AIP | Chaos
An Interdisciplinary Journal of Nonlinear Science
25TH ANNIVERSARY

Statistical scaling, Shannon entropy, and Generalized space-time q-entropy of rainfall fields in tropical South America

Germán Poveda and Hernán D. Salas

Citation: Chaos **25**, 075409 (2015); doi: 10.1063/1.4922595
View online: http://dx.doi.org/10.1063/1.4922595
View Table of Contents: http://scitation.aip.org/content/aip/journal/chaos/25/7?ver=pdfcov
Published by the AIP Publishing

CHAOS **25**, 075409 (2015)

Statistical scaling, Shannon entropy, and Generalized space-time *q*-entropy of rainfall fields in tropical South America

Germán Poveda[a)] and Hernán D. Salas
Department of Geosciences and Environment, Universidad Nacional de Colombia, Sede Medellín, Colombia

(Received 23 December 2014; accepted 3 June 2015; published online 18 June 2015)

We study diverse scaling and information theory characteristics of Mesoscale Convective Systems (MCSs) as seen by the Tropical Rainfall Measuring Mission (TRMM) over continental and oceanic regions of tropical South America, and 2-D radar rainfall fields from Amazonia. The bi-dimensional Fourier spectra of MCSs exhibit inverse power laws with respect to the spatial scale, whose scaling exponents, β, capture the type of spatial correlation of rainfall among the study regions, including those over the Andes of Colombia as well as over oceanic and Amazonian regions. The moment-scaling analysis evidences that the structure function deviates from simple scaling at order $q > 1.0$, thus signaling the multi-scaling nature of rainfall fields within MCSs in tropical South America, with departures from simple scaling associated with the physical characteristics of MCSs over the different study regions. Entropy is estimated for a large set of radar rainfall fields during the distinctive atmospheric regimes (Easterly and Westerly events) in this part of Amazonia. Results evidence that there are significant differences in the dynamics of rainfall among regimes. No clear-cut relationship is found between entropy and the first two statistical moments, but power fits in space and time, $S(\gamma) \sim \gamma^{-\eta}$ for skewness and, $S(\kappa) \sim \kappa^{-\epsilon}$ for kurtosis. The exponents η and ϵ are statistically different between Easterly and Westerly events, although the significance of fits is less when L-moments are used to estimate skewness and kurtosis. Interesting differences are identified between the time and space generalized q-entropy functions of Amazonian rainfall fields. In both cases, the functions are a continuous set of power laws (analogous to the *structure function* in turbulence), $S(T, q) \sim T^{\beta}$, and, $S(\lambda, q) \sim \lambda^{\beta}$, covering a broad range of temporal and spatial scales. Both time and space generalized q-entropy functions exhibit linear growth in the range $-1.0 < q < -0.5$, and saturation of the exponent β for $q \geq 1.0$. In the case of the spatial analysis, the exponent saturates at $\langle\beta\rangle \sim 1.0$, whereas at $\langle\beta\rangle = 0.5$ for the temporal case. Results are similar for time series extracted from the S-POL radar and time series of rainfall in tropical Andes. Additionally, differences in values of $\langle\beta\rangle$ for $q \geq 1.0$ between Easterly and Westerly events are not statistically significant.
[http://dx.doi.org/10.1063/1.4922595]

Understanding, modelling, and predicting tropical rainfall constitute a fundamental challenge of current geosciences. Its convective origin contributes to explain its strong variability and intermittence in space and time. Tropical South America houses one of the rainiest regions on Earth over western Colombia, but also over the Amazon River basin, of global hydro-climatological importance. Here, we apply diverse tools from the frameworks of Statistical Scaling and Information Theory to quantify the dynamics and variability of rainfall in tropical South America over a broad range of scales in space and time. For estimation purposes, we use information from two different data sets: (1) the largest and more intense Mesoscale Convective Systems (MCSs) developed over seven selected oceanic and continental regions of tropical South America, identified by the Tropical Rainfall Measuring Mission (TRMM) during the period 1998–2011 at 5-km resolution and (2) series of 10-min radar rainfall fields gathered during the January–February 1999 Amazon Wet Season Atmospheric Meso-scale Campaign, at 2-km resolution. Results show that the decay of the 2-D Fourier spectrum is best represented by a power fit, $E(k) \sim ck^{-\beta}$, between the spatial scales of 5 km and 160 km, with values of $\beta = [1.46, 1.64]$ that capture different features of the MCSs. Besides, moment-scaling results evidence the multi-scaling spatial structure of MCSs, a result that has important implications in rainfall downscaling models. On the other hand, we estimate the Shannon entropy for the radar rainfall fields during the Easterly (E) and Westerly (W) atmospheric regimes in central Amazonia (CA). Results show significant differences during both regimes. We also show that the relationship between entropy and skewness and kurtosis of the radar rainfall fields in space and time are best represented by a power fit, whose exponents also differ among the E and W regimes. We introduce and estimate the time and space generalized *q*-entropy (SGE) functions in these Amazonian rainfall fields. In both cases, the functions are a continuous set of power laws (analogous to the *structure function* in turbulence), such that $S(T, q) \sim T^{\beta}$, and, $S(\lambda, q) \sim \lambda^{\beta}$, covering a large range of temporal and spatial scales. Both the time and space generalized *q*-entropy functions exhibit a linear growth in the range $-1.0 < q < -0.5$, and a saturation of the exponent β for

[a)]gpoveda@unal.edu.co

$q \geq 1.0$. In the case of the spatial analysis, the exponent saturates at $\langle\beta\rangle \sim 1.0$, whereas at $\langle\beta\rangle = 0.5$ for the temporal case. Our results reinforce the need to linking these statistical analyses with the space-time (thermo-) dynamics of rainfall in tropical South America.

I. INTRODUCTION

A. Rainfall in tropical South America

Understanding and modelling tropical rainfall constitute a fundamental challenge of current geosciences, given its strong variability and intermittence in space and time, and given its multiple implications in hydrology, ecology, climatology, and water resources, among others. In particular, tropical South America provides an interesting geographical setting to advance such an understanding owing to (i) the presence of the Andes that induces strong topographic and orographic effects to control and focus shallow and deep convection;[21,51] (ii) the strong influence of the diurnal cycle;[38,51] (iii) the size, location, and the tropical rainforest of the Amazon River basin exert a strong control over the hydro-climatology of most of South America;[36,39,53] (iv) the latitudinal migration of the intertropical convergence zone,[52] and (v) the strong activity of MCS over Amazonia and the far eastern Pacific.[22,48,75,83,84]

B. Statistical scaling

The theoretical framework of statistical scaling has contributed to diagnose and model the space-time dynamics of rainfall, using diverse tools, such as the bi-dimensional Fourier power spectrum,[18] moment-scaling analysis,[16,18,38] universal multifractals,[8,66] random cascades,[40] multiscale fluctuations,[42,43] and tests for the validity of Taylor's hypothesis.[21,50,79]

C. Information theory

Diverse information theory tools have been used to characterize the strong space-time intermittency of rainfall, including mutual information and the scale of information,[47] and the classical Boltzmann-Gibbs-Shannon definition of *entropy*.[7,26,47,62,63,76]

A recent study[47] has shown that entropy of high resolution rainfall records in the tropical Andes of Colombia grows as a power law with aggregation time, $S(T) \sim T^{\beta}$ with $\langle\beta\rangle = 0.5$, up to a timescale, T_{MaxEnt} (70 h–202 h), at which entropy saturates and tappers off ($\beta = 0$) thereafter. Scaling exponents ($\beta = 0.5$) are independent of record length. Maximum entropy is reached through a dynamic generalized Pareto distribution, whose parameters are not statistically sensitive to record length. On the other hand, information entropy for daily streamflows behaves as $S(T) \sim T^{\beta}$ with $\beta \gtrsim 0$.[58] The different scaling behaviour of entropy in rainfall and streamflows is due to the presence of zeroes in rainfall series, and by the rate at which they vanish upon temporal aggregation.[12,13,15,32,45,57,73]

The generalized Tsallis *q-entropy*, introduced by Tsallis[68] in the framework of non-extensive statistical mechanics, has been applied recently to study the temporal dynamics of rainfall and streamflows.[47,58] Tsallis q-entropies for rainfall time series of hourly rainfall records in the Andes of Colombia also exhibit power laws with T, such that $S_q(T) \sim T^{\beta(q)}$, with $\beta(q) \leq 0$ for $q \leq 0$, and $\beta(q) \simeq 0.5$ for $q \geq 1$. Reviews of the theoretical framework and applications of the Tsallis q-entropy are provided in recent studies.[3,70,71]

D. Objectives

In this study, we aim at studying the space-time characteristics of rainfall over tropical South America, using diverse tools from statistical scaling and information theory. In particular, we aim at

(1) Quantifying the bi-dimensional Fourier spectra and estimating the moment-scaling function of meso-scale convective systems as seen by the TRMM over diverse continental and oceanic regions of tropical South America,
(2) estimating the temporal dynamics of the Shannon entropy in 2-D Amazonian radar rainfall fields, considering Easterly and Westerly climatic regimes prevailing in the region,[10,44] and comparing the results with those for hourly rainfall series in the Colombian Andes,[47]
(3) quantifying the relations between the first statistical moments of the distribution function and entropy of Amazonian rainfall fields,
(4) investigating the behavior of the scaling properties of Amazonian rainfall under time aggregation, through the Time Generalized q-Entropy (TGE) proposed by Poveda,[47] and introducing the SGE to study the scaling properties of rainfall fields during the aforementioned climatic regimes.

II. STUDY REGIONS AND DATA SETS

A. Meso-scale convective systems as seen by TRMM

We study the scaling properties of MCS over seven continental and oceanic regions of tropical South America, exhibiting different rainfall dynamics and controls, namely, Caribbean Sea (CS), Pacific Ocean (PO), Colombia, Western Amazonia (WA), Central Amazonia, Eastern Amazonia (EA), and Southern Amazonia. See Fig. 1 and Table I.

Our study of MCSs is based on the Precipitation Radar (PR) and TRMM Microwave Imager (TMI) instruments on board the TRMM satellite, for a period spanning from January 1998 to December 2011. The PR horizontal resolution is 5.0 km. This study uses near-surface reflectivity and near-surface rainfall products from the 2A25 algorithm (version 7). A detailed description of the PR and TMI instruments on board the TRMM satellite is provided by Kummerow *et al.*,[28,29] and a full description of the TRMM datasets can be found at http://pps.gsfc.nasa.gov/Documents/filespec.TRMM.V7.pdf.

For each of the seven study regions, the largest MCSs of each year were selected out of 280 events during 2003–2011,

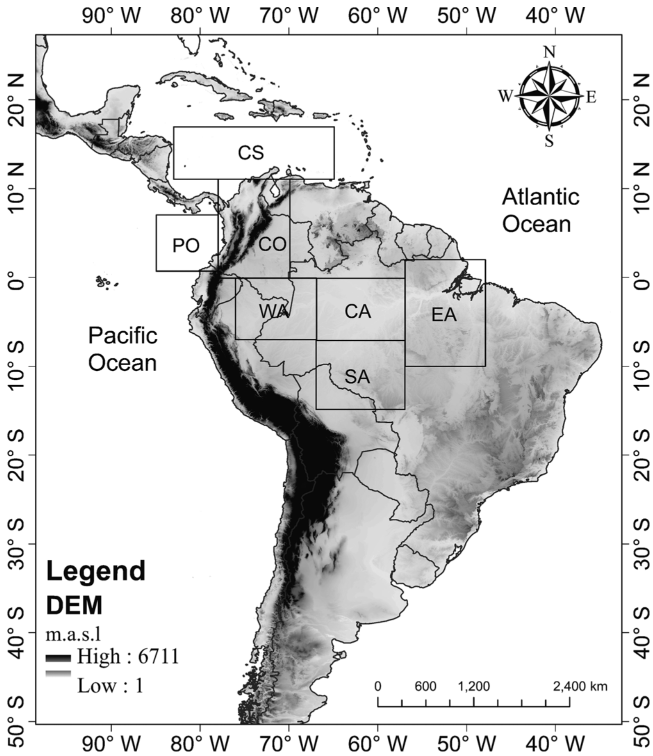

FIG. 1. Location of the study regions in tropical South America. CO: Colombia, PO: Pacific Ocean, WA: Western Amazonia, CA: Central Amazonia, EA: Eastern Amazonia, CA: Southern Amazonia, and CS: Caribbean Sea.

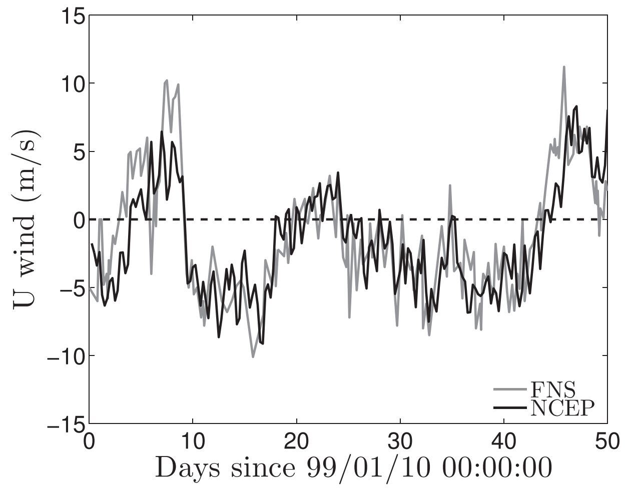

FIG. 2. Time series of the zonal component of the horizontal wind between 850 and 700 hPa in the region of the WETAMC/LBA Campaign, after two different data sources: NCEP/NCAR reanalysis project and the Fazenda Nossa Senhora (FNS) radiosonde.

which were classified by Jaramillo *et al.*[22] using criteria of maximum area (km^2) and maximum intensity (mm/h).

B. Radar rainfall fields

We study Amazonian rainfall fields using radar data gathered during the January–February 1999 Wet Season Atmospheric Meso-scale Campaign/LBA (WETAMC/LBA), which were designed to studying the dynamical, microphysical, electrical, and diabatic heating characteristics of tropical convection over southwestern Amazonia.[2,33,38,61] The WETAMC campaign was developed in the state of Rondônia (Brazil). Data used in this paper consisted on radar scans of storm intensities recorded by the S-POL radar (S-band, dual polarimetric) located at 61.9982°W, 11.2213°S. Data consist of 2-km resolution microwave band reflectivity which is directly related to rainfall intensity, over a circle of ~31 000 km^2. Scans produced by the Colorado State University Radar Meteorology Group were available every 7–10 min at the URL http://radarmet.atmos.colostate.edu/trmm-lba/rainlba.html.

Two distinctive predominant climate regimes have been identified in this part of Amazonia, which are characterized by the direction of wind speed between the 850 hPa and 700 hPa,[2,10,30] as[44]

(1) Easterly regime (E): Relatively dry atmospheric conditions, increased lightning activity, and presence of more intense and deeper convective systems.
(2) Westerly regime (W): Characterized by a diminished lightning activity, whereas convection is less deep and precipitation rates are less intense.

Additionally, information about zonal wind velocity at 700 hPa during the study period was obtained from the NCEP/NCAR Reanalysis (Fig. 2),[24] over the region between 61°W–62.8°W and 10.4°S–12.1°S, corresponding to the area covered by the S-POL radar. Data were obtained four times per day at 0000, 0600, 1200, and 1800 LST. Also, radiosonde data (62.37°W, 10.75°S) were used from the WETAMC campaign in Ouro Preto d'Oeste at the Fazenda

TABLE I. Rainfall intensity fields in meso-scale convective systems of largest areas during 2003–2010. (Columns 3 and 4) coordinates of the study regions; (columns 5 and 6) coordinates of the event centroid; (column 7) date: month/day/year; (column 8) hour of occurrence UTC; (column 9) minimum intensity (mm/h); (column 10) maximum intensity (mm/h); (columns 11–13) pre-factor, c; scaling exponent, β; and determination coefficient, R^2 of the power law fitted to the 2-D Fourier spectrum, respectively.

1 ID	2 Zone	3 Latitude	4 Longitude	5 Latitude	6 Longitude	7 Date	8 Hour	9 Min (mm/h)	10 Max (mm/h)	11 c	12 β	13 R^2
CA	Central Amazonia	7°S-0°N	67°W-57°W	−4.724	−66.880	01/25/2003	0000	0.861	58.2	0.007	1.487	0.957
EA	Eastern Amazonia	10°S-2°N	57°W-48°W	−0.826	−48.665	02/12/2010	0300	0.701	107.1	0.008	1.557	0.984
SA	Southern Amazonia	15°S-7°S	67°W-57°W	−9.106	−58.398	02/17/2009	0600	0.509	55.7	0.005	1.519	0.990
WA	Western Amazonia	7°S-0°N	76°W-67°W	−6.103	−67.182	11/20/2005	0600	0.423	85.4	0.005	1.644	0.991
PO	Pacific Ocean	1°N-7°N	85°W-78°W	5.534	−81.130	10/13/2006	0700	0.359	85.7	0.008	1.601	0.985
CO	Colombia	0°N-11°N	78°W-70°W	7.533	−75.929	06/25/2008	0400	0.549	75.6	0.007	1.464	0.966
CS	Caribbean Sea	11°N-17°N	83°W-65°W	13.598	−66.907	10/15/2008	1000	0.441	54.0	0.005	1.498	0.989

Nossa Senhora site, located inside the S-POL radar coverage region. Radiosonde data were obtained in the URL http://www.master.iag.usp.br/lba/.

III. METHODS

A. Bi-dimensional Fourier spectrum

The variability of rainfall can be investigated at multiple spatial scales through the 2-D Fourier power spectrum. We investigate whether the power spectrum exhibits a power law with respect to wave number, k, or whether 2-D rainfall behaves like a "1/f noise,"[35] such that

$$E(k) \sim ck^{-\beta}, \tag{1}$$

where c is the prefactor and β is the scaling exponent. Estimated values of β have a direct relationship with the Hurst exponent and with the kind of spatial persistence of the process.[9,11,17,27,35] The 2-D Fourier spectrum has been studied in Amazonia regarding the spatial scaling properties of vegetation indices,[49] and rainfall fields during the WETAMC/LBA campaign.[38]

The bi-dimensional power spectrum is computed using a standard 2-D Fast Fourier Transform (FFT) algorithm.[55] The Fourier power or energy spectrum, $E(k_x, k_y)$ of a two-dimensional field, is estimated multiplying the 2-D FFT by its complex conjugate, where k_x and k_y are the wave number components. To facilitate visualization and comparison, the 2-D power spectra of rainfall fields are averaged angularly about $k_x = k_y = 0$ to produce the isotropic energy spectrum, $cE(k)$, with $k = (k_x^2 + k_y^2)$. Such isotropic energy spectrum does not mean that the field is isotropic, but rather that the angular averaging about $k_x = k_y = 0$ integrates the anisotropy.[18]

B. Moment-scaling analysis

The spatial intermittency of a random field can be quantified through the moment-scaling function. It allows discerning the type of spatial scaling of the random field (simple or multiple), and provides important information about the statistical behavior of the process at different scales.[18,40] A random field *X(t)* exhibits simple scaling if

$$X(rt) \overset{d}{=} r^{\theta} X(t), \tag{2}$$

where the equality refers to all finite dimensional distribution functions. Given that a probability distribution function can be characterized by its statistical moments, defined as $E[X^q] = \int x^q f(x)dx$, with $q = 1, 2, 3,.,$ for a simple scaling random field, $X(t)$,

$$E[X^q(rt)] = r^{\theta q} E[X^q(t)], \quad q = 1, 2, 3, ..., \text{or} \tag{3}$$

$$\log E[X^q(rt)] = q\theta \log r + \log E[X^q(t)]. \tag{4}$$

Equation (2) requires two conditions for simple scaling: (i) log-log linearity and (ii) linear growth of the slope of the power fit with order moments, $s(q) = q\theta$, whereas multi-scaling holds for a non linear growth. In general, moment-scaling is defined as[16]

$$M_q(r) \sim r^{-\tau(q)}, \tag{5}$$

where $\tau(q)$ is the moment scaling exponent function resulting from a power law fit between the sample $M_q(r)$ and statistical moments, q. Values of $\tau(q)$ are related to rainfall characteristics, e.g., higher values are associated with rougher and more intermittent rainfall fields, resulting from the highly localized nature and structure of convective rainfall-triggering processes in Amazonia.[38]

C. Entropy and statistical moments

The classical Boltzmann-Gibbs-Shannon *entropy* is defined as a measure of uncertainty and disorder of a random variable (r.v.), X, based on its discrete n-bin probability mass function[59]

$$S(X) = -\sum_{i=1}^{n} p_X(i) \log_a p_X(i), \tag{6}$$

where $p_X(1), p_X(2),..., p_X(k)$ represents the probability mass function, such that $\sum_{i=1}^{n} p_X(i) = 1$ and $p_X(i) \geq 0, \forall i$.

Skewness and kurtosis, directly linked to the third and fourth central moments, are often used to quantify the degree to which the probability distribution function (PDF) departs from the Gaussian distribution. They are estimated, respectively, as

$$\gamma = \frac{\mu_3}{\mu_2^{3/2}} = \frac{E(x-\mu)^3}{\sigma^3}, \tag{7}$$

$$\kappa = \frac{\mu_4}{\mu_2^2} = \frac{E(x-\mu)^4}{\sigma^4}. \tag{8}$$

The third central moment, Eq. (7), is associated with the symmetry of the PDF, whereas the forth central moment, Eq. (8), is an indicator of the weight of the PDF tails, very much related with intermittent phenomena.[54]

Alternative estimates of skewness and kurtosis based on L-moments have been proposed by Hosking,[19,20] which are less sensitive than the conventional ones (Eqs. (7) and (8)) to sampling variability or measurement errors in extreme values, and may therefore be expected to yield more accurate and robust estimates of the underlying probability distribution. The L-moments are defined for a sample $x_1, x_2, \ldots, x_n$ by considering the ordered sample $x_{1:n} \leq x_{2:n} \leq \ldots \leq x_{n:n}$, then the rth sample L-moments are defined as

$$l_r = \binom{n}{r}^{-1} \sum_{1 \leq i_1} \sum_{\leq i_2} \cdots \sum_{\leq n} r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} x_{i_{r-k}:n}, \quad r = 1, 2, \ldots, n, \tag{9}$$

in particular,