Temporal Downscaling and Statistical Analysis of Rainfall across a Topographic Transect in Northwest Mexico

GIUSEPPE MASCARO
School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, Arizona

ENRIQUE R. VIVONI
School of Sustainable Engineering and the Built Environment, and School of Earth and Space Exploration, Arizona State University, Tempe, Arizona

DAVID J. GOCHIS
National Center for Atmospheric Research,* Boulder, Colorado

CHRISTOPHER J. WATTS AND JULIO C. RODRIGUEZ
Universidad de Sonora, Hermosillo, Sonora, Mexico

(Manuscript received 23 October 2013, in final form 2 January 2014)

ABSTRACT

In this study a temporal statistical downscaling scheme of rainfall is calibrated using observations from 2007 to 2010 at eight sites located along a 14-km topographic transect of 784 m in elevation in northwest Mexico. For this purpose, the rainfall statistical properties over a wide range of temporal scales (3 months–1 min) for the summer (July–September) and winter (November–March) seasons are first analyzed. Rainfall accumulation is found not to be significantly correlated with elevation in either season, and a strong diurnal cycle is found to be present only in summer, peaking in the late afternoon. Winter rainfall events are highly correlated between individual stations across the transect even at short aggregation times (<30 min), and summer storms are more localized in space and time. Spectral and scale invariance analyses showed the presence of three (two) scaling regimes in summer (winter), which are associated with typical meteorological phenomena of the corresponding time scales (frontal systems and relatively isolated convective systems). These analyses formed the basis for calibrating a temporal downscaling model to disaggregate daily precipitation to hourly resolution in the summer season, based on scale invariance and multifractal theory. In this downscaling scheme, a modulation function was used to reproduce the time heterogeneity introduced by the diurnal cycle. The model showed adequate performances in reproducing the small-scale observed precipitation variability. Results of this work are useful for the interpretation of storm-generation mechanisms in the region, and for creating hourly rainfall time series from daily rainfall data, obtained from observations or simulated by climate, meteorological, or other statistical models.

1. Introduction

The characterization of the statistical properties of precipitation across a wide range of time scales (from monthly to subdaily) is crucial for diagnostic and prognostic applications in several areas, spanning hydrology (Georgakakos and Kavvas 1987; Veneziano and Langousis 2010), climatology (Gochis et al. 2006; Arriaga-Ramírez and Cavazos 2010; Gutierrez-Ruacho et al. 2010), meteorology (Seo et al. 2000; Langousis and Veneziano 2009a,b), and ecology (Watts et al. 2007; Yépez et al. 2007; Méndez-Barroso et al. 2009). This information is also useful for developing and calibrating statistical downscaling tools that reproduce the finescale rainfall properties, starting from coarse predictions of numerical weather prediction (NWP) models and satellite products.
In semiarid regions of northwest Mexico, investigating the rainfall statistical features is also useful to support water resources management and water infrastructure design (Gochis et al. 2006; Langousis and Veneziano 2007; Robles-Morua et al. 2013). This is motivated by the complex features of the rainfall regime in this region, characterized by (i) low annual rainfall depth [mean ranging from 350 to 700 mm, as documented in Hallack-Alegria and Watkins (2007)] with large interannual variability; (ii) strong seasonality, with the majority (50%–80%) of the annual rainfall falling during the North American monsoon (NAM) season (from July to September), mainly in the form of diurnally modulated convective storms (e.g., Douglas et al. 1993; Gochis et al. 2006; Nesbitt et al. 2008), and less frequent frontal storm systems in winter associated with activity of the subtropical and polar jet streams (Brito-Castillo et al. 2003); and (iii) high spatiotemporal variability, due to the presence of other transient meteorological factors and the influence of complex topography (Johnson et al. 2007; Gebremichael et al. 2007; Rowe et al. 2008).

During the last decade, several U.S. and Mexican institutions have carried out efforts to understand the hydroclimatology of northwest Mexico, especially during the NAM season. Field campaigns, including the North American Monsoon Experiment in 2004 (NAME; Higgins et al. 2006), have been conducted to install and maintain hydrometeorological stations (Gochis et al. 2003, 2004; Vivoni et al. 2007, 2010b), launch meteorological soundings (Johnson et al. 2007), and acquire remotely sensed images (Bindlish et al. 2008). The synergistic use of these data in combination with numerical models has contributed to advancing our knowledge of the regional hydroclimatology, including (i) understanding the physical mechanisms influencing weather systems and their linkage with large-scale meteorological variables (Johnson et al. 2007; Douglas and Englehart 2007; Nesbitt et al. 2008), (ii) characterizing the statistical behavior of rainfall and soil moisture in time and space (Gebremichael et al. 2007, Gochis et al. 2004, 2006; Vivoni et al. 2007, 2008a; Mascaro and Vivoni 2010, 2012), and (iii) investigating and quantifying the interactions between soil, vegetation and the atmosphere, and the hydrological response of basins in this area (Gochis et al. 2006; Watts et al. 2007; Vivoni et al. 2008b, 2010a).

In one of these campaigns, a network of eight tipping-bucket rain gauges was installed across a topographic transect spanning a distance of ~14 km and an elevation difference of 748 m in the Sierra Los Locos basin of Sonora, Mexico (Fig. 1). In this paper, we use the high-resolution precipitation records collected by these gauges during the period 2007–10 to develop a downscaling scheme to reproduce in a stochastic fashion the subdaily variability of the summer precipitation. For this aim, we preliminarily conduct a set of analyses to investigate the rainfall statistical properties in the region. In contrast to previous studies that only considered the NAM period and used sparse networks of gauges covering large areas (size on the order of a few hundreds of kilometers), we focus on both the summer and winter seasons and analyze high-resolution data observed at short separation distance (less than 2 km on average). Winter precipitation in this region has not been systematically analyzed across a range of scales, with the exception of Brito-Castillo et al. (2003), who focused on monthly variability.

The sets of analyses on the rainfall time series, carried out separately for summer and winter events, are aimed at (i) studying the effect of elevation on rainfall accumulation, (ii) investigating the diurnal cycle, (iii) characterizing the correlation structures of the records sampled at several aggregation times to capture different types of weather systems occurring in the two seasons, and (iv) investigating scaling properties through spectral and scale invariance analysis across a wide range of time scales (from ~3 months to 1 min). The information acquired through these analyses, particularly the evidence of scaling regimes and the quantification of the diurnal cycle, is then used to calibrate an existing statistical multifractal model that is able to downscale daily rainfall in the summer season down to an hourly resolution. We hypothesize that the calibrated model will be applicable over the mountainous regions in northern Sonora, due to the relatively similar statistical properties of the rainfall time series in this area. Given the importance of the summer precipitation in this region and the lack of a permanent finescale, high-frequency observing system such as weather radar, this disaggregation tool is useful in practice for representing the subdaily rainfall variability required to improve hydrological, meteorological, and climatological simulations. This paper is organized as follows. In section 2, we describe the study region where the topographic transect is located and the rainfall dataset. The analyses on the time series are presented in section 3, while the calibration and validation of the downscaling scheme is illustrated in section 4. Conclusions are outlined in section 5.

2. Study area and dataset

Figure 1 shows the study region where the eight rain gauge stations were installed. The site is located in the state of Sonora in northwest Mexico. The rain gauges
are included within the basin of the Río San Miguel (3796 km$^2$), a major ephemeral river system that is part of the Sierra Madre Occidental, as represented in the digital elevation model (DEM) derived from INEGI (1998). Overall, the topography of the region is rather complex, with a high-elevation range, primarily due to the effects of channel incision (Coblentz and Riitters 2004). The location of the stations has been selected to cover a topographic transect within the Sierra Los Locos basin (92.5 km$^2$; Fig. 1c). The transect spans a distance of $\sim$14 km and an elevation difference of 748 m, thus including valley, midelevation, and ridge sites where rainfall-generation mechanisms in the summer and winter seasons are potentially affected by orography.

The identification (ID) codes, the coordinates in the universal transverse Mercator (UTM) system (x and y), and the elevations (z) of the stations are shown in Table 1, along with the dominant vegetation types derived from the classification of SIUE-IMADES (1998). These encompass three different, elevation-dependent ecosystems, including (from lower to higher elevations) desert scrubland, subtropical scrubland, and oak savanna. The type of instrument consists of a tipping-bucket (TB) gauge with a bucket diameter of 203.2 mm (8 in.) and resolution of 0.254 mm (0.01 in.), automatically recording the tipping instants in digital memory (TB4; Hydrological Services, Warwick Farm, Australia). For the analyses carried out here, the rainfall signal ($i_D$; mm h$^{-1}$)
at different time resolutions ($\Delta t$) was derived from the tipping instants following the procedure described in appendix A of Mascaro et al. (2013).

The dataset spans a period from June 2007 to November 2011 with different missing records at each station. Figure 2a reports the percentage of missing data of the signal sampled at $\Delta t = 1$-min resolution for each month and gauge. The records are almost complete from July 2007 to July 2010, apart for some intervals with missing data for gauges 2, 3, 6, and 7. Figure 2b displays the time series of the monthly rainfall depths, plotted as mean ± standard deviation (STD) of the observations measured at the stations that had complete observations for that month. It is apparent how, for each year, the precipitation observed during the summer months of July–September accounts for the majority of the annual total. Despite the relatively short period of observation, it is also possible to note the presence of a significant interannual variability, as well as a higher variability within the transect (higher STD) during the summer.

The interannual variability observed during this period is typical of the precipitation climatology of the region and was not due to unusual or extreme meteorological conditions.

### 3. Analyses of the rainfall time series

As documented in Brito-Castillo et al. (2003), precipitation mechanisms during the summer and winter seasons in northwest Mexico are different. As a result, the following analyses on the time series were performed separately for the summer period, including the months from June to September when the NAM occurs, and the winter period, defined as being from November to March. The months not included in these two seasons are usually extremely dry and they may be affected by different types of weather systems that cannot be classified as typically belonging to summer or winter. These analyses are informative for the calibration of the temporal downscaling scheme described in section 4.

#### a. Effect of elevation on rainfall accumulation

A number of studies have documented the effect of elevation on the rainfall in northwest Mexico during the summer monsoon. Gochis et al. (2004) analyzed the rainfall characteristics using data recorded during the summers of 2002 and 2003 by 81 gauges belonging to the NAME Event Rain Gauge Network (NERN), deployed over a large region of $\sim 400$ km × $\sim 800$ km in northwest Mexico. The stations were grouped into six topographic transects, each aligned along a southwest–northeast axis with a total separation distance of more than 300 km. The authors found that, in all the transects, the total rainfall depth during the NAM is higher at lower elevation.
close to the Pacific coast and decreases toward the ridges of the Sierra Madre Occidental. Gebremichael et al. (2007) carried out a study based on data from a network of 12 rain gauges installed over a smaller study region of 50 km × 75 km, finding instead that, in the area of the Rio San Miguel, the rainfall accumulation during the NAM season of 2004 decreased with decreasing elevation.

In this work, we investigate the effect of orography by focusing on a much shorter distance (~14 km) using ~3 yr of observations in both the summer and winter seasons. Figure 3 reports the relation between elevation and the rainfall depth observed in summer 2007, winter 2007–08, and the sum of the two seasons (W + S). The linear regression and the coefficient of determination $R^2$ are also plotted for each season.

The diurnal cycle was analyzed separately for summer and winter, according to three different methods applied to the rainfall signals sampled at Δτ = 1-h resolution. In the first and second methods, we computed the average over all available days of rainfall rate and frequency of occurrence for each hour, respectively. In the third approach, we calculated a modulation function defined as the ratio between the rainfall rate observed at a given hour and the mean rainfall rate of the corresponding day. The modulation functions obtained for each hour were then averaged across all days. To compare the three metrics, we standardized them; that is, we scaled their values, so that they have unitary mean across the 24 h. The first two approaches have already been adopted to study the diurnal cycle (e.g., Gochis et al. 2004; Gebremichael et al. 2007; Nesbitt et al. 2008), while the third method was introduced to deal with the problem of time heterogeneity in the downscaling model described in section 4.

Figure 4 shows the diurnal cycle analyzed through the three metrics for summer and winter. The metrics were averaged over all gauges, as we did not find appreciable differences across the stations. This result differs somewhat from those of Gochis et al. (2004, 2006) and Nesbitt et al. (2008), who found that the diurnal cycle of precipitation frequency and intensity was more strongly correlated with the elevation, although they focused on larger terrain gradients at locations farther south in western Mexico. In general, the diurnal cycle is apparent in the summer season with values of the metrics greater than 1 in the late evening, peaking around 2000 local time (LT). The rainfall rate is more variable and has a second major peak around 2200 LT, while the rain frequency and the modulation function have a smoother pattern of behavior. Overall, results for the NAM season confirm previous findings in the region (Gebremichael et al. 2007). In winter, we did not find the presence of a diurnal cycle, as revealed by the rain frequency and the modulation functions that oscillates around 1. The rainfall rate seems to indicate the presence of a diurnal cycle (higher values from 0000 to 0800 LT and lower values from 0900 to 1400 LT), but this metric is less effective in this specific case as it is highly sensitive to the characteristics of the few storms used to compute it.

Results of this analysis provide additional evidence that the physical features of the rainfall systems occurring

**Figure 3.** Relation between the rainfall depth observed during the summer (S: July–September 2007), winter (W: November 2007–March 2008), and the sum of the two seasons (W + S). The linear regression and the coefficient of determination $R^2$ are also plotted for each season.
in the two seasons are significantly different. Summer is dominated by short convective storms concentrated in the late evening, while more organized frontal precipitation systems occur in winter, without a preferential time. In our interpretation here we note that while frontal systems may indeed have embedded convective elements within them, they are characterized by much more widespread precipitation and that the precipitation-generation mechanisms are significantly modulated by large-scale dynamical lifting as opposed to more localized thermal convection. Similarly, we recognize that summer storm events may also possess nonconvective features (e.g., trailing stratiform formations) but we identify summer events by the dominant mechanism for storm initiation and organization, which is convection. Under these definitions we broadly use the terms convective and frontal for these two different precipitation regimes. In addition, previous studies analyzing the rainfall variability across a larger separation distance, from the coast up to the Sierra Madre Occidental, demonstrated that the diurnal cycle in summer is affected by large-scale topography (e.g., Gochis et al. 2006; Nesbitt et al. 2008). However, because of the relatively short separation distance of the transect considered here (14 km) and the relatively shallow elevation change (784 m), the effect of orography on summer diurnal rainfall is limited and harder to capture. The absence of a diurnal trend in winter is instead likely an indication that the local surface heating of terrain has a minimal influence in this season. These considerations will be further supported by the analyses described in the next sections.

c. Correlation structure

In this section, we use the correlation structures of the rainfall signals to investigate the spatial variability of storms and quantify the geographic area where the statistics computed for a given station can be considered representative. The correlation structure was analyzed by (i) computing the Pearson correlation coefficient (CC) between pairs of stations, (ii) plotting CC as a function of the separation distance, and (iii) fitting the analytical model $\rho(d) = \exp[-(d/m)^n]$, where $\rho(d)$ is the CC at separation distance $d$, and $m$ and $n$ are parameters. This was done for the rainfall signal aggregated at different resolutions $\Delta t$ ranging from 5 min to 24 h.

Figures 5a,b show results for $\Delta t = 30$ min and $\Delta t = 3$ h for both seasons, respectively. It is apparent that the correlation is significantly higher in winter, with CC always greater than 0.55 (0.88) for $\Delta t = 30$ min (3 h). This illustrates that frontal systems typical of this season occupy an area larger than the maximum separation distance of the transect. Parameters $m$ and $n$ of the analytical model are reported in Table 2 for all $\Delta t$. In summer, the CC reduces quickly when the rainfall signal is sampled at high resolution ($\Delta t = 30$ min; Fig. 5a), as a result of the high spatial variability and intermittency of the isolated convective rain cells. When the signal is sampled at lower resolution ($\Delta t = 3$ h; Fig. 5b), the correlation is significant across the entire separation distance (~14 km). This indicates that, within this time scale, rain cells have moved or have been created and dissipated in such a way as to cover the entire transect.

To summarize the results for all time resolutions, Fig. 5c reports the correlation distance (the distance where $\rho(d) = e^{-1}$, thus becoming insignificant) returned by the analytical model as a function of $\Delta t$. These coincide with the values of the $m$ parameter reported in Table 2. Gebremichael et al. (2007) performed a similar analysis using hourly data of summer 2004 with stations located at separation distances from 10 to 50 km, whereas we consider distances from 0.6 to 14 km. Those authors found a correlation distance of 17.0 km, which is very close to the value of 11.3 km found here (Table 2). These results, based on independent datasets and different ranges of separation distance, confirm that the statistics derived from rainfall signals sampled at hourly resolution are valid over an area whose size is comparable to the length of the transect.
d. Spectral analysis

As a next step, we investigated the scaling properties of the rainfall time series using spectral analysis. To compute the spectra, we used the standard fast Fourier transform on the rainfall intensity signal in the range of scales from 2 to 2^{16} min (~3 months) for the summer season, and from 2 to 2^{16} min (~45.5 days) for the winter season. Based on data availability at each station, we were able to extract from three to five summer seasons and from five to eight nonoverlapping winter events per gauge. A single spectrum was produced for each gauge and season by averaging, for each frequency \( f \), the energy \( E(f) \) of all available rainfall sequences. The presence of scaling regimes was investigated by inspecting the presence of the power law with exponent \( \alpha \):

\[
E(f) \sim f^{-\alpha}.
\]

In Eq. (1), \( \alpha \) was estimated as the slope of the linear relation in log–log space.

Figure 6 shows the spectra obtained for all gauges during the two seasons. In summer, the spectra of all stations are almost flat from the scale of ~3 months to ~3 days, while two scaling regimes can be detected (i) from ~3 days to 2.1 h and (ii) from 2.1 h to 2 min. This behavior is consistent with previous studies in different climates in terms of both time ranges and numerical values of \( \alpha \) (Molini et al. 2009; Verrier et al. 2011; Mascaro et al. 2013). The scaling regimes are a signature of the meteorological phenomena typical of the corresponding time resolution (Friedrich and Larnder 1993). The time range from ~3 months to ~3 days does not show evidence of a clear power law and can be considered to be a transition zone between the intraseasonal variability at larger time scales and the following regime at smaller time scales. The range from ~3 days to 2.1 h is representative of systems that occupy the area for a few days. These are typically not organized as classic frontal systems but instead are a broad group of transient disturbances, which propagate through the monsoon region such as inverted troughs, tropical moisture surges, or decaying tropical storms. This interpretation is also corroborated by the values of the spectral exponent \( \alpha \) in Table 3 (Purdy et al. 2001). In fact, lower values of \( \alpha \) found in this scaling regime are typical of larger systems where the energy of the spectrum tends to be similar across all frequencies, thus decreasing at a slower rate. The regime between 2.1 h and 2 min is instead associated with single convective systems with a shorter lifetime. Accordingly, the decay of the spectra is faster (\( \alpha \) is higher) in this temporal range due to the presence of single cells localized in time.

In winter, the flat portion of the spectra extends from ~45.5 days to 17.1 h, and a single regime with evidence of a power law emerges from 17.1 h to 2 min. This indicates the presence of a single type of weather system with a characteristic scale close to the daily resolution and without the visible signature of shorter-term convective cells. Interestingly, the spectral exponent \( \alpha \) has an average

\[
\begin{array}{cccccc}
\text{\( \Delta t \)} & \text{Summer} & \text{\( m \) (km)} & \text{\( n \)} & \text{Winter} & \text{\( m \) (km)} & \text{\( n \)} \\
5 \text{~min} & 0.6 & 0.21 & 6.8 & 1 \\
10 \text{~min} & 1.2 & 0.23 & 9.9 & 1 \\
30 \text{~min} & 4.2 & 0.28 & 22.3 & 1 \\
1 \text{~h} & 11.3 & 0.34 & 46.1 & 1 \\
3 \text{~h} & 20.6 & 0.71 & 107.7 & 1 \\
12 \text{~h} & 51.0 & 0.57 & 209.9 & 1 \\
24 \text{~h} & 52.0 & 0.58 & 194.0 & 1 \\
\end{array}
\]
value of 0.87 (Table 3), intermediate between those found in the two summer scaling regimes (mean of 0.35 and 1.47). This may suggest that winter events are characterized by a mixed contribution of stratiform and convective systems, which are typical of frontal systems. However, further analyses are required considering the small number of rainfall events observed per winter and the relatively short length of our records. Finally, we highlight that, in both seasons, no significant relation is apparent between the spectral exponents and the gauge elevation (Table 3), implying that this type of analysis does not support evidence of different rainfall properties within the topographic transect studied here.

e. Scale invariance analysis

The scale invariance analysis was conducted from the finescale $\tau_0 = 1$ min to the coarse scale $T = 2^{17}$ min ($\sim$3 months) in summer and $T = 2^{16}$ min ($\sim$45.5 days) in winter, using the same rainfall events identified for the spectral analysis. Following the approach of Deidda et al. (1999), we computed the structure function $S_q(\tau)$, defined as

$$S_q(\tau) = \frac{1}{N(\tau)} \sum_{k=1}^{N(\tau)} (i_{\tau,k})^q,$$

where $i_{\tau,k}$ is the rainfall signal aggregated at resolution $\tau = \tau_0 \times 2^j$ ($j = 0, \ldots, M$; with $M = 17$ in summer and $M = 16$ in winter) in the $k$th time step. $N(\tau)$ is the number of nonoverlapping steps $\tau$ included in $T$, and $q$ are the moments. Note that, in this analysis, the scales $\tau$ are linked through a binary disaggregation approach (i.e., $\tau = \tau_0 \times 2^j$). Thus, $S_q(\tau)$ is computed for the rainfall

![Figure 6](image-url)
TABLE 3. Values of the spectral exponent $\alpha$ in the scaling regimes identified through the spectral analysis at each gauge during the summer and winter seasons (see Fig. 6). The mean and the STD are also reported.

<table>
<thead>
<tr>
<th>Gauge</th>
<th>$z$ (m MSL)</th>
<th>Summer</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 min–2.1 h</td>
<td>2.1 h–3 days</td>
<td>2 min–17.1 h</td>
</tr>
<tr>
<td>1</td>
<td>663</td>
<td>1.46</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>720</td>
<td>1.45</td>
<td>0.44</td>
</tr>
<tr>
<td>3</td>
<td>806</td>
<td>1.68</td>
<td>0.26</td>
</tr>
<tr>
<td>4</td>
<td>906</td>
<td>1.54</td>
<td>0.35</td>
</tr>
<tr>
<td>5</td>
<td>1014</td>
<td>1.47</td>
<td>0.33</td>
</tr>
<tr>
<td>6</td>
<td>1160</td>
<td>1.18</td>
<td>0.54</td>
</tr>
<tr>
<td>7</td>
<td>1357</td>
<td>1.50</td>
<td>0.29</td>
</tr>
<tr>
<td>8</td>
<td>1411</td>
<td>1.45</td>
<td>0.40</td>
</tr>
<tr>
<td>Mean</td>
<td>—</td>
<td>1.47 (0.14)</td>
<td>0.36 (0.10)</td>
</tr>
</tbody>
</table>

A signal resampled at each scale $\tau$ and for different values of $q$, and the presence of scale invariance is investigated by inspecting the existence of the power law:

$$S_q(\tau) \sim \tau^{-K(q)},$$

where $K(q)$ are the multifractal exponents, which are estimated through linear regression in log–log space, as in the case of the spectral exponent $\alpha$.

Figure 7 shows results for all gauges and the moment $q = 3$ (results are identical for other $q$ values). In summer, three scaling regimes can be identified: (i) from $\sim 3$ months to $\sim 11.4$ days, (ii) from $\sim 11.4$ days to 32 min, and (iii) from 32 to 1 min. As found in other studies (Verrier et al. 2011; Mascaro et al. 2013), this type of analysis reveals the presence of different scaling regimes as compared to spectral analysis, likely due to different approaches adopted to sample the signal. However, the physical interpretation of the three ranges is similar, as they can be associated with (i) a transition zone accounting for the intraseasonal variability at the largest scales, (ii) frontal systems covering the region for multiple days at intermediate synoptic time scales, and (iii) convective storms at the highest time resolutions.

Table 4 reports the values of the multifractal exponents $K(3)$, whose analysis suggests the following considerations. First, $K(3)$ is lower in the scaling regime from 32 to 1 min (mean of 0.54) as compared with the time range from $\sim 11.4$ days to 32 min, indicating that the statistical properties of the rainfall signals of single convective cells have lower variability across the corresponding scales [i.e., the values of $S_q(\tau)$ tend to decrease at a lower rate for different $\tau$]. In contrast, in the regime from $\sim 11.4$ days to 32 min, the multifractal exponent is larger (mean of 1.52) and the rainfall signal is more intermittent, due to the presence of uneven peaks (high rainfall intensities) when it is sampled at smaller resolutions. As a result, the decrease of $S_q(\tau)$ with $\tau$ occurs at a faster rate. Overall, the scale invariance analysis provides additional insights as compared to the spectral analysis on the structure of the rainfall signals and the connection with the physical phenomena, as it better identifies the characteristic time scale of the convective cells and their role within larger systems.

An additional consideration is related to the effect of elevation on the intermittency properties of the rainfall records. The values reported in Table 4 and in Fig. 8a show that $K(3)$ in the range from $\sim 11.4$ days to 32 min tends to decrease with increasing elevation, indicating higher storm intermittency at lower heights. This result can be confirmed and better understood through a simple analysis. For each gauge, we sampled the rainfall signals at the characteristic scale of $\Delta t = 32$ min. Next, we isolated the consecutive time steps with nonzero rainfall and determined the event duration and intensity. Only the periods with concomitant data at all eight stations were used. Figures 8b,c present the relation between the mean duration and intensity versus elevation. We note that rainfall events at lower elevation tend to have shorter duration and greater intensity, resulting in higher intermittency. These results are in accordance with the findings of Gochis et al. (2004, 2007) and Rowe et al. (2008), who analyzed gauge and radar precipitation data spanning a wider range of elevations (0–3000 m). No significant trend with elevation emerges in the scaling regimes of convective storms (32–2 min), implying that orography seems not to significantly affect the statistical properties of the single convective cells within the transect. A similar result was obtained in the statistical analysis of Gochis et al. (2006), who found that there was no elevational relationship on short-term (e.g., 10 min) rain rates.

Results of the scale invariance analysis for the winter season (Fig. 7b) are in accordance with the spectral analysis, with the presence of the same two scaling regimes. The values of the multifractal exponents $K(3)$, reported for the scaling regime from 17.1 h to 1 min, do not have a significant correlation with elevation. Overall, the spectral and scale invariance analyses on the winter rainfall records indicate that these events are more organized in space and time and are less likely affected by local terrain characteristics, confirming what has been also inferred from the analysis of the diurnal cycle and the correlation structure. This result seems to contradict the elevational relationship that has been demonstrated in other mountain areas for wintertime rainfall. However, the absence of a significant link between precipitation and elevation found here may be due to the isolated nature of the mountain ranges in
northern Sonora, which are not sufficient to induce a stable orographic precipitation process.

4. Statistical downscaling of subdaily rainfall variability in summer

a. Model overview

In the previous section, we have demonstrated the presence of a scaling regime in summer from $\sim 11.4$ days to $32$ min. This result supports the use of a calibrated downscaling model to simulate the subdaily statistical variability of precipitation, required in many hydrological applications, starting from daily observations, collected by the Comisión Nacional de Agua, the federal agency in Mexico charged with hydrometeorological observations. For this purpose, we used the Space–Time Rainfall (STRAIN) downscaling model (Deidda et al. 1999; Deidda 2000), which is able to reproduce the scale-invariant and multifractal properties of the rainfall signal through binary multiplicative cascades obtained by means of a log-Poisson stochastic generator. The model provides a theoretical expectation for the multifractal exponents $K(q)$,

$$K(q) = c \frac{q(1 - \beta) - (1 - \beta^q)}{\ln 2},$$

as a function of two parameters, $c$ and $\beta$. These are estimated by fitting Eq. (4) to the observed $K(q)$ computed through the scale invariance and multifractal analysis. Once $c$ and $\beta$ are estimated for a set of rainfall events, and empirical calibration relations are identified between the parameters and one or more coarse predictors, then the model can be utilized operationally. Starting from the coarse predictors and the coarse rainfall observation, the model parameters are derived from the calibration relations and used to generate an
ensemble of equally probable disaggregated rainfall fields that should capture the finescale observed statistical distribution. In previous applications, the model was calibrated to reproduce the rainfall variability in time (Deidda et al. 1999), space (Deidda 2000), and space and time (Deidda et al. 2004, 2006; Badas et al. 2006). The steps required to calibrate the model are described next.

b. Selection of rainfall events

Model application requires first identifying the coarse and fine scales where the statistical downscaling is performed. Since it is of practical interest to disaggregate rainfall data at daily resolution, we adopted the coarse-scale $T = 24 \text{ h}$. Because of the binary downscaling approach, we selected the finescale $\tau_0 = T \times 2^{-M} = 0.75 \text{ h} = 45 \text{ min}$, with $M = 5$ downscaling levels. Note that the coarse and the fine scales are included in the scaling regime $\sim 11.4 \text{ days} - 32 \text{ min}$. If needed, the rainfall signal disaggregated at 45 min can be resampled at hourly resolution, which is often used in several hydrological applications. As a next step, for each gauge, we selected the rainfall events used to calibrate the model, by extracting the daily observations with mean $I > 0.1 \text{ mm h}^{-1}$ (total depth of 2.4 mm) and then resampling the signal at resolution of $\tau_0 = 45 \text{ min}$. The total number of selected events is included between 34 (gauge 6) and 50 (gauge 1).

c. Introduction of time heterogeneity due to the diurnal cycle

Since a diurnal cycle is present in the summer, the mean rainfall frequency and intensity can be higher or lower over certain times of the day. The STRAIN model is not able to deal with heterogeneity in time, as it only reproduces homogeneous fields where the expected value of the rainfall intensity is the same at all times. To introduce a heterogeneous component, we adopted an approach similar to Badas et al. (2006), who dealt with spatial heterogeneity in the rainfall distribution due to orography. For this aim, for each day $j$, we calculated the modulation function $\xi_{k,j}$ defined as

$$\xi_{k,j} = i_{\tau_0,k,j}/I_j,$$

where $i_{\tau_0,k,j}$ is the rainfall intensity at resolution $\tau_0 = 45 \text{ min}$ in the $k$th time of the day $j$ where $k = 1, \ldots, 32$ and $I_j$ is the mean rainfall intensity in the same day. The mean modulation function $\bar{\xi}_k$ for each time $k$ was then computed by averaging $\xi_{k,j}$ over all days and rescaling the values to obtain a unitary mean across the day. This led to a pattern of behavior similar to the hourly trend in

<table>
<thead>
<tr>
<th>Gauge</th>
<th>$z$ (m MSL)</th>
<th>Summer</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>663</td>
<td>0.36</td>
<td>1.62</td>
</tr>
<tr>
<td>2</td>
<td>720</td>
<td>0.54</td>
<td>1.54</td>
</tr>
<tr>
<td>3</td>
<td>806</td>
<td>0.44</td>
<td>1.53</td>
</tr>
<tr>
<td>4</td>
<td>906</td>
<td>0.62</td>
<td>1.51</td>
</tr>
<tr>
<td>5</td>
<td>1014</td>
<td>0.57</td>
<td>1.48</td>
</tr>
<tr>
<td>6</td>
<td>1160</td>
<td>0.70</td>
<td>1.45</td>
</tr>
<tr>
<td>7</td>
<td>1357</td>
<td>0.56</td>
<td>1.52</td>
</tr>
<tr>
<td>8</td>
<td>1411</td>
<td>0.52</td>
<td>1.51</td>
</tr>
<tr>
<td>Mean (STD)</td>
<td>0.54 (0.10)</td>
<td>1.52 (0.05)</td>
<td>1.02 (0.19)</td>
</tr>
</tbody>
</table>
The values of the modulation function $\xi_k$ for each time step $k = 1, \ldots, 32$. The range of values (min) of each time interval is also reported.

<table>
<thead>
<tr>
<th>$k$</th>
<th>Time interval</th>
<th>$\xi_k$</th>
<th>$k$</th>
<th>Time interval</th>
<th>$\xi_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0–45</td>
<td>0.98</td>
<td>17</td>
<td>720–765</td>
<td>0.52</td>
</tr>
<tr>
<td>2</td>
<td>45–90</td>
<td>0.98</td>
<td>18</td>
<td>765–810</td>
<td>0.62</td>
</tr>
<tr>
<td>3</td>
<td>90–135</td>
<td>0.34</td>
<td>19</td>
<td>810–855</td>
<td>0.82</td>
</tr>
<tr>
<td>4</td>
<td>135–180</td>
<td>0.27</td>
<td>20</td>
<td>855–900</td>
<td>0.90</td>
</tr>
<tr>
<td>5</td>
<td>180–225</td>
<td>0.98</td>
<td>21</td>
<td>900–945</td>
<td>1.06</td>
</tr>
<tr>
<td>6</td>
<td>225–270</td>
<td>0.78</td>
<td>22</td>
<td>945–990</td>
<td>0.98</td>
</tr>
<tr>
<td>7</td>
<td>270–315</td>
<td>0.70</td>
<td>23</td>
<td>990–1035</td>
<td>0.52</td>
</tr>
<tr>
<td>8</td>
<td>315–360</td>
<td>0.44</td>
<td>24</td>
<td>1035–1080</td>
<td>1.15</td>
</tr>
<tr>
<td>9</td>
<td>360–405</td>
<td>0.33</td>
<td>25</td>
<td>1080–1125</td>
<td>1.23</td>
</tr>
<tr>
<td>10</td>
<td>405–450</td>
<td>0.46</td>
<td>26</td>
<td>1125–1170</td>
<td>2.64</td>
</tr>
<tr>
<td>11</td>
<td>450–495</td>
<td>0.53</td>
<td>27</td>
<td>1170–1215</td>
<td>2.86</td>
</tr>
<tr>
<td>12</td>
<td>495–540</td>
<td>0.44</td>
<td>28</td>
<td>1215–1260</td>
<td>1.93</td>
</tr>
<tr>
<td>13</td>
<td>540–585</td>
<td>0.42</td>
<td>29</td>
<td>1260–1305</td>
<td>2.08</td>
</tr>
<tr>
<td>14</td>
<td>585–630</td>
<td>0.46</td>
<td>30</td>
<td>1305–1350</td>
<td>3.22</td>
</tr>
<tr>
<td>15</td>
<td>630–675</td>
<td>0.41</td>
<td>31</td>
<td>1350–1395</td>
<td>1.65</td>
</tr>
<tr>
<td>16</td>
<td>675–720</td>
<td>0.60</td>
<td>32</td>
<td>1395–1440</td>
<td>0.69</td>
</tr>
</tbody>
</table>

Fig. 9. Time series of mean modulation function $\xi_k$, where $k = 1, \ldots, 32$. The range of values (min) of each time interval is also reported. The values of $\xi_k$ are also listed in Table 5.

For reference, the values of $\xi_k$ are shown in Fig. 9 and reported in Table 5. The modulation function can be used to reproduce the heterogeneous rainfall field $i_{t,k}$ from a homogeneous field $\hat{i}_{t,k}^{\text{hom}}$ as

$$i_{t,k} = \xi_k \times \hat{i}_{t,k}^{\text{hom}}, \quad (6)$$

or to homogenize in time a heterogeneous rainfall series as

$$\hat{i}_{t,k}^{\text{hom}} = i_{t,k} / \xi_k. \quad (7)$$

### d. Model calibration

The selected daily rainfall events discretized at 45-min resolution were homogenized in time by using Eq. (7). Next, the scale invariance and multifractal analysis was performed on the homogeneous fields from the fine ($\tau_0$) to the coarse ($T$) scales, to estimate the parameters $c$ and $\beta$. As an example, Fig. 10 shows the outcomes of this analysis for two events observed at gauges 1 and 8. The scale invariance analysis was carried out for $q = 1.5, 2, 2.5,$ and $3$ (Figs. 10a,c), leading to estimates of the multifractal exponents $K(q)$ as slope of the regression lines. Figures 10b,d show the relation between the observed $K(q)$ and the $q$ moments. The slightly nonlinear behavior is evidence of multifractality. Model parameters $c$ and $\beta$ are estimated by fitting Eq. (4) to the observed $K(q)$. The resulting theoretical expectations are plotted in Figs. 10b,d with solid lines.

Once $c$ and $\beta$ have been estimated on all events available for each gauge, we searched for calibration relations. As shown in previous applications (e.g., Deidda et al. 1999), the parameter $\beta$ was found to be constant and close to the value of $e^{-1}$. The parameter $c$ was then estimated by assuming $\beta = e^{-1}$. Following the results of earlier model calibrations on spatiotemporal rainfall fields (Deidda et al. 2004), we inspected the relation between $c$ and the coarse mean precipitation intensity $I$ at each gauge. Despite the presence of a certain dispersion of the $c$ values (STD of 0.2), we did not identify any significant trend. As a result, for each gauge we adopted a single value of $c$, equal to the average across all the rainfall events, independently of $I$. Figure 11 shows the relation between the mean $c$ and the gauge elevation $z$. A decreasing pattern is apparent and confirms the results shown in Fig. 8: higher values of $c$ are associated with more intermittent rainfall (shorter duration and higher rainfall rate) at lower elevation. A regression line was fitted to the data to represent the decreasing trend. This line can be utilized as a calibration relation for the STRAIN model to disaggregate daily observations as a function of elevation of the considered gauge. The robustness of the relation was tested by repeating the calibration with rainfall data observed during each year, obtaining very similar parameter values of the linear regression.

### e. Model verification

The model performances were tested using the rainfall events selected for the calibration. For each mean daily intensity $I$, we used STRAIN to generate an ensemble of 100 homogeneous disaggregated fields, using the parameter $c$ provided by the linear relation as a function of the elevation $z$ of the corresponding station. Each homogeneous field was then transformed into a heterogeneous one by applying Eq. (6) to reproduce the mean diurnal cycle. Note that, in this process, the same coarse-scale mean intensity $I$ was preserved for all
of the synthetic time series. The capability of the downscaling model to reproduce the rainfall variability at the finescale $t_0 = 45\,\text{min}$ was evaluated in two ways. First, we compared the empirical cumulative distribution functions (ECDFs) of the observed rainfall time series at the fine resolution against the 90% confidence intervals derived from the ensemble of 100 disaggregated time series. Figure 12 shows two examples of this comparison for each gauge. One can note how the model is able to capture the entire distribution with relatively good accuracy in most cases. However, examples of poorer performances can be found in Figs. 12b,f,j,n, where the model fails in reproducing the lower or upper tails of the distribution. It is worth pointing out that the STRAIN model reproduces multifractal cascades of strictly positive values. However, the generated intensities may be extremely low, thus falling below the resolution of the rain gauges. Thus, to reproduce the lacunar character of the rainfall field (i.e., presence of zeroes), a threshold (e.g., 0.1 mm h$^{-1}$) may be adopted to filter out the low intensities, and the other positive rainfall values should be rescaled to preserve the daily mean.

In a second verification of the model skill, we tested the capability to reproduce the higher rainfall rates at hourly resolution, which is often adopted for hydrological applications. For this purpose, the disaggregated

![Figure 10](image1.png)  
**FIG. 10.** Evidence of (left) scale invariance and (right) multifractality in the range from $T = 24\,\text{h}$ to $T = 45\,\text{min}$, for two rainfall events observed at gauges (a),(b) 1 and (c),(d) 8. The lines in (b) and (d) are the theoretical multifractal exponents provided by the log-Poisson generator.

![Figure 11](image2.png)  
**FIG. 11.** Linear relation between the parameter $c$ (mean value across all events) and the gauge elevation $z$. The equation of the linear regression between $c$ and $z$ and the coefficient of determination $R^2$ are also reported.
fields were resampled from $\Delta t = 45\text{ min}$ to $\Delta t = 1\text{ h}$, assuming a constant rainfall rate within each 45-min time interval. For each day, we extracted the 90% quantile from the ECDF of the observed hourly rainfall rates and from the ECDF of each disaggregated field produced by the STRAIN model. Figure 13 shows the frequency distribution of the observed and simulated 90% quantiles for each gauge. The observations are plotted with gray bars, while simulated distributions are presented as the mean ± STD of the ensemble simulations. Note the fairly good agreement between the distributions, which gives us confidence in the model’s capability to reproduce the heavy rainfall events at hourly resolution across the elevation transect.

5. Conclusions

We calibrated a downscaling model to reproduce the subdaily temporal rainfall variability using observations from eight gauges installed across an elevation transect.
FIG. 13. Comparison between the observed and simulated distributions of the 90% quantiles of hourly rainfall rates, extracted from 24-h-long events. Observations are plotted as gray bars, and the simulations are plotted as mean ± STD, computed from an ensemble of 100 disaggregated fields. Each panel refers to a different gauge.
in northwest Mexico. To this aim, we preliminarily analyzed the statistical properties of the rainfall records across a wide range of scales (from months to 1 min). In contrast with past studies in this region based on sparse networks of gauges covering large areas (size of hundreds of kilometers), the transect spans a much shorter distance (~14 km) with an elevation difference of 748 m, thus allowing us to better focus on the statistical features of localized storms. As such, these results need to be kept in the context of the northwest Mexico–southwest U.S. basin and range region. It is not clear, and likely doubtful, that the same results hold for the southern Sierra Madre where the mountain front comes down to the sea and the total elevation change is much greater (3000 m). Investigations have been carried out for both summer (from June to September) and winter (from November to February) rainfall observations. The characterization of winter rainfall in this region represents an innovative contribution, as past studies have been almost entirely focused on the warm season precipitation.

The main findings of the analyses on the rainfall time series can be summarized as follows. (i) The total rainfall accumulation does not appear to be significantly correlated with elevation in both seasons; a larger dataset would be necessary to further support this result. (ii) As found in several former studies, a strong diurnal cycle was identified in summer, peaking around 2000 LT, while no significant cycle emerged in winter. (iii) The investigation of the spatial correlation structure showed a significant correlation up to the largest separation distance of the transect (~14 km) for the winter rainfall sampled at relatively high temporal resolution (between 10 and 30 min). In summer, it is necessary to aggregate the signal at more than 1 h to obtain the same result. (iv) The spectral and scale invariance analyses revealed the presence of three scaling regimes in summer and two regimes in winter, which are hypothesized to be associated with the dominant meteorological phenomena at the corresponding time scales.

Overall, we found that summer rainfall is characterized by relatively short-lived, isolated convective storms with a typical duration of ~30 min whose statistical properties do not markedly vary across the aggregation time scales (from ~30 to 1 min) or as a function of the gauge elevation across the 748-m gradient. When the summer rainfall features were investigated at aggregation scales larger than ~30 min, we identified a regime typical of summertime propagating transient disturbances within the monsoon with longer duration (up to 10 days) and, likely, larger spatial coverage. In this time range, (i) the signal variability increases with decreasing scales of aggregation, due to the effects of the high rainfall intensities of the convective cells and (ii) the intermittency of the signal is higher at lower elevation (heavier rainfall events, less frequent). In winter, a single regime was detected ranging from about the daily scale down to 1 min, which accounts for frontal systems with spatial coverage larger than the transect size and without an important effect of elevation. This lack of a relationship with elevation in winter suggests that the effective height of the orographic barrier or its longitudinal extent may not be sufficient to induce sufficient lifting to produce a significant precipitation response. In winter, the effect of convective storms is embedded within this larger system, without the formation of a distinct time regime, as is found in the summer.

The statistical analyses provided the information for calibrating a downscaling scheme able to statistically reproduce the subdaily variability of daily precipitation observations in the summer season. This is achieved by simulating the scale invariance and multifractal properties of the rainfall time series from the coarse scale of 24 h to the finescale of 45 min. A modulation function was used to reproduce the time heterogeneity introduced by the diurnal cycle. This is a new contribution in the application of this type of multifractal model in the time domain. The model showed adequate performance in reproducing the small-scale observed precipitation variability. Note that the disaggregated time series can be easily resampled from the scale of 45 min to the hourly resolution, which is often adopted in hydrological and ecological applications. The tool can be applied to disaggregate several types of daily rainfall data, including observations collected by the Comisión Nacional de Agua network; outputs of climate models, spatially corrected to account for the point-based nature of the gauges; and rainfall time series stochastically generated by weather simulators (e.g., Rodríguez-Iturbe and Porporato 2004). The relatively similar statistical properties of the rainfall time series of different time periods found in previous work in northwest Mexico (Gebremichael et al. 2007; Gochis et al. 2004, 2006, 2007) provide confidence in the robustness of the calibration relations, which we hypothesize are applicable to the mountainous regions in northern Sonora.

Acknowledgments. We thank three anonymous reviewers for their comments that helped to improve the quality of the manuscript. Funding from the NSF IRES Program (OISE 0553852, IIA 1130254) and the NOAA Climate Program Office (GC07-019; GC09-493) is acknowledged. The authors also thank Luis A. Mendez-Barroso and Leila El Ashmawy for data collection and
preliminary analysis. DJG also acknowledges support from the National Science Foundation, which supports NCAR through a cooperative agreement.

REFERENCES


