenomenon described by Hales (1972) and Brenner (1974). These surges are believed to contribute significantly to the widespread summer rainfall events over much of the desert Southwest. Considering that several gulf surges can occur in this area during a given NA monsoon season, they may account for a great deal of the variability in total summer precipitation (Brenner 1974). It is important to note here that gulf surges, weak or strong, have been assumed to result in increased convective activity and precipitation, but to date no study has been published in the formal literature quantifying this relationship.

A number of authors have examined changes in the synoptic-scale circulation and its effects on intraseasonal variability of the NA monsoon (Bryson and Lowry 1955a, b; Carleton 1986, 1987). Analyses of composite synoptic pressure patterns have identified atmospheric configurations that enhance or stifle
convective activity (bursts or breaks) in terms of cloudiness, precipitation, and cloud-to-ground lightning (Carleton 1986, 1987; Watson et al. 1994b). Carleton (1986, 1987) has shown that widespread convective activity in the southwestern United States is typically associated with passing upper-level troughs in the westerlies. In addition, the northward displacement of the subtropical ridge and the formation of a cutoff “four corners high” also results in increased thunderstorm activity in Arizona. A southerly displacement of the subtropical ridge over northern Mexico tends to result in (drier) break conditions over the desert Southwest. His conclusions highlight the importance of the location of the subtropical ridge in controlling convective activity in the southwestern United States.

In examining cloud-to-ground lightning, Watson et al. (1994a) found that similar synoptic configurations to those of Carleton led to bursts and breaks in total convective activity. Again, a southerly shift of the subtropical ridge axis places Arizona under the strong, drying effects of the westerlies, which promote break conditions. A northerly shift of the subtropical high is associated with increased inflow of moisture from the south and results in widespread shower activity.

Although the results of the studies by Carleton (1986, 1987) and Watson et al. (1994a) agree on the importance of the latitudinal position of the subtropical ridge in determining bursts and breaks, Maddox et al. (1995) point out that the composite 500-mb burst pattern of Watson et al. (1994a; their Fig. 18c) closely resembles Carleton’s (1986) composite break pattern (his Fig. 5b). Their definition of burst and breaks are based, however, on different criteria: Carleton (1986) on total cloudiness over a large portion of the western United States, and Watson et al. (1994a) on cloud-to-ground lightning only in Arizona. Such intricacies further demonstrate the complex nature of the relationship between convective activity and synoptic circulation during the NA monsoon.

d. Interannual variability

The sparseness of long-term instrumental precipitation records (both temporally and spatially) in the southwestern United States and northwestern Mexico has hindered research on NA monsoon variability at interannual timescales. Nonetheless, several studies have noted marked variations in variability of the NA monsoon since the turn of the century (Green and Sellers 1964; Bradley 1976; Tenharkel 1980; Brazel and Prasad 1984; Carleton et al. 1990; Hereford and Webb 1992).

Interannual variability in NA monsoon activity may result from the corresponding variability of certain synoptic-scale patterns, similar to those associated with intraseasonal variability. Carleton et al. (1990) have argued that latitudinal shifts in the midlevel subtropical ridge over the southwest United States, which accounts for a great deal of within-season monsoon variability, are also responsible for year-to-year and decadal-scale variability. Figure 8 shows that, in general, northerly displacement of the subtropical ridge is associated with wetter Arizona summers, while a southerly shift coincides with decreases in summer rainfall totals (there are exceptions; e.g., 1990 was a
very wet year even though there was a southerly shift). Carleton et al. (1990) found that latitudinal displacement of the subtropical ridge is tied to different phases of the Pacific–North American (PNA) teleconnection pattern, which in turn is linked to the El Niño–Southern Oscillation (ENSO). Wetter summers are associated with the meridional phase of the PNA and positive Pacific SST anomalies that enhance ridging over North America. Drier summers have tended to follow zonal phases of the PNA.

e. ENSO-related studies

Relationships between ENSO and NA monsoon variability have been examined for possible causality, but such links have proven to be elusive. Some studies have analyzed the strength and geographic distribution of rainfall during and following ENSO events. More frequent tropical perturbations off the southwestern coast of Mexico have tended to follow the mature ENSO stage (Northern Hemisphere winter) when the ITCZ is displaced farther north and there is increased northward transport of water vapor (Reyes and Cadet 1988; Reyes and Mejía-Trejo 1991). Under these conditions, parts of northwestern Mexico (particularly the Sonoran Desert) have experienced positive anomalies in summer rainfall (Reyes and Mejía-Trejo 1991). Within the southwestern United States, Andrade and Sellers (1988) found little correlation between ENSO and total summer rainfall in New Mexico and Arizona. Yet Harrington et al. (1992) found a correlation between ENSO warm and cold events and the geographic distribution of summer precipitation for the same areas. Likewise, Hereford and Webb (1992) suggest there is some relationship between increased summer precipitation in the Colorado plateau region and the warm ENSO phase. Thus, relationships between ENSO-related phenomena and the NA monsoon are not yet well understood.

f. Paleoclimate studies

The focus of this review is on the modern NA monsoon, but a number of studies utilizing proxy data include description of summer paleoclimatic in the southwestern United States (e.g., Wells 1979; van Devender and Spaulding 1979; Davis and Shafer 1992). Briefly, paleoenvironmental conditions consistent with a circulation similar to the modern NA monsoon were present about 10 000 to 8000 BP, or roughly 10 000 yr after the last glacial maximum (van Devender and Spaulding 1979; Davis and Shafer 1992). Prior to this time, the presence of the Laurentide ice sheet over eastern North America delayed the development of the Bermuda high and associated summer rains west of the Continental Divide (van Devender and Spaulding 1979). The period since the establishment of a monsoonlike circulation has included intervals considerably wetter and drier than those found in the instrumental record (Davis 1994; Petersen 1994). In general, the atmospheric controls on the paleomonsoon are assumed to be fundamentally the same as those for the modern monsoon (i.e., position and strength of subtropical high and associated latitudinal position of the westerlies), with changes in their relative location and intensity causing the observed long-term variability.

6. Severe storms

a. Convective environment

An important aspect of NA monsoon research deals with severe storms and their predictability. The southwestern United States and northwestern Mexico have experienced tremendous population growth in the last few decades, and severe thunderstorms and related phenomena can pose a serious hazard for residents of this region. A range of studies have been performed, including those on the severe storm environment (Hales 1975; Randerson 1986; Smith and Gall 1989; McCollum et al. 1995; Maddox et al. 1995), lightning (Reap 1986; Watson et al. 1994a; Watson et al. 1994b), dust storms (Idso et al. 1972; Nickling and Brazel 1984; Brazel and Nickling 1986), and flash flooding (Randerson 1976, 1986; Maddox et al. 1980). In recent years, there has been a greater focus on the convective environment of severe storms in the NA monsoon. Much of the latest research has resulted from the SWAMP, which is aimed at greater understanding of convective environments in central Arizona and northwestern Mexico (Maddox 1990; Meitin 1991; Reyes et al. 1994).

Two common approaches for severe storm analysis in the NA monsoon region have been analysis of synoptic patterns (Maddox et al. 1980; Randerson 1976; Brazel and Nickling 1986; Maddox et al. 1995) and of satellite imagery (Hales 1975; Smith and Gall 1989; Farfán and Zehnder 1994). Smith and Gall (1989) describe a characteristic MCS of the NA monsoon that shares a number of characteristics with tropical squall lines. Here, MCSs typically form over the elevated regions of southeastern Arizona and northern Sonora as the result of the merging of late afternoon thunderstorm cells. By evening, the storms tend
to move in a westerly direction toward the low desert via discrete propagation. Although their life span is typically less than 1 day, they often produce remnant cyclonic circulation that can stimulate convection the next day over southern California and Baja California. Interestingly, this behavior is not too far from the turn-of-the-century description of the “Sonora storm” (Campbell 1906). The necessary pattern for the formation of these squall lines appears to be a 500-mb ridge with its axis centered over the Utah–Nevada border. This configuration promotes strong flow from the east and northeast over the elevated terrain along the Continental Divide. Under these conditions, dry midtropospheric air is advected westward over moist warm air covering the lower terrain of Sonora and southern Arizona. If low-level wind shear is present, squall lines may form. Work by Farfán and Zehnder (1994) highlights the association between strong midlevel easterlies over the Sierra Madre Occidental in the formation of propagating MCS over northern Mexico.

The same anomalous ridge pattern has also been associated with severe storms in central Arizona (Maddox et al. 1995) and dust storm generation in Arizona (Brazel and Nickling 1986). Figures 9a and 9b illustrate an example of this pattern for a severe storm over central Arizona. Other patterns have also been identified as precursors to severe thunderstorm development in central Arizona (Dunn and Horel 1994a; Maddox et al. 1995), and although these patterns are not always positioned directly over Arizona, they do provide clues for the evolution of the central Arizona convective environment (Maddox et al. 1995). Yet subsynoptic features such as low-level moisture advection from the Gulf of California and localized convergence resulting from the complex terrain of central Arizona are also critical for development of the severe thunderstorm environment. As a result, forecasting of severe thunderstorms for this region is extremely difficult (Dunn and Horel 1994a,b; McCollum et al. 1995).

b. Forecasting and modeling

Because the convective environment is tied to particular synoptic patterns, to various subsynoptic features, and to complex topography, considerable difficulties must be overcome for successful model prediction and forecasting. A major problem is that nearly identical synoptic situations can produce days where thunderstorms are either confined to topographically favored locations or where thunderstorm activity spreads to lower elevations (Dunn and Horel 1994a).

In Arizona, the forecasting skill scores for prediction of severe storm events have been exceptionally
low in comparison to other regions of the continental United States (Maddox et al. 1995). Likewise, numerical weather prediction models have not performed well for summer convection in Arizona. Dunn and Horel (1994a,b) describe the difficulties experienced with the National Center for Environmental Prediction models (both the nested grid and Eta models) in predicting southwest United States precipitation. Overall, both models provided poor predictions of heavy rain episodes over Phoenix. Several mesoscale features prevented the Eta model from properly simulating severe convection over Phoenix. The model does not resolve the advection of low-level moisture from the Gulf of California into central Arizona, which is critical for destabilization of the local environment prior to convection. This may result from the failure of the model to produce the low-level jet above the Gulf of California that seems to be responsible for moisture transport into the low deserts (Douglas et al. 1993; McCollum et al. 1995). This, in turn, may result from inaccurate initial conditions for summertime SSTs over the Gulf of California (Stensrud et al. 1995). In addition, another critical factor in localized severe convection is that of thunderstorm outflows (McCollum et al. 1995), which are not included in the model’s convection scheme. Stensrud et al. (1995) found that many of the general features of the NA monsoon circulation could be reproduced using the PSU–NCAR mesoscale model. They demonstrate that model initialization of mesoscale features such as the low-level winds over the Gulf of California and the summertime warming of the Gulf are critical for dynamically consistent simulations of mesoscale structures.

7. Summary and conclusions

The NA monsoon is one of the more complex and interesting atmospheric circulation features over North America, and the associated research literature dates back to the early part of this century. Although the circulation covers a significant portion of the continent, research on the NA monsoon (as well as its broader understanding in the meteorological and climatological communities) has not enjoyed a matching prominence. The recent appearance of several papers from the SWAMP project has, however, somewhat improved this situation, especially in the context of its mesoscale manifestations.

It appears that a consensus has gradually emerged regarding major circulation mechanisms. The role of the westward expansion of the Bermuda high was recognized early on, following the availability of upper-air data. Synoptic climatological studies have provided further details on the subtropical ridge and the effects of its latitudinal displacement. More recently, topographic influences and mesoscale phenomena have been highlighted as critical components of the NA monsoon circulation. These findings have had important consequences for the debate over moisture sources. Most investigators are now in agreement that low-level moisture is advected from the eastern tropical Pacific–Gulf of California, while the Gulf of Mexico may contribute to upper-level moisture. In fact, recent work suggests that moisture from these two sources becomes mixed over the Sierra Madre Occidental, prior to being transported northward into the United States.

The spatial extent of the NA monsoon is large, covering much of the western United States and northwestern Mexico. Historically, most of the literature has focused on Arizona, but the NA monsoon is actually most spatially consistent over northwestern Mexico with greater variability to the north. Variability at a range of temporal scales is related to interactions between synoptic and mesoscale circulations, in particular the latitudinal position of the subtropical ridge and the low-level moisture surges from the Gulf of California. Interannually, these relationships are partly influenced by the PNA pattern (and perhaps to some degree by ENSO). Forecasting of the NA monsoon, especially the ability to distinguish between weak or strong severe storm outbreaks, has been poor because current models do not adequately resolve the mesoscale features. Recent research from the SWAMP project, together with mesoscale modeling, has further highlighted the integrated roles of synoptic circulation, topography, and mesoscale circulations in controlling NA monsoon activity.

Many of the relationships noted above are not yet well understood, and there are a number of important questions for future research.

1) What are the relative roles of convective mixing over the Sierra Madre Occidental and subsequent northward moisture transport compared to low-level moisture from the Gulf of California and upper-level moisture from the Gulf of Mexico, with respect to widespread rainfall across the NA monsoon region?

2) Additionally, what are the spatial and temporal variations in each of these kinds of moisture de-
livery and how are they related to seasonal and interannual circulation changes?

3) With respect to ongoing research into the central Arizona convective environment, what are the causes of the nocturnal maxima in thunderstorm activity over the lower deserts?

4) Furthermore, what are the complementary roles of low-level moisture advection and localized convergence at the mesoscale over these same areas?

5) What conditions trigger gulf surges (at the synoptic or mesoscale), and can their relationship to convective activity and precipitation be quantified?

6) The ENSO signal seems unclear in interannual NA monsoon variability, so what other causes are likely to influence interannual variation?

7) For improved forecasting and modeling, a better understanding of mesoscale features is required on a routine basis. What are the minimum data requirements and how might the data be acquired?

Regarding this last question, Douglas and Stensrud (1996) have called for an increased density of upper-air stations, which for the NA monsoon area would help to overcome some of the current problems with unresolved circulation features and complex topography.

Overall, there is clearly a need to better understand the complex links between synoptic circulation, topography, gulf surges, and mesoscale features within the NA monsoon. As mentioned above, this will most likely be achieved via improved data availability and modeling at the mesoscale. As with many other areas of atmospheric research, advances in understanding and forecasting depend to some degree on new measurements, in this case at a level of detail currently unavailable (except in field campaigns such as SWAMP). In conclusion, a somewhat cohesive body of knowledge has emerged from past and current NA monsoon research, but the picture is by no means complete. There are still many aspects of the NA monsoon that require further investigation, and these include a number of challenging meteorological and climatological questions.

Acknowledgments. We wish to thank Bob Maddox, Andrew Carleton, and the reviewers for their detailed comments on early versions of the manuscript. Also, thanks go to Sandra Brazel of the Office of Climatology at Arizona State University for providing the 1984 satellite image.

References


PACS, 1996: Pan-American Climate Studies (PACS) homepage. [Available online at http://taro.atmos.washington.edu/pacs.]


——, 1986: A mesoscale convective complex type storm over the Desert Southwest. NOAA Tech. Memo. NWS WR-196, 54 pp. [Available from NWS-NOAA, Western Region, 125 South State St., Salt Lake City, UT 84147.]


