

Research Article

Clustering of Rainfall Stations in RH-24 Mexico Region Using the Hurst Exponent in Semivariograms

Francisco Gerardo Benavides-Bravo,¹ F-Javier Almaguer,² Roberto Soto-Villalobos,² Víctor Tercero-Gómez,³ and Javier Morales-Castillo²

¹Instituto Tecnológico de Nuevo León, 67170 Guadalupe, NL, Mexico

²Universidad Autónoma de Nuevo León, 66451 San Nicolás de los Garza, NL, Mexico

³Tecnológico de Monterrey, 64849 Monterrey, NL, Mexico

Correspondence should be addressed to Francisco Gerardo Benavides-Bravo; fgbenavid@gmail.com

Received 15 August 2015; Revised 14 November 2015; Accepted 23 November 2015

Academic Editor: Costas Panagiotakis

Copyright © 2015 Francisco Gerardo Benavides-Bravo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An important topic in the study of the time series behavior and, in particular, meteorological time series is the long-range dependence. This paper explores the behavior of rainfall variations in different periods, using long-range correlations analysis. Semivariograms and Hurst exponent were applied to historical data in different pluviometric stations of the Río Bravo-San Juan watershed, at the hydrographic RH-24 Mexico region. The database was provided by the Water National Commission (CONAGUA). Using the semivariograms, the Hurst exponent was obtained and used as an input to perform a cluster analysis of rainfall stations. Groups of homogeneous samples that might be useful in a regional frequency analysis were obtained through the process.

1. Introduction

When limited observations of hydrological events are available, the ability to provide appropriate characterization, analysis, and predictions of a phenomenon gets compromised. However, the analysis can be improved by identifying homogeneous samples that can be used in combination to make better estimates of a probability model. This is one of the major concerns within the practice of regional frequency analysis (RFA), where the final output is the estimation of extreme events in a geographical area that can be used as input in risk analysis, water management, zoning, and land use applications, Hosking and Wallis [1]. However, the estimation of extreme events is considered a complex problem, mostly because the information is usually limited, serial correlation exists, multiple change-points might be present, and observations follow trends and seasonal patterns. To address these issues, hydrological time series studies have been successfully applied in the past, Machiwal and Jha [2]; however, most research efforts have been focused on trend detection tests, leaving aside other important properties such

as stationarity, homogeneity, periodicity, and persistence. By addressing these properties, a better selection of homogeneous samples might be possible, and as a consequence, practitioners might achieve better predictions.

Previous works on time series analysis in climatology with applications in precipitation go back to Bhuiya [3], with the development of a test for stationarity after periodic and trend components were subtracted from hydrologic series. Buishand [4] used trend tests to evaluate the difference in precipitation between rural and urban areas of Amsterdam and Rotterdam. Buishand [5, 6] constructed several tests of homogeneity in the mean of series with the use of cumulative sums, likelihood tests, and Bayesian inference. Kothyari et al. [7] evaluated three stations in India, Agra, Dehradun, and Dehli, to test for changes in rainfall and temperature, providing evidence of a change in the number of rainy days during monsoon season and an increment in temperature. Giakoumakis and Baloutsos [8] performed a trend analysis on historical series of annual precipitations from the basin of the Evinos Riven in Greece. By applying different tests of randomness, decreasing trends were found in the rainfall

records. Other authors dealing with trend analysis, homogeneity, and change-points found in the literature are Angel and Huff [9], Mirza et al. [10], Tarhule and Woo [11], Luís et al. [12], Kripalani and Kulkarni [13], Adamowski and Bougadis [14], Yu et al. [15], and Kumar et al. [16]. A comprehensive review of these works can be found in Machiwal and Jha [2] with descriptions of related developments in hydrological time series analysis.

Recent developments in hydrological analysis include the works of Golian et al. [17] with a classification and clustering approach of rainfall data using the natural-breaks classification method and the fuzzy c-means (FCM) algorithm. Shi et al. [18] analyzed variations in trends for precipitation data using a linear regression method, the Mann-Kendall test, and the Hurst exponent. The Hurst exponent, as part of a fractal analysis, was used to evaluate long-range dependence and the possibility of trends in the data. The following works around the Hurst exponent include the developments of Golder et al. [19], where the Hurst exponent is also used to explore long-term correlations, and cumulative rainfall observations were modeled using the alpha-stable probability law to deal with heavy-tailed distributions. Chang [20] extended the application of the Hurst exponent by developing a computation approach to estimate the exponent over time series that fits a discrete time fractional Brownian motion and fractional Gaussian noise. Yu et al. [21] also studied long-term correlations using the Hurst exponent and performed a multifractal analysis of rainfall series (see Kantelhardt [22]) based on a multiplicative cascade model and a multifractal detrended fluctuation analysis. Other recent works on time series analysis can be found in Carbone et al. [23] with the construction of a simulation model of storms using a double exponential distribution. Chou [24] investigated the complexity at different temporal scales of rainfall and runoff time series using the sample-entropy method, and finally, García-Marín et al. [25] performed a regional frequency analysis over rainfall data from Málaga, Spain, where the grouping of stations into homogenous regions has been done by following a cluster analysis with multifractal values of the different series.

In this paper, a clustering approach is used to group stations into homogeneous samples after summarizing the results of semivariogram analysis into a Hurst exponent. As a case study, a sample of pluviometric stations, the Río Bravo-San Juan watershed, at the RH-24 Mexico region was analyzed (Figure 1). A map of the Río Bravo-San Juan watershed is shown in Figure 2. This region is located in Mexico between the states of Nuevo León, Coahuila, and Tamaulipas covering an approximate area of 29420 km². Some of the rainfall stations are shown in Figure 3. The data used has been provided by CONAGUA, the local institution responsible for water management in the country.

2. Problem Description

In practice, to perform RFA, it is required to identify homogeneous regions where data follow similar patterns that can be analyzed together to improve the identification of probability models that in turn can be used to estimate extreme events



FIGURE 1: The watersheds of Mexico with Google Earth. In black edge, the Río Bravo-San Juan watershed. Source of the database: <http://www.conagua.gob.mx/>.

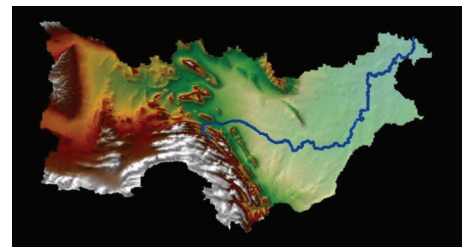


FIGURE 2: Río Bravo-San Juan watershed (San Juan river in blue). Image source: “Water Management in the Río San Juan Watershed, in the Southern Río Bravo Hydrologic Region of Mexico” at <http://earthzine.org/2012/08/13/>.

and their frequency in terms of return periods. This analysis is usually executed when dealing with droughts, pollution, wind movement, temperature, atmospheric pressure, and rainfall observations, to name a few. These researches deal with the problem of finding groups of rainfall stations that create homogeneous regions by considering fractal structures captured through semivariograms and Hurst exponents. Rainfall data from a sample of the hydrographic region RH-24 Mexico, the Río Bravo-San Juan watershed, are used as a case study to evaluate the proposed approach.

3. Methodology

Semivariograms, in the present study, are used to quantify long-range correlations of data from different pluviometric stations using monthly records. By considering the analysis of semivariograms of historical series, a rescaled range analysis R/S is performed to obtain a measure of the Hurst exponent [26]. The Hurst exponent is used as a metric of a particular pluviometric station. The process is repeated over each pluviometric station within the region under analysis. Hurst exponents are used as a reference to identify stations that exhibit similar patterns. As a consequence, a cluster analysis is applied to identify homogeneous samples. An advantage of the Hurst exponent is the simplicity of its algorithm that can be used to measure the condition of persistence or antipersistence of a process, and it provides a metric that can be used to classify different time series.

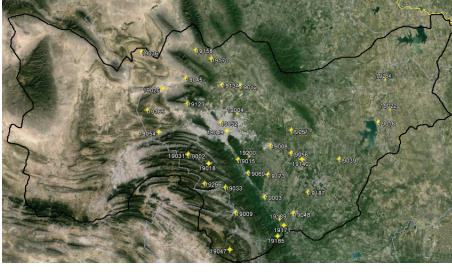


FIGURE 3: Geographical locations (from Google Earth) of the pluviometric stations at the Río Bravo-San Juan watershed. Source of the database: <http://www.conagua.gob.mx/>.

3.1. Semivariogram. The semivariogram or variogram $\gamma(h)$ is used to describe the relationship of paired observations separated by a distance h . It is a geostatistical technique that allows a quantitative measure of the long-range persistence in nonstationary time series Witt and Malamud [27], Haslett [28], and Dmowska and Saltzman [29]. Correlations over time and space create patterns that can be used to describe the behavior of a set of observations. Mathematically, the variogram estimates the expected squared difference between neighboring random variables. This calculation is performed over different h values. Given a time series or stochastic processes $\{X_t, t \geq 0\}$, the autocovariance function at the point $(t, t+h)$ is defined as $C_X(t, t+h) = E[X_t X_{t+h}] - E[X_t]E[X_{t+h}]$, with $E[X_t]$ as the mean of the process at time t . The semivariogram $\gamma(h)$ is given by half of the variance of the difference between pairs of observations at different “locations” in time:

$$\begin{aligned} \gamma(h) &= \frac{1}{2} \text{Var}(X_{t+h} - X_t) \\ &= \frac{1}{2} E[(X_{t+h} - X_t - E[X_{t+h} - X_t])^2] \\ &= \frac{\text{Var}(X_t) + \text{Var}(X_{t+h})}{2} - \text{Cov}(X_t, X_{t+h}) \quad (1) \\ &= \frac{\text{Var}(X_t) + \text{Var}(X_{t+h})}{2} \\ &\quad - \sqrt{\text{Var}(X_t) \text{Var}(X_{t+h})} \rho_X(t, t+h), \end{aligned}$$

where $-1 \leq \rho_X(t, t+h) \leq 1$ is the autocorrelation function (the autocovariance function normalized).

In the special case when $\rho(x_{t+h}, x_t) = 0, \forall(t, h)$, it is said that the stochastic processes $\{X_t, t \geq 0\}$ are uncorrelated and the semivariogram is reduced to the arithmetic mean of the variance of processes at times t and $t+h$:

$$\gamma(h) = \frac{\text{Var}(X_t) + \text{Var}(X_{t+h})}{2}. \quad (2)$$

If the random field $\{X_t, t \geq 0\}$ has constant mean $= E[X_t] = E[X_{t+h}] = \mu \forall t$, the semivariogram (1) adopts the simple form:

$$\gamma(h) = \frac{1}{2} E[(X_{t+h} - X_t)^2]. \quad (3)$$

If $X(t)$ and $X(t+h)$ are independent random variables $\forall t$, again, though for a different reason, the semivariogram is reduced to the special case (2).

In principle, given a stochastic process $\{X_t, t \geq 0\}$, the expected value of differences $E[X_{t+h} - X_t]$ at time t and with lag h is empirically estimated by the average over a “large enough” ensemble of realizations or paths in time. However, for a single time series $\{X_n, n = 1, 2, \dots, n\}$, the expected value can be estimated assuming an ergodic hypothesis, that is, a statistical principle of equivalence according to which *the average over time and the average over the ensemble are the same*, Lefebvre [30]. Thereby the differences $X_{t+h} - X_t$, which would be obtained with an infinitely reproducible process, are “simulated” or “cloned” from the “mother series.” Thus, the average value of differences $X_{t+h} - X_t$ is estimated by

$$E[X_{t+h} - X_t] = \frac{1}{n(h)} \sum_{i=1}^{n(h)} (x_{i+h} - x_i), \quad (4)$$

where $n(h)$ is the number of differences with a lag h . When $h = 1, 2, 3, \dots$, the averages (4) are, respectively,

$$\begin{aligned} E[X_{t+1} - X_t] &= \frac{x_n - x_1}{n-1}, \\ E[X_{t+2} - X_t] &= \frac{x_{n-1} + x_n - (x_1 + x_2)}{n-2}, \\ E[X_{t+3} - X_t] &= \frac{x_{n-2} + x_{n-1} + x_n - (x_1 + x_2 + x_3)}{n-3}, \\ &\vdots \\ E[X_{t+h} - X_t] &= \frac{1}{n(1-h/n)} \left(\sum_{j=0}^{k-1} x_{n-j} - \sum_{l=1}^k x_l \right). \end{aligned} \quad (5)$$

According to (5) for a maximum value of h “relatively moderate” or $h/n < 1$, except in the presence of isolated extreme outliers, the two summations in (5) are roughly of the same order, such that the empirical average value (4) can be approximated by $E[X_{t+h}] \approx E[X_t] = m = \text{constant}$. This is an observed characteristic in the time series of the pluviometric stations. Therefore, the corresponding estimator of (3) is simply

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} (x_{t+h} - x_t)^2. \quad (6)$$

3.2. Measurement of the Hurst Exponent H . To estimate the Hurst exponent from a temporal series $\{X_k\}$, with $k \in 1, 2, \dots, N$, the series is divided in a group of d -subseries of length m . Really, the size m is an average number. A standard way, though not the only, to obtain the m size of the subseries is partitioning the original series in powers of base 2. In doing so, in each of the successive partitions, the approximate value

of m is as follows: $N, N/2, N/2^2, N/2^3, \dots$. For each subseries $n = 1, 2, \dots, d$, do the following:

- (1) Calculate the mean E_n and the standard deviation S_n .
- (2) Calculate the deviation with respect to the mean by subtracting the mean of each element using

$$Z_{in} = X_{in} - E_n, \quad i = 1, 2, \dots, m. \quad (7)$$

- (3) Get the partial sums:

$$Y_{in} = \sum_{j=1}^i Z_{jn}, \quad i = 1, 2, \dots, m. \quad (8)$$

- (4) Calculate the range:

$$R_n = \max_{i=1:m} (Y_{in}) - \min_{i=1:m} (Y_{in}). \quad (9)$$

- (5) Normalize the range:

$$\frac{R_n}{S_n}. \quad (10)$$

- (6) For each subseries of length m take the average:

$$\left\langle \frac{R}{S} \right\rangle_m = \frac{1}{d} \sum_{n=1}^d \frac{R_n}{S_n}. \quad (11)$$

- (7) Hurst [26] found the relation of the statistical $\langle R/S \rangle_m$ given by the following power law:

$$\left\langle \frac{R}{S} \right\rangle_m \approx cm^H, \quad (12)$$

where H is the Hurst exponent and c is a positive constant.

Two factors involved in the determination of the Hurst coefficient are the way time series is divided into a group of subseries and the asymptotic behavior of the rescaled range. First, the range of values m are used to calculate the slope of $\log(\langle R/S \rangle_m)$ given the relationship

$$\log \left(\left\langle \frac{R}{S} \right\rangle_m \right) = \log(c) + H \log(m). \quad (13)$$

Second, the determination of H is the result of the asymptotic behavior of the rescaled range, that is, when the value m tends to infinity. The analysis of the rescaled $\langle R/S \rangle_m$ over some values of m is estimated using a log/log expression given in (13). To obtain H coefficient, the least-squares method is used. The slope of this line is the Hurst coefficient H .

This exponent is considered a fractal index, Mandelbrot and Wallis [31], and provides information about long-term correlations exhibited by a series of observations; for a theoretical review of the Hurst exponent, see Mandelbrot [32]. In practice, the Hurst exponent can take values between 0 and 1, where

- (i) $0 < H < 0.5$ indicates nonpersistence in a series; that is, an increment is more likely to be followed by a decrement and vice versa;
- (ii) $H = 0.5$ indicates lack of serial correlation (Gaussian white noise);
- (iii) $0.5 < H < 1$ indicates persistency; that is, an increment is likely to be followed by an increment and a decrement by another decrement.

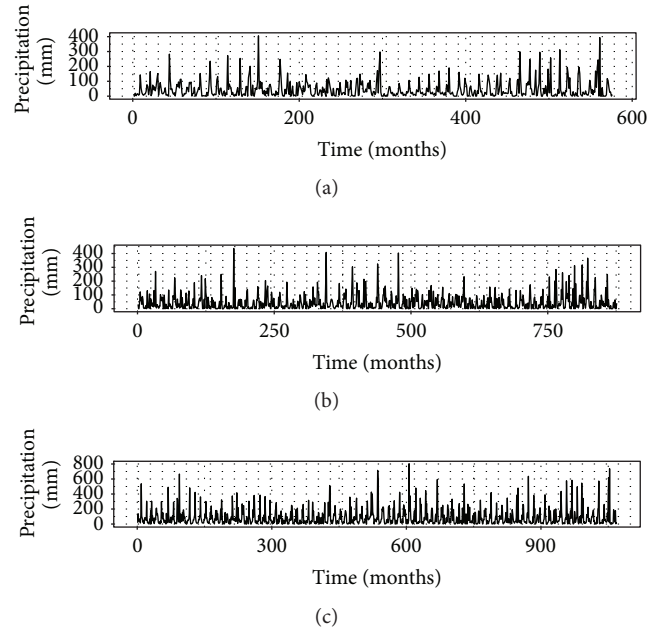


FIGURE 4: Time series for rainfall measurements from three stations at Río Bravo-San Juan watershed, Mexico. From the top to the bottom, respectively: Apodaca, station number 19004, 1940 January–2012 December; El Cuchillo, station number 19016, 1939 January–2012 December; La Boca, station number 19069, 1923 January–2012 December.

4. Results

To illustrate the procedure, only the analysis of three rainfall stations was selected to be presented in this section. The measured values of monthly precipitation in millimeters from the stations Apodaca, El Cuchillo, and La Boca are displayed in Figure 4, and results obtained taking into account all stations (following the same process) are presented at the end of this section.

As can be seen, patterns and relationships between different stations are difficult to assess only through “eyeball” analysis. However, when semivariograms are obtained, as shown in Figure 5, a footprint of the data becomes more evident. A closer inspection in every station shows a seasonal pattern that repeats every 12 observations in the semivariogram. This can be explained due to the fact that monthly observations were used in the analysis.

Once the semivariograms were obtained, the Hurst exponent for the series of the γ values was calculated. It can be seen that Hurst coefficients are close to 1, which indicates a positive long dependency of the data in the semivariograms. Hurst exponents from all stations under analysis are presented in Table 1, where the long dependency in all variograms becomes clear. Some of the coefficients that appear in the table exceed the interval established of possible values of the Hurst exponent, $0 < H < 1$; this is a known error due to estimation bias or a possible linear retrogression.

To address the issue of finding homogeneous samples, a cluster analysis is performed using estimates of the Hurst exponent. As shown in Figure 6, a histogram of frequencies

TABLE 1: Hurst coefficients for semivariogram (6) from pluviometric stations at Río Bravo-San Juan watershed.

Station	Name	Latitude	Longitude	Data	Hurst
19015	El Cerrito	25 30 36	100 11 36	1939–2012	0.71245
19039	Las Enramadas	25 30 05	099 31 17	1940–2012	0.78917
19018	El Pajonal	25 29 23	100 23 20	1955–2012	0.79764
19012	Ciénega de Flores	25 57 08	100 10 20	1940–2012	0.80106
19003	Allende	25 17 01	100 01 13	1940–2012	0.82492
19069	La Boca	25 25 46	100 07 44	1923–2013	0.8387
19189	El Pastor	25 09 06	099 55 36	1987–2011	0.84065
19009	Casillas	25 11 47	100 12 51	1956–2012	0.84258
19187	California	25 18 23	099 44 02	1982–2011	0.84406
19173	Palmitos	25 25 02	099 59 50	1982–2012	0.85304
19002	Agua Blanca	25 32 39	100 31 23	1958–2012	0.86171
19134	Salinas Victoria	25 57 33	100 17 34	1979–2011	0.86501
19031	La Cruz	25 32 47	100 31 23	1955–2011	0.86751
19036	La Popa	26 09 50	100 49 40	1956–2011	0.87324
19008	Cadereyta Jiménez	25 35 25	099 58 30	1995–2012	0.87335
19185	El Canada	25 02 48	099 56 29	1982–2011	0.87351
19056	San Juan	25 32 36	099 50 25	1944–2012	0.87678
19016	El Cuchillo	25 43 05	099 15 21	1960–2012	0.87778
19052	Monterrey(Obs)	25 44 01	100 16 01	1986–2008	0.87802
19026	Icamole	25 56 28	100 41 13	1954–2012	0.88481
19200	La Cienega	25 32 10	100 07 15	1984–2011	0.89216
19004	Apodaca	25 47 37	100 11 50	1964–2012	0.89506
19047	Mimbres	24 58 26	100 15 31	1957–2011	0.89837
19140	Tepehuaje	25 30 19	099 46 15	1979–2012	0.91092
19267	Santa Ma. La Floreña	25 10 59	99 46 00	1984–2011	0.91686
19048	Montemorelos	25 10 55	099 49 56	1940–2012	0.91836
19040	Los Aldama	26 03 52	099 11 48	1942–1994	0.92157
19022	General Bravo	25 48 05	099 10 32	1927–2012	0.92301
19158	Rancho de Gomas	26 10 11	100 27 52	1981–2011	0.93016
19045	Mina	26 00 08	100 32 00	1953–2012	0.93308
19054	Rinconada	25 40 52	100 43 03	1945–2012	0.93977
19165	Chupaderos del Indio	25 48 49	100 47 24	1982–2011	0.94288
19266	San Jose de Barranquillas	26 32 41	100 28 21	1978–2011	0.94911
19171	Lampacitos	25 06 38	099 53 57	1982–2011	0.95188
19264	Dr. Coss	25 51 16	099 56 36	1982–2011	0.96461
19170	El Hojase	26 06 55	100 21 38	1982–2011	1.00650

was used to separate stations into 6 clusters. These clusters and the result of fitting probability distributions over the data of each rainfall station are shown in Table 2.

Distributions were selected based on a goodness of fit analysis. After identifying a set of feasible distributions with p values bigger than a significant level of 0.05, in every case, the distribution with the highest average p value was selected for each station. Gamma and Generalized Extreme Value were the distributions that gave the best fit over the data analyzed. These functions are

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} e^{-x/\beta}, \tag{14}$$

$$f(x; k, \sigma, \mu) = \frac{1}{\sigma} e^{-(1+kz)^{-1/k}} (1+kz)^{-1-1/k}, \tag{15}$$

$$z = \frac{x - \mu}{\sigma},$$

respectively.

As a result of the analysis, estimated distributions do not mix within cluster. This input can be used in a posterior analysis to obtain maximum likelihood estimators of the distribution parameters using all data concurrently.

5. Conclusions

Variograms followed by analysis of R/S or Hurst exponent estimation were used as input of a cluster analysis. Hurst

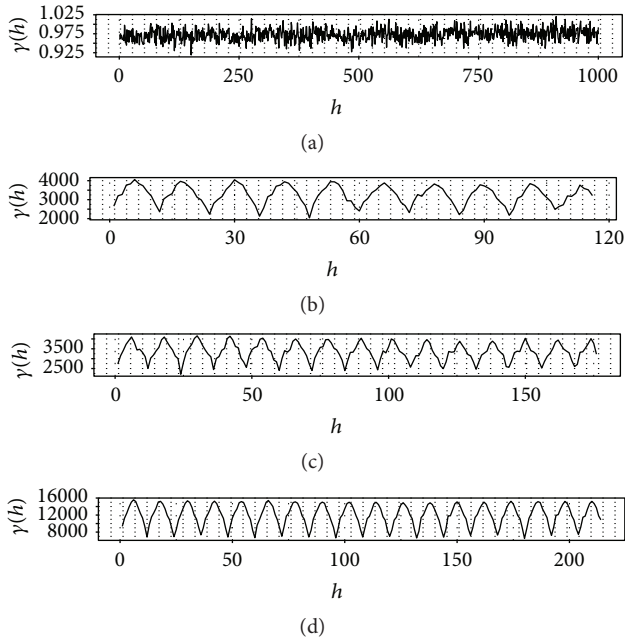


FIGURE 5: Semivariogram (6) corresponding to the time series of rainfall measurements from the three stations of the Río Bravo-San Juan watershed shown in Figure 4. In each case the maximum lag $h_{\max} = 0.20 * N$ was used, with N indicating the size of the time series. For comparison, the semivariogram corresponding to Gaussian white noise is included. From top to bottom, respectively: Gaussian white noise, Apodaca (station number 19004), El Cuchillo (station number 19016), and La Boca (station number 19069).

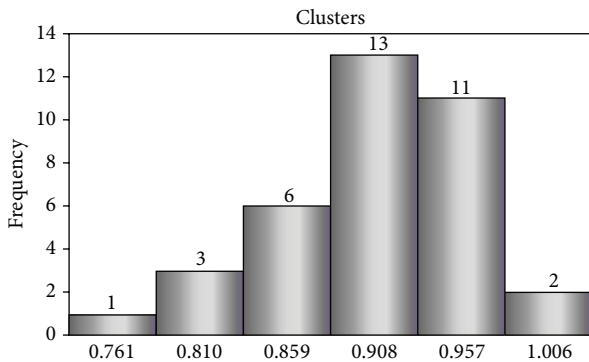


FIGURE 6: Histogram for the Hurst exponents of the analyzed pluviometric stations.

exponent provides a measure to determine if a time series is like a Gaussian white noise or has underlying trends, and it can be used to cluster the pluviometric stations according to the values of their semivariograms. As a case study to evaluate this approach, a sample of rainfall stations, those included in the Río Bravo-San Juan watershed from the hydrographic region RH-24 Mexico, was used in the analysis. Long-range dependency was found in every variogram evaluated with the Hurst exponent; however, it was still found useful as an input of a cluster analysis. A goodness of fit process was executed with every series, and the results showed that

TABLE 2: Clustering of pluviometric stations by the Hurst exponent of the semivariogram.

Cluster	Station	Hurst	Distribution	Parameters
1	19015	0.71245	Gamma(α, β)	(0.7022, 121.06)
	19039	0.78917		(0.8822, 101.75)
3	19018	0.79764	Gamma(α, β)	(0.5731, 78.07)
	19012	0.80106		(0.5402, 119.7)
6	19003	0.82492	GEV(k, σ, μ)	(0.349, 43.609, 38.143)
	19069	0.83870		(0.389, 43.321, 33.483)
	19189	0.84065		(0.453, 35.036, 22.915)
	19009	0.84258		(0.376, 25.309, 17.281)
	19187	0.84406		(0.377, 31.132, 24.640)
	19173	0.85304		(0.379, 32.826, 25.056)
	19002	0.86171		(0.351, 27.643, 20.471)
13	19134	0.86501	GEV(k, σ, μ)	(0.319, 21.703, 16.845)
	19031	0.86751		(0.313, 33.237, 23.081)
	19036	0.87324		(0.506, 9.5213, 5.0588)
	19008	0.87335		(0.407, 28.933, 22.421)
	19185	0.87351		(0.379, 22.415, 16.03)
	19056	0.87678		(0.372, 30.176, 22.56)
	19016	0.87778		(0.366, 24.295, 16.683)
	19052	0.87802		(0.429, 25.864, 19.559)
	19026	0.88481		(0.418, 8.7536, 5.424)
	19200	0.89216		(0.407, 33.194, 23.434)
11	19004	0.89506	GEV(k, σ, μ)	(0.368, 23.938, 17.563)
	19047	0.89837		(0.168, 33.305, 28.427)
	19140	0.91092		(0.369, 30.367, 23.162)
	19267	0.91686		(0.407, 20.642, 15.472)
	19048	0.91836		(0.356, 37.599, 30.006)
	19040	0.92157		(0.356, 20.614, 13.77)
	19022	0.92301		(0.383, 25.054, 16.587)
	19158	0.93016		(0.421, 16.537, 10.853)
	19045	0.93308		(0.457, 12.116, 7.9187)
	19054	0.93977		(0.485, 9.1504, 5.2866)
	19165	0.94288		(0.434, 10.814, 6.4381)
2	19266	0.94911	GEV(k, σ, μ)	(0.369, 32.361, 22.162)
	19171	0.95188		(0.435, 31.356, 22.367)
	19264	0.96461		(0.401, 20.168, 1.2395)
	19170	1.00650		(0.480, 20.046, 12.109)

existing dominant distributions within a feasible set (found independently in each station) do not overlap over clusters. The probability distributions were found nested within each cluster. This is indicative that homogeneous patterns were identified within groups, and groups were heterogeneous between themselves.

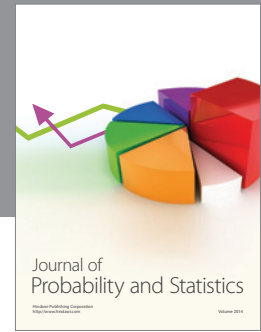
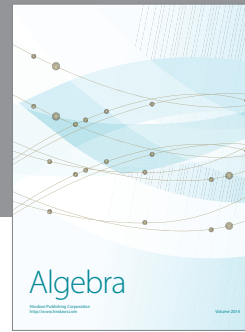
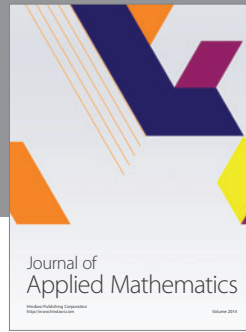
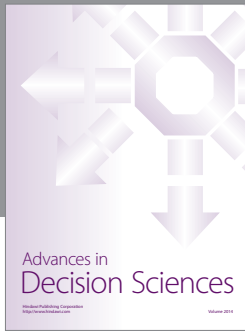
The study of the rainfall stations with semivariograms and R/S analysis provides a powerful tool that allows practitioners to analyze long-term correlations and clustering in hydrological time series. In future work, L -moments and spectral and wavelets analysis will be used to improve understanding of complex time series of pluviometric rainfall levels.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

- [1] J. R. M. Hosking and J. R. Wallis, *Regional Frequency Analysis: An Approach Based on L-Moments*, 2005.
- [2] D. Machiwal and M. K. Jha, "Current status of time series analysis in hydrological sciences," in *Hydrologic Time Series Analysis: Theory and Practice*, pp. 96–136, Springer Netherlands, 2012.
- [3] R. K. Bhuiya, "Stochastic analysis of periodic hydrologic process," *Journal of the Hydraulics Division*, vol. 97, no. 7, pp. 949–962, 1971.
- [4] T. A. Buishand, "Urbanization and changes in precipitation, a statistical approach," *Journal of Hydrology*, vol. 40, no. 3-4, pp. 365–375, 1979.
- [5] T. A. Buishand, "Some methods for testing the homogeneity of rainfall records," *Journal of Hydrology*, vol. 58, no. 1-2, pp. 11–27, 1982.
- [6] T. A. Buishand, "Tests for detecting a shift in the mean of hydrological time series," *Journal of Hydrology*, vol. 73, no. 1-2, pp. 51–69, 1984.
- [7] U. C. Kothyari, V. P. Singh, and V. Aravamuthan, "An investigation of changes in rainfall and temperature regimes of the Ganga Basin in India," *Water Resources Management*, vol. 11, no. 1, pp. 17–34, 1997.
- [8] S. G. Giakoumakis and G. Baloutsos, "Investigation of trend in hydrological time series of the Evinos River basin," *Hydrological Sciences Journal*, vol. 42, no. 1, pp. 81–88, 1997.
- [9] J. R. Angel and F. A. Huff, "Changes in heavy rainfall in Midwestern United States," *Journal of Water Resources Planning and Management*, vol. 123, no. 4, pp. 246–249, 1997.
- [10] M. Q. Mirza, R. A. Warrick, N. J. Ericksen, and G. J. Kenny, "Trends and persistence in precipitation in the Ganges, Brahmaputra and Meghna river basins," *Hydrological Sciences Journal*, vol. 43, no. 6, pp. 845–858, 1998.
- [11] A. Tarhule and M.-K. Woo, "Changes in rainfall characteristics in northern Nigeria," *International Journal of Climatology*, vol. 18, no. 11, pp. 1261–1271, 1998.
- [12] M. D. Luís, J. Raventós, J. C. González-Hidalgo, J. R. Sánchez, and J. Cortina, "Spatial analysis of rainfall trends in the region of Valencia (East Spain)," *International Journal of Climatology*, vol. 20, no. 12, pp. 1451–1469, 2000.
- [13] R. H. Kripalani and A. Kulkarni, "Monsoon rainfall variations and teleconnections over South and East Asia," *International Journal of Climatology*, vol. 21, no. 5, pp. 603–616, 2001.
- [14] K. Adamowski and J. Bougadis, "Detection of trends in annual extreme rainfall," *Hydrological Processes*, vol. 17, no. 18, pp. 3547–3560, 2003.
- [15] P.-S. Yu, T.-C. Yang, and C.-C. Kuo, "Evaluating long-term trends in annual and seasonal precipitation in Taiwan," *Water Resources Management*, vol. 20, no. 6, pp. 1007–1023, 2006.
- [16] V. Kumar, S. K. Jain, and Y. Singh, "Analysis of long-term rainfall trends in India," *Hydrological Sciences Journal*, vol. 55, no. 4, pp. 484–496, 2010.
- [17] S. Golian, B. Saghafian, S. Sheshangosht, and H. Ghalkhani, "Comparison of classification and clustering methods in spatial rainfall pattern recognition at Northern Iran," *Theoretical and Applied Climatology*, vol. 102, no. 3, pp. 319–329, 2010.
- [18] P. Shi, X. Ma, X. Chen, S. Qu, and Z. Zhang, "Analysis of variation trends in precipitation in an upstream catchment of Huai River," *Mathematical Problems in Engineering*, vol. 2013, Article ID 929383, 11 pages, 2013.
- [19] J. Golder, M. Joelson, M.-C. Neel, and L. Di Pietro, "A time fractional model to represent rainfall process," *Water Science and Engineering*, vol. 7, no. 1, pp. 32–40, 2014.
- [20] Y.-C. Chang, "Efficiently implementing the maximum likelihood estimator for Hurst exponent," *Mathematical Problems in Engineering*, vol. 2014, Article ID 490568, 10 pages, 2014.
- [21] Z.-G. Yu, Y. Leung, Y. D. Chen, Q. Zhang, V. Anh, and Y. Zhou, "Multifractal analyses of daily rainfall time series in Pearl River basin of China," *Physica A: Statistical Mechanics and Its Applications*, vol. 405, pp. 193–202, 2014.
- [22] J. W. Kantelhardt, "Fractal and multifractal time series," in *Encyclopedia of Complexity and Systems Science*, pp. 3754–3779, Springer, New York, NY, USA, 2009.
- [23] M. Carbone, M. Turco, G. Brunetti, and P. Piro, "A cumulative rainfall function for subhourly design storm in Mediterranean Urban areas," *Advances in Meteorology*, vol. 2015, Article ID 528564, 10 pages, 2015.
- [24] C.-M. Chou, "Complexity analysis of rainfall and runoff time series based on sample entropy in different temporal scales," *Stochastic Environmental Research and Risk Assessment*, vol. 28, no. 6, pp. 1401–1408, 2014.
- [25] A. P. García-Marín, J. Estévez, M. T. Medina-Cobo, and J. L. Ayuso-Muñoz, "Delimiting homogeneous regions using the multifractal properties of validated rainfall data series," *Journal of Hydrology*, vol. 529, pp. 106–119, 2015.
- [26] H. E. Hurst, "Long-term storage capacity of reservoirs," *Transactions of the American Society of Civil Engineers*, vol. 116, no. 1, pp. 770–799, 1951.
- [27] A. Witt and B. D. Malamud, "Quantification of long-range persistence in geophysical time series: conventional and benchmark-based improvement techniques," *Surveys in Geophysics*, vol. 34, no. 5, pp. 541–651, 2013.
- [28] J. Haslett, "On the sample variogram and the sample autocovariance for non-stationary time series," *Journal of the Royal Statistical Society—Series D: The Statistician*, vol. 46, no. 4, pp. 475–485, 1997.
- [29] R. Dmowska and B. Saltzman, Eds., *Advances in Geophysics, Long-Range Persistence in Geophysical Time Series*, Academic Press, 1999.
- [30] M. Lefebvre, *Applied Stochastic Processes*, Springer, New York, NY, USA, 2007.
- [31] B. B. Mandelbrot and J. R. Wallis, "Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence," *Water Resources Research*, vol. 5, no. 5, pp. 967–988, 1969.
- [32] B. Mandelbrot, "Statistical methodology for nonperiodic cycles: from the covariance to R/S analysis," *Annals of Economic and Social Measurement*, vol. 1, no. 3, pp. 259–290, 1972.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

