

# An Evaluation of the Statistics of Rainfall Extremes in Rain Gauge Observations, and Satellite-Based and Reanalysis Products Using Universal Multifractals

XIAOMING SUN AND ANA P. BARROS

*Department of Civil and Environmental Engineering, Pratt School of Engineering, Duke University, Durham, North Carolina*

(Manuscript received 22 December 2008, in final form 6 July 2009)

## ABSTRACT

Confidence in the estimation of variations in the frequency of extreme events, and specifically extreme precipitation, in response to climate variability and change is key to the development of adaptation strategies. One challenge to establishing a statistical baseline of rainfall extremes is the disparity among the types of datasets (observations versus model simulations) and their specific spatial and temporal resolutions. In this context, a multifractal framework was applied to three distinct types of rainfall data to assess the statistical differences among time series corresponding to individual rain gauge measurements alone—National Climatic Data Center (NCDC), model-based reanalysis [North America Regional Reanalysis (NARR) grid points], and satellite-based precipitation products [Global Precipitation Climatology Project (GPCP) pixels]—for the western United States (west of 105°W). Multifractal analysis provides general objective metrics that are especially adept at describing the statistics of extremes of time series. This study shows that, as expected, multifractal parameters estimated from the NCDC rain gauge dataset map the geography of known hydrometeorological phenomena in the major climatic regions, including the strong orographic gradients from west to east; whereas the NARR parameters reproduce the spatial patterns of NCDC parameters, but the frequency of large rainfall events, the magnitude of maximum rainfall, and the mean intermittency are underestimated. That is, the statistics of the NARR climatology suggest milder extremes than those derived from rain gauge measurements. The spatial distributions of GPCP parameters closely match the NCDC parameters over arid and semiarid regions (i.e., the Southwest), but there are large discrepancies in all parameters in the midlatitudes above 40°N because of reduced sampling. This study provides an alternative independent backdrop to benchmark the use of reanalysis products and satellite datasets to assess the effect of climate change on extreme rainfall.

## 1. Introduction

Changes in the risk and magnitude of extreme precipitation events rank among the most studied effects and as indicators (symptoms) of climatic variations and climate change (Barros and Evans 1997; Easterling et al. 2000; Schnur 2002; Milly et al. 2002). Extreme precipitation translates generally into extreme flooding and, consequently, loss of life, collapse of lifeline infrastructure, and the breakdown of public health services, among others. In approaching the problem of quantifying the risk and magnitude of extreme precipitation events, there are two major challenges: 1) it is difficult to characterize “observed” (twentieth and twenty-first century) conditions because of the lack of

long-term observations—that is, short and incomplete historical records and limited-duration model simulations; and 2) it is difficult to characterize “predicted” (twenty-first century) conditions or generally changes with regard to the reference situation, because of the lack of skill of precipitation forecasts at spatial and temporal scales meaningful for impact studies (Palmer and Räisänen 2002; Fekete et al. 2004; Hegerl et al. 2006; Zwiers and Hegerl 2008). The first challenge translates in estimating the probability of occurrence (rare) and magnitude (very large) of events that may have not happened yet—the tail (of the probability distribution) estimation problem (TEP), either in the observational record or in model simulations (see, e.g., Hunt 2006). The second challenge is that of quantifying uncertainty and separating climatic variability and change from model error.

Common approaches to address these challenges include fitting a known probability distribution to the

---

*Corresponding author address:* Xiaoming Sun, Duke University, Box 90287, 2457 CIEMAS Fitzpatrick Bldg., Durham, NC 27708.  
E-mail: xiaoming.sun@duke.edu

available observations and trend analysis of empirical indices of observations classified according to preset thresholds (Wilby and Wigley 2002; Groisman et al. 2004, 2005; Wehner 2004; Perkins et al. 2007; among others). Concerns with the fidelity of climate model simulations and imperfect forecasts have been addressed by using model ensembles understood as independent realizations of climate change and by focusing on seasonal time scales instead of extreme events per se (Zwiers and Kharin 1998; Palmer and Räisänen 2002; Wilby and Wigley 2002; Wehner 2004). Because of well-documented difficulties with the parameterization of precipitation processes in climate models and the lack of predictive skill at short time scales and for high intensity events, precipitation extremes are usually defined with reference to a seasonal threshold instead of the actual risk and magnitude of specific events (e.g., Kharin and Zwiers 2002; Wehner 2004).

At the regional scales relevant for climate assessment studies, the characterization of extreme rainfall requires time series with high spatial and temporal resolution. Traditionally, such datasets are obtained from rain gauge networks, which are point-based measurements often unevenly and sparsely distributed, especially in remote areas (away from population centers and in regions of complex terrain). From 1980 onward, satellite-based rainfall products have become increasingly available, and despite still coarse temporal and spatial resolution, datasets such as the Global Precipitation Climatology Project (GPCP; among others) have radically changed the landscape of rainfall studies by providing spatially consistent estimates of rainfall as a field as opposed to a point, and at global scale. In fact, because of these treats, satellite datasets are becoming the observational record of choice for climate modeling studies (e.g., Allan and Soden 2008). Recently, the development of data assimilation systems allows for a physically based, optimal merger of rainfall point and spatial measurements to generate high-resolution rainfall fields even in regions and at times where and when ground-based or satellite observations are missing (Mesinger et al. 2006).

To assess the magnitude and character of rainfall extremes in climate fingerprinting and climate variability studies, it is important to establish a baseline understanding of their statistical properties for current climate. Because the statistics are determined by the data, the question of whether and how such differences have an impact on the estimation of extreme values is especially relevant for global- and regional-scale assessments outside of North America and western Europe due to the dearth of reliable ground-based observa-

tions. This requires that a statistical model (probability distribution) that fits the data in an optimal fashion according to some predefined criteria be selected, and the implication is that changes in model parameters for different climate scenarios are indicative of the corresponding changes in precipitation extremes. One limitation is that such statistical models typically rely on the hypothesis that the observations are statistically independent, whereas rainfall results from highly nonlinear, complex dynamical processes that interact with each other, and extremes often occur in clusters, implying strong correlations and potentially long-range dependencies. Furthermore, rainfall time series are typically short, usually less than one century (in this study, no more than 60 years; see section 4), and there are very few extreme observations with which to fit the model, thus generating uncertainty in parameter estimation (Douglas and Barros 2003). Because of the difficulty in constraining the optimal fitting process by specifying meaningful, physically based conditions, ad hoc goodness-of-fit criteria are used instead, thus leading to ambiguity in the interpretation of parameter estimates (Bernardara et al. 2008). To describe extreme statistical behavior independently of model fit, there is therefore a need for objective quantitative metrics that can be generally applied.

Nevertheless, even if data availability is ideal, the space–time distributions of precipitation, and hydrometeorological variables generally, exhibit characteristics of complexity (neither random, nor fully deterministic) and intermittency (and nonstationarity) that cannot be described by standard second-order statistical models (Dooge 1986; Seuront et al. 1999). The scaling behavior of precipitation has been studied over a wide range of spatial scales representative of physically relevant processes, from rain microphysics to storm dynamics [see, e.g., reviews by Lovejoy and Schertzer (1995)]. These studies suggest that precipitation fields exhibit distinct but statistically invariant scaling behavior (multifractals) across different ranges of scales that can be described by only three universal parameters (Schertzer and Lovejoy 1987). This simplicity is appealing for extensive climate studies in contrast with alternative combinations of statistical point process models with complex structure and extensive parameterization (e.g., Waymire and Gupta 1981a,b,c; Cowpertwait 1995, 1998).

Building on previous work, we rely on multifractals (power-law scaling) to describe the high-order statistics of the tails of probability distributions of observed and simulated precipitation from three different climate datasets over the continental United States: rain gauge data from National Climatic Data Center (NCDC) records, satellite-based data from the GPCP (Xie and

Arkin 1996, 1998; Huffman et al. 2001; Adler et al. 2003; Xie et al. 2003), and spatial fields from the North American Regional Reanalysis (NARR) project (Mesinger et al. 2006). Both GPCP and NARR use rain gauge data for algorithm corrections and for data assimilation, respectively, and direct satellite observations (e.g., brightness temperatures) used to derive GPCP rainfall are also used in the NARR data assimilation system. Although there are many limitations associated with the sampling representativeness and measurement error of rain gauges, the NCDC rainfall data provide the observational reference (direct measurement without post-processing) against which the GPCP and the NARR data are compared, and limitations detected over the continental United States are indicative of limitations elsewhere for similar types of data and models. Douglas and Barros (2003) conducted similar work for the continental United States east of the 105°W meridian to characterize the magnitude and return period of the probable maximum precipitation (PMP). In this work, we focus now on the western United States (west of 105°W).

Section 2 briefly reviews the theoretical basis of universal multifractal analysis and its relationship with rainfall and extremes. Section 3 reports on numerical simulations that explain the physical linkages between multifractal parameters and rainfall characteristics. The datasets are described in detail in section 4, and lastly, the multifractal analysis results are presented and examined in detail in section 5. Conclusions and a discussion are presented in section 6. The appendix provides details on the estimation of the universal multifractal parameters.

## 2. Multifractals, rainfall, and extremes

Multifractals are analogous to the multiplicative scaling “cascade” processes in turbulence, whereby a large-scale eddy of characteristic length  $L$  is hierarchically broken up into “daughter” subeddies whose length is equal to  $L/\lambda_s^i$ . The intensity is the product of the “parents’ intensity” multiplied by a random variable, where  $i = 1, \dots, n$  denotes the step of the cascade process and  $\lambda_s$  is the scale ratio between two adjacent steps of the cascade process [i.e.,  $\lambda_s = l_n/l_{n+1}$ , where  $l_n$  and  $l_{n+1}$  are the size of eddies at step  $n$  and  $n + 1$ , respectively; see Figs. 1 and 2 of Schertzer and Lovejoy (1987) for more details]. According to the principle of scale invariance, the variability at any given scale reflects a large dynamical range of the phenomena determined by long-range dependencies between smaller and larger scales, which is quite different from the independence hypothesis in classical statistics.

The fundamental property of multifractal fields (Schertzer and Lovejoy 1987) can be described by the following probability distribution:

$$\Pr(\varepsilon_\lambda > \lambda^\gamma) \sim \lambda^{-c(\gamma)}, \quad (1)$$

where  $\varepsilon_\lambda$  is the intensity of a field (e.g., rainfall rate) at scale ratio  $\lambda$  (i.e., the ratio of the largest external scale  $L$  over the scale in consideration: for a 1-yr long, daily rainfall time series,  $\lambda = 365$  when  $l = 1$  day and  $L = 365$  days);  $\gamma$  is the order of singularity (maximum) corresponding to  $\varepsilon_\lambda$  and describes how fast the field diverges ( $\varepsilon_\lambda > \lambda^\gamma$ ); and  $c(\gamma)$  is the convex codimension function characterizing the sparseness of the  $\gamma$ -order singularities. For a rainfall time series,  $c(\gamma)$  measures the fraction of the probability space occupied by the rainfall intensities whose order of singularity exceeds  $\gamma$ . The tilde sign ( $\sim$ ) in Eq. (1) and hereafter implies equality within slowly varying constants. An equivalent relationship can be expressed by the scaling of the statistical moments,

$$\langle \varepsilon_\lambda^q \rangle \sim \lambda^{K(q)}, \quad (2)$$

where  $\varepsilon_\lambda^q$  is the  $q$ th moment of the field,  $\langle \cdot \rangle$  denotes the ensemble average, and the scaling exponent  $K(q)$  is related to  $c(\gamma)$  by the Legendre transform (Frisch and Parisi 1985):

$$\begin{cases} c(\gamma) = \max_q [q\gamma - K(q)] \\ K(q) = \max_\gamma [q\gamma - c(\gamma)] \end{cases} \quad (3)$$

Equation (3) establishes a one-to-one correspondence between the order of the singularity  $\gamma$  and the moment order  $q$ . Knowledge of either one,  $K(q)$  or  $c(\gamma)$  in Eq. (3), is sufficient to characterize the statistics of the multifractal field.

The scaling behavior of precipitation has been documented over a wide range of spatial scales representative of physically relevant processes, from rain microphysics to storm dynamics (e.g., Lovejoy and Schertzer 1995). These studies suggest that precipitation fields exhibit distinct but statistically invariant scaling behavior (multifractal) across different ranges of scales and that rain rate is a conservative field (the ensemble mean is the same at all scales) is well established (Hubert et al. 1993; Tessier et al. 1993; Fraedrich and Larnder 1993; de Lima and Grasman 1999; among others). Compared with classical statistical models that rely on assumptions of strong regularity and homogeneity and tend to display exponential behavior of the probability tails, Lovejoy et al. (2007) showed that multifractals capture extreme variability consistent with power-law behavior of the probability tails. In this case, the contributions of small-scale activity to much larger

scales are nonnegligible, and the corresponding variability has the most extreme manifestation (Schertzer and Lovejoy 1992).

A pure power law ( $\Pr(X > x) \sim x^{-k}$ , where  $k > 1$ ), however, diverges as  $x \rightarrow 0$ , and therefore it cannot be used to relate high-frequency (low rainfall) and low-frequency (extreme rainfall) events. Indeed, multifractals are more general than a power law in that the codimension function,  $c(\gamma)$  in Eq. (1), implies an infinite hierarchy of exponents rather than a single value linking moderate to extreme behavior by scaling, as noted by Bernardara et al. (2008). This is especially critical in the case of rainfall, because the statistics of extremes are usually quite poor due to the lack of observations, and thus, to produce robust statistics, it is necessary to extract information from events on record, typically frequent events of small magnitude. Multifractals inherently provide a strategy to bridge this scale gap. Moreover, the universal multifractal framework not only captures the scaling behavior of rainfall across a wide range of scales but in addition, it sidesteps the ambiguity inherent to the selection of specific statistical point process models, or a combination thereof (e.g., Kavvas and Delleur 1975; Waymire and Gupta 1981a,b,c).

Despite their apparent simplicity, Eqs. (1) and (2) are actually unmanageable for practical purposes because an infinite number of exponents is implied. Schertzer and Lovejoy (1987) showed that statistically scale invariant processes are stable and converge to some universal attractor, and thus can be defined by a small number of relevant parameters, specifically three, in the context of the universal multifractal framework. This should be particularly true in the case of complex systems with many perturbations and interactions that cooperate to emphasize essential features and smooth away peculiarities (Schertzer and Lovejoy 1987). For a conservative field, that is a field whose ensemble mean is preserved at different scale ratios (i.e.,  $\langle \varepsilon_\lambda \rangle = \text{constant}$ );  $c(\gamma)$  and  $K(q)$  depend only on two multifractal parameters,  $C_1$  and  $\alpha$ , and satisfy the following equations (Schertzer and Lovejoy 1987):

$$c(\gamma) = \begin{cases} C_1 \left( \frac{\gamma}{C_1 \alpha'} + \frac{1}{\alpha} \right)^{\alpha'}, & \alpha \neq 1 \\ C_1 \exp\left(\frac{\gamma}{C_1} - 1\right), & \alpha = 1 \end{cases} \quad \text{and} \quad (4)$$

$$K(q) = \begin{cases} \frac{C_1(q^\alpha - q)}{\alpha - 1}, & \alpha \neq 1 \\ C_1 q \log(q), & \alpha = 1 \end{cases}, \quad (5)$$

where the Levy index  $\alpha$ , ranging from 0 to 2, indicates how fast the process deviates from simple scaling (i.e.,

monofractal with  $\alpha = 0$ );  $\alpha'$  satisfies  $(1/\alpha) + (1/\alpha') = 1$ ; and  $C_1$  is the index of intermittency, a measure of the sparseness of the data (Davis et al. 1994).

For extreme events, besides  $\alpha$  and  $C_1$ , it is necessary to estimate the third parameter, the effective maximum order of the singularity  $\gamma_s$ , which is the upper bound of the observed singularities (maximum amplitude) in a finite sample with  $0 < \alpha \leq 1$ . For a time series, it can be estimated as follows (Hubert et al. 1993):

$$\gamma_s = \gamma_o \left[ 1 - \alpha \left( \frac{C_1}{D} \right)^{-1/\alpha'} \right], \quad (6)$$

where  $\gamma_o = C_1/(1 - \alpha)$  is the maximum order of the singularity for (ideal) time series with infinite length. Note that  $\gamma_s$  describes the amplitude of the variations of rainfall extremes, but it depends on both  $\alpha$  and  $C_1$ . Details regarding the estimation of  $\alpha$  and  $C_1$  are presented in the appendix.

Consistent with the common practice in hydrological and climate modeling studies of using bilinear interpolation to visualize the spatial variability of the statistical properties of rainfall extremes, the multifractal parameters were interpolated bilinearly across the entire study region. Although fractal disaggregation may capture small-scale variability and orographic effects better (Bindlish and Barros 1996; Kim and Barros 2002), this type of analysis requires careful evaluation of spatial scaling properties and ultimately multifractal objective analysis (Salvadori et al. 2001) if the rain gauge network itself is also multiscaling (Tessier et al. 1994). This is beyond the scope of this study, given the coarse and heterogeneous characteristics of the rain gauge network.

### 3. Interpretation of universal multifractal parameters

Using the algorithm recently proposed by Lovejoy and Schertzer (2009a,b), numerical simulations of precipitation intensity  $\varepsilon(x)$  for different values of  $\alpha$  but the same  $C_1$  ( $C_1$  was arbitrarily chosen as 0.5) and for different values of  $C_1$  but the same  $\alpha$  ( $\alpha$  was arbitrarily chosen as 0.5), and the corresponding empirical probability density functions (PDFs) are displayed respectively in Figs. 1–4. Compared with other cascade models, which are either discrete in scale (e.g., Novikov and Stewart 1964; Yaglom 1966; Mandelbrot 1974) as compared to continuous-scale physical processes or are computationally inefficient (Schertzer and Lovejoy 1987), the new algorithm stems from a continuous formulation of

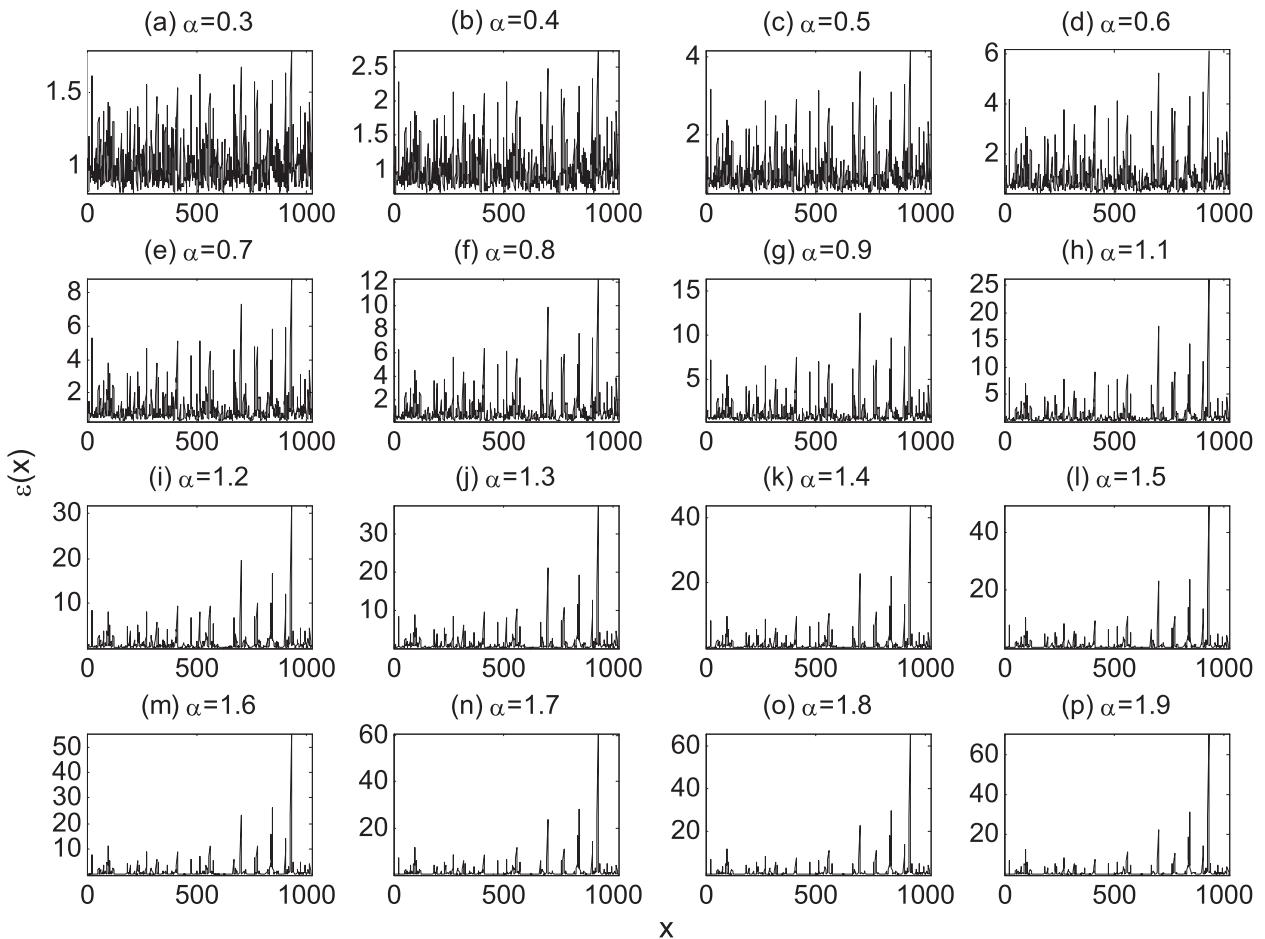


FIG. 1. One-dimensional isotropic multifractal simulations of rainfall intensity with same  $C_1$  (arbitrarily chosen as 0.5) but different  $\alpha$  using the Mathematica code from Lovejoy and Schertzer (2009a,b). Same random seeds were used to facilitate comparison. Note the increase of vertical axis scales as  $\alpha$  increases.

the cascade generation process and significantly improves the simulation of statistical properties at small scales, leading to a great increase of computation efficiency. To facilitate the comparison and therefore the interpretation of the physical meaning of multifractal parameters, the same random seeds were used in Figs. 1 and 3.

As shown in Fig. 1, for a specified  $C_1$  ( $C_1 = 0.5$ ), the magnitude of extreme events (associated with  $\gamma_s$ ) increases with increasing  $\alpha$ , which is also captured by the longer tails of corresponding empirical probability density plots (see Fig. 2). Furthermore, Fig. 1 shows there are fewer fluctuations as  $\alpha$  increases, consistent with the strong decrease in the probability of large fluctuations from a central value in Fig. 2 (note that these large fluctuations are not the same as the extreme events, but rather of large magnitude compared to the mean). This also confirms the findings by Carvalho et al. (2002), who argued that  $\alpha$  controls the fall off rate of probability density function.

The role of the intermittency index  $C_1$  (Davis et al. 1994) can be easily inferred from Figs. 3 and 4. When  $C_1$  increases, the occurrence of large fluctuations becomes less and less frequent (see increase in sparseness in Fig. 3). For a given amount of rainfall, the magnitude of individual events increases as the number of events decreases, therefore increasing the magnitude of extreme values as illustrated by the elongated PDF tails in Fig. 4.

The effective maximum singularity  $\gamma_s$ , which is the maximum singularity (extreme event) associated with the dataset (Schertzer and Lovejoy 1987; Hubert et al. 1993; Douglas and Barros 2003), integrates the effects of  $\alpha$  and  $C_1$  and increases with both the increase of  $\alpha$  and  $C_1$ . The codependency of the magnitude of extremes on  $\alpha$  and  $C_1$  is amplified if these two parameters vary in the same direction; otherwise, the response depends on the dominant influence. Royer et al. (2008) proposed  $\gamma_s$  to be proportional to the relative magnitude of extreme events as compared to the mean, that is,  $R_{\max}/R_{\text{mean}}$ ,

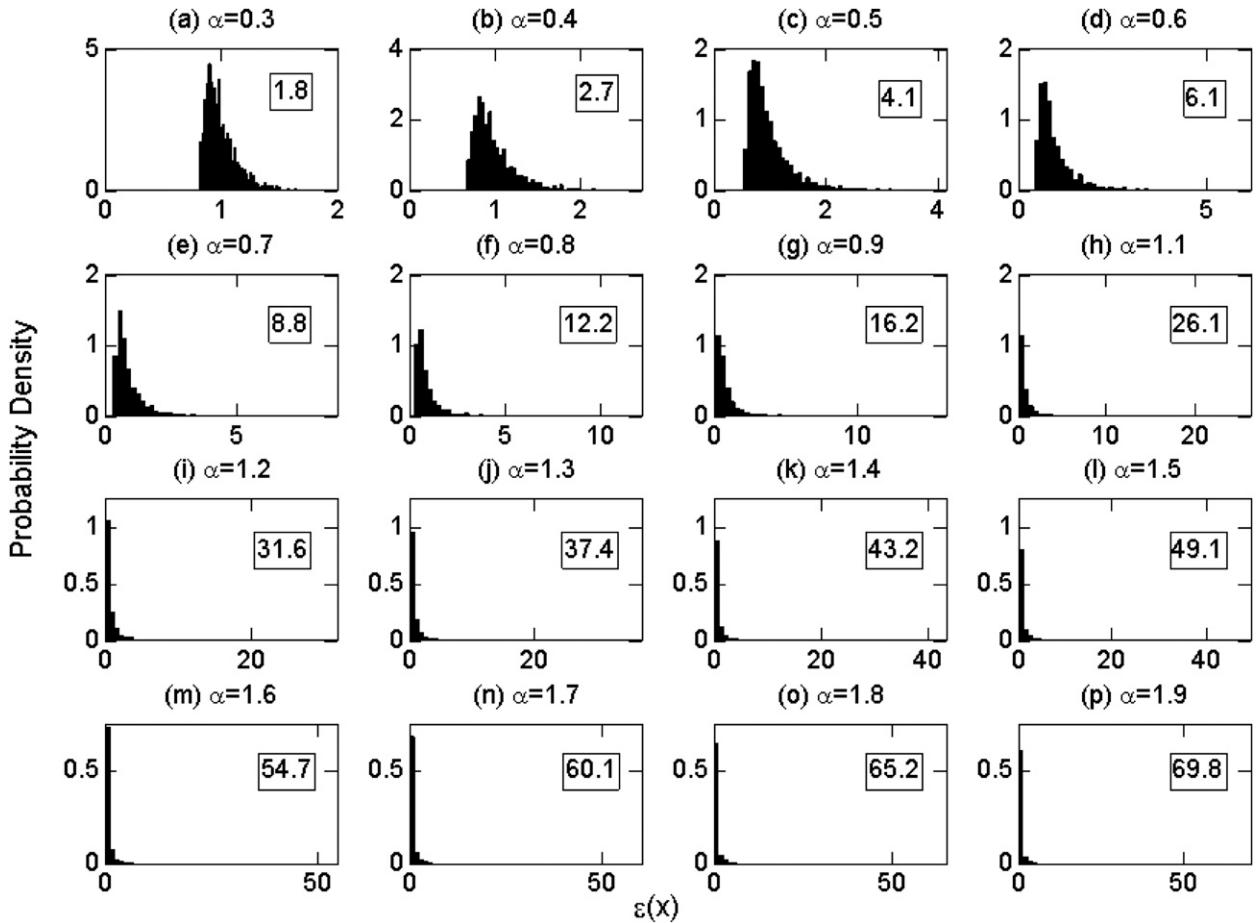


FIG. 2. Empirical probability density plots correspond to the numerical simulation in Fig. 1. Note that the value in the box of each panel corresponds to the maximum  $\epsilon(x)$ .

where  $R_{\max}$  is the maximum rainfall intensity and  $R_{\text{mean}}$  is the average rainfall intensity as follows:

$$\frac{R_s(\lambda)}{R} \sim \lambda^{\gamma_s}, \tag{7}$$

where  $R_s(\lambda)$  is the maximum probable rainfall at scale ratio  $\lambda$  and  $R$  is the rainfall average at a large scale. Analysis in terms of “relative” extremes is very useful because it allows for intercomparison of statistics from different climatic regions, for example (i.e., with different average rainfall).

It is often in the interest of climate studies to frame climate extremes in probabilistic terms focusing on the  $X$ -year event (event with a return period of  $X$  years) and/or exceedance probabilities. Although this is not directly relevant for the specific objectives of this work, such perspective can be directly derived from the results presented here [see Eq. (1)], as shown by Hubert et al. (1993) and Douglas and Barros (2003).

#### 4. Datasets

Although there are approximately 6000 NCDC rain gauge stations (available online at <http://www.ncdc.noaa.gov/oa/climate/stationlocator.html>) available in the western United States (west of 105°W), many rainfall records are either short or have long gaps (more than 5% observations of rainfall time series missing). The effect of missing data points in time series of observational records has long been recognized in the hydrology and hydrometeorology communities, and different strategies have been proposed, generally with an eye to preserve basic statistical properties of the record, such as the mean and the standard deviation (e.g., Kuligowski and Barros 1998, among others). However, in this case, because we are focusing on the tails of the probability distribution, it is important that no spurious information be added; therefore, to maximize both the spatial density and temporal continuity of rainfall time series, only 692 stations collecting data between 1950 and 2006 were used. The study domain and the location of the selected

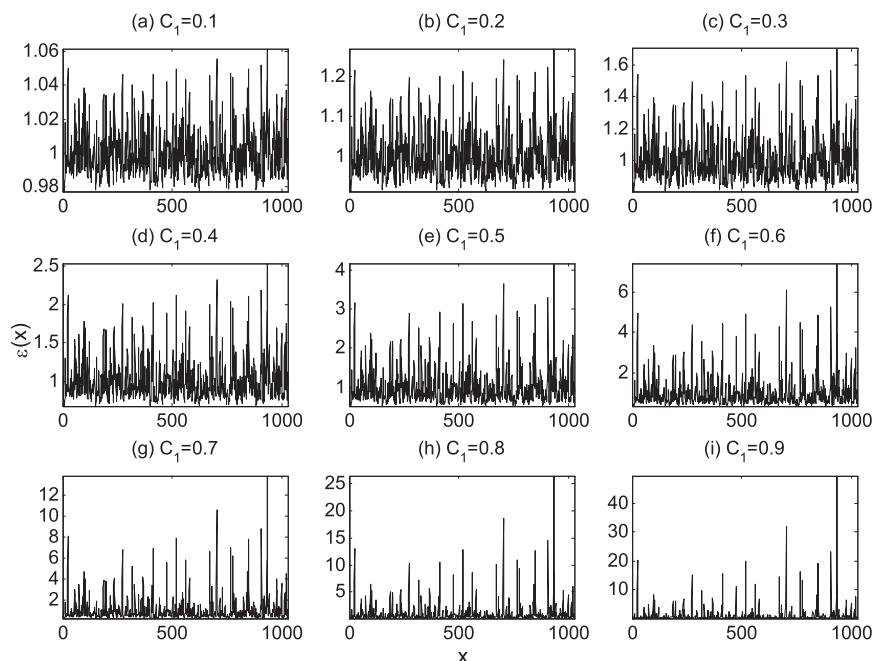


FIG. 3. One-dimensional isotropic multifractal simulations of rainfall intensity with same  $\alpha$  (arbitrarily chosen as 0.5) but different  $C_1$  using the Mathematica code from Lovejoy and Schertzer (2009a,b). Same random seeds were used to facilitate comparison. Note the increase of vertical scales as  $\alpha$  increases.

NCDC stations are shown in Fig. 5 with the topography as background.

Because the estimation of the codimension  $c(\gamma)$  is sensitive to the normalization method (Lovejoy and Schertzer 1995), the regional mean from all time series was employed to normalize the data such that the interpretation of spatial and temporal comparisons will not be confounded by the nondimensionalization process (Douglas and Barros 2003). Values for 29 February were ignored. Missing daily data were filled with the average of the preceding and following days if they exist, otherwise with the time series average. Because all the NCDC stations used do not have more than 5% data missed, the empirical probability density does not show significant differences before and after filling with averages.

The GPCP rainfall products were originally obtained by merging infrared and microwave satellite rainfall estimates with rain gauge data (available online at <http://www.ncdc.noaa.gov/oa/wmo/wdcamet-ncdc.html>). Although monthly GPCP (from 1979 to present, with a spatial resolution at  $2.5^\circ \times 2.5^\circ$ ) has a longer time span than daily GPCP precipitation (from 1997 to present at  $1^\circ$  resolution), the latter were used in this study because it is not possible to characterize the statistical behavior of extremes in the 16-day scaling regime using the monthly time-scale data.

The 3-hourly NARR reanalysis fields have a much finer spatial resolution ( $32 \text{ km} \times 32 \text{ km}$ ; available online at [http://nomads.ncdc.noaa.gov/data.php?name=access#narr\\_datasets](http://nomads.ncdc.noaa.gov/data.php?name=access#narr_datasets)). For the sake of facilitating the inter-comparison and reducing the effects of time series length, all data were analyzed at the daily time scale (the 3-hourly NARR was simply aggregated to daily values) within the same time window. To examine the extreme properties of each dataset separately, multifractal analysis was also conducted on each dataset at their own resolutions using the entire length of the historical record available. Each grid cell or pixel in the NARR and GPCP datasets is treated as one “station” for the purpose of our analysis. Chen and Knutson (2008) alert to the importance of not interpreting model precipitation output as a point estimate, and state that one should rather focus on recognizing that the output represents area averages and therefore should be compared with point observations only after aggregated to the model resolution. Although this is correct, the challenge to achieve a fair comparison is how to extrapolate area averages from point data while preserving the space-time scaling properties (higher-order statistics) of the point observations (see example in Bindlish and Barros 1996; Kuligowski and Barros 2001; Barros and Tao 2008). Here we focus on illustrating the differences among the different datasets at their nominal temporal and spatial

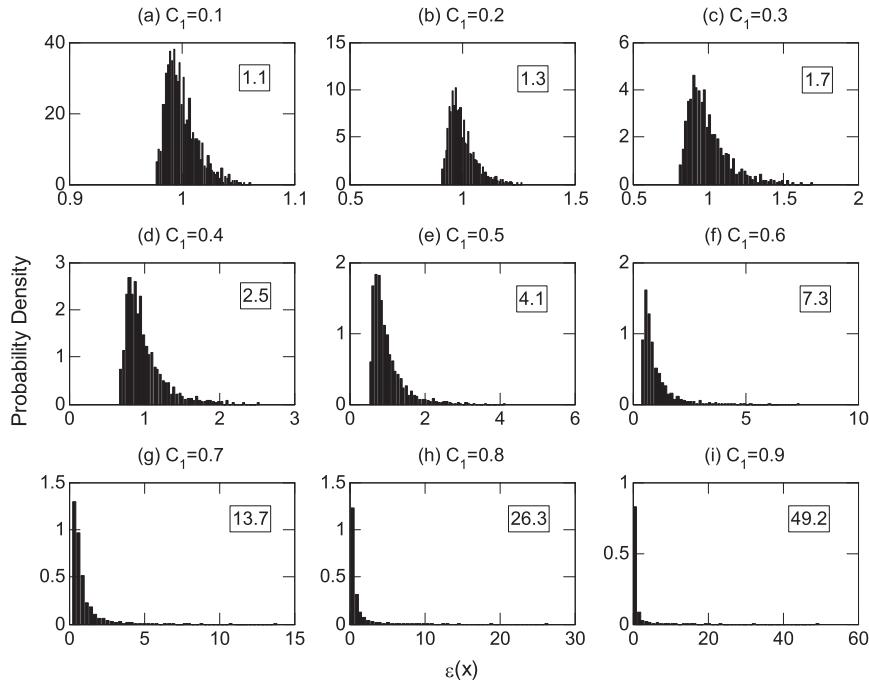


FIG. 4. Empirical probability density plots correspond to the numerical simulation in Fig. 3. Note that the value in the box of each panel corresponds to the maximum  $\epsilon(x)$ .

resolutions toward proving insight for their use in climate change detection and attribution studies proper, which are out of the scope of this work.

**5. Results**

The top row in Fig. 6 shows the range of multifractal parameters estimated from the datasets at their original temporal resolutions within their respective time windows: (a) daily NCDC (1950–2006), (b) 3-hourly NARR (1979–2007), and (c) daily GPCP (1997–2007). The estimated  $\alpha$  for NCDC is in the range (0, 1) with a median around 0.4, which is in reasonable agreement with other studies (e.g., Hubert et al. 1993; Tessier et al. 1993; Olsson 1995; Douglas and Barros 2003). However,  $\alpha$  is higher for the other two datasets, especially for NARR (Fig. 6b). Because higher  $\alpha$  indicates less frequency of large magnitude fluctuations (i.e., small event-to-event variations), these results suggest that these two datasets may miss large magnitude events that are characteristic of seasonal-to-interannual variability, and that they generally do not capture the right-hand-side tail of the underlying statistical distribution. Conversely, the lower values of  $C_1$  for NARR and GPCP imply that these two datasets are less sparse (fewer zeros) than NCDC (the higher the  $C_1$ , the more intermittency of the mean process). In other words, the NCDC rain gauge observations exhibit a higher frequency of dry periods at different scales (and longer interstorm arrival times).

A similar contrast among the multifractal parameters holds among NCDC, NARR, and GPCP within the same time window (1997–2006) at daily resolution (see the bottom row in Fig. 2). Comparing Fig. 6a with Fig. 6d, Fig. 6b with Fig. 6e, and Fig. 6c with Fig. 6f, no significant difference can be detected for  $C_1$  and  $\gamma_s$ , although there are some differences in  $\alpha$ , especially for the case of NARR. This is due to the aggregation from 3-hourly to daily rainfall accumulations, which can strongly reduce the amplitude of the event-to-event variations, leading to different  $\alpha$  estimates. This effect should be more apparent in regions where convective rainfall dominates

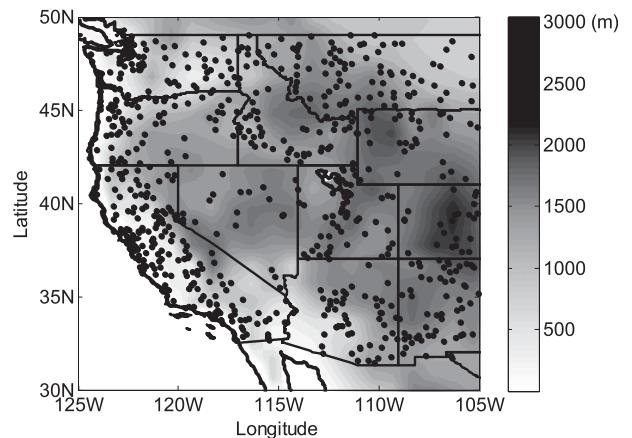


FIG. 5. Study domain and locations of selected NCDC daily stations with topography as the background.

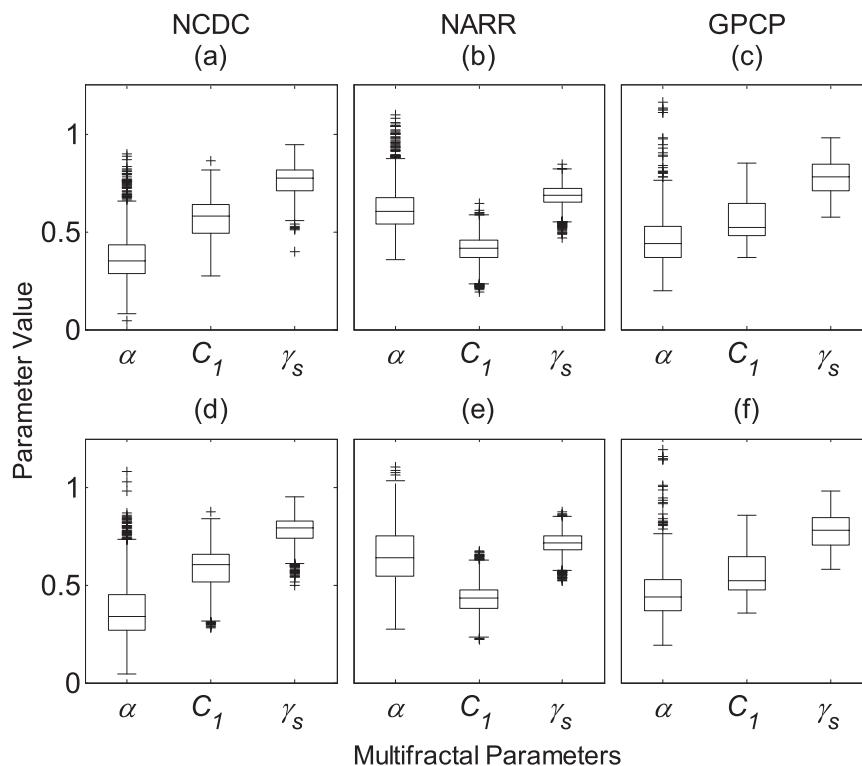


FIG. 6. Multifractal parameters estimated for (a) daily NCDC (1950–2006), (b) 3-hourly NARR (1979–2007), (c) daily GPCP (1997–2007), (d) daily NCDC (1997–2006), (e) aggregated daily NARR (1997–2006) and (f) daily GPCP (1997–2006). The ranges of multifractal parameters are for all stations (NCDC) and grid points (NARR and GPCP).

(short duration, high intensity) as opposed to regions where stratiform or orographic rainfall dominates (longer duration, light intensity).

These results further corroborate Tessier et al. (1993) and Boufadel et al. (2000), who argued that parameters other than  $\alpha$  can be more accurately estimated independently of the length and temporal resolution of the data. The small difference between parameter estimates using time series with different lengths is consistent with the fact that for  $\alpha < 1$ , finite sampling bias is not severe (Veneziano and Furcolo 1999).

Figure 7 displays the spatial distribution of the estimated multifractal parameters obtained by bilinear interpolation. For the NCDC dataset, the high  $\alpha$  values (lower probability of large magnitude fluctuations) and low  $C_1$  values (low intermittency) clearly delineate the marine climate and orographic controls along the West Coast from Washington through Oregon to northern California (Figs. 7a and 7b). In particular, note how both  $C_1$  and  $\alpha$  capture the rain shadow in the Columbia and Spokane River basins between the Cascades and the Blue Mountains. In the Southwest, there is also a clear demarcation of the desert and semiarid regions, al-

though the spatial interpolation of  $\alpha$  and  $C_1$  estimates over Arizona and south Nevada is likely affected by the limited number of rain gauges and their uneven distribution in that region. Nevertheless, the signature of the North American monsoon (NAM) seems to be reproduced, as implied by the low  $\alpha$  (i.e., large rainfall variations as compared to the very low mean precipitation). The considerably high values of  $C_1$  (high intermittency or large sparseness) agree well with the corresponding desert and semiarid climate in the region. According to the spatial distribution of effective maximum singularity  $\gamma_s$  (Fig. 7c),  $C_1$  is dominant over most of the study domain. The spatial patterns in the relative (or normalized) rainfall map in Fig. 8 are close to those in Figs. 7b and 7c, which lend more confidence to our estimation of  $\gamma_s$ . In particular, note the contrast between the Pacific Northwest and the Southwest: although annual rainfall accumulation is higher in the Northwest, the magnitude of rainfall extremes at temporal scales below 16 days is dramatically larger in the Southwest. Hereafter, we use the NCDC multifractal parameter estimates as the reference values for the analysis of the NARR and GPCP datasets.

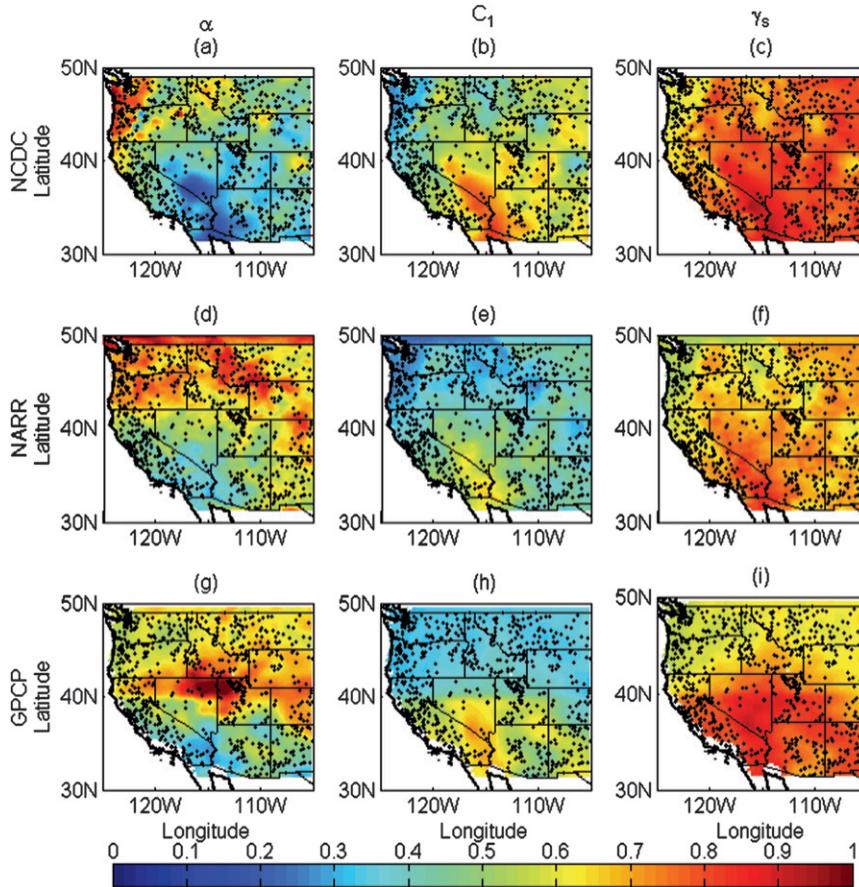


FIG. 7. Spatial pattern of estimated multifractal parameters for NCDC, NARR, and GPCP within their respective time windows at their own resolutions.

Interestingly, although the spatial distribution of  $\alpha$  from NARR does capture the broad differences between the Northwest and Southwest regions (Fig. 7d), its magnitude is much higher, which means that the NARR data do not capture intense fluctuations very well. Also note the disappearance of the rain shadow. The much lower  $C_1$  values in Fig. 7e explain the lower  $\gamma_s$  values in Fig. 7f, so that it can be seen that the NARR precipitation field underestimates large events in general, and extreme events in particular. Nevertheless, the pattern morphology for all multifractal parameters from NARR agrees quite well with NCDC, which may be explained in two ways: 1) the physically based data assimilation models used for generating NARR can reproduce the contrast among major climate regimes and orographic gradients and 2) averaging in the assimilation process and/or lack of appropriate model physics lowers the magnitude of the maximum singularity and decreases intermittency. This is consistent with Royer et al. (2008), who also found dominant  $C_1$  effects in their application with model-simulated rainfall in Europe, specifically over France.

The third row in Fig. 7 displays the estimated multifractal parameters for GPCP. The extremely high  $\alpha$  values in the central part of the study domain (Fig. 7g) may be an artifact of bilinear interpolation resulting from its relative coarse spatial resolution ( $1^\circ \times 1^\circ$ ); however, they more likely reflect the fact that over semiarid mountainous terrain, there is generally very little precipitation and that most precipitation occurs in the form of snow not rainfall, which is not well captured by current satellite observations. In addition, the standard bias correction applied to GPCP may be insufficient due to the dearth of rain gauges in the region. The high probability of large event-to-event variations, high intermittency, and large effective maximum singularity over the Southwest regions affected by NAM are in close agreement with NCDC, as illustrated in Figs. 7g–7i, and by comparison better than those for NARR. However, the rainfall regime along the West Coast, including orographic effects in the states of Washington and Oregon, go largely missed. One potential explanation for these results is the coarse resolution of the GPCP dataset (from 1997 to 2007 at  $1^\circ \times 1^\circ$ ), which would severely

limit any terrain modulation effects on the satellite-based rainfall estimates.

For an overall assessment, the scatterplots of multifractal parameters for NCDC versus NARR and GPCP are recorded in Fig. 9, where the nearest grid cell to each NCDC station was selected as the counterpart. As illustrated in Fig. 9a,  $\alpha$  is larger in both NARR and GPCP. Given that  $\alpha$  also indicates the degree to which the underlying process deviates from simple scaling (i.e., monofractal with  $\alpha = 0$ ), the larger the  $\alpha$ , the stronger the multifractality (scaling regimes). This result is expected because zero values in both NARR and GPCP are replaced with very small values, thus introducing an artificial threshold. The intermittency and effective maximum singularity are underestimated, as shown in Figs. 9b and 9c, and they further confirm the results shown in Figs. 6 and 7. More importantly, NARR parameters exhibit a bias as compared to NCDC; whereas the scatter is large for GPCP parameters, and there is a large number of points that show no correlation to NCDC (see the  $C_1$  plot in the [0.3, 0.4] range for GPCP). We mapped the geographical location of those grid cells and found they correspond to the upper half of the region of the study, suggesting there is little useful information on rainfall extremes in the GPCP dataset north of the 40°N circle. This may be mainly attributed to the spotty (less frequent) satellite sampling between 40° and 50°N over the western United States (west of 105°W), as illustrated in Fig. 5.19 of Tanré et al. (2008). By contrast, the close agreement with NCDC over the Southwest regions, as illustrated in Figs. 7g–7i, corresponds to the regions where satellite revisits are frequent (that is, the sampling is good; see Fig. 5.19 of Tanré et al. 2008). The GPCP rainfall products were originally obtained by merging infrared and microwave satellite estimates with rain gauge data (details of the data used in GPCP are available online at [http://precip.gsfc.nasa.gov/gpcp\\_daily\\_comb.html](http://precip.gsfc.nasa.gov/gpcp_daily_comb.html)). The low-

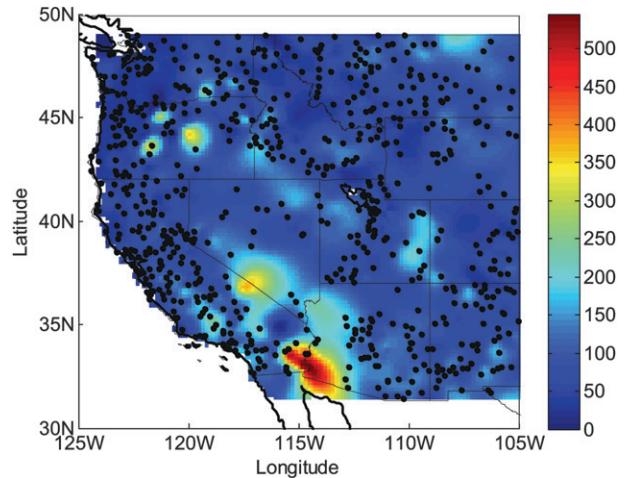


FIG. 8. Normalized extreme NCDC rainfall (maximum daily rainfall over mean rainfall) map in the study domain from 1950 to 2006 (unitless).

sampling issue in the low-earth-orbit infrared (leo-IR) is most apparent around the Indian Ocean before MeteoSat 5 began operations in July 1998, whereas for the Television and Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS)/Atmospheric Infrared Sounder (AIRS), both of which are also used in GPCP products, the issue is most visible along 40°N and 40°S (G. J. Huffman 2009, personal communication), which may also explain the highly biased values of  $\alpha$  in Figs. 7g and 10g. Moreover, the differences between the IR and TOVS–AIRS datasets required that the One-Degree Daily (1DD) GPCP be formulated in two parts, with smoothing over the latitude band 40°–50° in each hemisphere to patch the data boundary (available online at [http://precip.gsfc.nasa.gov/gpcp\\_daily\\_comb.html](http://precip.gsfc.nasa.gov/gpcp_daily_comb.html)). This smoothing can potentially render the GPCP dataset farther north of 40°N in our study domain of little use to characterize rainfall extremes. Note the north boundary of our study domain does not go beyond 50°N. Because

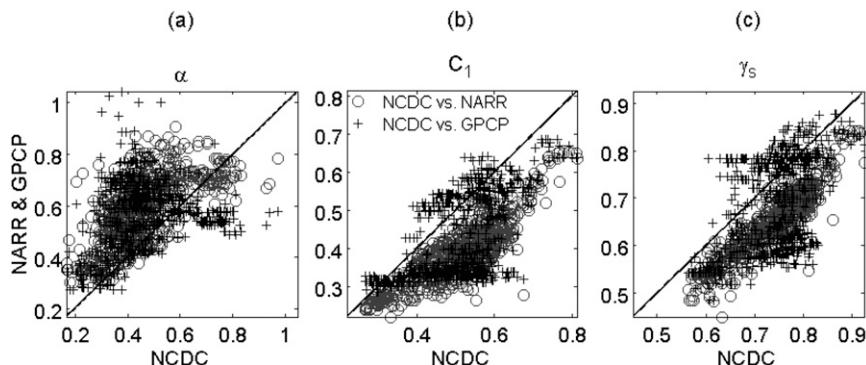


FIG. 9. Scatterplots of estimated parameters from NCDC vs NARR and GPCP.

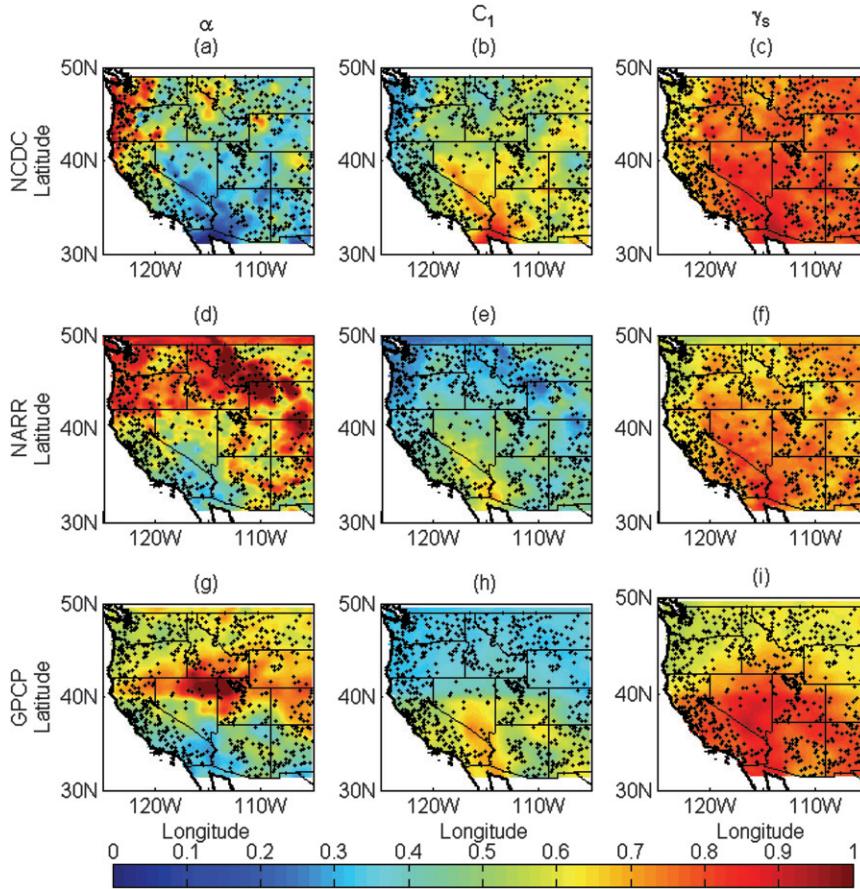


FIG. 10. Same as Fig. 7, but within the same time window (1997–2006) at the same resolution (daily).

rainfall measurement from space is the only path to consistent global-scale monitoring, this study illustrates the critical need for higher spatial and temporal resolution as well as middle and higher latitude satellite-based observations.

Finally, the spatial patterns of the multifractal parameter estimates in the same time window (1997–2008) at the same temporal resolution (daily) are shown in Fig. 10. There is generally good agreement between Figs. 7 and 10, except in Fig. 10d, which displays much higher NARR  $\alpha$  estimates throughout the domain. As mentioned earlier, this could be explained by the aggregation process, or by generic difficulties on accurately estimating  $\alpha$ , further exacerbated to some extent by using a much shorter time series, although finite sampling bias is not severe for  $\alpha < 1$ , as discussed in section 2. To explore the effect of spatial averaging on the values of the universal multifractal parameters, GPCP grid cells with at least eight NCDC rain gauges were selected for further analysis (Table 1). Specifically, parameters were calculated for the time series average of

all rain gauges (box mean), and these parameters were compared against the mean of the parameters calculated for each rain gauge separately (station mean). The GPCP parameters for each grid cell are also displayed in Table 1. The box mean  $C_1$  and  $\gamma_s$  are lower than the station mean, as expected, because of averaging, but they are also substantially lower than the GPCP values, which shows that there is added statistical value in the satellite-based data even if gauge-based corrections are necessary (Huffman 1997). The mix of data assimilated in the NARR varies as a function of time, depending on data availability (Mesinger et al. 2006); therefore, one might expect changes in the statistics of rainfall fields that reflect the changes in underlying data feed. Given that the most recent results deviate more from the NCDC control in the central and north-central sectors of the region of study, this suggests that the assimilated product appears to have deteriorated there in recent years. This provides an example of how to use multifractal parameters as diagnostics to evaluate models and satellite algorithms, for example.

TABLE 1. Parameter comparison among station means, box means, and GPCP.

Grid cell or box		$\alpha$			$C_1$			$\gamma_s$		
Latitude (°)	Longitude (°)	Station mean	Box mean	GPCP	Station mean	Box mean	GPCP	Station mean	Box mean	GPCP
34.5	-118.5	0.3161	0.4088	0.3089	0.5803	0.4992	0.6249	0.7649	0.718	0.8066
36.5	-121.5	0.4561	0.5649	0.3225	0.4833	0.3956	0.6535	0.7178	0.6578	0.8373
38.5	-121.5	0.4407	0.55	0.4238	0.4737	0.3856	0.6379	0.7020	0.6408	0.8524
38.5	-122.5	0.5047	0.5938	0.3947	0.4311	0.3712	0.6211	0.6778	0.6384	0.831
39.5	-111.5	0.3815	0.592	0.6545	0.5694	0.4164	0.6689	0.7790	0.6903	0.9112
39.5	-120.5	0.4201	0.5673	0.5866	0.4461	0.3482	0.6179	0.6621	0.6006	0.8702
40.5	-111.5	0.4551	0.6284	0.633	0.5055	0.3879	0.506	0.7401	0.6692	0.7908
45.5	-123.5	0.7350	0.8494	0.4472	0.2936	0.23	0.4436	0.5839	0.5276	0.6711
47.5	-121.5	0.7734	0.9069	0.4787	0.2992	0.2254	0.4342	0.6030	0.5369	0.6722

Following Douglas and Barros (2003), who found significant differences in the codimension and multifractal indices of time series from rainfall gauges before and after 1970 east of 105°W, the NCDC and NARR datasets were subsequently divided in 10- and 20-yr subsets to investigate whether interdecadal variability could be detected, but statistically significant differences were not found. Although there might be an explanation related to the limited resolution of the datasets, or the narrower range of variability spanned by these parameters in the western United States, these results are consistent with previous results by Groisman et al. (2004) and Smith (2008).

## 6. Conclusions and discussion

In this study, the statistics of extreme rainfall of NCDC, NARR, and GPCP datasets in the western United States were evaluated using the framework of universal multifractals. It was demonstrated (not unexpectedly) that the NCDC multifractal parameters exhibit spatial patterns that are consistent with the characteristics of regional climates, including orographic enhancement and rain shadow effects that are the dominant mechanism of rainfall modulation in the northwest and along the West Coast. Therefore, we use the NCDC results as the control for the evaluation of the NARR and GPCP datasets. The spatial patterns of multifractal parameters for NARR closely reproduce the spatial patterns for NCDC, but the frequency of large magnitude events, the magnitude of extremes, and the intermittency are underestimated overall. In particular, NARR values of  $\alpha$  north of 40°N significantly higher than those for NCDC indicate less frequency of large magnitude fluctuations, and thus they suggest that major storm events are either missed or strongly underestimated. For GPCP, there is good agreement with NCDC in the Southwest, but there is no match for the Northwest and central-west regions north of 40°N, which was attributed to satellite sampling limitations. The immediate implication, of course, is that caution must be

exercised in using NARR or GPCP data as reference in climate impact studies, especially with regard to the analysis of changes in the frequency and magnitude of extreme events.

One implication of this study for evaluating–verifying both data assimilation and satellite rainfall retrieval systems is that besides using standard categorical and statistical measures of quality in comparing rainfall fields against ground observations (rain gauge and radar data), which focus on mean statistical behavior (e.g., Barros and Tao 2008; Tao and Barros 2008), there is a role to be played by multifractal parameters as a means to incorporate quantitative metrics for statistical extremes. Accounting for the multiscaling properties of the spatial rain gauge networks by multifractal objective analysis (Salvadori et al. 2001) should further improve our understanding of the statistics of extremes.

Moreover, from the point of view of assessing the utility of prognostic simulations of climate change scenarios, it might be more important that climate models can reproduce the statistics of extremes of current climate rather than the mean behavior of rainfall. This study provides an alternative statistical narrative to inform the use of reanalysis products and satellite datasets to assess the effect of climate change on extreme rainfall.

*Acknowledgments.* The authors thank three anonymous reviewers for their helpful comments on an earlier version of this paper. This research was funded in part by the National Aeronautics and Space Administration (NASA) under Grant NNX07AK40G and by the National Science Foundation (NSF) Grant EAR-0711430.

## APPENDIX

### On the Parameter Estimation of Universal Multifractals

Consistent with Eq. (1), the probability distribution multiple scaling (PDMS) technique using histograms to

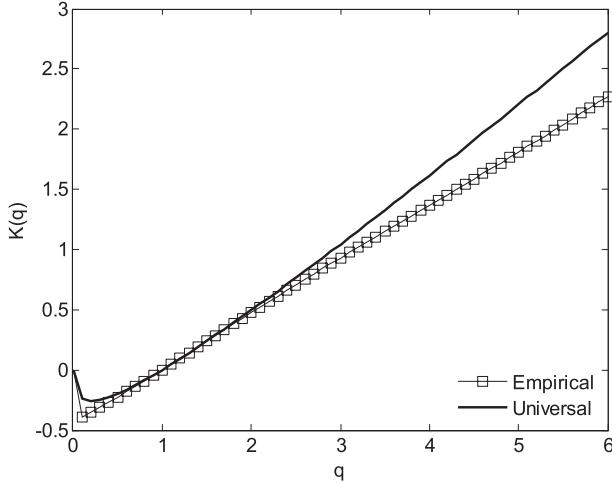


FIG. A1. Example  $K(q)$  function for one NCDC daily station.

directly estimate  $c(\gamma)$  as the slope of  $\log[\text{Pr}(\varepsilon_\lambda > \lambda^\gamma)]$  versus  $\log(\lambda)$  with  $\gamma = \log(\varepsilon_\lambda)/\log(\lambda)$  was used in earlier multifractal studies (Schertzer and Lovejoy 1989; Lavallée et al. 1991a). However, PDMS can fail to converge to a solution of  $\alpha$  and  $C_1$  when nonlinear curve fitting algorithms are applied to Eq. (4) as a result of the high degree of correlation between these two parameters (Tessier et al. 1994).

Taking the conserved multifractal density  $\varepsilon_\lambda$  at the homogeneous resolution with scale ratio  $\lambda'$  (the ratio of the largest scale to the scale of homogeneity), the (single) trace moment (Schertzer and Lovejoy 1987) method estimates  $K(q)$  using the following relationship:

$$Tr_\lambda(\varepsilon_\lambda)^q = \left\langle \sum_i \prod_{\lambda'}^q (B_{\lambda,i}) \right\rangle \sim \lambda^{K(q)-(q-1)D}, \quad (\text{A1})$$

where  $Tr_\lambda(\varepsilon_\lambda)^q$  is the trace moment of order  $q$ ,  $\prod_{\lambda'}(B_\lambda) = \int_{B_\lambda} \varepsilon_\lambda d^D x$ ,  $\lambda < \lambda'$  is the integration over all intervals  $B_\lambda$  (e.g., “balls” or cubes of characteristic scale ratio  $\lambda$ ) that cover the region of the embedding space occupied by the field, and  $D$  is the dimension of the embedding space.

In practice, the multifractal analysis of empirical data is constrained by insufficient sampling (finite number of observations), or by the magnitude of the maximum singularity measured over the period of observation if there are enough samples. Consequently, Eq. (A1) does not hold for  $q > q_{\text{cr}}$ , where the critical value corresponds to the order of the moment associated with the *maximum measured* singularity beyond which the statistical moments diverge (the so-called first-order phase transition  $q_{\text{cr}} = q_D$ ), or it corresponds to the maximum *measurable* singularity for the case of finite samples (the second-order phase transition  $q_{\text{cr}} = q_s$ ; Schertzer and Lovejoy 1987; Schertzer et al. 1993; Seuront et al. 1999).

Formally, the first- and second-order phase transitions correspond to first- and second-order continuity breaks of  $K(q)$ , respectively. The upper bound of  $q_D$  can be obtained by solving

$$K(q_D) = (q_D - 1)D, \quad (\text{A2})$$

and  $q_s = c'(\gamma_s) = [(D + D_s)/C_1]^{1/\alpha}$  with  $\gamma_s = c^{-1}(D + D_s)$  and  $D_s = \log(N_s)/\log(\lambda)$  ( $N_s$  is the number of independent samples; Tessier et al. 1993). Thus, although this method allows for the determination of  $K(q)$  for small  $q$ , poor nonlinear regression will be involved for relatively large  $q$  values (therefore, the estimation of  $\alpha$  and  $C_1$ ).

The double trace moment (DTM; Lavallée et al. 1991b) technique avoids this problem by introducing a second moment  $\eta$  with the transformation of the high-resolution field  $\varepsilon_\lambda$  to  $\varepsilon_\lambda^\eta$  in Eq. (A1) as follows:

$$Tr_\lambda(\varepsilon_\lambda^\eta)^q = \left\langle \sum_i \left[ \prod_{\lambda'}^{(\eta)} (B_{\lambda,i}) \right]^q \right\rangle \sim \lambda^{K(q,\eta)-(q-1)D}, \quad (\text{A3})$$

where  $\prod_{\lambda'}^{(\eta)}(B_\lambda) = \int_{B_\lambda} \varepsilon_\lambda^\eta d^D x$  and the double scaling exponent  $K(q, \eta) = K(q\eta) - qK(\eta) = \eta^\alpha K(q)$  with  $K(q, 1) = K(q)$ . Therefore,  $\alpha$  is the slope of  $\log[K(q, \eta)]$  against  $\log(\eta)$ . Compared with other empirical methods to estimate extremes from a small number of isolated values in a time series, the DTM technique is more robust because the parameters are determined by fitting over the entire scaling range. Furthermore, double trace moments can be calculated for different  $\eta$  values, and thus the parameter estimation will be more accurate. Using 25 independent numerical simulations with 1024 points in each realization, Lavallée (1991) showed that  $\alpha$  can be approximated to an accuracy within the range  $[-0.1, 0.1]$ .

Nevertheless, the DTM approach only yields unbiased estimates of  $\alpha$  for the homogeneous field at high temporal and spatial resolution to capture all relevant scale interactions (Veneziano and Furcolo 1999)—while experimentally measured quantities are necessarily averaged or integrated over a finite inhomogeneous scale that depends on the (limited) resolution of geophysical sensors—that is the measurement scale. To address this problem, Veneziano and Furcolo (1999) proposed a modified double trace moment (MDTM) method that theoretically scales well independently of measurement scale. However, although the theoretical underpinnings of MDTM suggest its application should be more general, and that the MDTM was used in the past for a related study (Douglas and Barros 2003), recent extensive testing of the method shows that it does not yield consistently reliable results for actual rainfall observations, which are inherently noisy and with relative coarse temporal

resolutions [more details are discussed in Sun and Barros (2008)]. The DTM method was therefore employed here as an approximation within the following constraints:

- (i) Because  $C_1$  may be underestimated because of the negative bias of  $K(q, \eta)$  if  $\eta < 1$  (Veneziano and Furcolo 1999),  $\eta = 1$  was used as the lower bound in this study.
- (ii) Because  $K(q, 1)$  can be approximated almost without bias by DTM (Veneziano and Furcolo 1999),  $C_1$  should be calculated using  $K(q, 1)$  and Eq. (5) instead of  $K(q, \eta) = \eta^\alpha K(q)$ , with  $\eta$  as any other values.
- (iii) The relationship  $\max(\eta, q\eta) \leq q_{cr} = \min(q_s, q_D)$  must be verified to avoid either moment divergence or spurious scaling, resulting from the breakdown of Eqs. (A1) and (A3), that is, the two phase transitions (Schertzer and Lovejoy 1987). On the other hand, as illustrated in Fig. A1 and reported elsewhere (Tessier et al. 1994; Olsson 1995; Harris et al. 1997),  $q$  also cannot be too small to avoid a linear offset of  $K(q)$  due to the significance of measurement noise, or limitation to low-order moments. Because  $q_{cr} \approx 2.5$  and  $q_{min} \approx 0.7$  in this study (see Fig. 1 for an example),  $q$  was set as 1.2 with the upper bound of  $\eta$  at 2, such that  $\max(\eta, q\eta) < q_{cr}$  and  $\min(\eta, q\eta) > q_{min}$ .
- (iv) Because of the asymptotic behavior of  $K(q, \eta)$  as  $\eta \rightarrow \infty$  caused by finite sampling [see Veneziano and Furcolo (1999) for a theoretical proof and numerical validation],  $\alpha$  should be estimated from the first nearly linear part of the log—log plot of  $K(q, \eta)$  versus  $\eta$ .
- (v) Although large aggregation-scale ranges may improve the accuracy of DTM (Veneziano and Furcolo 1999), scaling breaks either associated with actual physical processes or artificially introduced by measurement limitation (lack of ability to detect weak signals) need to be considered (Fraedrich and Larnder 1993). Because the scaling break at 16 days identified in previous studies of rainfall time series (e.g., Tessier et al. 1996; Royer et al. 2008) is physically consistent with the characteristic period of a planetary Rossby wave, this study was also carried out within this scaling regime, that is, from 1 to 16 days, which spans the time scales of interest for a risk assessment of natural hazards (Barros and Evans 1997; Douglas and Barros 2003).

#### REFERENCES

- Adler, R. F., and Coauthors, 2003: The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.*, **4**, 1147–1167.
- Allan, R. P., and B. J. Soden, 2008: Atmospheric warming and the amplification of precipitation extremes. *Science*, **321**, 1481–1484.
- Barros, A. P., and J. L. Evans, 1997: Designing for climate variability. *J. Prof. Issues Eng. Educ. Pract.*, **123**, 62–65.
- , and K. Tao, 2008: A space-filling algorithm to extrapolate narrow-swath instantaneous TRMM microwave rain-rate estimates using thermal IR imagery. *J. Atmos. Oceanic Technol.*, **25**, 1901–1920.
- Bernardara, P., D. Schertzer, E. Sauquet, I. Tchiguirinskaia, and M. Lang, 2008: The flood probability distribution tail: How heavy is it? *Stochastic Environ. Res. Risk Assess.*, **22**, 107–122.
- Bindlish, R., and A. P. Barros, 1996: Aggregation of digital terrain data using a modified fractal interpolation scheme. *Comput. Geosci.*, **22**, 907–917.
- Boufadel, M. C., S. Lu, F. J. Molz, and D. Lavallée, 2000: Multifractal scaling of the intrinsic permeability. *Water Resour. Res.*, **36**, 3211–3222.
- Carvalho, L. M. V., D. Lavallée, and C. Jones, 2002: Multifractal properties of evolving convective systems over tropical South America. *Geophys. Res. Lett.*, **29**, 1743, doi:10.1029/2001GL014276.
- Chen, C. T., and T. Knutson, 2008: On the verification and comparison of extreme rainfall indices from climate models. *J. Climate*, **21**, 1605–1621.
- Cowpertwait, P. S. P., 1995: A generalized spatial-temporal model of rainfall based on a clustered point process. *Proc. Roy. Soc. London*, **A450**, 163–175.
- , 1998: A Poisson-cluster model of rainfall: High-order moments and extreme values. *Proc. Roy. Soc. London*, **A454**, 885–898.
- Davis, A., A. Marshak, W. Wiscombe, and R. Cahalan, 1994: Multifractal characterizations of nonstationarity and intermittency in geophysical fields: Observed, retrieved, or simulated. *J. Geophys. Res.*, **99**, 8055–8072.
- 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.
- Dooge, J. C. I., 1986: Looking for hydrologic laws. *Water Resour. Res.*, **22**, 46S–58S.
- Douglas, E. M., and A. P. Barros, 2003: Probable maximum precipitation estimation using multifractals: Applications to the eastern United States. *J. Hydrometeorol.*, **4**, 1012–1024.
- Easterling, D. R., J. L. Evans, P. Ya. Groisman, T. R. Karl, K. E. Kunkel, and P. Ambenje, 2000: Observed variability and trends in extreme climate events: A brief review. *Bull. Amer. Meteor. Soc.*, **81**, 417–425.
- Fekete, B. M., C. J. Vörösmarty, J. O. Roads, and C. J. Willmott, 2004: Uncertainties in precipitation and their impacts on runoff estimates. *J. Climate*, **17**, 294–304.
- Fraedrich, K., and C. Larnder, 1993: Scaling regimes of rainfall time series. *Tellus*, **45A**, 289–298.
- Frisch, U., and G. Parisi, 1985: Fully developed turbulence and intermittency. *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, M. Ghil, R. Benzi, and G. Parisi, Eds., Elsevier, 84–88.
- Groisman, P. Ya., R. W. Knight, T. R. Karl, D. R. Easterling, B. Sun, and J. H. Lawrimore, 2004: Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations. *J. Hydrometeorol.*, **5**, 64–85.
- , —, D. R. Easterling, T. R. Karl, G. C. Hegerl, and V. N. Razuvaev, 2005: Trends in intense precipitation in the climate record. *J. Climate*, **18**, 1326–1350.

- Harris, D., A. Seed, M. Menabde, and G. Austin, 1997: Factors affecting multiscaling analysis of rainfall time series. *Nonlinear Processes Geophys.*, **4**, 137–155.
- Hegerl, G. C., T. R. Karl, M. Allen, N. L. Bindoff, N. Gillett, D. Karoly, X. Zhang, and F. Zwiers, 2006: Climate change detection and attribution: Beyond mean temperature signals. *J. Climate*, **19**, 5058–5077.
- Hubert, P., and Coauthors, 1993: Multifractals and extreme rainfall events. *Geophys. Res. Lett.*, **20**, 931–934.
- Huffman, G. J., 1997: Estimates of root-mean-square random error for finite samples of estimated precipitation. *J. Appl. Meteor.*, **36**, 1191–1201.
- , R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.*, **2**, 36–50.
- Hunt, B. G., 2006: Climatological extremes of simulated annual mean rainfall. *J. Climate*, **19**, 5289–5303.
- Kavvas, L., and J. W. Delleur, 1975: The stochastic and chronological structure of rainfall sequences: Application to Indiana. Purdue University, Water Resources Research Center Rep. 57, 199 pp.
- Kharin, V. V., and F. W. Zwiers, 2002: Climate predictions with multimodel ensembles. *J. Climate*, **15**, 793–799.
- Kim, G., and A. P. Barros, 2002: Downscaling of remotely-sensed soil moisture with a modified fractal interpolation method using contraction mapping and ancillary data. *Remote Sens. Environ.*, **83**, 400–413.
- Kuligowski, R. J., and A. P. Barros, 1998: Using artificial neural networks to estimate missing rainfall data. *J. Amer. Water Resour. Assoc.*, **34**, 1437–1447.
- , and —, 2001: Blending multiresolution satellite data with application to the initialization of an orographic precipitation model. *J. Appl. Meteor.*, **40**, 1592–1606.
- Lavallée, D., 1991: Multifractal analysis and simulation technique and turbulent fields. Ph.D. thesis, McGill University, 133 pp.
- , S. Lovejoy, and D. Schertzer, 1991a: On the determination of the codimension function. *Non-Linear Variability in Geophysics: Scaling and Fractals*, D. Schertzer and S. Lovejoy, Eds., Kluwer Academic, 99–109.
- , —, and —, 1991b: Universal multifractal theory and observations of land and ocean surfaces, and of clouds. *Wave Propagation and Scattering in Varied Media II*, V. K. Varadan, Ed., International Society for Optical Engineering (SPIE Proceedings, Vol. 1558), 60–75.
- Lovejoy, S., and D. Schertzer, 1995: Multifractals and rain. *New Uncertainty Concepts in Hydrology and Water Resources*, Z. W. Kundzewicz, Ed., Cambridge University Press, 62–103.
- , and —, 2009a: On the simulation of continuous in scale universal multifractals, Part I: Spatially continuous processes. *Comput. Geosci.*, in press.
- , and —, 2009b: On the simulation of continuous in scale universal multifractals, Part II: Space-time processes and finite size corrections. *Comput. Geosci.*, in press.
- , —, and Y. Tchigirinskaya, 2007: Scaling and extremes in precipitation and streamflow. *Eos, Trans. Amer. Geophys. Union*, Abstract H41A-06.
- Mandelbrot, B. B., 1974: Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid Mech.*, **62**, 331–350.
- Mesinger, F., and Coauthors, 2006: North American regional reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360.
- Milly, P. C. D., R. T. Wetherald, K. A. Dunne, and T. L. Delworth, 2002: Increasing risk of great floods in a changing climate. *Nature*, **415**, 514–517.
- Novikov, E. A., and R. Stewart, 1964: Intermittency of turbulence and spectrum of fluctuations in energy-dissipation. *Izv. Akad. Nauk. SSSR., Ser. Geofiz.*, **3**, 408–412.
- Olsson, J., 1995: Limits and characteristics of the multifractal behavior of a high-resolution rainfall time series. *Nonlinear Processes Geophys.*, **2**, 23–29.
- Palmer, T. N., and J. Räisänen, 2002: Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature*, **415**, 512–514.
- Perkins, S. E., A. J. Pitman, N. J. Holbrook, and J. McAneney, 2007: Evaluation of the AR<sub>4</sub> climate model's simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J. Climate*, **20**, 4356–4376.
- Royer, J. F., A. Biaou, F. Chauvin, D. Schertzer, and S. Lovejoy, 2008: Multifractal analysis of the evolution of simulated precipitation over France in a climate scenario. *C. R. Geosci.*, **340**, 431–440.
- Salvadori, G., D. Schertzer, and S. Lovejoy, 2001: Multifractal objective analysis: Condition and interpolation. *Stochastic Environ. Res. Risk Assess.*, **15**, 261–283.
- Schertzer, D., and S. Lovejoy, 1987: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *J. Geophys. Res.*, **92**, 9693–9714.
- , and —, 1989: Nonlinear variability in geophysics: Multifractal analysis and simulation. *Fractals' Physical Origin and Properties*, L. Pietronero, Ed., Ettore Majorana International Science Series, Vol. 45, Plenum Press, 49–79.
- , and —, 1992: Hard and soft multifractal processes. *Physica A*, **185**, 187–194.
- , —, and D. Lavallée, 1993: Generic multifractal phase transitions and self-organized criticality. *Cellular Automata: Prospects in Astrophysical Applications*, J. M. Perdag and A. Lejeune, Eds., World Scientific, 216–227.
- Schnur, R., 2002: Climate science: The investment forecast. *Nature*, **415**, 483–484.
- Seuront, L., F. Schmitt, Y. Lagadeuc, D. Schertzer, and S. Lovejoy, 1999: Universal multifractal analysis as a tool to characterize multiscale intermittent patterns: Example of phytoplankton distribution in turbulent coastal waters. *J. Plankton Res.*, **21**, 877–922.
- Smith, R. L., cited 2008: Climate extremes and global warming: A statistician's perspective. [Available online at <http://apha.confex.com/apha/136am/webprogram/Paper189287.html>.]
- Sun, X., and A. P. Barros, 2008: Practical challenges on the estimation of universal multifractal parameters of rainfall observations. *Eos, Trans. Amer. Geophys. Union*, **89** (Fall Meeting Suppl.), Abstract NG33A-1209.
- Tanré, D., P. Artuxo, S. Yuter, and Y. Kaufman, 2008: In situ and remote sensing techniques for measuring aerosols, clouds, and precipitation. *Aerosol Pollution Impacts on Precipitation: A Scientific Review*, Z. Levin, and W. R. Cotton, Eds., Springer, 143–203.
- Tao, K., and A. P. Barros, 2008: Metrics to describe the dynamical evolution of atmospheric moisture: Intercomparison of model (NARR) and observations (ISCCP). *J. Geophys. Res.*, **113**, D14125, doi:10.1029/2007JD009337.
- Tessier, Y., S. Lovejoy, and D. Schertzer, 1993: Universal multifractals: Theory and observations for rain and clouds. *J. Appl. Meteor.*, **32**, 223–250.

- , —, and —, 1994: Multifractal analysis and simulation of the global meteorological network. *J. Appl. Meteor.*, **33**, 1572–1586.
- , —, P. Hubert, D. Schertzer, and S. Pecknold, 1996: Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions. *J. Geophys. Res.*, **101**, 26 427–26 440.
- Veneziano, D., and P. Furcolo, 1999: A modified double trace moment method of multifractal analysis. *Fractals*, **7**, 181–195.
- Waymire, E. C., and V. K. Gupta, 1981a: The mathematical structure of rainfall representation: 1. A review of the stochastic rainfall models. *Water Resour. Res.*, **17**, 1261–1272.
- , and —, 1981b: The mathematical structure of rainfall representation: 2. A review of the point process theory. *Water Resour. Res.*, **17**, 1273–1285.
- , and —, 1981c: The mathematical structure of rainfall representation: 3. Some applications of the point process theory to rainfall processes. *Water Resour. Res.*, **17**, 1286–1294.
- Wehner, M. F., 2004: Predicted twenty-first-century changes in seasonal extreme precipitation events in the parallel climate model. *J. Climate*, **17**, 4281–4290.
- Wilby, R. L., and T. M. L. Wigley, 2002: Future changes in the distribution of daily precipitation totals across North America. *Geophys. Res. Lett.*, **29**, 1135. doi:10.1029/2001GL013048.
- Xie, P., and P. A. Arkin, 1996: Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Climate*, **9**, 840–858.
- , and —, 1998: Global monthly precipitation estimates from satellite-observed outgoing longwave radiation. *J. Climate*, **11**, 137–164.
- , and Coauthors, 2003: GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *J. Climate*, **16**, 2197–2214.
- Yaglom, A. M., 1966: The influence of fluctuations in energy dissipation on the shape of turbulence characteristics in the inertial interval. *Sov. Phys. Dokl.*, **11**, 26–29.
- Zwiers, F., and V. V. Kharin, 1998: Changes in the extremes of the climate simulated by CCC GCM2 under CO<sub>2</sub> doubling. *J. Climate*, **11**, 2200–2222.
- , and G. Hegerl, 2008: Attributing cause and effect. *Nature*, **453**, 296–297.

## CORRIGENDUM

XIAOMING SUN AND ANA P. BARROS

*Department of Civil and Environmental Engineering, Pratt School of Engineering,  
Duke University, Durham, North Carolina*

---

There was a typographic error in the title of Sun and Barros (2010) in which an extra “and” was included. The correct title as it should have appeared is shown below.

**An Evaluation of the Statistics of Rainfall Extremes in Rain Gauge  
Observations, Satellite-Based and Reanalysis Products  
Using Universal Multifractals**

The staff of the *Journal of Hydrometeorology* regrets any inconvenience this error may have caused.

### REFERENCE

Sun, X., and A. P. Barros, 2010: An evaluation of the statistics of rainfall extremes in rain gauge observations, satellite-based and reanalysis products using universal multifractals. *J. Hydrometeor.*, **11**, 388–404.

---

E-mail: xiaoming.sun@duke.edu