Physica A 423 (2015) 11–26

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Scaling of entropy and multi-scaling of the time generalized q-entropy in rainfall and streamflows

Hernán D. Salas *, Germán Poveda

Department of Geosciences and Environment, Universidad Nacional de Colombia, Sede Medellín, Colombia

HIGHLIGHTS

- Entropy of rainfall series exhibits a power law with aggregation time with exponent 0.5 before saturation.
- Scaling exponents for rainfall do not depend on record length.
- Entropy of streamflow series exhibits a power law with time with exponent marginally larger than zero.
- Scaling of entropy differs for rainfall and streamflows due to zeros in rainfall series.
- The time generalized q-entropy exhibits multi-scaling for rainfall and streamflows.

ARTICLE INFO

Article history:
Received 6 March 2014
Received in revised form 29 September 2014
Available online 31 December 2014

Keywords:
Hydrology
Tropical rainfall
Streamflows
Scaling
Shannon entropy
Tsallis entropy

ABSTRACT

We investigate the behavior of the Shannon entropy, $S(T)$, and the time generalized q-entropy, $S_q(T)$, at increasing aggregation intervals, T, using series of 15-min and hourly rainfall records in the tropical Andes of Colombia, spanning from 21 months to 40 years, as well as average daily streamflows in Colombia, the Amazon River basin, and USA, spanning from 34 to 69 years. Results for rainfall show that $S(T) \sim T^{\beta}$ with $\beta = 0.5$, valid up to a timescale $\langle T_{MaxEnt} \rangle = 83$ h, and $\beta = 0$ for $T > T_{MaxEnt}$. Scaling exponents ($\beta = 0.5$) are statistically independent of record length, although not so the values of T_{MaxEnt}, owing to the greater amount of zeros in rainfall series during El Niño in Colombia. Maximum entropy is reached through a dynamic Generalized Pareto distribution, whose parameters are not statistically affected by record length. Entropy for daily streamflows behaves as $S(T) \sim T^{\beta}$ with $\beta \gtrsim 0$, consistently with the theoretical limit. The scaling behavior of entropy differs between rainfall and streamflows due to the presence of zeros in rainfall series, and by the rate at which they vanish upon temporal aggregation. The amount of zeros exhibit two different scaling regimes with T, separated at $\langle T_b \rangle = 24$ h. Rainfall series get devoid of zeros at $\langle T_{noz} \rangle = 164$ h, but later on for longer record lengths. A power law relates both timescales: $T_{MaxEnt} \sim T_{noz}^{0.62}$. Besides, $S_q(T)$ consists of a continuous set of power laws with respect to T, for increasing values of q, whose scaling exponents vary nonlinearly with q, thus implying multi-scaling of the time generalized q-entropy, which in turn inherits all the conclusions from the Shannon entropy, since it is recovered at $S_{q=1}(T)$.

1. Introduction

The three branches of the tropical Andes of Colombia constitute an adequate geographical setting to study the highly intermittent dynamics of tropical mountain rainfall in space and time, owing to Ref. [1]: (i) strong topographic gradients,

* Corresponding author.
E-mail address: hdsalas@gmail.com (H.D. Salas).

http://dx.doi.org/10.1016/j.physa.2014.12.015
0378-4371/

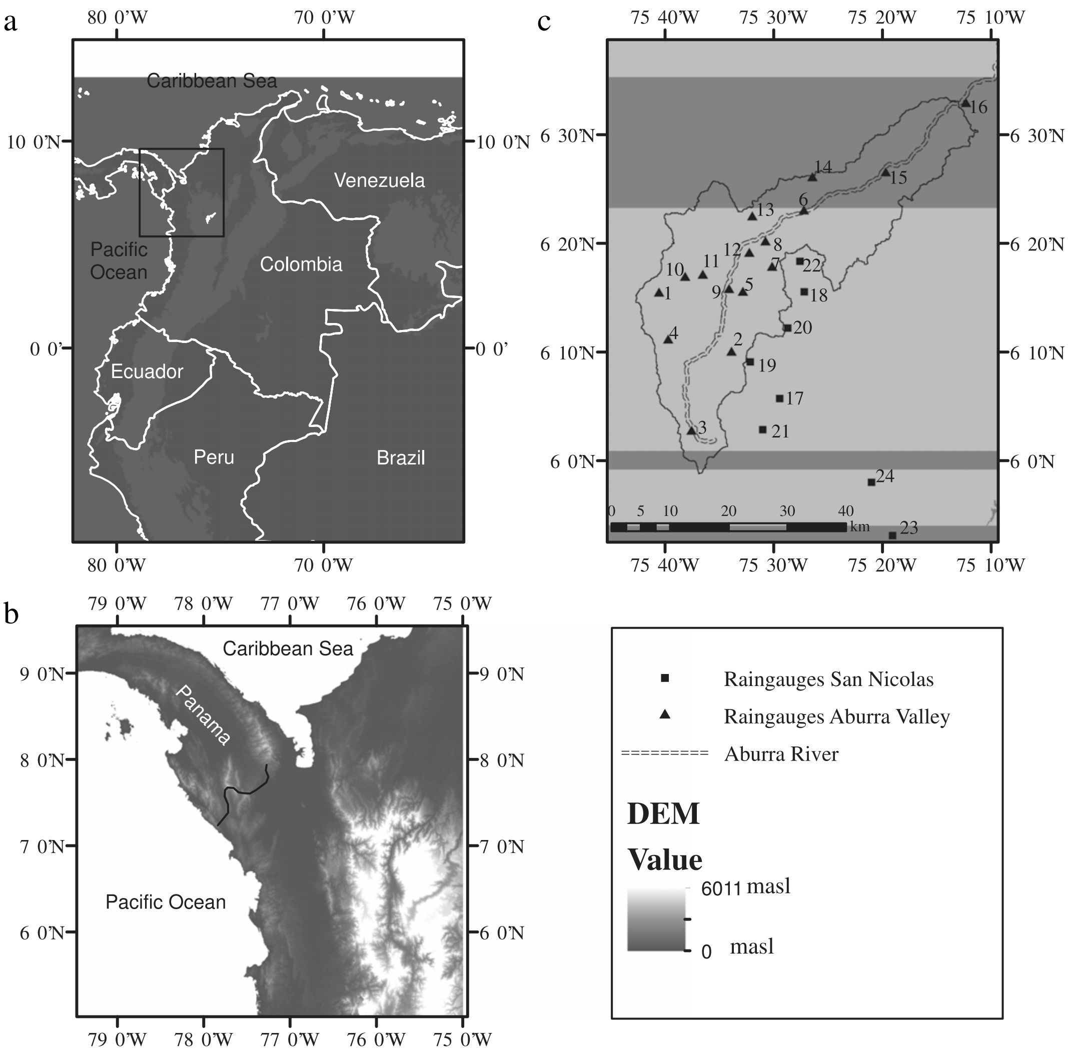

Fig. 1. Location of the Aburrá Valley, the San Nicolás Plateau to the southeast, and the set of raingauges used in this study, at different spatial scales: (a) continental, (b) regional, and (c) local.

slopes and aspects along the three branches of the Andes crossing from southwest to northeast, defining highly localized wind circulations that focus and enhance convection, as well as complex temperature and specific humidity gradients (Fig. 1); (ii) the influence from the hydro-climatic dynamics of the surrounding Amazon and Orinoco River basins, (iii) the (thermo-)dynamics of atmospheric circulation patterns and low-level jets interacting with mesoscale convective systems over the tropical Pacific and the Caribbean Sea [2], which contribute to explaining the existence of Lloró, one of the wettest place on Earth [3], and (iv) strong land-atmosphere feedbacks.

Besides, the small horizontal extent of convective clouds associated with tropical mountain rainfall contributes to its space–time intermittence, such that series of high resolution Andean rainfall records are mostly (approx. 90%) constituted by zeros [4]. The presence of zeros in rainfall series constitutes a fundamental issue in hydrology nowadays [5–10]. A recent study by Ref. [11] found that the continuous analysis of rainfall series, which contain zeros owing to the alternation of dry and wet periods, produce spurious results in the multifractal characterization of rainfall, and concluded that analysis should be restricted to within-storm intensities rather than continuous records.

On the other hand, river discharges result from the combined spatially-integrated dynamics of rainfall, evapotranspiration, vegetation, soil moisture, and the interactions between groundwater and surface water throughout the entire river basin and the embedded channel network, which in turn filter out the intermittence of rainfall. In particular, during dry periods the base flow is maintained by groundwater flow from both sides of the channel network, and thus the highly improbable presence of zeros in streamflows records in perennial river flows [12].

At present geosciences have difficulties in understanding and modeling tropical mountain rainfall and streamflows due to their geophysical complexity, added to the lack of detailed land-surface and atmospheric measurements [1]. There is still no unifying, coherent and satisfactory theory to explain the mechanisms of streamflow generation process [13], and therefore it is necessary to improve our understanding of mountain rainfall and streamflows toward their modeling, beyond the traditional descriptive statistics.

The information theoretic framework of entropy provides important elements to understand the complexity, nonlinearities, self-organizing and scaling properties of hydrological processes across a broad range of scales in space and time. See Ref. [14], and the references listed in the URL http://www.stahy.org/Activities/ICSHReferences/ReferencesonEntropyfunctioninhydrology/tabid/106/Default.aspx . In particular, diverse information theory metrics such as Shannon entropy, mutual Information, Tsallis q-entropy and time generalized q-entropy, have been used to quantify dependence or invariance of hydrological processes in time. Some of such relevant metrics are discussed next.

1.1. Shannon entropy

The classical Boltzmann–Gibbs–Shannon *entropy* measures the uncertainty and disorder of a random variable X, based on its discrete n-bin probability mass function as [15],

$$S(X) = -\sum_{i=1}^{n} p(x_i) \log_a p(x_i) \tag{1}$$

where $p(x_1), p(x_2), \ldots, p(x_k)$ represents the probability mass function, such that $\sum_{i=1}^{n} p(x_i) = 1$, and $p(x_i) \geq 0,\ \forall i$. This definition is valid for any logarithmic base, which in turn defines the units of entropy, e.g. a bit corresponds to base-2 logarithms, and a nat to natural logarithms.

The study of Ref. [1] investigated the behavior of the Shannon entropy of rainfall series upon aggregation increasing time scales, using 21-month (04/01/2005–12/31/2006) long high-resolution (15 min) records from the tropical Andes of Colombia. The said study found that entropy and aggregation interval are related through a power law, $S(T) \sim T^{\beta}$, with $\beta = 0.5$, and that the scaling regime holds up until some aggregation interval T_{MaxEnt}, at which entropy reaches a maximum value, and tappers off thereafter ($\beta = 0$ for $T > T_{MaxEnt}$). Furthermore, [1] showed that entropy reaches its maximum value in the form of a Generalized Pareto distribution, thus making it a maximum dynamic entropy (MDE) in the sense of Ref. [16], for tropical mountain rainfall in the Andes. In the present study, we aim to investigate the robustness of diverse scaling properties of entropy, using longer rainfall records and a larger set of raingauges over the Colombian Andes.

1.2. Tsallis q-entropy

The concept of *nonadditive entropy*, S_q, (q-entropy hereafter) was introduced by Ref. [17] within the framework of *nonextensive statistical mechanics* as,

$$S_q = \frac{1 - \sum_{i=1}^{n} p^{\,q}(x_i)}{q-1} \quad \left(\sum_{i=1}^{n} p(x_i) = 1;\ q \in \Re\right). \tag{2}$$

In the limit $q \to 1$, this expression recovers the usual Shannon entropy (Eq. (1)). Maximization of the q-entropy under a prescribed mean leads to a Pareto probability distribution exhibiting power-law tail [18–20], which belongs to the family of Lévy stable distributions [21], in particular to type II generalized Pareto distributions.

1.3. Time generalized q-entropy

Estimation of entropy (or any statistical parameter) for increasing values of T refers to the convergence of sums of random variables (r.v.). The central limit theorem applies for independent identically distributed (i.i.d.) r.v. exhibiting finite variance, whose sum converges to a Gaussian distribution, but no so for i.i.d. long-tailed r.v. The possible existence of q-versions of the standard central limit theorem for long-tailed r.v. was conjectured by Ref. [22], and proved for $1 \leq q < 3$ [23]. The attractor is given by distributions referred to as q-Gaussians [24, p. 295].

A dynamic time generalized q-entropy (Eq. (2)), was introduced as a function of order, q, and aggregation interval, T, as [1]:

$$S_q(T) = \frac{1 - \sum_{i=1}^{n} p^{\,q}(x_i, T)}{q-1} \quad \left(\sum_{i=1}^{n} p(x_i) = 1;\ q \in \Re\right). \tag{3}$$

The study by Ref. [1] estimated the function $S_q(T)$ for a set of 16 raingauges in Colombia using the said common 21-month record length. The present study aims to investigate whether the estimation of the time generalized q-entropy is sensitive to record length. Our sensitivity analysis has two purposes: (i) To examine the spatial distribution of the behavior (scaling exponents and time to maximum entropy) of the generalized Tsallis entropy and their relation with orography, owing to its role on the genesis, evolution and dynamics of rainfall, and given its relevance to focus and enhance (shallow and deep) convection, and to constrain wind circulations along intra-Andean valleys. (ii) To quantify the sensitivity of the scaling exponents, β, with respect to record length, due to the almost complete lack of long-term rainfall records over the study region, and its incidence on the estimation of multi-scale statistics in tropical mountain rainfall.

The purpose of this study is manifold:

1. To investigate whether the number of bins employed in the estimation of the probability mass function, $p(x_i)$ (Eq. (1)), affect the estimation of entropy under aggregation of the process at increasing values of T.
2. To test whether the scaling exponents found for rainfall entropy ($\beta = 0.5$), and the time to reach maximum entropy, T_{MaxEnt}, are sensitive to record length.
3. To test whether the Generalized Pareto distribution keeps standing out as a maximum dynamic entropy using longer rainfall records.
4. To study the relationship between entropy and aggregation interval using series of daily average streamflows in Colombia, the Amazon River basin, and USA.
5. To study and explain the possible differences between the scaling characteristics of entropy upon temporal aggregation for rainfall and streamflows, given the presence (absence) of zeros in rainfall (streamflows) records, in terms of the possible existence and estimates of scaling exponents, time to reach maximum entropy, and saturation of the scaling regimes at some T.
6. To compare the behavior of the time generalized q-entropy, $S_q(T)$, for rainfall and daily streamflows.

Reaching the above stated objectives will contribute to shed light to a suite of practical issues including: (i) developing synthetic models of high-resolution series of mountain rainfall, (ii) downscaling large-scale rainfall models, (iii) parameterizing surface hydrological processes, (iv) contributing to solve the PUB (Prediction in Ungaged Basins) problem, (v) quantifying hydrologic perturbations of river flows in intervened basins, and (vi) designing hydrologic monitoring networks. Toward those ends, Section 3 discuss methods and data, while results and discussions are presented in Section 4, and conclusions are drawn in the final section.

2. Data

Two rainfall data sets will be used herein. One consists of 15-min intensity records at 24 raingauges, 16 of which are located along the Aburrá Valley, and 8 in the contiguous San Nicolás Plateau, both located in the north-western department of Antioquia on the central range of the Colombian Andes, among the intra-Andean valleys of the Cauca and Magdalena Rivers, 6.0°N–6.6°N, and 75.7°W–75.2°W (see Fig. 1). The Aburrá Valley houses the city of Medellin and its metropolitan area, with terrain elevations ranging from 1000 to 2600 m a.s.l. The valley exhibits two main directions, with a predominant south–north orientation in the first half of the pathway, and southwest–northeast in the second half. The San Nicolás Plateau is located southeast of the Aburrá Valley, and exhibits hilly terrains with an average elevation of 2200 m a.s.l. (1900 m a.s.l. to 2600 m a.s.l.). Rainfall data were provided by IDEAM and Empresas Públicas de Medellín (EPM). Table 1 summarizes information of raingauges.

It is worth noting that the set of 16 raingauges located in the Aburrá Valley is the same one studied by Ref. [1] using a 21-month common period for all raingauges (Apr 01/2005 to Dec 31/2006). The present study will use the longest series available for each of those raingauges (columns 6–8 of Table 1), with the aim of quantifying the sensitivity of results to record length.

The second rainfall data set consists of hourly records at 5 raingauges located on the Central Andes of Colombia, some 150 km south of the Aburrá Valley, spanning from 8 years (Jan 01/1972 to Jun 23/1979) to 40 years (Apr 4/1960 to Feb 28/1999), without missing data, provided by the National Center for Coffee Research (Cenicafe), hereafter referred to as Cenicafe No. 1 to Cenicafe No. 5.

Daily average streamflows records belong to rivers in Colombia, the Amazon River basin, and USA, with no missing records, and having record lengths from 34 to 68 years. Details and data sources are shown in Table 2.

Entropy, S, will be estimated using Eq. (1) for increasing aggregation intervals from $T = 15$ min to 64 days, each 15 min. Regarding the estimation of entropy, an important issue is the sensitivity to record length and to the number of bins used to estimate the probability mass function, $p(x_i)$ (Eq. (1)). As a rule of thumb, using between 10 and 20 bins and more than 500–1000 data are suggested to estimate entropy, mutual information, and other information theory statistics in finite-interval bin-counting probability mass function estimation schemes [25,26]. These issues will be studied and discussed as well.

3. Results and discussions

3.1. Scaling of entropy for rainfall

3.1.1. Sensitivity to the number of bins

The sensitivity of rainfall entropy estimates to the number of bins is shown in Fig. 2 for raingauge Cenicafe No. 1. Results indicate that the scaling exponent ($\beta = 0.5$) is not significantly affected by the number of bins, although not so the prefactor c. For estimation purposes, 100 bins will be used, as per the average obtained through the method introduced by Ref. [27].

Table 1
Information of raingauges along the Aburrá Valley and the San Nicolas Plateau separated by the double horizontal line, and estimates of statistical parameters. (Columns 3, 4 and 5) Geographic coordinates; (Columns 6, 7, and 8) Entire record lengths; (Columns 9 and 10) Scaling exponents, β, between entropy and aggregation time, for the Common Period (CP), and the Entire Record Length (ER); (Columns 11 and 12) Timescales to reach maximum entropy (T_{MaxEnt}) (hours); (Columns 13–18) Parameters k, σ, and μ of the generalized Pareto distribution fitted at $T = T_{MaxEnt}$ for the Common Periods (CP), and the Entire Record Lengths (ER); (Columns 19 and 20) First set of scaling exponents, α_1, between amount of zeros and aggregation time; (Columns 21 and 22) Second set of scaling exponents α_2, between amount of zeros and aggregation time; (Columns 23 and 24) Timescales of scale breaking in the relation between amount of zeros and aggregation time (hours); (Columns 25 and 26) Timescales for the vanishing of zeros (hours).

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
ID	Gauge	Lon W	Lat N	Height m a.s.l.	Initial Date	Final Date	ER Months	CP β	ER β	CP T_{MaxEnt}	ER T_{MaxEnt}	CP k	ER k	CP μ	ER μ	CP σ	ER σ	CP α_1	ER α_1	CP α_2	ER α_2	CP T_b	ER T_b	CP T_{noz}	ER T_{noz}
1	Astillero	75.68	6.26	2420	08/01/2001	11/30/2007	76	0.49	0.48	85	106	−0.150	−0.209	−5.608	−3.115	44.670	38.930	1.229	1.332	1.561	1.894	19	41	163	433
2	Ayurá	75.56	6.16	1750	03/01/2001	08/31/2008	90	0.53	0.49	77	123	−0.216	−0.252	−1.810	−1.631	42.760	39.753	1.245	1.284	1.608	1.787	18	27	156	266
3	Caldas	75.63	6.04	1930	03/01/2001	10/31/2003	32	0.50	0.47	66	82	−0.580	−0.208	−1.195	−1.150	75.500	28.472	1.400	1.268	1.922	1.693	27	18	138	250
4	San Antonio	75.66	6.19	2000	12/01/2004	12/31/2008	49	0.50	0.47	70	85	−0.380	−0.213	−1.597	−0.478	51.240	33.006	1.233	1.396	1.672	2.098	16	38	145	211
5	Villa Hermosa	75.54	6.25	1690	02/01/2001	12/31/2007	83	0.51	0.47	75	135	−0.228	−0.261	−3.916	−1.575	31.210	36.018	1.253	1.362	1.630	1.849	22	51	163	332
6	Girardota	75.45	6.38	1350	11/01/2001	04/30/2007	66	0.52	0.50	87	134	−0.129	−0.099	−3.018	−1.416	27.290	29.858	1.189	1.280	1.574	1.871	17	42	168	386
7	Chorrillos	75.50	6.29	2370	09/01/2003	04/30/2007	44	0.50	0.48	90	88	−0.276	−0.126	−3.400	−1.686	40.830	25.868	1.273	1.321	1.724	1.884	21	29	182	210
8	El Convento	75.51	6.34	1580	10/01/2005	12/31/2008	39	0.56	0.48	81	97	−0.149	−0.290	−3.010	−1.658	31.550	30.564	1.253	1.354	1.612	1.681	24	21	179	274
9	Aguinaga	75.57	6.26	1549	10/01/2004	12/31/2006	27	0.50	0.50	83	102	−0.196	−0.021	−2.616	−0.058	25.450	21.858	1.241	1.543	1.543	1.562	25	25	199	274
10	San Cristobal	75.63	6.28	1890	01/01/2004	12/31/2008	60	0.50	0.48	92	106	0.008	−0.200	−2.802	−1.742	24.400	34.920	1.209	1.254	1.606	1.657	18	24	209	327
11	Cucaracho	75.61	6.28	1830	04/01/2001	12/31/2005	57	0.53	0.51	93	160	−0.056	−0.182	−2.884	−1.918	28.610	37.947	N.A.	1.215	N.A.	1.634	N.A.	24	N.A.	492
12	Manantiales	75.53	6.31	2043	08/01/2005	12/31/2008	41	0.51	0.54	87	131	0.189	0.038	−1.778	−2.078	10.140	19.044	1.250	1.156	1.674	1.507	19	22	171	283
13	Niquia	75.53	6.37	2150	03/01/2001	06/30/2003	28	0.53	0.50	79	134	−0.192	−0.206	−3.559	−1.791	31.890	29.020	N.A.	1.264	N.A.	1.803	N.A.	41	N.A.	333
14	San Andres	75.44	6.43	2240	01/01/2004	12/31/2008	60	0.50	0.48	87	91	−0.115	−1.141	−2.774	−0.884	27.190	27.955	1.258	1.295	1.769	1.796	20	26	155	227
15	Barbosa	75.33	6.44	1290	07/01/2003	12/31/2005	30	0.52	0.50	81	72	−0.215	−0.010	−2.734	−0.738	42.060	22.569	1.253	1.794	1.794	1.743	17	27	147	240
16	Gabino	75.20	6.54	1080	01/01/2002	09/30/2003	21	0.51	0.45	77	89	−0.217	−0.077	−7.483	−3.867	60.020	40.651	N.A.	1.313	N.A.	1.863	N.A.	17	N.A.	189
17	La Fe	75.49	6.10	2150	12/01/2003	10/31/2007	47	0.46	0.45	97	97	−0.128	−0.217	−3.348	−2.652	29.450	34.694	1.237	1.393	1.602	2.097	17	47	159	212
18	La Severa	75.45	6.26	2170	03/01/2003	09/30/2008	67	0.47	0.47	72	86	−0.067	−0.151	−2.548	−1.572	21.170	28.683	1.210	1.459	1.735	1.696	21	43	165	177
19	Las Palmas	75.54	6.15	2495	05/01/2001	11/30/2006	67	0.44	0.46	68	103	−0.034	−0.208	−1.936	−1.994	19.621	35.810	1.198	1.268	1.650	1.688	14	18	546	186
20	Vasconia	75.48	6.20	2510	06/01/2001	07/31/2006	62	0.46	0.47	85	103	−0.084	−0.107	−2.063	−2.144	23.674	32.040	1.298	1.295	1.755	1.900	26	22	141	171
21	El Retiro	75.52	6.05	2190	09/01/2001	06/30/2004	34	0.46	0.45	89	91	−0.145	−0.167	−1.634	−1.315	25.038	27.030	1.183	1.264	1.623	1.792	13	21	165	165
22	La Mosca	75.46	6.31	2155	05/01/2001	08/31/2003	28	0.47	0.47	87	85	−0.102	−0.105	−2.054	−1.918	24.504	23.743	1.199	1.209	1.507	1.647	16	19	170	213
23	Mesopotamia	75.32	5.88	2410	02/01/2004	11/31/2006	34	0.46	0.39	91	48	−0.098	−0.110	−3.004	−0.554	20.530	25.559	1.178	1.368	1.680	2.027	21	15	154	103
24	La Union	75.35	5.97	2500	12/01/2003	08/31/2006	33	0.48	0.43	96	68	−0.079	−0.335	−2.573	−1.222	25.640	28.364	1.235	1.394	1.780	1.993	23	23	169	143

Table 2
Location of stream gauges and description of daily streamflow series used in this study.

Country	River	Gauging station	Agency	Code	Period of record
Colombia	Atrato	Bellavista	IDEAM	1 107 701	01/01/1965–12/01/2002
Colombia	Cali	Bocatoma	CVC	N.A.	01/01/1973–12/31/2006
Brazil	Tijuco	Ituiutaba	ANA	60 845 000	01/01/1942–12/01/2006
Brazil	Correste	Campo Alegre	ANA	60 940 000	07/01/1972–04/01/2007
USA	Little Sioux	Turin, IA	USGS	6 607 500	07/05/1942–11/05/2010

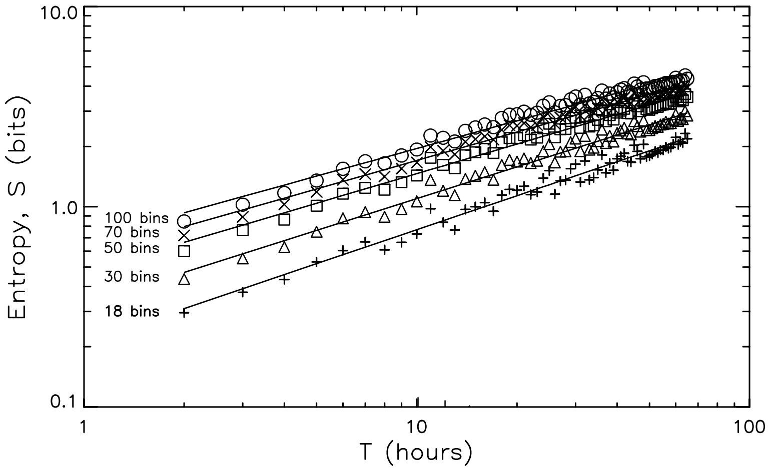

Fig. 2. Sensitivity of entropy estimates to the number of bins as a function of the time aggregation interval, for the 40-yr long series of hourly rainfall at raingauge Cenicafe No. 1. Notice that for all bins, $S \sim T^{\beta}$, with $\beta \approx 0.5$.

3.1.2. Sensitivity to record length

This section investigates the effect of record length on the scaling properties of both rainfall data sets. In particular, we aim to investigate the sensitivity of scaling exponents, β, the time to reach maximum entropy, T_{MaxEnt}, and the validity of the Generalized Pareto distribution as a maximum dynamic entropy (MDE) for rainfall in the Andes.

- **Scaling Exponents**. Fig. 3 shows the estimation of entropy as a function of aggregation interval, for raingauges located on the Aburrá Valley using the 21-month common period (left panels), and the entire record lengths (right panels) at raingauges Astilleros and Ayurá (numbered 1 and 2 in Fig. 1 and Table 1). Scaling exponents (column 9 of Table 1) are statistically equal to 0.5 at the 95% confidence level, according a t-Student test [28].

  For raingauges located along the San Nicolás Plateau, two common periods were considered. Raingauges No. 17, 18, 23, and 24 (Table 1) share a common period of 44 months (Dec 01/2003 to Jul 31/2006), whereas raingauges No. 19, 20, 21, and 22 share a common period of 24 months (Sep 01/2001 to Aug 31/2003). The mean value of scaling exponents using the common periods is $\beta = 0.46$ (standard deviation 0.01). For the entire record lengths (column 10 of Table 1), the mean value is $\beta = 0.46$ (standard deviation 0.03). Likewise, all values of the scaling exponents are statistically equal to $\beta = 0.50$, at the 95% confidence level. Fig. 4 summarizes all estimated values of β, for the common periods and the entire record lengths.

  A detailed analysis of the relation between entropy and T for the Cenicafe data set indicates that $S \sim T^{\beta}$, with $\beta = 0.5$, as shown in both panels of Fig. 5 for the 40-yr long record at Cenicafe No. 1, and the 28-yr long record at Cenicafe No. 2. Likewise, this result confirms that the scaling regime $S \sim T^{\beta}$, with $\beta = 0.5$, is not significantly affected by record length, which confirms the robustness and consistency of previous results [1].
- **Time to Reach Maximum Entropy**. Estimates of the timescales to reach maximum entropy using the common periods are $\langle T_{MaxEnt} \rangle = 83$ (66–96) h (column 11 of Table 1), while using the entire record lengths, $\langle T_{MaxEnt} \rangle = 100$ (48–134) h (column 12). An explanation for the higher (smaller) values of T_{MaxEnt} for the entire (common) record lengths will be provided in Section 3.3. Regarding the Cenicafe data set, Fig. 5 shows the same scaling regime, with $T_{MaxEnt} = 64$ h for Cenicafe No. 1, while $T_{MaxEnt} = 89$ h for Cenicafe No. 2.

3.1.3. The generalized Pareto distribution as a maximum dynamic entropy

Regarding the PDF through which entropy reaches a maximum value, a set of 50 probability distribution functions (PDF) were fitted to all rainfall data sets, for values of $T = 15$ min to $T = T_{MaxEnt}$. For fitting purposes, the EasyFit software was used. It contains a large suite of PDFs including Generalized Pareto, Generalized Extreme Value, Wakeby, Weibull, Frechet, Lognormal, etc., and diverse goodness-of-fit tests that include both the standard and maximum likelihood tests of Kolmogorov–Smirnov, Anderson–Darling, χ^2, and Cramer–von Mises [29].

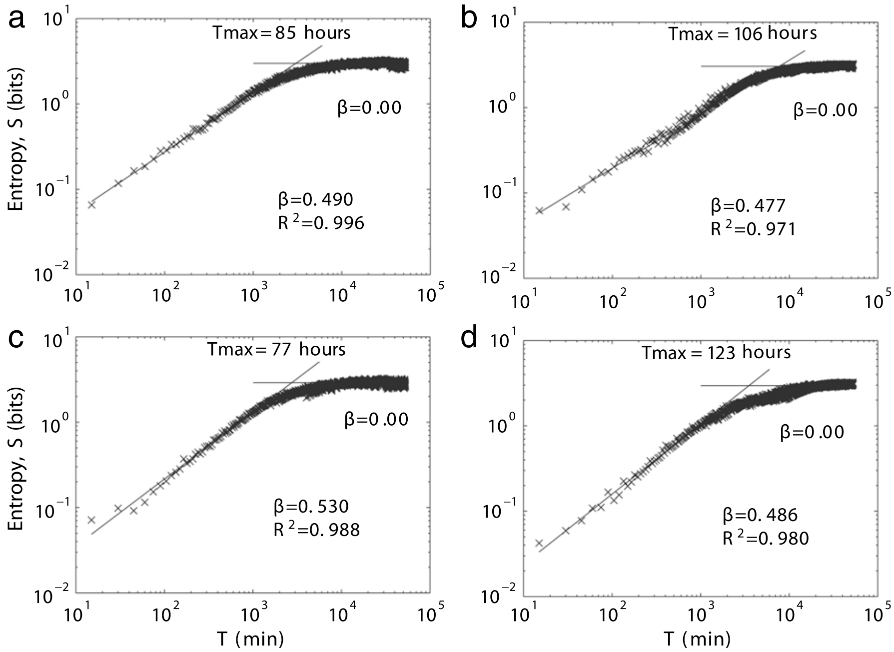

Fig. 3. Sensitivity of entropy as a function of aggregation interval, with respect to record length at raingauges on the Aburrá Valley: (a) Astilleros (21 months), (b) Astilleros (76 months), (c) Ayurá (21 months), and (d) Ayurá (90 months).

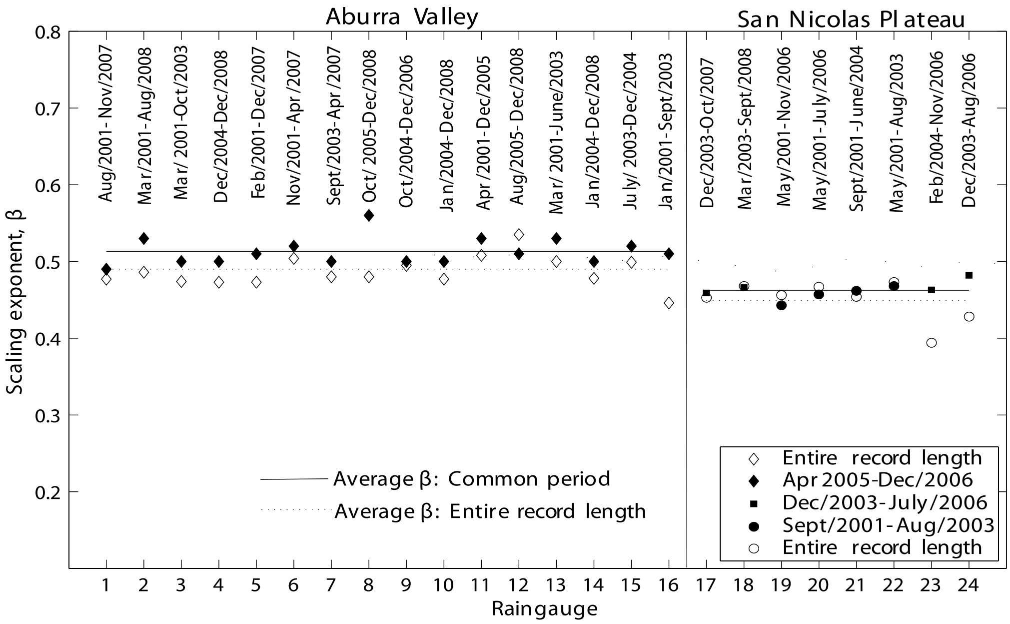

Fig. 4. Estimates of the scaling exponent, β, for the whole set of raingauges along the Aburrá Valley (left), and the San Nicolás Plateau (right), using the entire record length. Denoted are also the averages using the 21-month common period for Aburrá Valley, and 24-month common period for the San Nicolás Plateau.

Results obtained using the entire record lengths confirm that the Generalized Pareto PDF stands out as a maximum dynamic entropy for most values of T and most raingauges. Fig. 6 shows the dynamics of the sample and fitted Generalized Pareto PDFs for increasing values of T at raingauges Astilleros (top) and Ayurá (bottom) on the Aburrá Valley, with the sample and fitted PDFs shown for values of T ranging from $T = 1$ h to $T = 96$ h. Columns 13–18 of Table 1 show the parameters k, σ, and μ of the Generalized Pareto PDFs using the common periods and the entire record lengths.

A Monte Carlo experiment was performed using 10,000 simulated series with the parameters κ, σ, and μ estimated at each raingauge, with the aim of testing whether a Generalized Pareto r.v. with the parameters of the common periods are

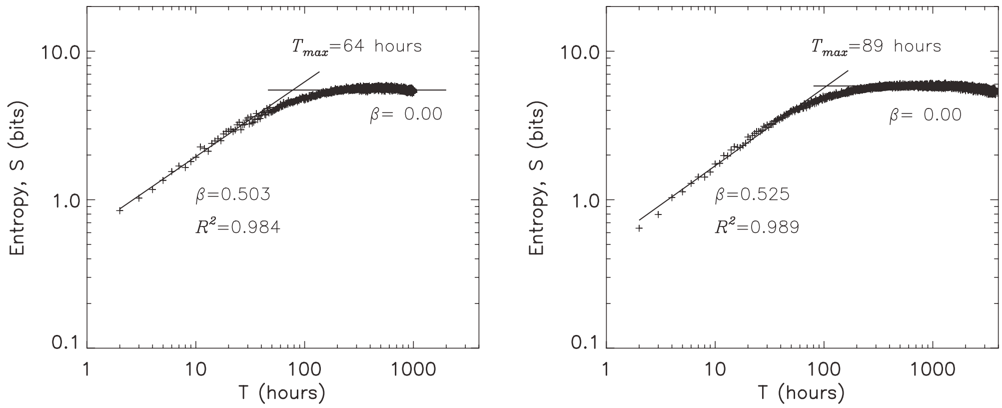

Fig. 5. Detailed scaling relationships between entropy, S, and time aggregation interval, T, for the 40-yr long series of hourly rainfall at raingauge Cenicafe No. 1 (left), and the 28-yr long series at raingauge Cenicafe No. 2 (right).

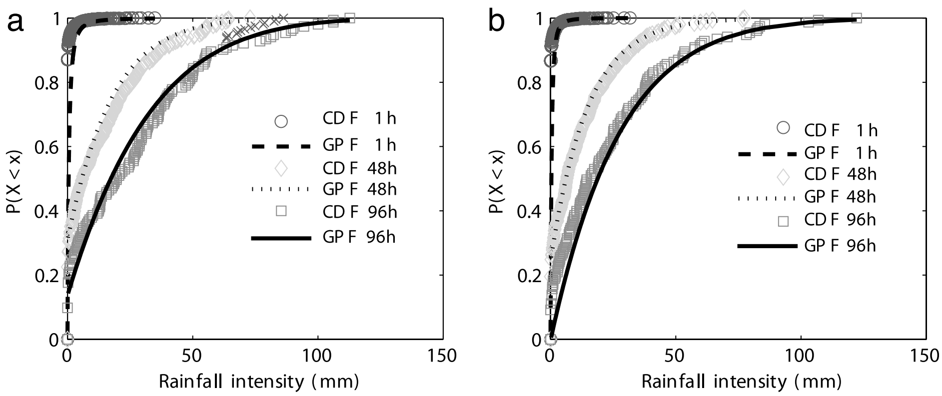

Fig. 6. Behavior of the sample and fitted Generalized Pareto cumulative distribution functions (CDF), for increasing values of T at raingauge Ayurá, using the 21-month common period (panel (a)), and the entire record length 90 months panel (b). Values of T range from $T = 1$ h to $T = 96$ h.

statistically equal (Smirnov–Kolmogorov test) to a Generalized Pareto r.v. with κ, σ, and μ of the entire record lengths. Results indicate that, with very few exceptions, the parameters κ, σ, and μ of both samples are statistically equal, at the 99% significance level.

Similar analysis performed on Cenicafe's hourly records confirm that the Generalized Pareto exhibits the best fit for all raingauges and most values of T. All these results confirm that the Generalized Pareto constitutes a maximum dynamic entropy (MDE) in the sense of Ref. [16], for high-resolution tropical mountain rainfall. This is a very remarkable result insofar the Generalized Pareto distribution satisfies two fundamental conditions: (i) it is a maximum entropy probability distribution function (PDF) under the constraint of heavy tails, and (ii) it is a stable PDF in the sense that it is preserved under aggregation.

3.2. Scaling of entropy for daily streamflows

First, we explore the sensitivity of entropy to the number of bins. Fig. 7 shows the estimated values of $S(T)$ for the Little Sioux River, using 100, 300, 500, 800, and 1000 bins. The scaling exponent is not significantly affected by the number of bins, although not so the prefactor. Using the algorithm of Ref. [27], an optimal number between 400 and 600 bins was estimated for all streamflow series. Regarding the scaling exponent, the inset at Fig. 7 shows that entropy of the Little Sioux River grows very slowly with time aggregation, or that $S(T) \sim T^{\beta}$, with $\beta \gtrsim 0$ (e.g., scaling exponent positive but slightly greater than zero). This behavior is consistent and satisfies Shannon's theoretical entropy inequality [30],

$$H(T) \leq \log \sqrt{2\pi eV(T)}, \tag{4}$$

where $V(T)$ is the variance of the process at T, an inequality that is saturated only for the Normal distribution. The thick line in Fig. 7 denotes such theoretical limit. Results are similar for the remaining time series of daily streamflows in USA, Colombia and Brazil.

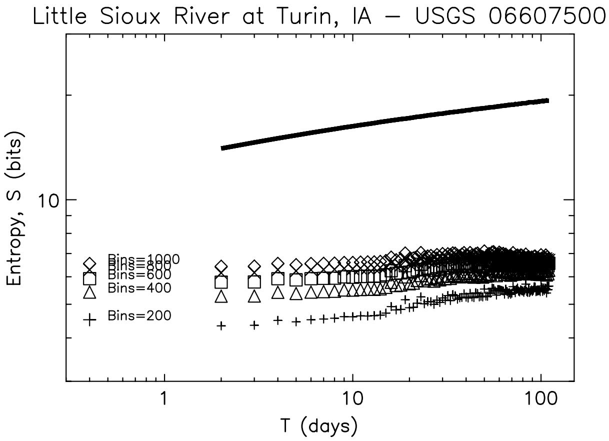

Fig. 7. Sensitivity of $S(T)$ with respect to the number of bins (100, 300, 500, 800, and 1000) employed to estimate the probability mass function, $p(x_i)$ (Eq. (1)), for the Little Sioux River daily streamflows. The thick line denotes the theoretical limit given by Eq. (4).

3.3. On the different scaling of entropy for rainfall and streamflows: the presence of zeros in rainfall

In light of our previous results, an explanation is in order for the difference between the scaling exponents of the power law relating S with T, of rainfall ($\beta = 0.5$), and streamflows ($\beta = 0$). The most notable difference between high-resolution time series of rainfall and river discharges is the presence of zeros in the former one. Even for tropical climates, hourly (let alone 15-min) rainfall is a rare phenomenon, such that series are mostly (∼90%) constituted by zeros [31,1,4]. Regarding the measurements of rainfall, raingauges gather information from minutes to hours and beyond. For a daily time resolution, a raingauge will record a positive rainfall value for a certain day (24-h) regardless of the duration of the event (1, 2, 3, . . . hours) within the same day. For an hourly time resolution, if no rain is observed during one complete day, there will be 24 straight zeros in the record.

Zeros constitute highly important information to understand, diagnose and forecast the dynamics of rainfall. We argue that the presence of zeros (inter-storm periods) in rainfall is associated with the timescales required by nature to build-up the dynamic and thermodynamic conditions of the next storm, as an atmospheric analog of the time of energy build-up between earthquakes, avalanches and many other relaxational processes in nature [32]. This explanation connects the dynamical nature of rainfall with the presence of zeros. Obviously, zeros of high-resolution rainfall series become rarer upon aggregation at longer timescales. An important question arises at this point: at what aggregation interval do zeros vanish from high-resolution rainfall series?

Meanwhile, as we pointed out at the Introduction, streamflows result from the basin-integrated dynamics of diverse hydro–geo-biological processes filtering out the intermittence of rainfall, and thus the unlikely presence of zeros in records of perennial rivers at all timescales.

This section studies the influence of zeros on estimates of entropy of rainfall under aggregation. In particular, we aim at quantifying the probability of finding zeros, and how the amount of zeros vanish from rainfall series at increasing values of T, using the common periods and the entire record lengths. Also, we aim at estimating the time aggregation interval at which zeros vanish from the series, T_{noz}, and the relationship between T_{noz} and T_{maxEnt}.

Fig. 8 shows both estimates of entropy and the probability of finding zeros, as a function of aggregation interval, T, for the whole set of raingauges along the Aburrá Valley and the San Nicolás Plateau, using the common periods and the entire record lengths. Besides the discussed direct power law relating entropy with T, Fig. 8 shows an inverse power law relating the probability of finding zeros and T. In fact, there seems to be two scaling regimes separated at around $T = 1000$ min (16 h). At any rate, entropy reaches its maximum at values of T corresponding to a probability equal to or less than 0.0001 of finding zeros in the series. The relationship between the amount of zeros $\zeta(T)$ and T show two different scaling regimes,

$$\zeta(T) = a_1 T^{-\alpha_1}, \tag{5}$$

$$\zeta(T) = c + a_2 T^{-\alpha_2}, \tag{6}$$

with scaling exponents $\langle\alpha_1\rangle = 1.29$ [1.16–1.80], and $\langle\alpha_2\rangle = 1.74$ [1.51–2.10], as shown in columns 19–22 of Table 1. The two distinctive scaling regimes are separated at $T_b = 17$ h for the common periods, and at $T_b = 28$ h for the entire record lengths. The break in the scaling regime was identified using a t-Student test [33,34] of equality between scaling exponents, α_1, for the common periods (column 19 of Table 1), and the entire record lengths (column 20). The test confirms that the scaling exponents α_1 are statistically equal ($P > 95\%$) using the common periods and the entire record lengths. Likewise,

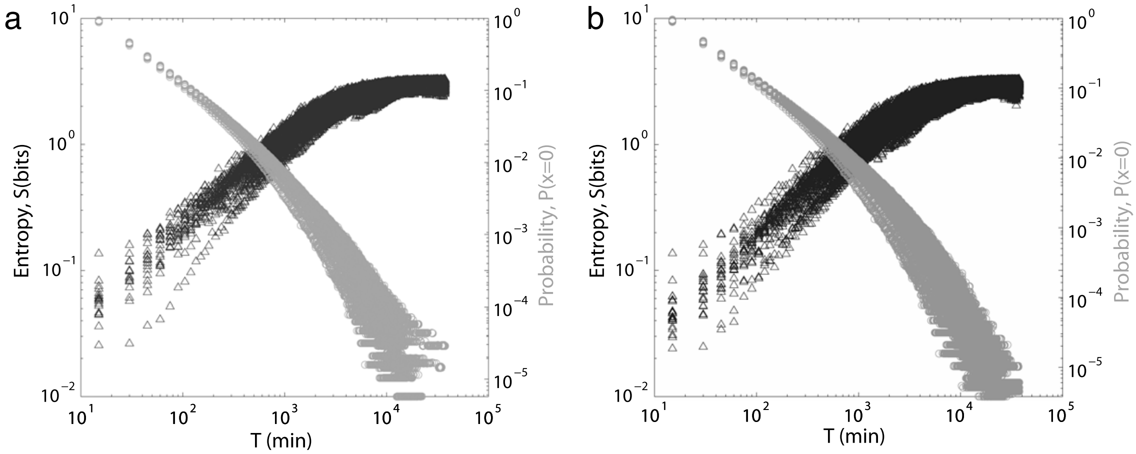

Fig. 8. Entropy (triangles) and probability of finding zeros (circles), as a function of aggregation interval T, for the entire set of raingauges along the Aburrá Valley and the San Nicolás Plateau, for (a) the common periods, and (b) the entire record lengths.

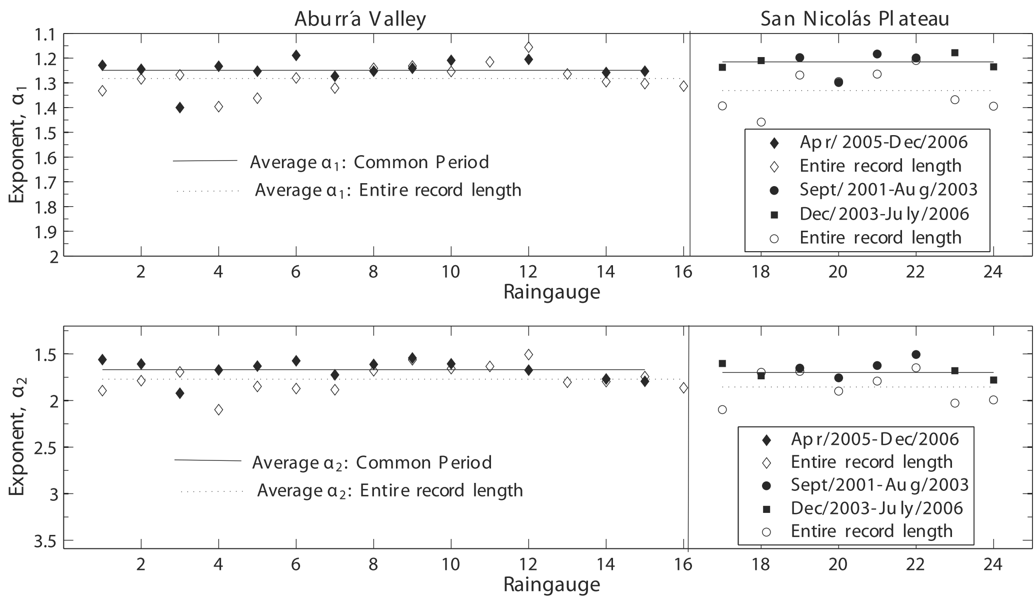

Fig. 9. Estimated values of the scaling exponents relating the amount of zeros, $\zeta(T)$, with time aggregation interval, T, for the whole set of raingauges along the Aburrá Valley and the San Nicolás Plateau.

results indicate that the estimated values of α_2 using the common periods (column 21 of Table 1) are statistically equal to those using the entire record lengths (column 22), as shown in Fig. 9. The timescales of scaling break for both record lengths are summarized in columns 23 and 24 of Table 1.

3.3.1. *Relation between T_{MaxEnt}, and T_{noz}*

Rainfall series get devoid of zeros at $\langle T_{noz}\rangle = 164\,[138–210]$ h for the common periods, while at $\langle T_{noz}\rangle = 254\,[103–492]$ h for the entire record lengths, as shown in columns 25 and 26 of Table 1. For all raingauges in the Aburrá Valley and the San Nicolás Plateau, the timescales to reach maximum entropy, T_{MaxEnt}, are smaller than those at which zeros vanish from rainfall series, T_{noz}, but both timescales are related through the power law $T_{MaxEnt} \sim T_{noz}^{0.62}$ ($R^2 = 0.57$), as shown in Fig. 11.

The smaller entropy value belongs to the original 15-min series, since most of the mass is concentrated in the lower bins of the histogram, and therefore very few bins contribute to entropy (Eq. (1)). Once rainfall is aggregated at increasing values of T, the probability mass function undergoes an "erosion" process from the lower to the higher bins, owing to the accumulation of more rain events and (likely) more intense ones, which in turn contribute to augment entropy until it reaches its maximum value at T_{MaxEnt}. Such dynamics is illustrated in Fig. 10, for raingauge Mazo for values of $T = 60$ min to $T = 10{,}185$ min.

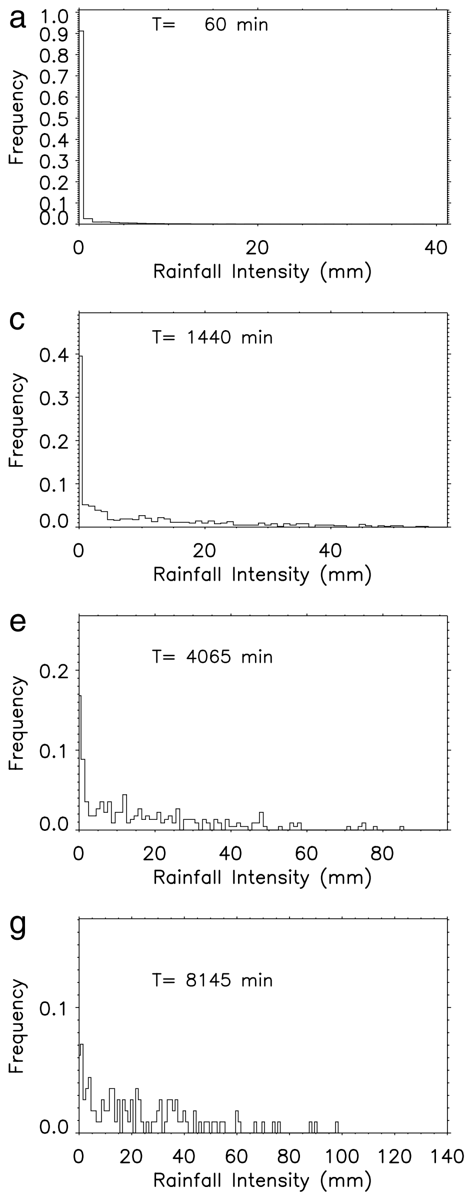

Fig. 10. Evolution of the probability mass function at raingauge Mazo, for increasing values of the time aggregation interval, from $T = 60$ min to $T = 10{,}185$ min $\gg T_{MaxEnt}$. Notice how the histogram gets eroded upon aggregation of rainfall at increasing values of T.

The previous analysis allow us to conclude that the difference between scaling exponents for rainfall ($\beta = 0.5$) and streamflows ($\beta = 0$) can be attributed to the presence of zeros in rainfall.

3.3.2. Values of T_{MaxEnt} increase with record length owing to interannual (ENSO) variability

Section 3.1.2 showed that the time scales to reach maximum entropy, T_{MaxEnt}, are longer for the entire record lengths than for the shorter common periods. This result can be explained by the presence of zeros, combined with the interannual variability of rainfall, which is mostly driven by El Niño/Southern Oscillation (ENSO). In general, Colombia experiences negative anomalies in annual, seasonal, monthly, and hourly rainfall during El Niño, and positive anomalies during La Niña [35,36,4,37,1]. Therefore, it is expected that entropy and the amount of zeros, at increasing aggregation intervals, depend on the number and intensity of El Niño and La Niña events during the record length. In fact, zeros vanish at a lower pace during El Niño than during La Niña, as a consequence of the longer El Niño-driven dry spells. In other words, T_{noz} (El Niño) $>$ T_{noz} (La Niña). Fig. 12 for raingauge Ayurá confirms that such is the case for El Niño event of 06/2006–06/2007, and La Niña of 06/2007–06/2008. Panel (a) shows that T_{MaxEnt} (El Niño) $>$ T_{MaxEnt} (La Niña), while panel (b) shows that T_{noz} (El Niño) $>$ T_{noz} (La Niña).

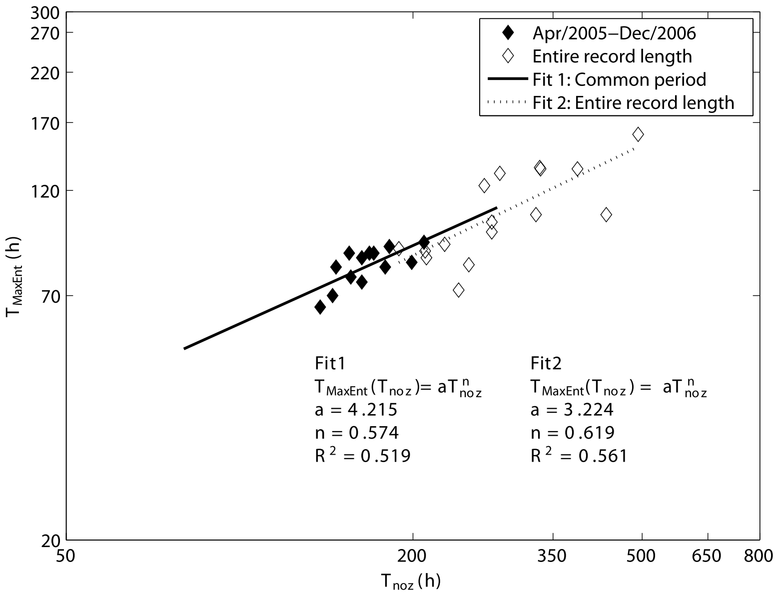

Fig. 11. Relation between the timescales at which zeros vanish from 15-min rainfall series (T_{noz}), and to reach maximum entropy (T_{MaxEnt}), using the common periods and the entire record lengths, for raingauges along the Aburrá Valley and the San Nicolás Plateau.

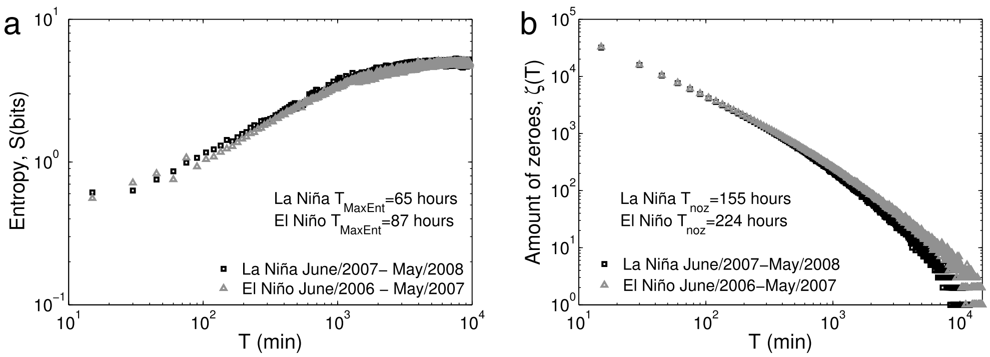

Fig. 12. (a) Entropy, $S(T)$; and (b) Amount of zeros, $\zeta(T)$, both as a function of time aggregation interval, T, for raingauge Ayurá during El Niño 2006–2007 (triangles), and La Niña 2007–2008 (squares). Panel (a) evidences that T_{MaxEnt} (El Niño) $>$ T_{MaxEnt} (La Niña), while panel (b) that T_{noz} (El Niño) $>$ T_{noz} (La Niña).

3.4. Time generalized q-entropy for rainfall and streamflows

Fig. 13 shows the time generalized q-entropy corresponding to the 40-year long series of hourly rainfall at raingauge Cenicafe No. 1. Panel (b) shows the projection of $S(q, T)$ vs. q for different values of T. Panel (c) shows $S(q, T)$ vs. T, for different values of q, or the time structure function for entropy. Panel (d) shows that the regression slopes of the time structure function for entropy, $\beta(q)$, as a function of q, exhibit a non-linear growth such that $\beta \approx 0.5$ for $q \geq 1$, while panel (a) illustrates the function $S(q, T)$ (Eq. (3)). Notice that values of $S_q(T)$, for $q = 1$ correspond to those of the Shannon entropy for raingauge Cenicafe No. 1, shown on the left panel of Fig. 5.

Several features of the $S(q, T)$ function for rainfall are worth noticing:

1. $S_q(T)$ decreases monotonically with q for all values of T.
2. For a given value of q, estimates are inversely related to T for $q < 0$, but directly related for $q \geq 0$.
3. Estimates of $S_q(T)\,|_{q=1}$ recover the standard entropy for different values of T.
4. Estimates of $S_q(T)$ increase with T for values of $q \geq 0$, up to a certain saturation value (maximum q-entropy).
5. The function $S(T)$ vs. T in log – log space, for different values of q, can be considered a (time) entropy analog of the (space) *structure function* in turbulence [38].
6. The scaling exponents, $\beta(q)$, or the slope of the relation $S(T)$ vs. T in log – log space, for different values of q, exhibit a non-linear growth with q, such that $\beta \approx 0.5$ for $q \geq 1$. Such non-linear behavior of the scaling exponents with respect to q implies that the time generalized q-entropy exhibits multiscaling, analogous to the spatial multi-scaling behavior of random fields [39,40].

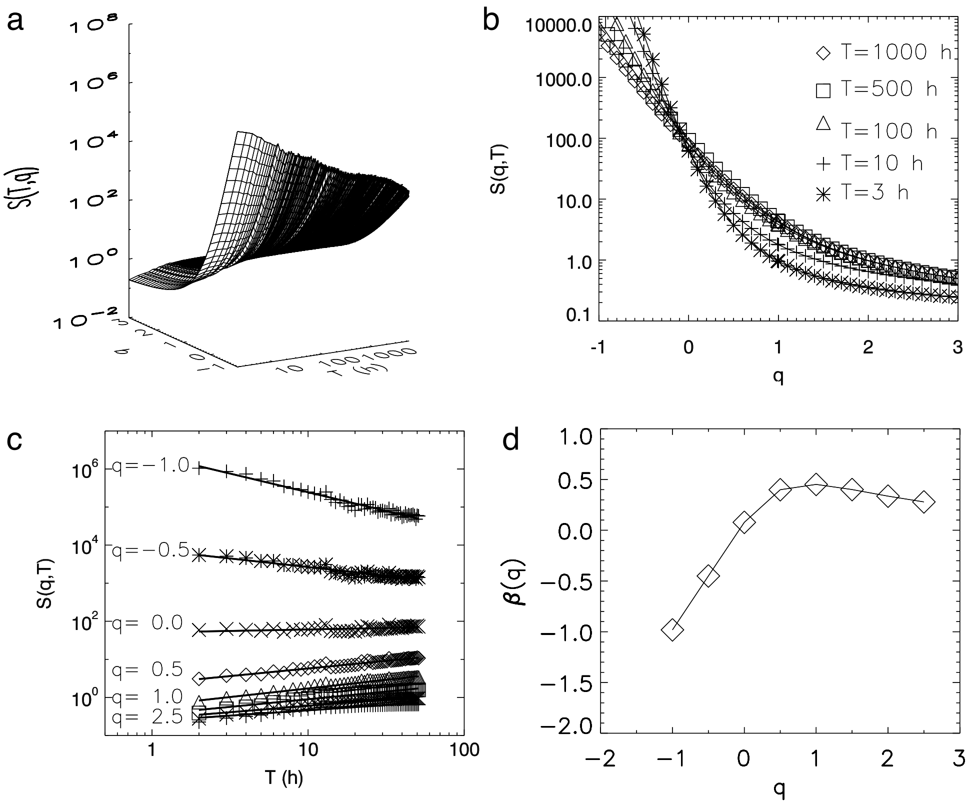

Fig. 13. Time generalized q-entropy for the 40-year long series of hourly rainfall at raingauge Cenicafe No. 1. (a) Surface $S(q, T)$ (Eq. (3)). (b) Projection of $S(q, T)$ vs. q for different values of T. (c) Projection of $S(q, T)$ vs. T, for different values of q, or time structure function for entropy. (d) Values of the regression slopes of the time structure function for entropy, $\beta(q)$, as a function of q, exhibiting a non-linear growth up to $\beta \approx 0.5$ for $q \geq 1$.

Panel (a) of Fig. 14 shows the time generalized q-entropy, $S_q(T)$ (Eq. (3)), for daily streamflows of the Atrato River in Colombia, along with their bi-dimensional projections (panels b and c), and the values of the regression slopes of the time structure function for entropy, $\beta(q)$ (panel d), as a function of q. Mathematically, the non-linear behavior of this function is characterized by $S_q(T) \sim T^{\beta(q)}$, with $\beta(q) = kq$ for $q \leq 0$, with k a positive constant, and $\beta(q) = 0$ for $q > 0$. In turn, the Shannon entropy ($q = 1$), behaves as $S(T) \sim T^{\beta}$, with $\beta \gtrsim 0$, as shown in Section 3.2. Similar results were obtained for streamflows of the Atrato and Cali Rivers in Colombia, the Tijuco and Correste Rivers in Brazil, and the Little Sioux River at Turin, IA, USA (Fig. 15). No major differences were found in the behavior of $S_q(T)$ for streamflows among the three regions.

Beyond the differences in the scaling exponents relating entropy with aggregation interval, $S_{q=1}(T)$, there is a remarkable difference between the behavior of $S_q(T)$ for rainfall and streamflow series, which has to do with the scaling exponents for different values of q, shown in panels (d) of Figs. 13–15. For rainfall, the relationship between q and $\beta(q)$ saturates at $\beta \approx 0.5$ for $q \geq 1$, whereas for streamflows it saturates at $\beta \cong 0$ for $q \geq 1$. This is a topic of ongoing research.

4. Conclusions

We have studied the behavior of the Shannon entropy of rainfall when series are aggregated at increasing timescales, $S(T)$, using data from the tropical Andes of Colombia, as well as of average daily streamflows of rivers in Colombia, the Amazon River basin, and USA. All rainfall series exhibit a power law $S(T) \sim T^{\beta}$ with $\beta = 0.5$, valid up to a certain timescale at which entropy reaches an upper bound, T_{MaxEnt}, and tappers off thereafter such that $\beta = 0$ for $T > T_{MaxEnt}$. The scaling exponents ($\beta = 0.5$) are nearly insensitive to changes in the number of bins employed to estimate the probability mass function, as well as to variations in record lengths from 21–24 months to 50–90 months and up to 40 years, although the values of T_{MaxEnt} tend to increase (∼50%) when record lengths increase from 21–24 months to 50–90 months. This last result can be explained by the presence of zeros in high-resolution rainfall records, in combination with the occurrence of more frequent El Niño events (in comparison to La El Niña events), and the associated rainfall negative anomalies, and the concomitant longer dry spells throughout most of Colombia.

Maximum rainfall entropy is reached through a dynamic Generalized Pareto distribution, thus making it a maximum dynamic entropy (MDE), in the sense of Ref. [16], for rainfall in the tropical Andes. Furthermore, for most raingauges, the

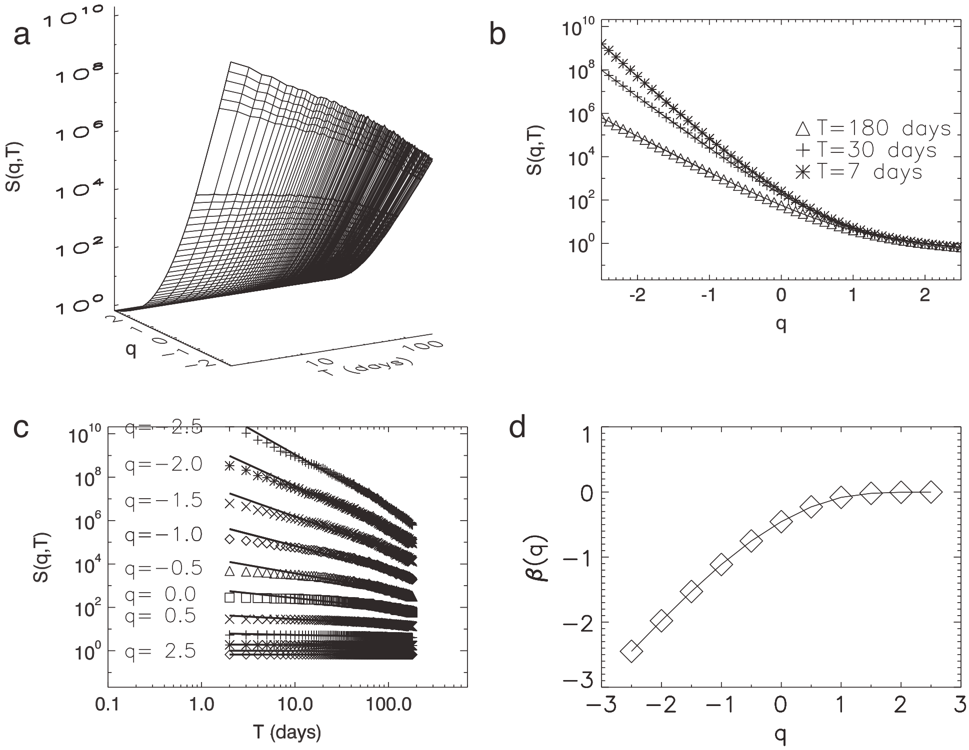

Fig. 14. As in Fig. 13 for the Atrato River in Colombia.

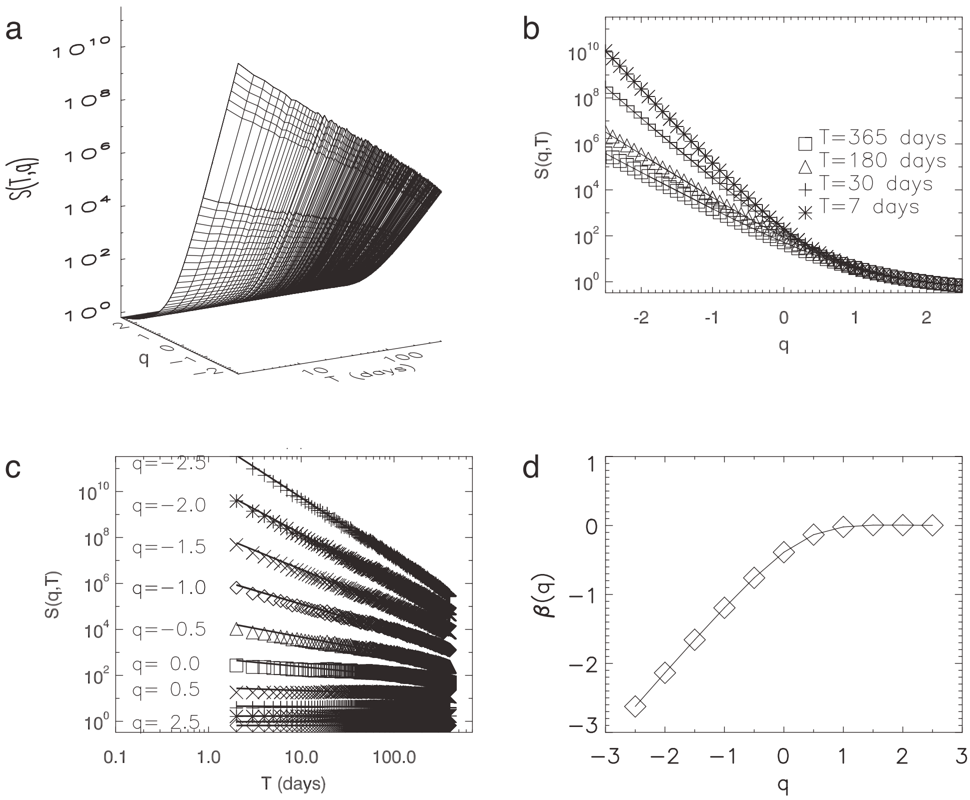

Fig. 15. As in Fig. 14 for the Little Sioux River at Turin, IA, USA.

three parameters of the fitted Generalized Pareto distributions (κ, σ, and μ) estimated at $T \approx T_{MaxEnt}$, do not exhibit significant statistical changes when record lengths are augmented from 21–24 months to 50–90 months.

Regarding daily streamflows, results show that $S(T) \sim T^{\beta}$ with $\beta \gtrsim 0$, and satisfying the theoretical limit. The difference between the scaling exponents of rainfall ($\beta = 0.5$) and streamflows ($\beta \gtrsim 0$) is explained by the presence of zeros in rainfall, and by the rate at which they vanish under aggregation at increasing intervals, in the form of two power laws: (i) $\zeta(T) = aT^{-\alpha_1}$, with $\langle\alpha_1\rangle = 1.24$ for the common periods, while $\langle\alpha_1\rangle = 1.34$ for the entire record lengths, and (ii) $\zeta(T) = aT^{-\alpha_2} + c$, with $\langle\alpha_2\rangle = 1.67$ for the common periods, whereas $\langle\alpha_2\rangle = 1.80$ for the entire record lengths. The two scaling regimes are separated at $T_b = 17$ h for the common periods, and at $T_b = 28$ h for the entire record lengths.

Rainfall series get devoid of zeros at $\langle T_{noz}\rangle = 164$ h for the common periods, while $\langle T_{noz}\rangle = 254$ h for the entire record lengths. In turn, the two timescales for maximum entropy and for the vanishing of zeros are related through a power-law, $T_{MaxEnt} \sim T_{noz}^{0.62}$ ($R^2 = 0.57$). These results are explained by the presence of zeros, combined with the interannual variability of rainfall, mostly driven in Colombia by El Niño/Southern Oscillation (ENSO).

The time generalized q-entropy, $S_q(T)$, for rainfall and streamflows consists of a continuous set of power laws with respect to T, for increasing values of q, whose scaling exponents vary nonlinearly with q, thus implying multiscaling of the time generalized q-entropy. Also, $S_q(T)$ exhibits a monotonic decrease with q for all values of T. Mathematically, $S_q(T) \sim T^{\beta(q)}$, with $\beta(q) = kq$ for $q \leq 0$, k is a positive constant, and $\beta(q) = 0$ for $q > 0$. It is also shown that $S(q, T)\,|_{q=1}$ recovers the standard entropy for all values of T. The standard Shannon entropy, $S_{q=1} \sim T^{\beta}$, with $\beta \gtrsim 0$ for daily streamflows, whereas $\beta = 0.5$ for series of 15-min to weekly rainfall records. No major differences were found among the estimated $S_q(T)$ for streamflows in Colombia, the Amazon basin, and USA.

Acknowledgments

The work of H.D. Salas was partly supported by AMAZALERT, a research programme funded by the European Commission under grant agreement No. 282664. G. Poveda's work has been supported and is a contribution of Universidad Nacional de Colombia at Medellin to the same research programme. We thank Oscar J. Mesa for insightful discussions, in particular for pointing us out to Eq. (4). Empresas Públicas de Medellín, Area Metropolitana del Valle de Aburrá (AMVA) environmental authority, and Cenicafé kindly provided rainfall data sets. IDEAM and CVC of Colombia, ANA of Brazil and the USGS of USA provided access to streamflow data.

References

[1] G. Poveda, Mixed memory, (non) Hurst effect, and maximum entropy of rainfall in the tropical Andes, Adv. Water Resour. 34 (2) (2011) 243–256.
[2] G. Poveda, L. Jaramillo, L.F. Vallejo, Seasonal precipitation patterns along pathways of South American low-level jets and aerial rivers, Water Resour. Res. 50 (2014) 98–118. http://dx.doi.org/10.1002/2013WR014087.
[3] G. Poveda, O.J. Mesa, On the existence of Lloró (the rainiest locality on earth): enhanced ocean–atmosphere–land interaction by a low-level jet, Geophys. Res. Lett. 27 (11) (2000) 1675–1678.
[4] G. Poveda, O.J. Mesa, L.F. Salazar, P.A. Arias, H.A. Moreno, S.C. Vieira, P. Agudelo, V.G. Toro, J.F. Alvarez, Diurnal cycle of precipitation in the tropical Andes of Colombia, Mon. Weather Rev. 133 (1) (2005) 228–240.
[5] A. Gires, I. Tchiguirinskaia, D. Schertzer, S. Lovejoy, Influence of the zero-rainfall on the assessment of the multifractal parameters, Adv. Water Resour. 45 (2012) 13–25.
[6] A. Gires, I. Tchiguirinskaia, D. Schertzer, S. Lovejoy, Development and analysis of a simple model to represent the zero rainfall in a universal multifractal framework, Nonlinear Processes Geophys. 20 (2013) 343–356.
[7] D. Schertzer, I. Tchiguirinskaia, S. Lovejoy, P. Hubert, Multifractal extreme value theory (MEV), in: AGU Chapman Conference on Complexity and Extreme Events in Geosciences, 2010.
[8] D.K. Nykanen, Linkages between orographic forcing and the scaling properties of convective rainfall in mountainous regions, J. Hydrometeorol. 9 (2008) 327–347. http://dx.doi.org/10.1175/2007JHM839.1.
[9] S. Lovejoy, D. Schertzer, Scaling and multifractal fields in the solid earth and topography, Nonlinear Process. Geophys. 14 (2007) 465–502.
[10] W. Gong, D. Yang, H.V. Gupta, G. Nearing, Estimating information entropy for hydrological data: one-dimensional case, Water Resour. Res. 50 (2014) 5003–5018. http://dx.doi.org/10.1002/2014WR015874.
[11] D. Veneziano, C. Lepore, The scaling of temporal rainfall, Water Resour. Res. 48 (2012) W08516. http://dx.doi.org/10.1029/2012WR012105.
[12] I. Andrés-Doménech, R. García-Bartual, A. Montanari, J.B. Marco, Climate and hydrological variability: the catchment filtering role, Hydrol. Earth Syst. Sci. Discuss. 11 (2014) 10411–10430.
[13] W. Brutsaert, Hydrology: An Introduction, Cambridge University Press, 2005.
[14] V.P. Singh, Hydrologic synthesis using entropy theory: review, J. Hydrol. Eng. 16 (5) (2011) 421–433.
[15] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948) 379–423.
[16] M. Asadi, N. Ebrahimi, G.G. Hamedani, S.E. Soofi, Maximum dynamic entropy models, J. Appl. Probab. 41 (2004) 379–390.
[17] C. Tsallis, Possible generalization of Boltzmann–Gibbs statistics, J. Stat. Phys. 52 (1988) 479–487.
[18] J.-F. Bercher, Tsallis distribution as a standard maximum entropy solution with 'tail' constraint, Phys. Lett. A 372 (2008) 5657–5659.
[19] C. Tsallis, R.S. Mendes, A.R. Plastino, The role of constraints within generalized nonextensive statistics, Physica A 261 (1998) 534–554.
[20] A. Plastino, Why Tsallis' statistics? Physica A 344 (2004) 608–613.
[21] P.N. Rathie, S. Da Silva, Shannon, Lévy, and Tsallis: a note, Appl. Math. Sci. 2 (2008) 1359–1363.
[22] C. Tsallis, Nonextensive statistical mechanics, anomalous diffusion and central limit theorems, Milan J. Math. 73 (2005) 145–176. http://dx.doi.org/10.1007/s00032-005-0041-1.
[23] S. Umarov, C. Tsallis, S. Steinberg, A generalization of the central limit theorem consistent with nonextensive statistical mechanics, Milan J. Math. 76 (2008) 307. http://dx.doi.org/10.1063/1.2828756.
[24] D. Sornette, Critical Phenomena in Natural Sciences, Chaos, Fractals, Self-Organization and Disorder: Concepts and Tools, Springer, 2000.
[25] K.H. Knuth, A. Gotera, C.T. Curry, K.A. Huyser, K.R. Wheeler, B. Rossow, Revealing relationships among relevant climate variables with information theory. in: Paper Presented at Earth–Sun System Technology Conference, NASA Adelphi Md, 2005.
[26] J.M. Nichols, Inferences about information flow and dispersal for spatially extended population systems using time-series data, Proc. R. Soc. Lond. Ser. B 272 (2005) 871–876. http://dx.doi.org/10.1098/rspb.2004.2889.

[27] K.H. Knuth, Optimal data-based binning for histograms, 2006. arxiv:physics/0605197v1.
[28] J.O. Rawlings, S.G. Pantula, D.A. Dickey, Applied Regression Analysis: A Research Tool, Springer-Verlag, 1988.
[29] J. Zhang, Y. Wu, *k* Samples test based on the likelihood ratio, Comput. Statist. Data Anal. 51 (2007) 4682–4691.
[30] I. Bialynicki-Birula, J. Mycielski, Uncertainty relations for information entropy in wave mechanics, Comm. Math. Phys. 44 (2) (1975) 129–132.
[31] A.F. Hurtado, G. Poveda, Linear and global space–time dependence and Taylor hypotheses for rainfall in the tropical Andes, J. Geophys. Res. 114 (2009) D10105. http://dx.doi.org/10.1029/2008JD011074.
[32] O. Peters, K. Christensen, Rain: relaxations in the sky, Phys. Rev. E 66 (2002) 036120.
[33] J.D. Gibbons, Nonparametric Statistical Inference, Marcel Dekker, 1985.
[34] M. Hollander, D.A. Wolfe, Nonparametric Statistical Methods, John Wiley and Sons, 1999.
[35] G. Poveda, O.J. Mesa, Feedbacks between hydrological processes in tropical South America and large scale oceanic atmospheric phenomena, J. Clim. 10 (1997) 2690–2702.
[36] G. Poveda, A. Jaramillo, M.M. Gil, N. Quiceno, R.I. Mantilla, Seasonality in ENSO related precipitation, river discharges, soil moisture, and vegetation index (NDVI) in Colombia, Water Resour. Res. 37 (2001) 2169–2178.
[37] G. Poveda, P.R. Waylen, R. Pulwarty, Modern climate variability in northern South America and southern Mesoamerica, Palaeogeogr. Palaeoclimatol. Palaeoecol. 234 (2006) 3–27.
[38] U. Frisch, Turbulence: The Legacy of A.N. Kolmogorov, Cambridge University Press, 1995.
[39] V.K. Gupta, E. Waymire, Multiscaling properties of spatial rainfall and river flow distributions, J. Geophys. Res. 95 (1990) 1999–2009.
[40] V.K. Gupta, E. Waymire, A statistical analysis of mesoscale rainfall as a random cascade, J. Appl. Meteorol. 32 (1993) 251–267.