

## Submesoscale Spatiotemporal Variability of North American Monsoon Rainfall over Complex Terrain

MEKONNEN GEBREMICHAEL AND ENRIQUE R. VIVONI

*Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, New Mexico*

CHRISTOPHER J. WATTS

*Departamento de Física, Universidad de Sonora Hermosillo, Sonora, Mexico*

JULIO C. RODRÍGUEZ

*Instituto del Medio Ambiente y Desarrollo Sustentable del Estado de Sonora Hermosillo, Sonora, Mexico*

(Manuscript received 2 December 2005, in final form 22 March 2006)

### ABSTRACT

The authors analyze information from rain gauges, geostationary infrared satellites, and low earth orbiting radar in order to describe and characterize the submesoscale (<75 km) spatial pattern and temporal dynamics of rainfall in a 50 km × 75 km study area located in Sonora, Mexico, in the periphery of the North American monsoon system core region. The temporal domain spans from 1 July to 31 August 2004, corresponding to one monsoon season. Results reveal that rainfall in the study region is characterized by high spatial and temporal variability, strong diurnal cycles in both frequency and intensity with maxima in the evening hours, and multiscaling behavior in both temporal and spatial fields. The scaling parameters of the spatial rainfall fields exhibit dependence on the rainfall rate at the synoptic scale. The rainfall intensity exhibits a slightly stronger diurnal cycle compared to the rainfall frequency, and the maximum lag time between the two diurnal peaks is within 2.4 h, with earlier peaks observed for rainfall intensity. The time of maximum cold cloud occurrence does not vary with the infrared threshold temperature used (215–235 K), while the amplitude of the diurnal cycle varies in such a way that deep convective cells have stronger diurnal cycles. Furthermore, the results indicate that the diurnal cycle of cold cloud occurrence can be used as a surrogate for some basic features of the diurnal cycle of rainfall. The spatial pattern and temporal dynamics of rainfall are modulated by topographic features and large-scale features (circulation and moisture fields as related to geographical location). As compared to valley areas, mountainous areas are characterized by an earlier diurnal peak, an earlier date of maximum precipitation, closely clustered rainy hours, frequent yet small rainfall events, and less dependence of precipitation accumulation on elevation. As compared to the northern section of the study area, the southern section is characterized by strong convective systems that peak late diurnally. The results of this study are important for understanding the physical processes involved, improving the representation of submesoscale variability in models, downscaling rainfall data from coarse meteorological models to smaller hydrological scales, and interpreting and validating remote sensing rainfall estimates.

### 1. Introduction

The North American monsoon (also referred to as the southwest, Mexican, or Arizona monsoon) is a subcontinental-scale climate feature that produces a sig-

nificant increase in rainfall during the summer months in northwestern Mexico and the southwestern United States (Douglas et al. 1993; Adams and Comrie 1997; Fuller and Stensrud 2000). It is the most important source of water in the region, as it accounts for 50%–70% of the annual precipitation. The monsoon impacts semiarid areas, which are generally characterized by low annual rainfall and large interannual variability. The large-scale variability of monsoon rainfall and its relationship to teleconnective, and synoptic- and meso-scale forcing mechanisms has been a focus of many

---

*Corresponding author address:* Dr. Mekonnen Gebremichael, Dept. of Earth and Environmental Science, New Mexico Institute of Mining and Technology, MSEC 254, 801 Leroy Place, Socorro, NM 87801.

E-mail: mekonnen@nmt.edu

studies (e.g., Douglas et al. 1993; Adams and Comrie 1997; Higgins and Shi 2001). However, work on the submesoscale (<75 km) variability of rainfall is noticeably absent in the literature, mainly due to a paucity of ground observations and the complexity of the terrain. Information on the submesoscale variability is of great interest to water resource managers in this water-scarce region. It is essential to address basic research questions on the link between topography and monsoon rainfall variability. Improved large-scale numerical simulations also depend on the proper characterization of the submesoscale variability (Gutzler et al. 2005).

In this paper, we present an in-depth examination of the submesoscale monsoon rainfall variability in the periphery of the core monsoon region using a variety of rainfall data. The rainfall data consist of (i) in situ observations from a network of tipping-bucket rain gauges deployed during a summer field experiment (the Soil Moisture Experiment 2004, SMEX04), (ii) space-based radar observations from the Tropical Rainfall Measuring Mission Precipitation Radar (TRMM-PR), and (iii) space-based cloud-top temperature observations from the Geostationary Operational Environmental Satellites- Infrared (GOES-IR). These sensors have complementary features that allowed us to look at different aspects of the temporal and spatial variability of rainfall using a multitude of techniques, and also to perform a limited validation assessment.

This paper is organized as follows. Section 2 presents the regional topographical and meteorological features. Section 3 describes the study area and the different datasets used. The spatial variability of marginal statistics, joint statistics, the diurnal cycle, and fractals in temporal rainfall fields are presented in section 4, on the basis of hourly rain gauge observations. Section 5 provides the spatial features and multifractal properties of spatial rainfall fields, obtained from TRMM-PR data. Section 6 presents the spatial variability of the diurnal cycle of cloud-top temperature obtained from GOES-IR data, and compares it with that of rainfall obtained from rain gauge observations. Finally, section 7 closes the paper with conclusions.

## 2. Regional characteristics

The North American monsoon region is bounded to the west by the Pacific Ocean, including the Gulf of California, and to the east by the Gulf of Mexico and by the central plains of the United States (Fig. 1). The interior of the region is characterized by complex topography. Noteworthy features are the north-south-aligned mountain ranges through Nevada, southwest-

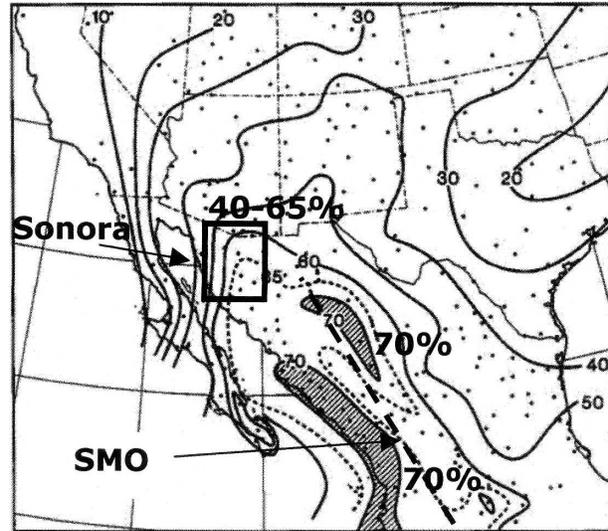


FIG. 1. Analysis of the contribution of the precipitation during July–September to the annual total, expressed in percent (adapted from Douglas et al. 1993). Greater than 70% is cross-hatched. Stations used in the analysis are shown as dots. The thick dashed line indicates the SMO. The inset contains the study region located in the state of Sonora. The crosshatched region along the coast indicates the monsoon core region.

ern Arizona, and northwestern Sonora; the Mexican plateau that is defined by the Sierra Madre Occidental (SMO) and Sierra Madre Oriental to the west and east, respectively; the Sonora Desert located along the boundary between Sonora and Arizona; and the Chihuahua Desert between the SMO and Sierra Madre Oriental.

The development of the North American monsoon is characterized by heavy rainfall in late May or early June over southern Mexico, which quickly spreads northward along the western slopes of the SMO into northwestern Mexico by early July (Higgins et al. 2003). Precipitation increases over northwestern Mexico coincide with the increased vertical transport of moisture by convection (Douglas et al. 1993) and southerly winds flowing up the Gulf of California (Baden-Dagan et al. 1991). Rainfall in this North American monsoon region appears to be associated with transients (e.g., gulf surges, easterly waves, tropical storms) rather than the mean flow (Fuller and Stensrud 2000; Englehart and Douglas 2001; Berbery 2001; Ellis et al. 2004).

Figure 1 (adapted from Douglas et al. 1993) shows the contribution of the summer (July–September) monsoon rainfall to the annual precipitation total. The maximum contribution of the monsoon rainfall is located along the western slope of the SMO in northwestern Mexico. This part of Mexico is the core monsoon region and receives up to 70% of its annual rainfall in

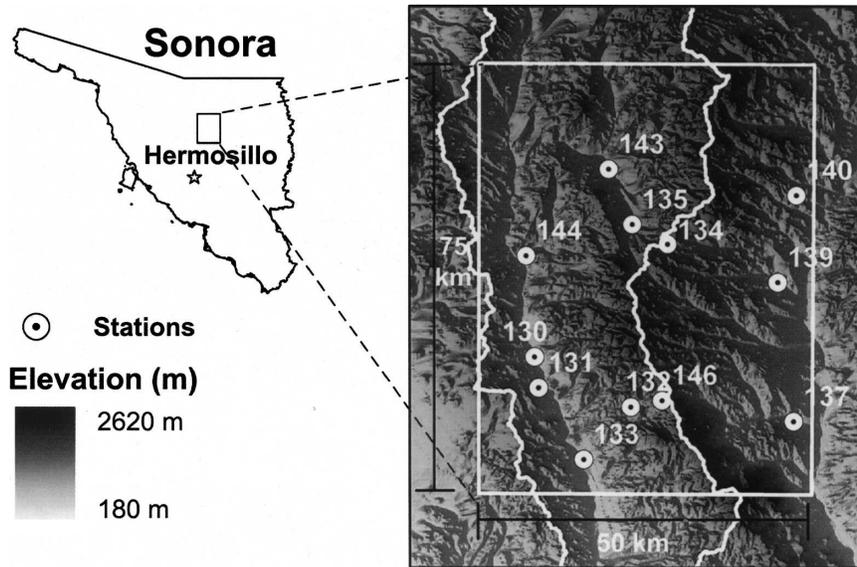


FIG. 2. Layout of the rain gauge stations in the study region, a 50 km  $\times$  75 km  $\times$  50 km box in northern Sonora, and its topographic characteristics from a 90-m digital elevation model (DEM).

the months of July–September (Gutzler et al. 2005). The contribution of rainfall decreases as one moves away from this region. Our study area in the state of Sonora lies in the periphery of the core monsoon region, and receives 40%–65% of its annual rainfall from the monsoon. Notice that the monsoon signature shows large spatial variability in Sonora.

### 3. Study region and data

#### a. Study region

Figure 2 presents our study region and its topographic variability. Our study region is located in northern Sonora. It is bounded by 30.50°N to the north, 29.83°N to the south, 110.75°W to the west, and 110.23°W to the east. Studies have demonstrated that rainfall to the north of, approximately, 29.0°N lies in the periphery of the core North American monsoon rainfall regime (e.g., Gochis et al. 2004; Gutzler et al. 2006).

Our study area is roughly 50 km (east–west) by 75 km (north–south). Note the north–south-trending mountain ranges and river valleys in the study area that form part of the SMO. The topographic distribution is characterized by a high mean elevation and a large elevation range, which are primarily due to the effects of channel incision (Coblentz and Riitters 2004). Two major ephemeral (seasonal) rivers flow north–south through the region: the Río San Miguel (west) and Río

Sonora (east), with the former draining into the latter south of the study area.

#### b. Rain gauge network

Our dataset consists of hourly rainfall rates at 12 rain gauge stations, with records spanning from 1 July through 31 August 2004. The dataset has no periods of missing data. Although 14 gauges were installed in this region, we excluded two of them from our analysis because they had some missing data that affected our analysis. As shown in Fig. 2, four of the stations are located to the east of the watershed divide (Sierra Aconchi), whereas the rest are located to the west. Stations 134 and 146 are located near the Sierra Aconchi mountaintop, stations 132 and 135 are located on the slopes within a short distance (<7 km) from the mountaintop, and the rest are located on the foot slopes or valleys far from the mountaintop. The elevations represented by the stations range from 660 to 1375 m (see Table 1 for the geolocation information of the stations).

The rain gauge types are 6-in. tipping buckets with a resolution of 0.2 mm (TE525, Texas Electronics, Dallas, Texas). Habib et al. (2001) and Ciach (2003) showed that the random errors in tipping-bucket measurements average out at 15-min accumulation, and so our choice of an hourly time scale is large enough to filter out the random errors. Since our goal is to assess the relative variability of rainfall over this region, absolute accuracy of the rain gauge observations is not as important in this

TABLE 1. Sonora rain gauge network site geolocation information. Note: Lat–lon, spheroid Clarke 1866, datum NAD27. Universal Transverse Mercator (UTM), spheroid Clarke 1866, zone 12, datum NAD27.

Site	Lat (°)	Lon (°)	Northing (m)	Easting (m)	Elevation (m)	Regime
130	30.0403	–110.6736	3323293	531471	724	III, Río San Miguel valley
131	29.9908	–110.6671	3317811	532109	719	III, Río San Miguel valley
132	29.9604	–110.5202	3314498	546296	905	
133	29.8768	–110.5954	3305202	539068	642	III, Río San Miguel valley
134	30.2200	–110.4612	3343285	551854	1180	I, Sierra Aconchi
135	30.2516	–110.5177	3346767	546401	1044	IV, Cucurpe
137	29.9371	–110.2617	3312043	571255	660	II, Río Sonora valley
139	30.1590	–110.2867	3336621	568684	758	II, Río Sonora valley
140	30.2982	–110.2568	3352065	571467	1017	II, Río Sonora valley
143	30.3397	–110.5551	3356513	542767	960	IV, Cucurpe
144	30.2016	–110.6869	3341164	530135	799	III, Río San Miguel valley
146	29.9705	–110.4704	3315634	551093	1375	I, Sierra Aconchi

study as are the relative differences spatially and temporally. All stations had recorded between 57 and 85 total hours of rainfall, for the two monsoon months (July and August) used in this study. Time series of hourly rainfall averaged over 12 stations (Fig. 3) show the following features: the monsoon onset occurs on 7 July, the peak rainfall ( $13 \text{ mm h}^{-1}$ ) occurs on 13 July around midnight, the tall spikes in rain rates decrease over time, for the most part the region receives rain somewhere on a daily basis, and the longest periods without any rain in the region are in the late weeks of August.

### c. TRMM-PR

The TRMM satellite is  $\sim 400 \text{ km}$  above the earth's surface, and orbits between  $38^\circ\text{N}$  and  $38^\circ\text{S}$  in a non-sun-synchronous orbit. It is equipped with the Precipitation Radar (PR), among other sensors. The most unique characteristic of the TRMM-PR is its capability

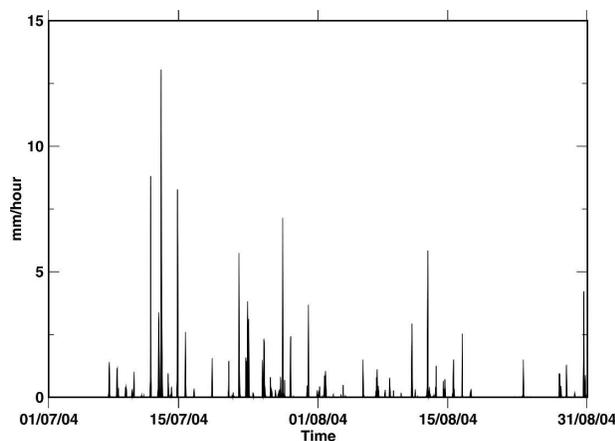


FIG. 3. Hourly rain rate averaged over 12 gauges, 1 Jul–31 Aug 2004.

to observe the three-dimensional structure of rain from space. The minimum detectable signal that can be observed by the TRMM-PR above the noise level is about  $16\text{--}18 \text{ dBZ}$  in the absence of attenuation, which roughly translates to a rainfall rate of about  $0.5 \text{ mm h}^{-1}$  (Kummerow et al. 1998).

Product 2A25 is the principal instantaneous TRMM-PR dataset detailing the rain structure (Iguchi et al. 2000). Among the several variables in 2A25, we used the version-6 near-surface rainfall rate products processed by the TRMM Science Data and Information System (TSDIS). These products represent “instantaneous” rainfall-rate maps with a horizontal resolution of  $5.0 \text{ km}$  at nadir. The first step in creating 2A25 was to correct for the effects of attenuation and the non-uniform beam-filling effect (NUBF) in the original reflectivity values. The NUBF effect correction method is described in Kozu and Iguchi (1999). The attenuation correction method is a hybrid method between the traditional Hitschfeld–Bordan (Hitschfeld and Bordan 1954) path-integrated attenuation correction method and the surface reference technique correction method (Iguchi et al. 2000; Meneghini et al. 2000). The rain rate is then estimated from the corrected reflectivity values using a reflectivity–rain-rate power law in which the parameters are functions of the horizontal and vertical structures of the rainfall and attenuation.

### d. GOES-IR

We used the *GOES-West* (i.e., *GOES-10*) infrared ( $10.2\text{--}11.2 \mu\text{m}$ ) cloud-top temperatures, available from the National Oceanic and Atmospheric Administration Comprehensive Large Array-data Stewardship System (NOAA-CLASS) database. The data have a spatial resolution of  $4 \text{ km}$  at nadir and were available every half-an-hour. Cloud-top temperature is often used as a

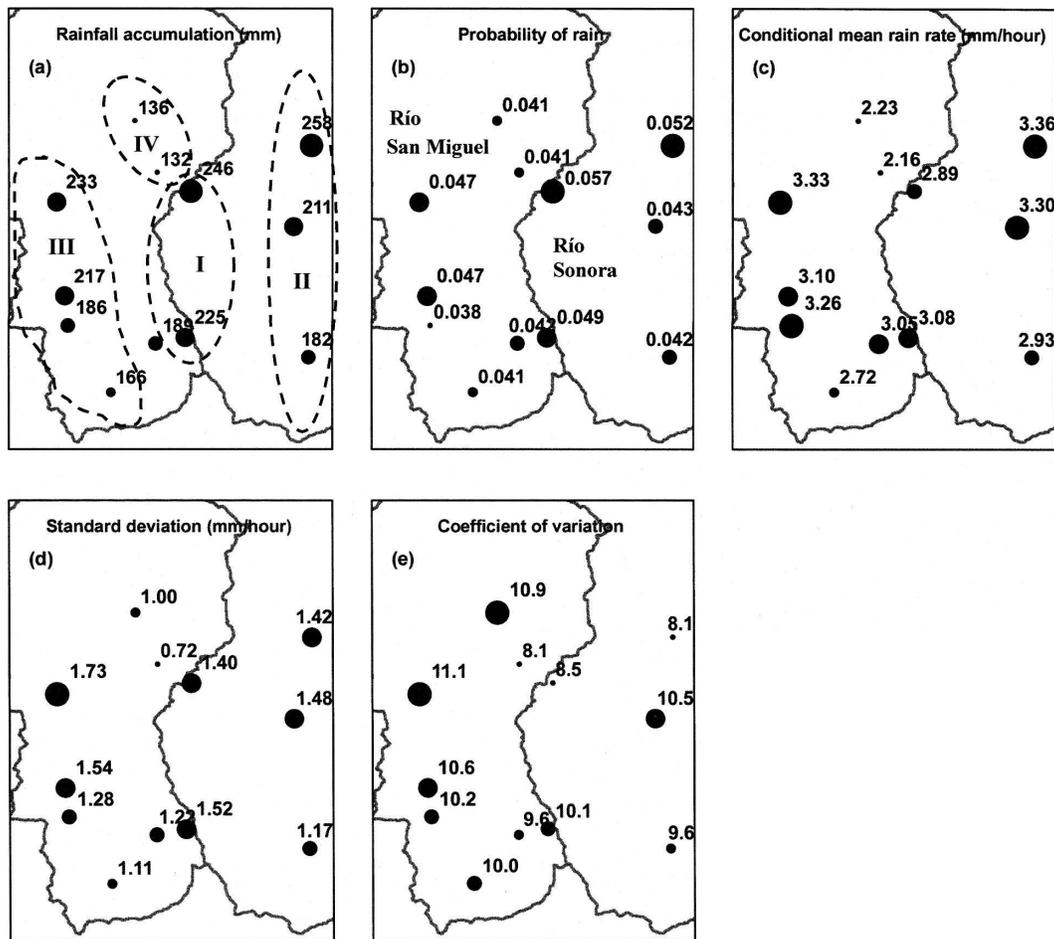


FIG. 4. Spatial variability of various hourly rain-rate marginal statistics: (a) rainfall accumulation (mm), (b) probability of rain, (c) conditional mean rain rate defined as mean of positive rain rates ( $\text{mm h}^{-1}$ ), (d) standard deviation ( $\text{mm h}^{-1}$ ), and (e) coefficient of variation. Note that all the statistics used, except the conditional mean rain rate, are not conditioned on the rain rates being positive.

proxy for deep convection, with the assumption that a colder cloud-top temperature implies a higher cloud top and thus a more intense precipitation than a warmer cloud top. We set three threshold temperatures to define cold cloudiness: 235, 225, and 215 K, corresponding to low, middle, and high clouds, respectively. We point out that Arkin and Meisner (1987) used 235 K as the threshold temperature for convection when they developed the GOES Precipitation Index (GPI), which is used in the derivation of the Global Precipitation Climatology Project (GPCP) precipitation products (Huffman et al. 2001; Gebremichael et al. 2005). For each pixel, we counted the number of cases (over a 2-month period) that have temperatures below the thresholds. During the counting, if there were one or two cases that satisfied the criteria for a given hour (recall there are two scans per hour), we counted them

as one. This ensures that the results are comparable with those of hourly gauge observations.

#### 4. Results obtained with hourly rain gauge observations

##### a. Spatial variability of marginal statistics

We computed selected rain statistics for each station, and show the results in Fig. 4. The rainfall total during the 2-month period varies from 132 to 246 mm (i.e., a factor of 2) depending on the location (Fig. 4a). We delineated four rainfall regimes based on spatial coherence: I, Sierra Aconchi; II, Río Sonora valley; III, Río San Miguel valley; and IV, Cucurpe. We did not incorporate station 132 in one of the rainfall regimes because it is the only station that is not located either along valley areas or close to mountaintops. The first regime

(I, Sierra Aconchi) includes stations 134 and 146; both are near the Sierra Aconchi mountaintop. The Sierra Aconchi regime receives large rainfall totals (225–250 mm) that decrease in the north–south direction at a gradient of  $0.76 \text{ mm km}^{-1}$  distance. The second regime (II, Río Sonora valley) lies along the river system in the Río Sonora watershed, to the east of the watershed divide. The rainfall totals in the Río Sonora valley regime decrease in the north–south direction, which coincides with the drop in elevation, at a gradient of  $0.2 \text{ mm km}^{-1}$  elevation drop or  $2.1 \text{ mm km}^{-1}$  distance. The third regime (III, Río San Miguel valley) lies along the river system in the Río San Miguel watershed, to the west of the watershed divide, and runs parallel to the second regime. The Río San Miguel valley regime also shows a decrease in rainfall total in the north–south direction, coinciding with the drop in elevation, at a gradient of  $0.4 \text{ mm km}^{-1}$  elevation drop or  $1.8 \text{ mm km}^{-1}$  distance. The fourth regime (IV, Cucurpe) is confined to a relatively small area in the narrow upper valley that is bounded by the Sierra Aconchi to the right and by a ridge to the left that separates it from the Río San Miguel valley regime. The Cucurpe regime receives the smallest rainfall total, which does not show an appreciable change with distance or elevation. However small it may be, the rainfall total also decreases in the north–south direction.

It may be concluded that the 2004 monsoon rainfall exhibited high spatial variability in the region, and its distribution was influenced by topography. We have identified four rainfall regimes that have a distinct geographical setting and spatial rainfall pattern. All regimes share the pattern that the rainfall total systematically decreases from north to south. However, the rate of decrement and its behavior with respect to elevation vary from regime to regime. Along the two major river systems (II, Río Sonora valley; III, Río San Miguel), rainfall accumulation shows a marked dependence on elevation: it decreases with decreasing elevation. However, this pattern is lost in other regimes. For the regimes along the river systems, we obtained a (Spearman's) rank correlation coefficient of 0.96 between elevation and rainfall accumulation, suggesting a strong linkage between the two. If we combine all regimes, the resulting rank correlation coefficient drops to 0.36, telling us that combining different rainfall regimes destroys the apparent dependence structure between elevation and rainfall accumulation observed along the river systems.

Along the two rivers, elevations of rain gauges decrease as one goes from north to south. This raises the following questions: Is the high correlation obtained between elevation and rainfall rate merely a conse-

quence of the fact that the elevations are decreasing in the north–south direction? Or is there indeed a strong relationship between elevation and rainfall rate, after the effect of the north–south direction (characterized by interstation distance) is removed? To investigate this, we used the partial correlation statistic. The formula for the partial correlation between two random variables  $Y$  and  $X$  with the effect of the third random variable  $W$  removed from both, denoted as  $r_{YX|W}$ , is

$$r_{YX|W} = \frac{r_{XY} - r_{XW}r_{YW}}{\sqrt{1 - r_{XW}^2}\sqrt{1 - r_{YW}^2}}, \quad (1)$$

where  $r_{XY}$ ,  $r_{XW}$ , and  $r_{YW}$  are Pearson's coefficients. We found partial correlations of 0.51 and 0.99 between the elevation and the rainfall accumulation, after the effect of the interstation distance among them is removed, for the Río San Miguel (III) and Río Sonora (II) regimes, respectively. This suggests that there is indeed a strong dependence of rainfall accumulation on elevation, which cannot be explained by the interstation distance alone. However, given the small sample sizes used in this analysis (four for Río San Miguel; three for Río Sonora), further investigation with more stations is needed to reach a definite conclusion.

The probability of rainfall occurrence (Fig. 4b) and conditional rainfall (Fig. 4c) maps reveal differences in rainfall-producing mechanisms between mountainous and valley sites. The mountainous sites receive frequent, small events, while the valley sites are dominated by larger, infrequent storms. This is consistent with Gochis et al. (2004) who showed that single convection cells were common at high elevations, while a higher fraction of low-elevation rains were from more organized systems. The standard deviation varies from 0.72 to 1.73 (i.e., a factor of 2.4) over the region (Fig. 4d). The standard deviation shows larger values for the stations that receive large rainfall totals, indicating large hour-to-hour variability in these regimes. As can be seen from the coefficient of variation (CV = standard deviation divided by mean) values in Fig. 4d, the standard deviations amount to 800%–1100% of the unconditional mean rain rate, suggesting a high variability in the rainfall amount from event to event. The temporal variability at the mountainous sites exhibits less hour-to-hour variability than those at the valley sites, which is probably due to the relatively infrequent rain events at the valley sites and a corresponding smaller sample size.

In Fig. 5, we show the spatial distribution of the maximum hourly rainfall rate and its Julian day of observation. There is a marked difference between the maximum rainfall rates at various stations, ranging from 13

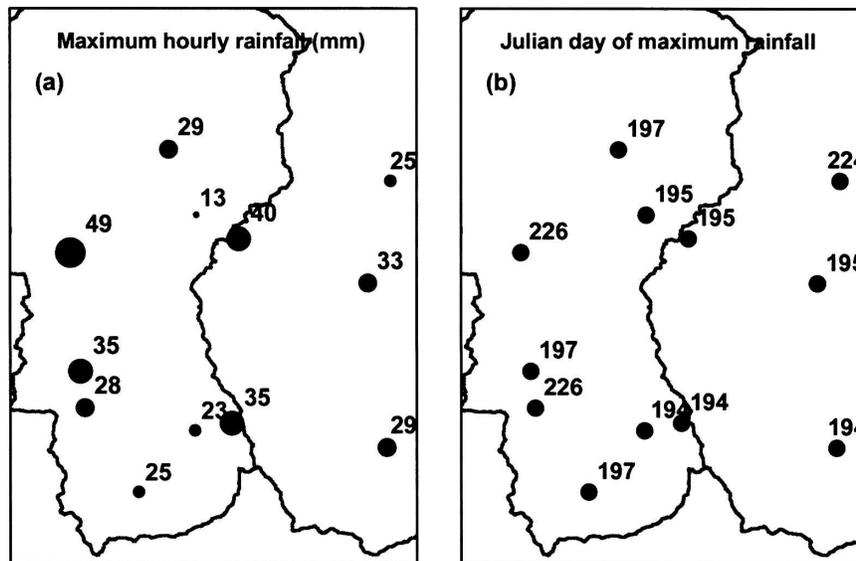


FIG. 5. Spatial distribution of (a) maximum hourly rainfall rate and (b) its corresponding Julian observation day. Note that single events on Julian days 194, 195, 197, 224, and 226 account for maximum rates over the region.

to  $49 \text{ mm h}^{-1}$  (i.e., a factor of 3.8). The highest maximum rain rate ( $49 \text{ mm h}^{-1}$ ) occurs in the high elevations of the Río San Miguel valley (III). This value is comparable to the highest maximum rain rate ( $41.7 \text{ mm h}^{-1}$ ) reported by Gochis et al. (2003) for the summer 2002 monsoon rainfall based on 10 rain gauges deployed at elevations ranging from 500 to 1000 m. The maximum precipitation occurs earlier on the mountainous and nearby sites, which is consistent with the results of Douglas et al. (1993). Figure 6 presents the entire variability in terms of the histogram of the normalized frequency of hourly positive rain rates, with a bin width of  $0.5 \text{ mm h}^{-1}$ , focusing on the rain rate below  $25 \text{ mm h}^{-1}$  for the sake of clarity. All the distributions have tall narrow spikes for hourly rainfall accumulations of 0.5 mm and less, and are skewed with a level of skewness varying from 2 to 5. Regions that receive large rainfall have generally high skewness and kurtosis. We found a strong correlation (0.98) between the skewness and kurtosis values.

#### b. Spatial variability of joint statistics

The previous section focused on how the marginal distributions of rainfall at station locations vary spatially. Here, our focus is on the joint distribution (i.e., covariation) of rainfall measured at two stations, in terms of the Pearson correlation coefficient and the critical success index (CSI). This analysis is important to answer questions such as the following: How large a geographical area can one assume is well represented

by rainfall statistics derived from the stations? The answer to this question is important in the design of rain gauge networks and areal rainfall estimation (Rodríguez-Iturbe and Mejía 1974), data assimilation (e.g., Krajewski 1987) and determining the uncertainty in areal rainfall averages (Morrissey et al. 1995; Ciach and Krajewski 1999; Gebremichael et al. 2003), among other applications.

Combining all rainfall regimes, we calculated the Pearson correlation coefficient between pairs of stations and show the results in Fig. 7a, along with a fitted analytical function. In general, the correlations decay with increasing interstation distance. There is however a large scatter in the correlations; for example, at an interstation distance of 30 km, the correlation may vary from 0.10 to 0.45 depending on the stations. This could be partly explained by the standard error associated with the correlation estimates, and partly by the mix of different rainfall regimes. For example, Gebremichael and Krajewski (2004) found that the standard error in correlation estimated from 15-min rainfall data over a 2-month period could reach up to 0.1 for Florida, dominated by small-scale summer convection. The sample sizes are much smaller for hourly data, and so the standard error could be larger than 0.1. In this study, we did not attempt to calculate the standard error because of the difficulty associated with obtaining reliable estimates for small sample sizes (see the discussion in Krajewski et al. 2000; Habib et al. 2001; Gebremichael and Krajewski 2004). The best-fitted function shown in Fig.

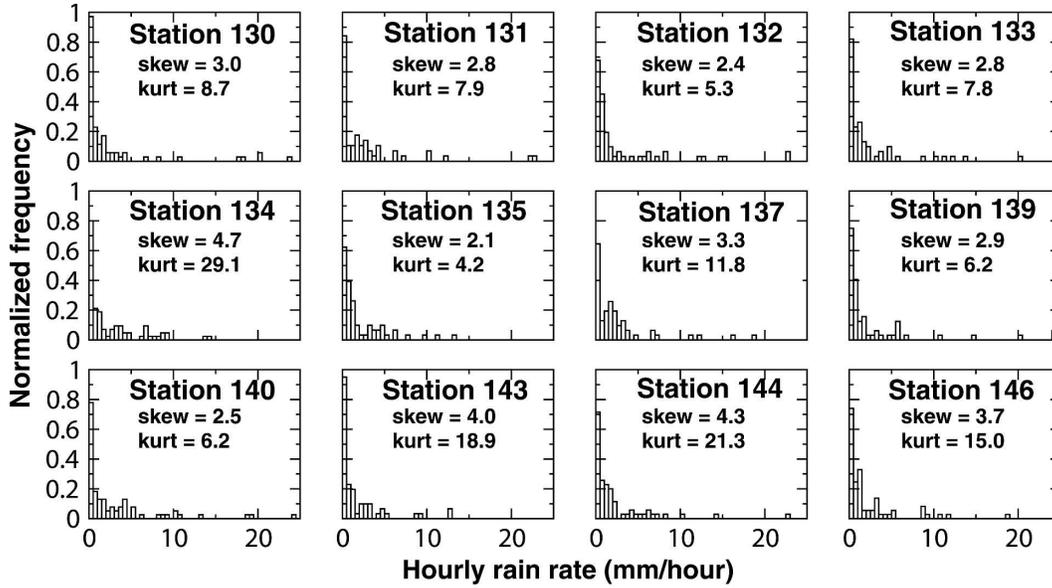


FIG. 6. Frequency distribution histograms for hourly positive rain rates at the stations used in this study, focusing on rain rates below  $25 \text{ mm h}^{-1}$ . The histograms are normalized to 1 (hence producing a probability distribution function). The skewness and kurtosis values are shown for each distribution.

7a is  $\exp[-(h/17)^{0.80}]$ , where  $h$  is the interstation distance. The correlation distance (i.e., the distance at which the correlation becomes 0.37 or insignificant) is 17 km. This implies that a negligible portion (less than

15%) of the total variance in one of the stations is explained by the variance in the other station located more than 17 km away. The exponent in the fitted function is less than unity, indicating a sharper drop in correlation at small separation distances ( $<17 \text{ km}$ ) than that predicted by the commonly used exponential functions (i.e., with an exponent of unity).

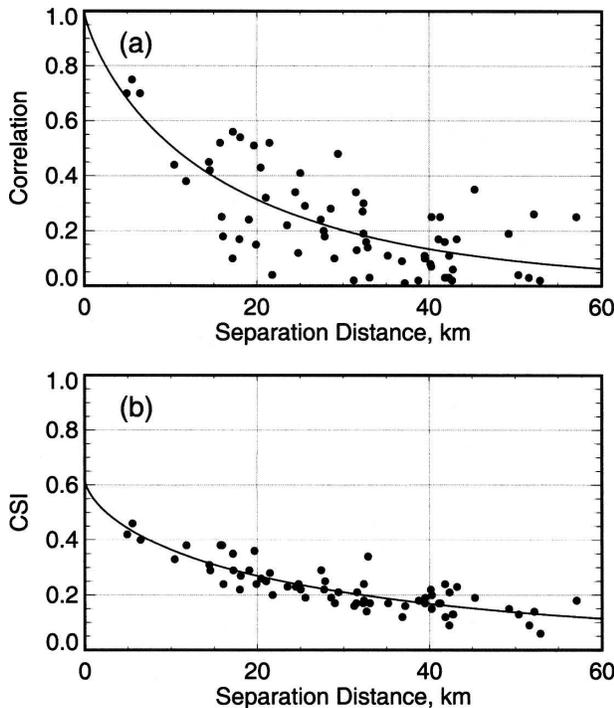


FIG. 7. Spatial statistics of hourly rain rate as a function of interstation distance, and fitted functions. The statistics are the (a) Pearson's correlation coefficient and (b) CSI.

Let us now examine the intermittence (rain–no rain) covariability of rainfall given the occurrence of rain as recorded by either gauge. We used the critical success index (CSI) as a measure of this covariability. Let  $B(x, y)$  represent the number of hours when there was rain at station  $x$  and none at station  $y$ ,  $C(x, y)$  represent the number of hours when there was no rain station  $x$  and rain at station  $y$ , and  $A(x, y)$  represent the number of hours when there was rain at both  $x$  and  $y$  stations. The CSI is then defined as

$$\begin{aligned}
 \text{CSI} &= \text{CSI}(x, y) = \text{CSI}(y, x) \\
 &= \frac{A(x, y)}{A(x, y) + B(x, y) + C(x, y)}. \quad (2)
 \end{aligned}$$

CSI measures the presence–absence of rain at two stations given the presence of rain at one or both stations. A CSI value of zero indicates rain at one station is accompanied by no rain at the other station within that hour, whereas a value of one indicates rain at one station is accompanied by rain at the other station within that hour. Note that the CSI does not measure the co-

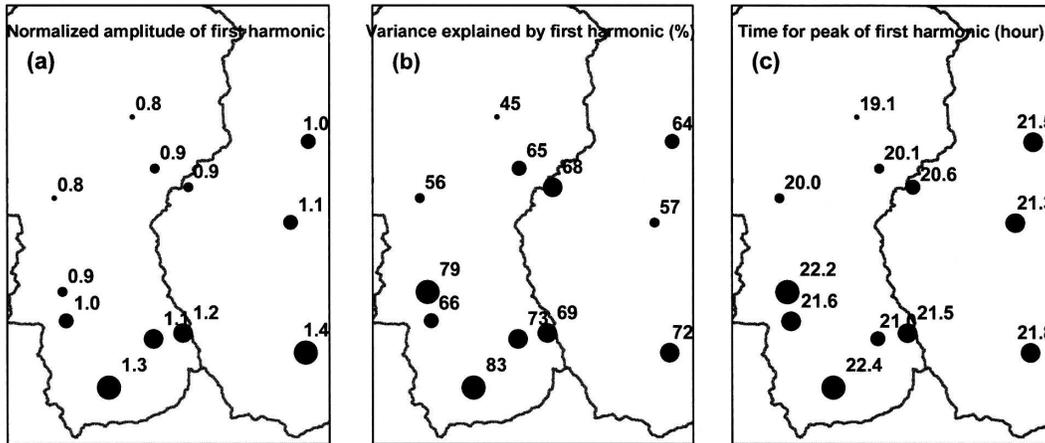


FIG. 8. (a) Normalized amplitude of first harmonic, (b) variance explained by the first harmonic (%), and (c) local solar time (h) for peak of the first harmonic, derived from hourly rainfall frequency.

variability of no-rain events. In Fig. 7b, we show the CSI between pairs of stations as a function of the distance between them, as well as the analytical function. For two stations located 5 km apart, half of the time rain at one station is not accompanied by rain at the other station within that hour. The CSI generally decays with distance, and the decay can be captured by an exponential function of this form:  $0.62 \exp[-(h/26)^{0.65}]$ . The nugget effect (i.e., obtained as  $1-0.62$  from the fitted equation) shows a significantly large natural variability at small separation distances (Journel and Huijbregts 1978; de Marsily 1986). Notice also that the scatter around the CSI fitted function is very small, compared to that of the correlation. It may therefore be concluded that the region is dominantly characterized by localized, convective, cells with radii smaller than 5 km. This implies that coarse rainfall products (such as those obtained from remote sensing data) could be subject to large nonuniform beam-filling problems, and validation of these products using rain gauges requires a dense network.

### c. Spatial variability of diurnal cycle

Establishing the diurnal cycle in rainfall is important to understanding the physical processes involved on this time scale and to producing accurate forecasts. Currently, the diurnal cycle is poorly represented in models over the North American monsoon regions (Li et al. 2004; Gutzler et al. 2005). The diurnal cycle is also essential in interpreting satellite rainfall estimates, since low earth orbiting satellites view a given area only intermittently, and interpolating between the measurements should be adjusted according to the time of the day. We examined the diurnal cycle using the method

of harmonic analysis, a method used in many other investigations of diurnal rainfall patterns (e.g., Balling and Brazel 1987; Bell and Reid 1993; Dai 2001). In this method, the diurnal cycle is expressed by Fourier decomposition:

$$\hat{P}(t) = P_0 + P_1 \cos(\omega t - \phi_1) + P_2 \cos(2\omega t - \phi_2) + \dots, \quad (3)$$

where  $t$  is the hour of the day;  $\hat{P}$  is the fitted (i.e., estimated) statistic;  $\omega$  equals  $2\pi/24$ , where 24 indicates the number of hourly intervals per day; and  $P$  and  $\phi$  are the amplitude and phase angle of the cosine function. The zeroth ( $P_0$ ), first ( $P_1$ ), and second ( $P_2$ ) harmonic components correspond to the mean, diurnal, and semi-diurnal cycles. We used the method of least squares to obtain these parameters. The portion of the variance explained by the  $r$ th harmonic component can be computed as  $P_r^2/2\sigma^2$ , where  $\sigma$  is the standard deviation of the 24 hourly values.

The rainfall frequency is the statistic most frequently used in rainfall diurnal cycle studies, because it provides a much cleaner signal for assessing the occurrence of rainfall on the diurnal cycle (Nesbitt and Zipser 2003; Gochis et al. 2004). We obtained the rainfall frequency by counting the number of rainfall events that occurred in a specific hourly interval, and applied the harmonic analysis to the rainfall frequency. In Fig. 8, we show the normalized amplitude of the first harmonic, the variance explained by the first harmonic, and the time of the diurnal maximum derived from the phase of the first harmonic, for the rainfall frequency statistic. We obtained the normalized first harmonic amplitude by dividing the amplitude of the first harmonic by the am-

plitude of the zeroth harmonic, following Balling and Brazel (1987) and Dai et al. (1999). This result quantifies the peakedness of the hourly time series, ranging from zero for a flat time series to two in the case of all zero values, except one rainfall peak during the day.

As can be seen from Fig. 8a, the normalized amplitudes (ranging from 0.8 to 1.4) are fairly high; thus, a strong diurnal cycle is indicated for the region. Therefore, a rainfall estimation scheme that samples conditions only once a day will tend to significantly bias the rainfall over this region. The highest values of normalized amplitudes are observed for the southern portion of the region. There is no significant difference between the values over the mountaintop and over the slopes. The percentage of variance accounted for by the diurnal cycle ranges from 45% in the northernmost part of the region to 83% in the southernmost part of the region (Fig. 8b), indicating that the diurnal cycle gets more pronounced as one goes from north to south. This pattern may indicate that the signature of the monsoon increases in the north–south direction, consistent with the results shown in Fig. 1 (i.e., contribution of summer rainfall to annual rainfall increases in the north–south direction). There is no significant difference in terms of the variance explained between the rainfall over the mountaintop and over the slopes.

The time of maximum, as interpreted from the phase angle of the first harmonic, ranges from 1910 to 2240 LST (Fig. 8c). These results are consistent with those of Gochis et al. (2004) and Li et al. (2004), who found that the maximum rainfall frequency occurs in the afternoon over the high terrain of the SMO, and is shifted later in the evening as one goes to the northwest (in the direction of our study region). The maximum rainfall frequency occurs later as one moves away from the mountain toward the valley, in the direction of downslope winds. Along the rivers, the maximum rainfall frequency starts early in the northern (higher elevation) region and moves toward the southern (lower elevation) region in the late evening.

What is the cause of the nocturnal rainfall? Studies of the monsoon rainfall have suggested several mechanisms; however, none of these mechanisms appears, by itself, to explain the majority of the rainfall variability pattern. Berbery (2001) analyzed the Eta Model's moisture flux at 950 hPa and suggested that the transients rather than the mean flow play a dominant role in bringing moisture flux into this region. Tropical storms are one transient phenomenon that brings abundant rainfall to the region (Englehart and Douglas 2001). Analyzing 9 yr of radiosonde observations and National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalyses,

Douglas and Leal (2003) found that the highest rainfall in Sonora is associated with gulf surges, reinforcing the notion that surges are associated with the presence of convective cloud masses over the southern gulf, and that these in some manner develop northward with time. However, the relative contributions of the transient phenomena (e.g., gulf surges and tropical storms) are not clear (e.g., Douglas and Leal 2003). Numerous studies have also highlighted the importance of the southerly nocturnal low-level jet from the Gulf of California (e.g., Tucker 1993; Douglas 1995; Fawcett et al. 2002; Li et al. 2004), in contributing to nighttime boundary layer convergence that favors nocturnal convection in this region.

The second harmonic (not shown here) explains less than 10% of the variability in all stations except two. For stations 137 and 139, the second harmonic explains 20% and 15% of the variability, respectively, and the corresponding times of the semidiurnal peaks derived from the second harmonic are 2100 and 2300 LST, respectively. This suggests that the two harmonics are phase-locked and appear to reinforce the primary maximum depicted by the first harmonic fit. In other words, the predominant diurnal peak and the asymmetry about this in the time series adds some power to the semidiurnal harmonic.

Up to this point, we have described the diurnal cycle of rainfall frequency. How much do the frequently occurring events contribute to the total rainfall? If the contribution is small, the usefulness of the results may be limited. To answer this question, we sorted the rainfall rates in ascending order and constructed a running accumulation from the smallest to the largest amounts. In Fig. 9, we present a plot of the percentage of the running accumulation to the total accumulation against the percentage of observations to the total observations, for each station. The figure shows that the large contribution of the total rainfall comes from infrequent yet heavy rains. For example, all stations show that 50% of the total rainfall is provided by the heaviest 10%–15% of the observations. These results reveal that although heavy rainfall events, associated with deep convection, are relatively rare, they contribute disproportionately to the total rainfall.

We applied the harmonic analysis to the rainfall intensity and show the results in Fig. 10. Overall, the diurnal cycle in rainfall intensity follows that of rainfall frequency, with high amplitude and a late afternoon or evening maximum. These findings provide evidence that the heavy yet infrequent events as well as the light yet frequent events occur in the late afternoon and evening hours. The basic spatial variability features of rainfall intensity are similar to those of rainfall frequency.

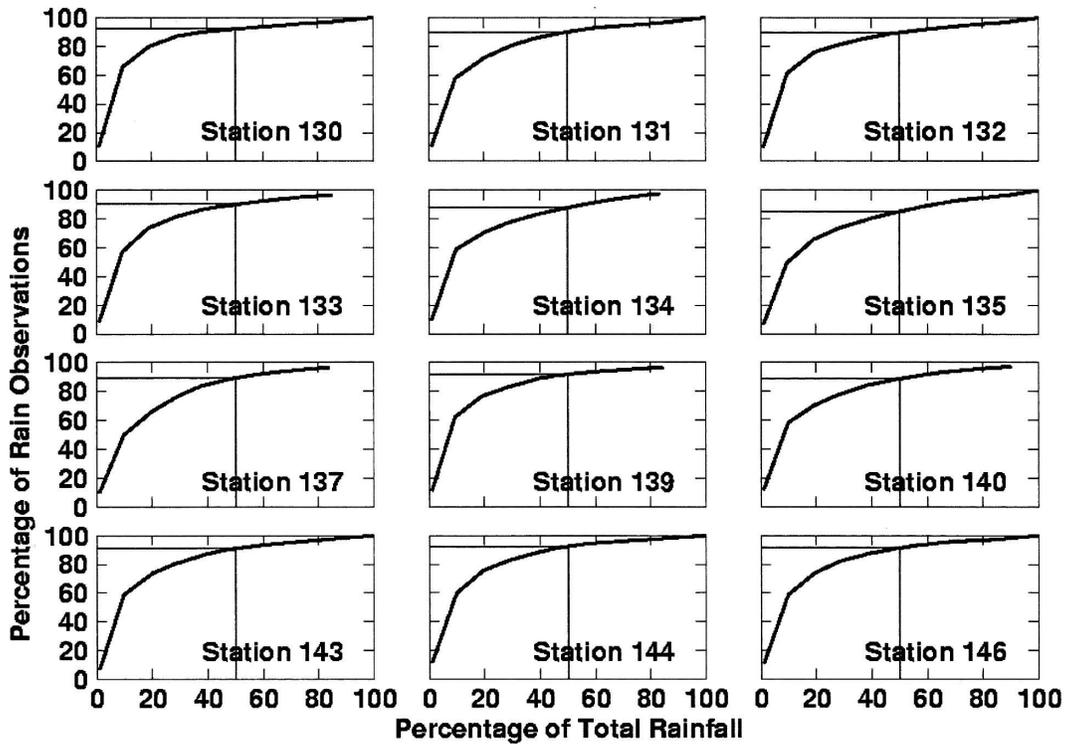


FIG. 9. Percentage of observations (nonraining observations were discarded) vs percentage of total rainfall, for each gauge station.

However, there are some differences between the actual values of the two diurnal cycle parameters. There is a larger hourly variation in rainfall intensity, of which the amplitude amounts to 110%–150% of the daily mean. The amplitude of the diurnal cycle of rainfall frequency amounts to 80%–140% of the daily mean. The lag time between the rainfall intensity and rainfall frequency peak hours is within 2.4 h. In most cases, the

peak rainfall intensity precedes the peak rainfall frequency, suggesting a nonsymmetric typical storm temporal structure (early sharp peak and more slowly falling tail).

*d. Fractals in temporal rainfall pattern*

The (multi)fractality in the structure of the rainfall process may lead to a better understanding of its vari-

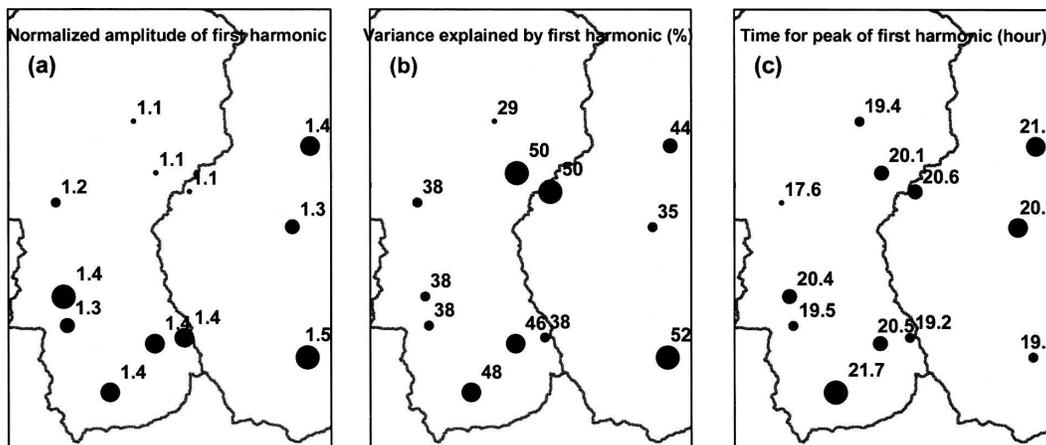


FIG. 10. Same as in Fig. 8 but for rainfall intensity.

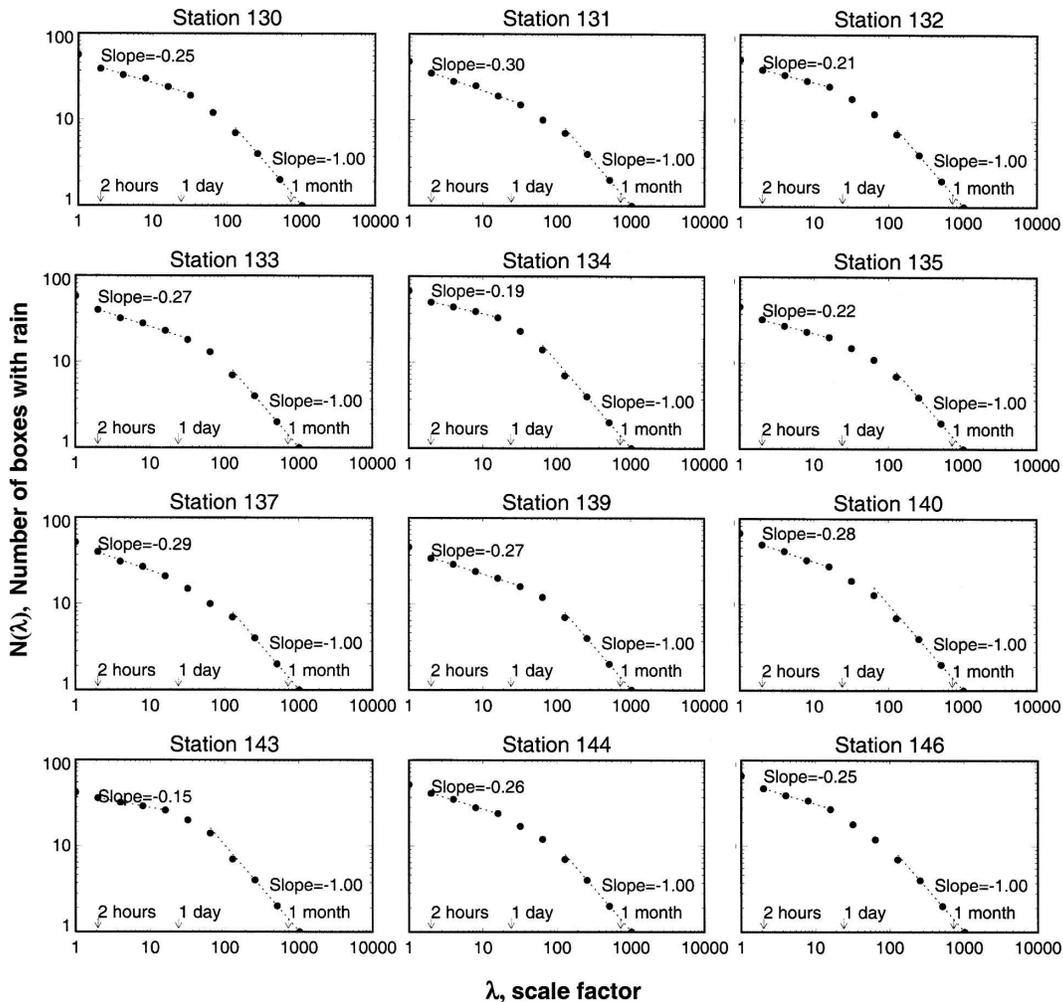


FIG. 11. Box-counting log-log plot obtained with hourly rainfall. The plots display time scales from 1 h up to 42.7 days. Two distinct scaling regimes are shown, with the corresponding slope parameters.

ability that cannot be grasped from other descriptions of the complex dynamics of this process. Fractal geometry (Mandelbrot 1982) is an extension of classical geometry and concerns the analysis of subsets of metric spaces that are typically geometrically complicated. The fractal set is defined by some relation between the structures observed in the set at various levels of resolution (e.g., Barnsley 1993). This relation is formulated quantitatively by the concept of fractal dimensions. In this study, we will focus on the fractal dimension of the intermittence (rain–no rain) of the rainfall time series. The box-counting method is usually used to investigate the fractality of the intermittence (de Lima and Grasmann 1999). The general procedure is to progressively divide the space of an observation into nonoverlapping boxes (it is common for the size to be decreased gradually by a factor of 2) of side  $\lambda$ . For every grid size, the incidences of boxes that contain rain  $N(\lambda)$  are counted.

If the set is fractal (or scale invariant), then it can be expressed by the expression

$$N(\lambda) \propto \lambda^{-D}, \quad (4)$$

where  $D$  is the fractal dimension. The fractal dimension measures how densely the set occupies the metric space in which it lies.

Figure 11 shows the box-counting plots obtained with the hourly rainfall data. The plots display time scales from 1 h to 42.7 days (we used data from 1 July through 12 August 2004 for the purpose of scaling analysis). The absolute value of the slope of the log-log plot over a range of scales gives an estimate of the fractal dimension of the set of rainy periods observed in the period of time. We have indicated the fractal dimensions of two scaling regimes in the plots. The results imply that the rainfall distribution in time is scale invariant but with

different fractal dimensions for different ranges. All the stations show a scaling regime extending from 2 to 16 h, and another scaling regime extending from 2.7 days onward. The fractal dimensions in the former scaling regime exhibit spatial variability: the stations in the mountain have small fractal dimension, whereas the stations in the major river systems have higher fractal dimensions. This indicates that the stations in the river systems (compared to those in the mountains) are characterized by a denser structure with rainy hours closely clustered together. We found a Spearman's correlation of  $-0.61$  between elevation and fractal dimension, suggesting that lower elevations are more likely to have larger fractal dimensions and, hence, closely clustered rainfall events. For the scaling regime above 2.7 days, the fractal dimension for all stations is 1 (i.e., dimension of a line), implying saturation of the process, meaning rainfall is always occurring within a period of at least 2.7 days (recall that the period under investigation is 1 July–12 August).

## 5. Results obtained with TRMM-PR observations

The TRMM-PR can offer a unique vantage point for examining the spatial patterns of rainfall due to its characteristics of wide areal coverage, and fairly reasonable accuracy and resolution. We performed the spatial characterization of the PR rainfall fields by means of (multi)fractal or (multi)scaling analysis. The spatial scaling properties of rainfall fields have recently attracted much attention in the research community. One reason for this increased interest is the need to fill the gap between the large scales of meteorological model outputs and the smaller hydrological scales. Discrepancies in scale also arise when remote sensing estimates are compared to point measurements for validation. Studies have documented the importance of small-scale rainfall variability on runoff simulation (Ogden and Julien 1993, 1994; Winchell et al. 1998), radiative transfer computations (Harris et al. 2003), estimation of land–atmosphere fluxes (Nykanen et al. 2001), and water balance in land surface schemes (Lammering and Dwyer 2000). The impact of ignoring the small-scale rainfall variability and the propagation of this variability via the nonlinear equations of hydrological models can result in significant biases of the predicted variables. Rainfall downscaling models often require only two to three parameters to reproduce rainfall over a large range of scales and, hence, could serve as a possible bridge for the transfer of information from large scales to small scales. However, before implementing scale-invariance transformation methods, there are research questions that need to be addressed: Is the North

American monsoon rainfall scale-invariant? Can the scaling parameters be predicted from large-scale observables like large-scale average rain rate? In this section, we attempt to answer these questions.

To begin, let us look at the large-scale spatial variability of rainfall obtained from the TRMM-PR overpasses. We found a total of eight cases of TRMM-PR overpasses, over the entire 2-month period, that crossed the study region when rain was present. Figure 12 presents these cases; each grid shown is  $5^\circ \times 5^\circ$ . The 28 July case was a mesoscale convective system (MCS) with a much more extensive area of rainfall. This event occurred a few minutes after midnight, and it consisted of stratiform rain with embedded convective cells. The cases of 1, 3, and 7 August have large, scattered, convective cells. The remaining four cases consist of small, scattered, convective cells.

A more detailed view of the TRMM-PR data over the study region is shown in Fig. 13. The results shown are resampled to  $0.1^\circ$  resolution over a domain of  $2^\circ \times 2^\circ$  ( $29.0^\circ$ – $31.0^\circ$ N,  $111.5^\circ$ – $109.5^\circ$ W); the inset shows our study region (box) and the location of the rainfall gauges (points). We calculated the spatial scale of each system as the diameter of a circle that has an area equivalent to the average contiguous area (excluding events occurring on single pixels). Figure 13 shows a large contiguous rain area around midnight on 28 July, due to the MCS event. The spatial scale of this rain system is 160 km. For the other events occurring in the evening or late night, the spatial scales amount to 78, 65, 52, and 29 km, corresponding to 3, 7, 15, and 23 August. The spatial scales for the events occurring in the morning or afternoon are 21, 58, and 48 km, corresponding to the cases of 10 July, and 1 and 18 August. In summary, the TRMM-PR overpasses during the 2004 monsoon season indicate that the largest contiguous rain areas occur in the evening, while smaller, localized, events occur either in the afternoon or in the evening. None of the storm events covered the entire area during the period.

Let us now focus on the scaling properties of the TRMM-PR rainfall fields. The scaling characteristics of a geophysical field can be parameterized in several ways. In this section, we perform a scaling analysis in the manner of Over and Gupta (1994, 1996). The spatial scaling is best described by starting with the largest-scale  $L_0$ . Consider a two-dimensional ( $d = 2$ ) region with dimensions  $L_0 \times L_0$ , which is successively divided into  $b$  equal parts ( $b = 2d$ ) at each step, and the  $i$ th subregion after  $n$  levels of subdivision is denoted by  $\Delta_n^i$ . At the first level, the region is subdivided into  $b = 4$  subregions denoted by  $\Delta_1^i$ ,  $i = 1, 2, \dots, 4$ . At the second level, each of the above subregions is further subdivided

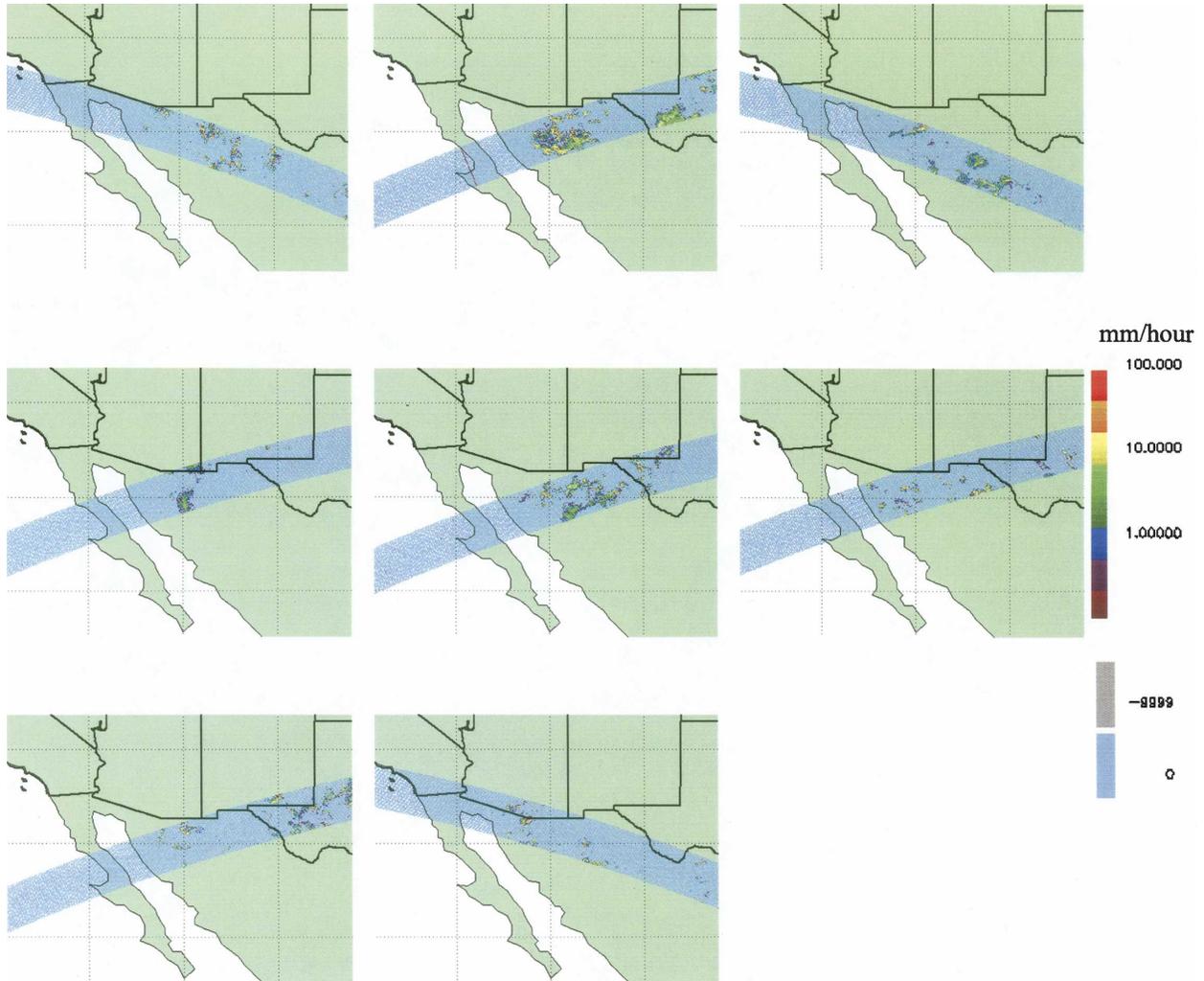


FIG. 12. The PR swaths over the study region. Date (yymmdd) and time are shown for each overpass. Each grid is  $5^\circ \times 5^\circ$ .

vided into  $b = 4$  subregions, which are denoted by  $\Delta_2^i$ ,  $i = 1, 2, \dots, 16$ , for a total of  $b^2 = 16$  subregions. At the  $n$ th level, we have a total of  $b^n$  subregions. Denoting the side length at the  $n$ th level as  $L_n$ , the scale factor at level  $n$  is given by

$$\lambda_n = L_n/L_0 = b^{-n/d}. \quad (5)$$

For the subregion  $\Delta_n^i$ , we denote the *volume* of water falling in this subregion as  $\mu(\Delta_n^i)$ .

We define the spatial moments of the volume of water as

$$M_n(q) = \sum_{i=1}^{b^n} \mu^q(\Delta_n^i), \quad (6)$$

where  $q$  is the moment order (e.g.,  $q = 0$  is the rain-no-rain intermittency, and  $q = 1$  is the mean). The

scaling analysis in space can be performed by investigating the behavior of spatial moments (6) for different spatial scales  $\lambda_n$ . The rainfall intensity is considered to exhibit spatial scale invariance at moment order  $q$  if the following relationship holds:

$$M_n(q) \propto \lambda_n^{-\tau(q)}, \quad (7)$$

in the limit as  $n$  approaches infinity. Therefore, for scale invariance to hold, the parameters  $\tau(q)$ , referred to as (multi)scaling parameters, should not depend on the spatial scale  $\lambda_n$ . This presupposes the existence of a finite scaling range between two scales referred to here as the smallest scale ( $L_{\min}$ ) and the largest scale ( $L_0$ ).

We analyzed each TRMM-PR scene separately. We had to first select  $L_{\min}$  and  $L_0$  for which the scaling law would be investigated. We used  $L_{\min} = 5$  km and  $L_0 = 160$  km. Our choices of the largest and smallest scales

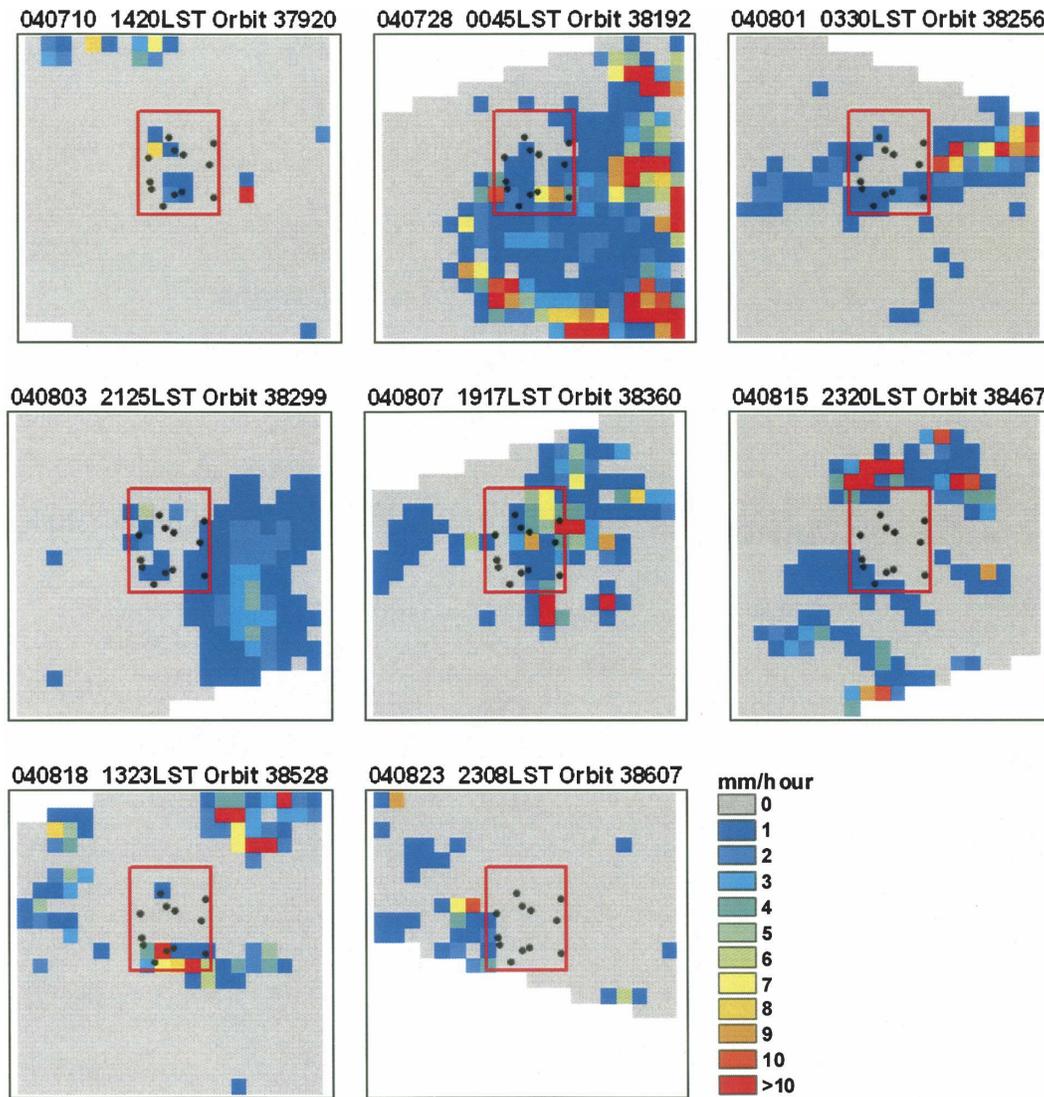


FIG. 13. The PR swaths over the study region outlined by red. Date (yymmdd), time, and orbit number are indicated for each overpass. Dots represent locations of rain gauge stations.

were dictated by the PR's swath width (247 km) and resolution (5 km), respectively. There are four levels in between the smallest and largest scales, and the corresponding scale factors are 1,  $\frac{1}{2}$ ,  $\frac{1}{4}$ ,  $\frac{1}{8}$ ,  $\frac{1}{16}$ , and  $\frac{1}{32}$ . Here,  $\lambda_0 = 1$  corresponds to  $L = L_0$ , and  $\lambda_5 = 1/32$  corresponds to  $L = L_{\min}$ . Estimation begins with deriving rainfall maps from each scene at different spatial scales. The PR rainfall data are available at  $L_{\min}$  scale. We aggregated the pixels simply by averaging to obtain the data at different spatial scales up until  $L_0$ . From each scene of data, we estimated  $\tau(q)$  as a slope of the regression equation  $[\ln M_N(q)]$  versus  $-\ln \lambda_n$ , obtained by log-transforming (7), and applying evenly weighted least squares regression.

In Fig. 14, we show the scaling of the moments  $M_n(q)$  for  $q = 0, 0.5, \dots, 4$ . The log-log linearity for all the TRMM-PR scans is reasonably good at moment orders ( $0 \leq q \leq 3$ ), indicating that the spatial rainfall fields are scale invariant at these moment orders. The perceived failure of scale invariance for moment orders exceeding three might be perhaps due to the small sample size used in estimating higher-moment orders (Troutman and Vecchia 1999).

In Fig. 15, we show the  $\tau(q)$  estimates that correspond to the scaling of moments shown in Fig. 14. For  $q < 1$ , the  $\tau(q)$  estimates were the highest for the 28 July case and the lowest for the 10 July case, implying that the increment of rainy areas with increasing scale is

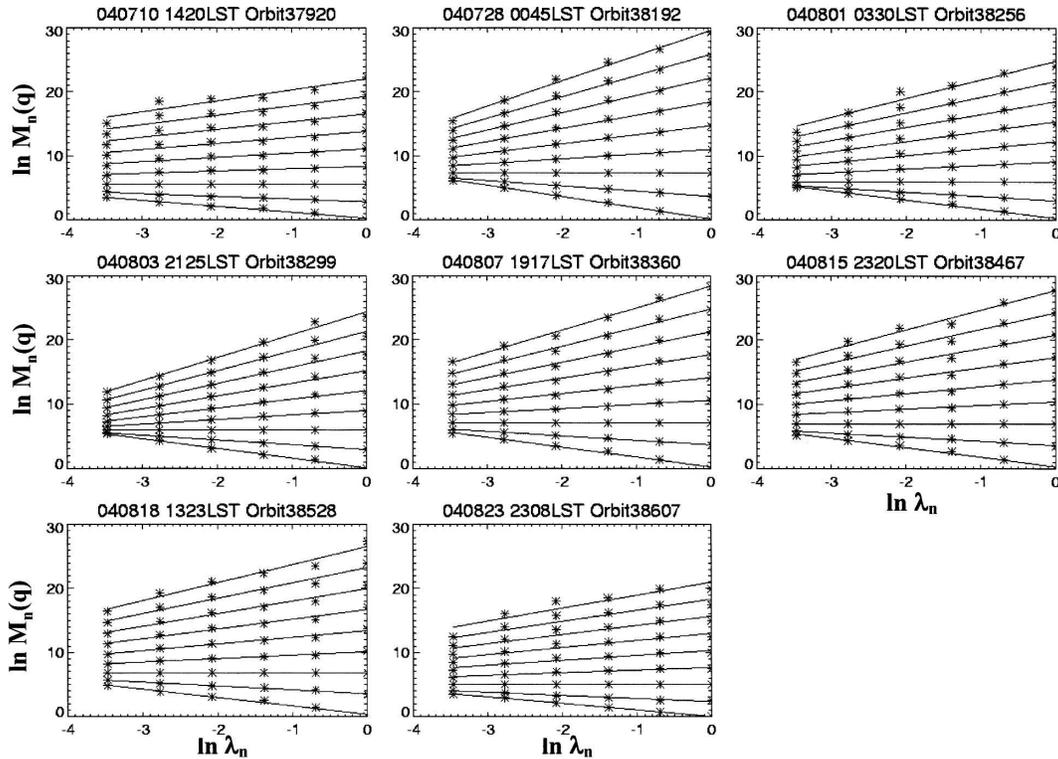


FIG. 14. Scaling of the marginal moments with moment orders. Within each panel, the moment orders  $q = 0.0, 0.5, 1.0, \dots, 4.0$  are organized from the bottom of the plots up. Each panel represents one overpass; date (yymmdd), time, and orbit number are indicated for each PR overpass.

more pronounced on 28 July than on 10 July. For  $q > 1$ , the  $\tau(q)$  estimates were the highest for the 10 July case and the lowest for the 28 July case, implying that the spatial rainfall intensity field is more uniform on 28 July than on 10 July. Recall that the 28 and 10 July cases correspond to the highest (160 km) and lowest (21 km) spatial scales, respectively, observed during the summer 2004 TRMM-PR overpasses.

The scaling parameters  $\tau(0)$  and  $\tau(2)$  are usually sufficient to parameterize some commonly used models of  $\tau(q)$ , and consequently may be the only parameters required to simulate scale-invariant fields. For example, the beta-lognormal cascade model proposed by Over and Gupta (1996) requires only an estimate of  $\tau(0)$  to parameterize the intermittency and an estimate of  $\tau(2)$  to fully describe the scaling properties of positive rain rate, that is,  $\tau(q > 0)$ . Below, we will discuss the interpretations of the two scaling parameters and decipher their relationships, if any, with the rainfall rate at the synoptic scale. This helps to address the question, Can the scaling parameters be derived from large-scale observables that can be obtained from meteorological models?

The intermittence scaling parameter  $\tau(0)$  is the fractal dimension of the support of  $\mu$  and measures the rate

of growth of the fraction of the rainy areas with scale (Hentschel and Procaccia 1983). Here,  $\tau(0) = 0$  indicates a single box with rain at each scale, whereas  $\tau(0) = 2$  indicates rain everywhere (Gebremichael et al. 2006). So  $\tau(0)$  is theoretically bounded by zero and two, with higher values indicating increasingly large rainy areas. In Fig. 16a, we present a scatterplot of  $\tau(0)$  estimated from each TRMM-PR overpass versus the corresponding large-scale average rain rate  $\bar{R}$ . Our results show a strong one-to-one relationship of the following functional form between  $\tau(0)$  and  $\bar{R}$ :

$$\hat{\tau}(0) = s \ln \bar{R} + i. \quad (8)$$

where  $s$  (slope) and  $i$  (intercept) are the fit parameters. Over (1995) and Gebremichael et al. (2006) also found a similar functional relationship between  $\tau(0)$  and  $\bar{R}$ , for different datasets. Over (1995) and Over and Gupta (1994,1996) based their analysis on ground-based radar rainfall from the Global Atmospheric Research Program (GARP) Atlantic Tropical Experiment (GATE) conducted in the tropical Atlantic. Gebremichael et al. (2006) based their analysis on both TRMM-PR and ground-based radar rainfall data at the oceanic Kwajalein site, and coastal Melbourne, Florida, and Houston,

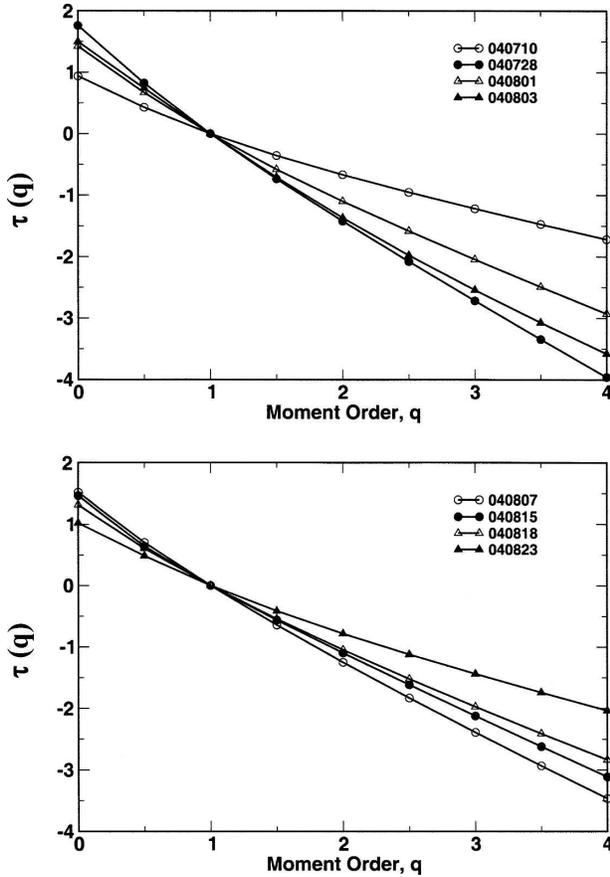


FIG. 15. The  $\tau(q)$  vs  $q$  plots, corresponding to the scaling of the moments shown in Fig. 14, for each TRMM-PR image.

Texas, sites. Table 2 compares our parameter estimates (mean  $\pm$  standard deviation) with those of these studies. We estimated the uncertainty (i.e., the standard error) in our parameter estimate that resulted from the small number of points, using a bootstrapping resampling experiment. Our “best estimate” slope parameter is higher than those for the other sites, whereas our intercept lies within the range reported for the other sites.

What does this imply about monsoon convection?

Over (1995) related the slope parameter to the number of levels  $N$  between the largest scale at which scale invariance holds and the scale at which the probability distribution of rain rate is independent of  $\bar{R}$ . This relation follows analytically [see Over (1995) for details] from two observations: (a) the particular relation between  $\tau(0)$  and  $\bar{R}$  expressed in (8), and (b) the independence of the scaling of the positive rain rates on the large-scale rain rate (result not shown here). At larger scales, rain rates increase with  $\bar{R}$ , and at smaller scales, rain rates decrease with  $\bar{R}$ . Since the latter seems physically unlikely, this scale was interpreted by Over (1995) as the minimum scale at which scaling invariance can hold. Following Over’s approach, we found

$$N = 2/(s \ln b). \tag{9}$$

For  $s = 0.3241$ , (9) gives  $N \approx 4$ . Note that this is not a prediction of either the smallest or largest scales alone, but only the number of levels between them that includes the minimum (5 km) and maximum (160 km) spatial scales used in this study. For the oceanic and coastal sites (see Table 2),  $N \approx 5 - 7$ . If we assume the same size of convective cells for all sites, our results suggest that the MCSs have smaller areas of rainfall over the semiarid Sonora than over the oceanic and coastal sites. This agrees with Nesbitt et al. (2000) who showed larger areas of MCSs over ocean than over land. However, the uncertainty associated with our slope parameter estimate (see Table 2) is very large, suggesting that further investigation with larger samples of TRMM-PR overpasses and information on the size of convective cells is required to reach a definite conclusion.

We point out that whereas  $\tau(0)$  represents how the rain–no-rain areas vary with the spatial scale, for a fixed temporal scale, the fractal dimension ( $D$ ) in the box-counting power-law relation (4) represents how the hourly rain–no-rain events vary with the temporal scale, for a fixed spatial scale. There are studies that attempt to link these two scaling parameters in space–

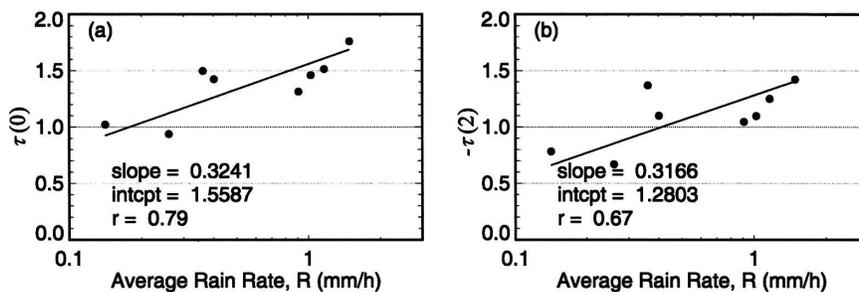


FIG. 16. Dependence of the scaling parameters on the spatial average rain rate  $\bar{R}$  estimated from the PR overpasses. The scaling parameters are (a)  $\tau(0)$  and (b)  $\tau(2)$ .

TABLE 2. A summary of  $\tau(0)$  vs  $\bar{R}$  regression fits obtained from PR and ground-based radar (GR) scans, reported by previous and current studies.

Site	Sensor used	Parameters obtained		Reference
		Slope	Intercept	
GATE phase I	GR	0.2040	1.5356	Over (1995)
GATE phase II	GR	0.2064	1.5104	
Kwajalein	PR	0.2634	1.6409	Gebremichael et al. (2006)
Kwajalein	GR	0.2252	1.5097	
Houston, TX	PR	0.2185	1.5156	
Houston, TX	GR	0.1816	1.6934	
Melbourne, FL	PR	0.2275	1.5097	
Melbourne, FL	GR	0.1925	1.4797	
Sonora, Mexico	PR	0.3241 ( $\pm$ 0.1396)	1.5587 ( $\pm$ 0.0709)	This study

time rainfall downscaling schemes (e.g., Deidda et al. 2004).

The second-order moment scaling parameter  $\tau(2)$  measures the variability (in the second-order sense) of positive rain rate with scale within the rainy areas. Here,  $\tau(2) = 0$  implies the single rainy box case and  $\tau(2) = -2$  implies the uniform rain field case. The more negative  $\tau(2)$  becomes, the less intense the rain becomes at each smaller scale. In Fig. 16b, we present a scatterplot of  $-\tau(2)$  estimated from each TRMM-PR overpass versus the corresponding large-scale average rain rate  $\bar{R}$ . It is clear that  $\tau(2)$  depends on  $\bar{R}$  in the same functional form as  $\tau(0)$  depends on  $\bar{R}$ , though the parameter values differ. However, the statistical relationship between  $\tau(2)$  and  $\bar{R}$  is weaker than that of  $\tau(0)$  and  $\bar{R}$ . This suggests the need for exploring other large-scale variables that could explain the variability in  $\tau(2)$  that was left unexplained by  $\bar{R}$ .

## 6. Results obtained with GOES-IR observations

Geostationary infrared data are suitable for studying the diurnal variation of cold cloud occurrence, because of their high temporal sampling frequency. We performed a harmonic analysis on the hourly total number of cases with temperatures below three brightness temperature thresholds. The thresholds 235, 225, and 215 K correspond to low, middle, and high clouds, respectively. In Fig. 17, we show the results resampled to  $0.1^\circ$  resolution over a domain of  $2^\circ \times 2^\circ$ , with the inset showing our study region. For all thresholds, the diurnal variations of cold cloud occurrences show a strong diurnal cycle (normalized amplitude exceeding 0.7, and variance explained varying between 40% and 85%), with a maximum in the late afternoon and evening hours (1700–2300 LST), at any location across the domain. The time of maximum cold cloud occurrence does not vary with the threshold temperature used,

while the normalized amplitude and the variance explained by the diurnal cycle do vary. The normalized amplitude is higher for high clouds than for low clouds, indicating that deep convective cells have stronger diurnal cycles. As opposed to the other parameters, the time of maximum cold cloud occurrence parameter shows a clear regional coherency, with delayed peaks occurring in the southern portion of the domain (or the inset). This (along with the rainfall results; see Figs. 8 and 10) indicates that cloud and rainfall systems start from the northern region in the early evening and move toward the southern region in the late evening. This leads to the finding that the nighttime maximum in the southern region tends to be the result of deep, organized systems, while the late afternoon/early evening maximum in the northern region is related to relatively shallow, less organized convection. This implies a separate population of rain systems (with different life cycles of convective systems) in these regions.

Can the diurnal cycle of cold cloud occurrence be used as a surrogate for the diurnal cycle of rainfall? To address this question, we compared the diurnal cycles of cold cloud occurrence obtained with GOES-IR to the diurnal cycle of rainfall frequency obtained with rain gauges, at the station locations (Fig. 18). Regardless of the IR thresholds used in this study, the GOES-IR results agree with the gauge results in that (i) both show strong diurnality, with normalized amplitude ranging from 0.7 to 1.4 and variance explained by the diurnal cycle exceeding 40%; (ii) both show that the diurnal peak occurs in the evening between 1900 and 2200 LST; and (iii) the diurnal peak time for the cold cloud occurrence closely follows that for the rainfall: the Pearson correlation coefficient between the two is 0.7–0.8. Among the three cloud types, the low clouds have diurnal cycles that are more similar to those of the rainfall frequency.

However, there are also differences between the

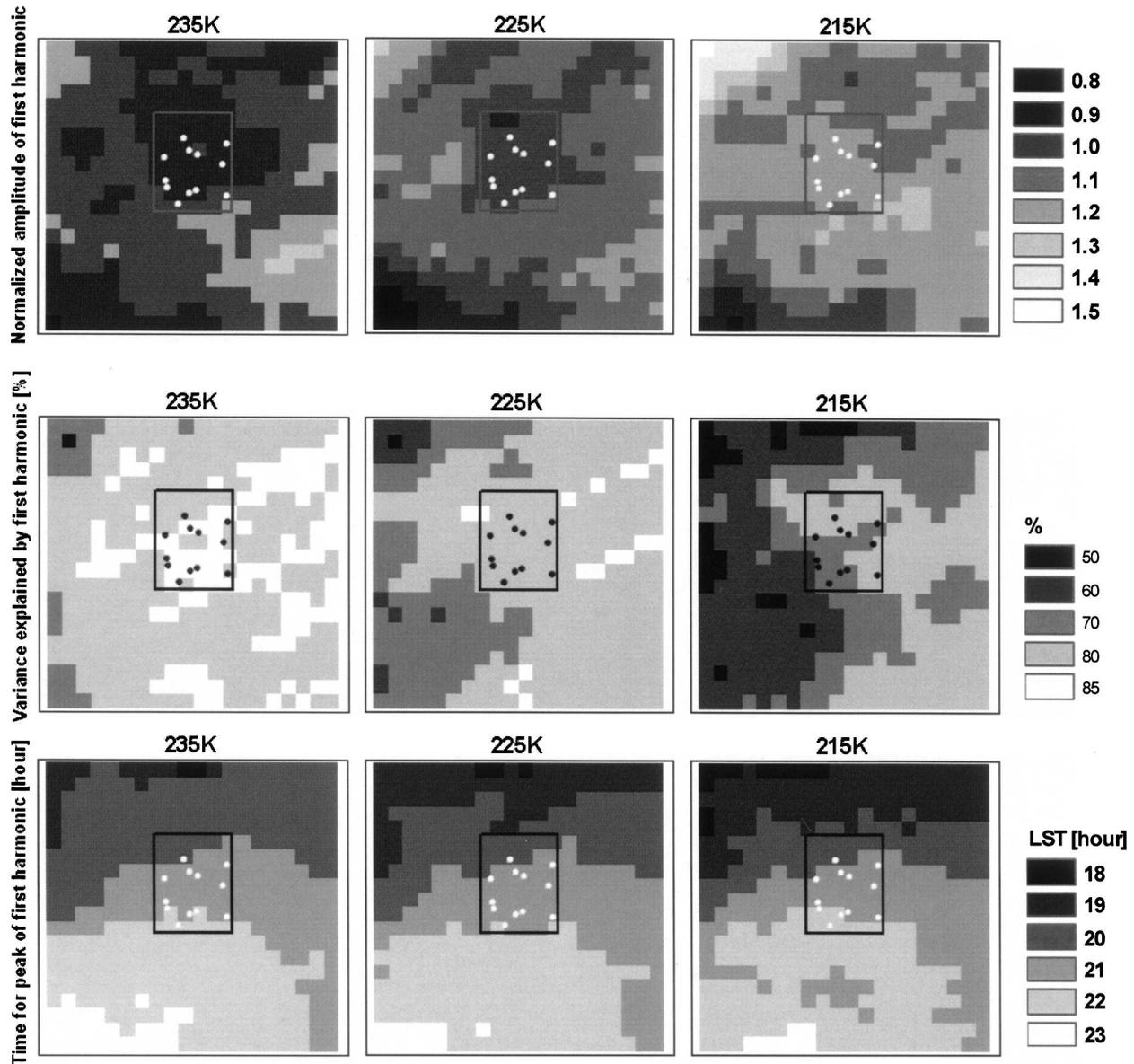


FIG. 17. First harmonic of the number of pixels colder than (left) 235, (middle) 225, and (right) 215 K for (top to bottom) normalized amplitude, variance explained, and time for peak. The legends show the values (and the associated bins) used in assigning colors.

cloud and rainfall diurnal variations. The diurnal cycle of cold cloud occurrence exhibits less spatial variability than the diurnal cycle of rainfall. This could be partly attributed to the coarse resolution of the GOES-IR that could smear small-scale fluctuations, and partly perhaps to the smoother distribution of clouds than rainfall. The amplitudes of low and middle clouds appear to be higher (lower) than that for the rainfall, when the rainfall amplitude is less than (greater than) one. This suggests that there are likely more anvils that may either contain light stratiform precipitation, or be non-precipitating dissipating stratiform rain. The ampli-

tudes of high clouds are higher than those for the rainfall, at almost all amplitudes, indicating that deep convective cells have stronger diurnal cycles. The variance explained by the diurnal cycle is much higher for the low clouds than for the rainfall, indicating that most of the low clouds that do not rain much occur around the diurnal peak time. The cloud diurnal cycle peaks late (early) when the rainfall diurnal cycle peaks early (late), with the maximum difference in peak times being 1 h. This systematic bias may be explained by the smoother distribution of the cloud diurnal peak times (see Fig. 17).

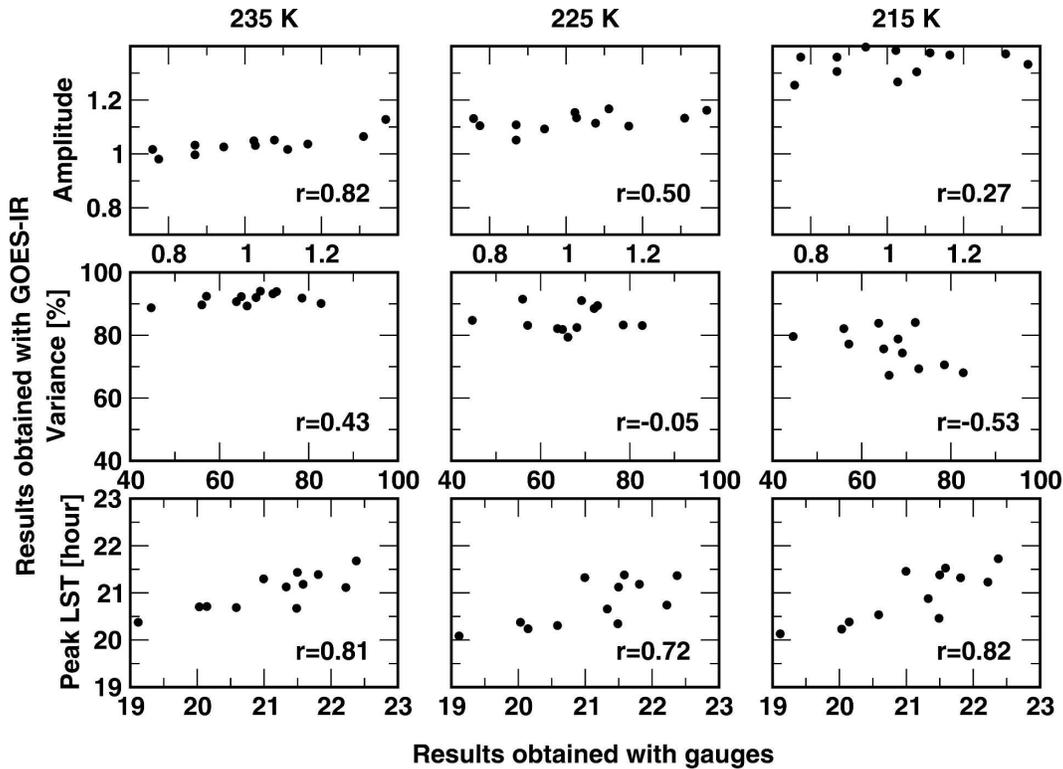


FIG. 18. Comparison of the first harmonic derived from rain gauges, rainfall frequency, and GOES-IR data. Threshold temperatures for the GOES-IR data are (left to right) 235, 225, and 215 K for (top to bottom) normalized amplitude, variance explained, and time for peak.

## 7. Summary and conclusions

We have explored in detail the submesoscale spatial pattern and temporal dynamics of rainfall in a 50 km  $\times$  75 km study area located in Sonora, Mexico, in the periphery of the North American monsoon system core region. We used data from rain gauges, GOES-IR, and TRMM-PR over a period spanning from 1 July to 31 August 2004, corresponding to one monsoon season. The time scales we considered are hourly for the gauge and IR data, and  $\sim$ 15 min (corresponding to a column-average snapshot) for the TRMM-PR data. The main findings of our study from the analysis of July–August 2004 rainfall and cloud data may be summarized as follows.

- 1) Rainfall exhibits high spatial and temporal variability in the region. The rainfall total and standard deviation vary by a factor of 2, whereas the maximum hourly rainfall varies by a factor of 4 over the region. The distance at which the Pearson correlation coefficient becomes 0.37 or insignificant is 17 km. For two stations located 5 km apart, half of the time rain at one station is not accompanied by rain at the other station within that hour. The standard deviations amount to 800%–1100% of the mean rain rate, indicating a high variability in the rainfall amounts from event to event.
- 2) Diurnal variations of cold cloud occurrence frequency, rainfall frequency, and rainfall intensity are dominated by the diurnal cycle, peaking in the evening hours. The amplitudes of the diurnal cycles of rainfall intensity and rainfall frequency amount to 110%–150% and 80%–140% of the daily mean, respectively. The corresponding figures for the low cloud and high cloud occurrence frequencies are 100%–113% and 126%–140%, respectively. The lag time between the rainfall intensity and rainfall frequency peak hours is within 2.4 h, with earlier peaks observed for rainfall intensity. The basic spatial variability features of the diurnal cycle of rainfall intensity are similar to those of rainfall frequency.
- 3) Deep convective cells have stronger diurnal cycles. The time of maximum cold cloud occurrence does not vary with the infrared threshold temperature used (215–235 K), while the normalized amplitude and the variance explained by the diurnal cycle do vary.
- 4) An evaluation of the diurnal cycle of cold cloud oc-

currence reveals that it agrees well with the diurnal cycle of rainfall frequency in terms of strong diurnality, large variance explained by the diurnal cycle, and evening maximum hours. The low clouds (with a threshold of 235 K) have diurnal cycles that are more similar to the rainfall frequency. The cloud diurnal cycle peaks late (early) when the rainfall frequency diurnal cycle peaks early (late), with the maximum difference in peak times being 1 h.

- 5) Topography plays an important role in the spatiotemporal variability of rainfall. As compared to valley areas, mountainous areas are characterized by an earlier diurnal peak, an earlier date of maximum precipitation, closely clustered rainy hours, frequent yet small rainfall events, and less dependence of precipitation accumulation on elevation.
- 6) The geographical location (south versus north) also plays an important role in the spatiotemporal variability. As compared to the northern section of the study area, the southern section is characterized by strong convective systems that peak late diurnally and have smaller rainfall totals.
- 7) The temporal rainfall fields are scale invariant at the moment order zero (rain–no rain) but with different fractal dimensions for different regimes. The two distinct scaling regimes include one extending from 2 to 16 h and another extending from 2.7 days onward. The spatial rainfall fields are scale invariant at moment orders ranging from zero to three. There is a one-to-one relationship between the scaling parameters and the large-scale spatial average rain rate. Multifractals models can therefore be used for estimation/simulation purposes in this region.

Using a variety of statistics, the above results have identified the key sources of monsoon submesoscale variations, and characterized the variability in the periphery of the North American monsoon core region. These results are important in improving the prediction of the North American monsoon rainfall in the region. The results have also provided evidence that the diurnal cycle of cold cloud occurrence can be used as a surrogate for the diurnal cycle of rainfall. The existence of scale-invariant properties in both the spatial and temporal rainfall fields indicates that the rainfall-producing mechanisms could be characterized by a multiplicative cascade process. It also implies that the outputs from meteorological models could be down-scaled to any scale needed for hydrological studies. However, information on the topographic features, large-scale features, and diurnal cycle need to be incorporated for accurate results. As pointed out by Gochis et al. (2006) and Vivoni et al. (2006), the most signifi-

cant source of uncertainty in the estimation of hydrologic responses and understanding of land–atmosphere interaction is the accuracy of rainfall data. Our results have therefore an implication for hydrologic responses to estimation accuracy.

Finally, we note that our analysis has been based on data from two summer months during 2004. Incorporation of data from additional seasons, through ongoing network measurements, will help build confidence in the climatology of the rainfall characteristics discussed herein.

*Acknowledgments.* We would like to acknowledge funding from the NOAA North American Monsoon Experiment (NAME) and Soil Moisture Experiment 2004 (SMEX04), as well as the NSF EPSCOR program. We are grateful to Tom Yoksas of Unidata User Support, UCAR/Unidata Program Center, for his relentless technical advice on the McIDAS software that was used for processing GOES-IR data. We thank the anonymous reviewers whose comments improved the quality of this paper.

#### REFERENCES

- Adams, D. K., and A. C. Comrie, 1997: The North American monsoon. *Bull. Amer. Meteor. Soc.*, **78**, 2197–2213.
- Arkin, P. A., and B. N. Meisner, 1987: The relationship between large-scale convective rainfall and cold cloud over the Western Hemisphere during 1982–84. *Mon. Wea. Rev.*, **115**, 51–74.
- Baden-Dagan, A., C. E. Dorman, M. A. Merrifield, and C. D. Winant, 1991: The lower atmosphere over the Gulf of California. *J. Geophys. Res.*, **96**, 16 877–16 896.
- Balling, R. C., Jr., and S. W. Brazel, 1987: Diurnal variations in Arizona monsoon precipitation frequencies. *Mon. Wea. Rev.*, **115**, 342–346.
- Barnsley, M. F., 1993: *Fractals Everywhere*. 2d ed. Academic Press, 394 pp.
- Bell, T. L., and N. Reid, 1993: Detecting the diurnal cycle of rainfall using satellite observations. *J. Appl. Meteor.*, **32**, 311–322.
- Berberly, E. H., 2001: Mesoscale moisture analysis of the North American monsoon. *J. Climate*, **14**, 121–137.
- Ciach, G. J., 2003: Local random errors in tipping-bucket rain gauge measurements. *J. Atmos. Oceanic Technol.*, **20**, 752–759.
- , and W. F. Krajewski, 1999: On the estimation of radar rainfall error variance. *Adv. Water Resour.*, **22**, 585–595.
- Coblentz, D. D., and K. H. Riitters, 2004: Topographic controls on the regional-scale biodiversity of the south-western USA. *J. Biogeogr.*, **31**, 1125–1138.
- Dai, A. D., 2001: Global precipitation and thunderstorm frequencies. Part II: Diurnal variations. *J. Climate*, **14**, 1112–1128.
- , F. Giorgi, and K. E. Trenberth, 1999: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States. *J. Geophys. Res.*, **104**, 6377–6402.
- Deidda, R., M. G. Badas, and E. Piga, 2004: Space–time scaling in high-intensity Tropical Ocean Global Atmospheric Coupled Ocean–Atmosphere Response Experiment (TOGA-COARE)

- storms. *Water Resour. Res.*, **40**, W02506, doi:10.1029/2003WR002574.
- de Lima, M. I. P., and J. Grasman, 1999: Multifractal analysis of 15-min and daily rainfall from a semi-arid region in Portugal. *J. Hydrol.*, **220**, 1–11.
- de Marsily, G., 1986: *Quantitative Hydrogeology: Groundwater Hydrology for Engineers*. Academic Press, 440 pp.
- Douglas, M. W., 1995: The summertime low-level jet over the Gulf of California. *Mon. Wea. Rev.*, **123**, 2334–2347.
- , and J. C. Leal, 2003: Summertime surges over the Gulf of California: Aspects of their climatology, mean structure, and evolution from radiosonde, NCEP reanalysis, and rainfall data. *Wea. Forecasting*, **18**, 55–74.
- , R. Maddox, K. Howard, and S. Reyes, 1993: The Mexican Monsoon. *J. Climate*, **6**, 1665–1677.
- Ellis, A. W., E. M. Saffell, and T. W. Hawkins, 2004: A method for defining monsoon onset and demise in the southwestern USA. *Int. J. Climatol.*, **24**, 247–265.
- Englehart, P. J., and A. V. Douglas, 2001: The role of eastern North Pacific tropical storms in the rainfall climatology of western Mexico. *Int. J. Climatol.*, **21**, 1357–1370.
- Fawcett, P. J., J. R. Stalker, and D. S. Gutzler, 2002: Multistage moisture transport into the interior of northern Mexico during the North American summer monsoon. *Geophys. Res. Lett.*, **29**, 2094, doi:10.1029/2002GL015693.
- Fuller, R. D., and D. J. Stensrud, 2000: The relationship between easterly waves and surges over the Gulf of California during the North American monsoon. *Mon. Wea. Rev.*, **128**, 2983–2989.
- Gebremichael, M., and W. F. Krajewski, 2004: Characterization of the temporal sampling error in space–time-averaged rainfall estimates from satellites. *J. Geophys. Res.*, **109**, D11110, doi:10.1029/2004JD004509.
- , —, M. Morrissey, D. Langerud, G. J. Huffman, and R. Adler, 2003: Error uncertainty analysis of GPCP monthly rainfall products: A data-based simulation study. *J. Appl. Meteor.*, **42**, 1837–1848.
- , —, —, G. Huffman, and R. Adler, 2005: A detailed evaluation of GPCP 1° daily rainfall estimates over the Mississippi River basin. *J. Appl. Meteor.*, **44**, 665–681.
- , T. M. Over, and W. F. Krajewski, 2006: Comparison of the scaling characteristics of rainfall derived from space-based and ground-based radar observations. *J. Hydrometeorol.*, **7**, 1277–1294.
- Gochis, D. J., J.-C. Leal, W. J. Shuttleworth, C. J. Watts, and J. Garatuza-Payan, 2003: Preliminary diagnostics from a new event-based precipitation monitoring system in support of the North American Monsoon Experiment. *J. Hydrometeorol.*, **4**, 974–981.
- , A. Jimenez, C. J. Watts, J. Garatuza-Payan, and W. J. Shuttleworth, 2004: Analysis of 2002 and 2003 warm-season precipitation from the North American Monsoon Experiment event rain gauge network. *Mon. Wea. Rev.*, **132**, 2938–2953.
- , L. Brito-Castillo, and W. J. Shuttleworth, 2006: Hydroclimatology of the North American monsoon region. *J. Hydrol.*, **316**, 53–70.
- Gutzler, D. S., and Coauthors, 2005: The North American Monsoon Model Assessment Project. *Bull. Amer. Meteor. Soc.*, **86**, 1423–1429.
- Habib, E., W. F. Krajewski, and A. Kruger, 2001: Sampling errors of tipping-bucket rain gauge measurement. *J. Hydrol. Eng.*, **6**, 159–166.
- Harris, D., E. Foufoula-Georgiou, and C. Kummerow, 2003: Effects of underrepresented hydrometeor variability and partial beam filling on microwave brightness temperatures for rainfall retrieval. *J. Geophys. Res.*, **108**, 8380, doi:10.1029/2001JD001144.
- Hentschel, H. G. R., and I. Procaccia, 1983: The infinite number of generalized dimensions of fractals and strange attractors. *Physica D*, **8**, 435–444.
- Higgins, R. W., and W. Shi, 2001: Intercomparison of the principal modes of interannual and intraseasonal variability of the North American monsoon system. *J. Climate*, **14**, 403–417.
- , and Coauthors, 2003: Progress in pan-American climate variability research: The North American monsoon system. *Atmosfera*, **16**, 29–65.
- Hitschfeld, W., and J. Bordan, 1954: Errors inherent in the radar wavelength measurement of rainfall at attenuating wavelengths. *J. Meteor.*, **11**, 58–67.
- Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at 1° daily resolution from multisatellite observations. *J. Hydrometeorol.*, **2**, 36–50.
- Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, and K. Okamoto, 2000: Rain-profiling algorithm for the TRMM precipitation radar. *J. Appl. Meteor.*, **39**, 2038–2052.
- Journel, A. G., and C. J. Huibregts, 1978: *Mining Geostatistics*. Academic Press, 600 pp.
- Kozu, T., and T. Iguchi, 1999: Nonuniform beamfilling correction for spaceborne radar rainfall measurement: Implication for TOGA COARE radar data analysis. *J. Atmos. Oceanic Technol.*, **16**, 1722–1735.
- Krajewski, W. F., 1987: Cokriging radar–rainfall and rain gage data. *J. Geophys. Res.*, **92**, 9571–9580.
- , G. J. Ciach, J. R. McCollum, and C. Bacotiu, 2000: Initial validation of the Global Precipitation Climatology Project over the United States. *J. Appl. Meteor.*, **39**, 1071–1086.
- Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.*, **15**, 809–817.
- Lammering, B., and I. Dwyer, 2000: Improvement of water balance in land surface schemes by random cascade disaggregation of rainfall. *Int. J. Climatol.*, **20**, 681–695.
- Li, J., X. Gao, R. A. Maddox, and S. Sorooshian, 2004: Model study of evolution and diurnal variations of rainfall in the North American monsoon during June and July 2002. *Mon. Wea. Rev.*, **132**, 191–211.
- Mandelbrot, B. B., 1982: *The Fractal Geometry of Nature*. W. H. Freeman, 480 pp.
- Meneghini, R., T. Iguchi, T. Kozu, L. Liao, K. Okamoto, J. A. Jones, and J. Kwiatkowski, 2000: Use of the surface reference technique for path attenuation estimates from the TRMM precipitation radar. *J. Appl. Meteor.*, **39**, 2053–2070.
- Morrissey, M. L., J. A. Maliekal, J. S. Greene, and J. Wang, 1995: The uncertainty in simple spatial averages using raingage networks. *Water Resour. Res.*, **31**, 2011–2017.
- Nesbitt, S. W., and E. J. Zipser, 2003: The diurnal cycle of rainfall and convective intensity according to three years of TRMM measurements. *J. Climate*, **16**, 1456–1475.
- , —, and D. J. Cecil, 2000: A census of precipitation features in the Tropics using TRMM: Radar, ice scattering, and lightning observations. *J. Climate*, **13**, 4087–4106.
- Nykanen, D. K., E. Foufoula-Georgiou, and W. M. Lapenta, 2001: Impact of small-scale rainfall variability on larger-scale spa-

- tial organization of land–atmosphere fluxes. *J. Hydrometeor.*, **2**, 105–121.
- Ogden, F. L., and P. Y. Julien, 1993: Runoff sensitivity to temporal and spatial rainfall variability at runoff plane and small basin scale. *Water Resour. Res.*, **29**, 2589–2597.
- , and —, 1994: Runoff model sensitivity to radar rainfall resolution. *J. Hydrol.*, **158**, 1–18.
- Over, T. M., 1995: Modeling space-time rainfall at the mesoscale using random cascades. Ph.D. thesis, University of Colorado, 238 pp.
- , and V. K. Gupta, 1994: Statistical analysis of mesoscale rainfall: Dependence of a random cascade generator on large-scale forcing. *J. Appl. Meteor.*, **33**, 1526–1542.
- , and —, 1996: A space–time theory of mesoscale rainfall using random cascades. *J. Geophys. Res.*, **101**, 26 319–26 331.
- Rodriguez-Iturbe, I., and J. M. Mejia, 1974: On the transformation of point rainfall to areal rainfall. *Water Resour. Res.*, **10**, 729–735.
- Troutman, B. M., and A. V. Vecchia, 1999: Estimation of Renyi exponents in random cascades. *Bernoulli*, **5**, 191–207.
- Tucker, D. F., 1993: Diurnal precipitation variations in south-central New Mexico. *Mon. Wea. Rev.*, **121**, 1979–1991.
- Vivoni, E. R., and Coauthors, 2007: Variation of hydrometeorological conditions along a topographic transect in northwestern México during the North American monsoon. *J. Climate*, **20**, 1792–1809.
- Winchell, M., H. V. Gupta, and S. Sorooshian, 1998: On the simulation of infiltration- and saturation-excess runoff using radar-based rainfall estimates: Effects of algorithm uncertainty and pixel aggregation. *Water Resour. Res.*, **34**, 2655–2670.