Analysis of persistence in fluctuation of the Cauca river through the Hurst coefficient

D A Prada¹, M P Sanabria¹, A F Torres¹, A Acevedo¹ and J Gómez²
¹Universidad Pontificia Bolivariana, Bucaramanga, Colombia
²Universidad Industrial de Santander, Bucaramanga, Colombia

E-mail: duwamg.prada@upb.edu.co

Abstract. Study the continuous changes in the fluctuations in the levels of watersheds, it is of great importance because it allows you to adjust predictions about behaviors that can lead to floods or droughts. The Cauca River is one of the most important rivers in Colombia due to its 1350 km long, its drainage area of 59.074 km² which represents 5% of the national territory. The Government entity Cormagdalena records daily levels of the Cauca River to the height of La Mojana in the rods. From these data, we developed a series of time on which normal test were applied to calculate the coefficient of Hurst and the fractal dimension to determine the persistence associated with this behavior.

1. Introduction
The basins are highly complex systems, their formation process is linked to factors as variables in time and space, such as climate, geology, vegetation etc. However, its evolution is subject to the combination of these forces with the climatic factors.

The measures the water level in a river contribute to understanding the dynamics of the download at a specific point in the basin. With this type of information, it is possible to predict and consider future flood events. The flow in the rivers also depends directly on the speed that it has, if the speed is small then the flow is regular, however, when you pass a stone, the river surrounding it, and if your speed increases becomes irregular, creating swirls and eddies in these eddies. The hydrodynamics describes these movements of initial form through the laws of Newton for fluids, in this way was found evidence of a non-linear equation, the equation of the Navier-Stokes Equations [1]. In the particular case of drainage networks the famous British hydrologist Harold Edwin Hurst (1880-1978), study the fluctuations in the level of the River Nile, in a large part of its extension, to be able to project the capacities of the reservations and take precautionary measures in times of drought [2].

2. Cauca's river
The Cauca river is one of the most important water resources in Colombia. It has a length of 1350 km, with a basin area of approximately 63300 km². It goes across the country from south to north through nine departments and a number of cities and towns without appropriate wastewater treatment plants [3].

The Cauca river is born in the Colombian Massif between the Western and Central Cordilleras of the Andes with a drainage area of 59.074 km² that represents the 5% of the national territory. In the course of its first 578 km (48% of the total) conforms the basin in its top-half the Colombian southwest [4]. The periods where there is rain in the months of March-April-May and September-October-November, and periods of low precipitation are December-January-February and June-July-August [5]. The region
has an average annual rainfall of 1597\text{mm} and 21.3°C average air temperature [4]. The Government entity Cormagdalena records every day of the Cauca river levels to the height of the sector of \textit{La Mojana}, place that includes the municipalities of Sucre, Guaranda and Majagual in the southern part of Department of Sucre, as seen in the Figure 1.

![Figure 1. La Mojana region [6].](image)

Study the continuous changes in the fluctuations in the levels of watersheds, it is of great importance because it allows you to adjust predictions about behaviors that can lead to floods or droughts. The first to recognize the characteristics of invariance of the drainage network to establish the laws of scale was Horton, and denoted as: Relationship of fork ($R_B$), a ratio of length ($R_L$) and ($R_A$) [7]. $R_L,R_A$Natural objects that have a complex spatial structure can be studied as fractals. Fractal geometry to describe irregular patterns and fragmented, which are repeated at different scales, usually in isotropic [8]. Fractal geometry as a branch of mathematics has been allowed to carry out research in relation to the similarity and dynamic systems, using notions of dimension and orbits in various disciplines such as medicine, engineering, psychology, music, biology, among others. In the particular case of drainage networks two fractal dimensions can be calculated, a fractal dimension (d), which describes the sinuosity of a current and the fractal dimension of the entire network (D), related to the characteristics of branching or compactness of the system [9,10]. With the analysis data between the length of the currents and the area of drainage basins, he found that the Mandelbrot fractal dimension of water courses (d) was equal to 1.2, while the fractal dimension of the entire system (D) was equal or close to 2 [11].

3. Mathematical method
The Fractal Geometry is a branch of mathematics, very attractive fractal objects that are possible to construct the notion of self-similarity, dynamic systems, chaos, orbits and dimension, both topological and Hausdorff spaces, which have become useful tools in numbers disciplinary fields such as
mathematics, science, art, etc [11-13]. Besides the self-similarity, the fractal objects have an idea out of the common; the fractal dimension. The dimension that is assigned by convention to certain geometrical and physical objects, is associated with a number infinite variable, for example, to cube is assigned to triple defined directly by the thickness, the width and height thereof, and then the size of this object is three. This type of dimension is known as the topological dimension [11,13].

The fractal dimension, as properly its name suggests, is a fractional dimension and is determined by a rational number. Long-range power-law correlations are traditionally measured by a scaling parameter or fractal dimension (D). If the time series is self-similar and self-affine, the parameter d is related to the Hurst exponent (H) through the expression \( D = 2 - H \) [14,15]. Thus, the Hurst exponent is a measure of the long-range correlation in time series data and allows to distinguish the persistence (correlation), anti-persistence (anti-correlation) or randomness of the data [16]. The original estimation of the Hurst exponent was first performed in hydrology by Harold Edwin Hurst in 1951 [17], by introducing an empirical relationship called the Rescaled-Range (R/S). Posteriorly, this relationship became the start point to establish the Classical R/S (CR/S) method developed by Mandelbrot and Wallis into the context of the fractal geometry [16,18,19]. Although the CR/S is one of the most popular methods to calculate the Hurst exponent, it has shown some serious limitations to study long-range correlation when the time series is not large enough [19,20].

4. Respect to the test of normality
Because of the behaviour variant that presents the time series of the fluctuating levels of depth of the river Cauca, were used Anderson-Daling Smirnov-Kolmogorov test and to ensure that the data do not present a Gaussian distribution, that is to say that the data are not standardized. The study of these phenomena in front of your distribution may present asymmetry and Homosedasticidad, because your behaviour platikurtico mesokurtico leptokurtico, or for these tests has been raised as a null hypothesis H0: there is a presence of normality (there is no presence of trend in the series) and how alternative hypothesis H1: There is a trend in the time series. In Figures 2 and 3 shows the analysis of normality for the time series using the test mentioned above, under the condition that if the value \( \rho < \) significance (0.05 usually) reject the null hypothesis and if the value \( \rho > \) significance (0.05 usually) accepts the null hypothesis.

![Figure 2. Analysis of normal](image-url)
With the data obtained in previous tests, we reject the null hypothesis and therefore it can be seen that there is a behaviour of similarity in the time series analysed.

**Figure 3.** Smirnov-Kolmogorov Test Anderson-Daling Test.

For the analysis of fractal dimension of the time series, specifically, to value $0 < H < 0.5$ corresponds to anti-eje correlated released (anti-persistence behaviour), a value $0.5 < H < 1$ corresponds to correlated axis released (persistence behaviour and the value $H = 0.5$ corresponds to random data (uncorrelated behaviour) [16]. The definition of H can be extended to values larger than 1 [18]. In this way, the case $H = 1.5$ corresponds to Brownian motion, the case $H=2$ corresponds to brown noise and the case $H > 2$ corresponds to black noise [21].

The following is CR/S method the time series described in [22] under consideration $X\{x_i\}$ is composed by N values. The full-time series is divided into windows of size M. The number of windows is defined by $s \equiv N/M$ and therefore there are s windows of data $Y_j$, with $j = 1, 2, ..., s$. Defining the vector $k = (j - 1)M + 1, (j - 1)M + 2, (j - 1)M + 3, ..., (j - 1)M + M$, the average over each window is calculated as:

$$\bar{Y}_j = \frac{1}{M} \sum_k x_k$$ (2)

The profile or sequence of partial summations $Z_j\{z_n\}$, with $n = 1, 2, ..., M$, is defined as the cumulative summation minus the average of the corresponding window

$$z_n = \sum_k^{nM} x_k - \bar{Y}_j$$ (3)

The range $R_j$ of the window is defined as the maximum minus the minimum data point of the profile

$$R_j = \max\{ Z_j \} - \min\{ Z_j \}$$ (4)

The standard deviation of each window $\sigma_j$ is given as

$$\sigma_j = \left[ \frac{1}{M} \sum_k (x_k - \bar{Y}_j)^2 \right]^{\frac{1}{2}}$$ (5)
The rescaled range is described by the quantity \((R/S)_M\), which is defined as

\[(R/S)_M \equiv \text{mean } (R_j / \sigma_j) \tag{6}\]

For the case in which a stochastic process associated to the data sequence under study is rescaled over a certain domain \(M \in \{M_{\text{min}}, M_{\text{max}}\}\), the \(R/S\) statistics follows the power law

\[(R/S)_M = aM^H \tag{7}\]

Herein, is a constant and \(H\) is the Hurst exponent which represents a fractal measure of long-range correlations in the analysed released.

5. Results and discussion

The Government entity Cormagdalena recorded 897 data from the Rio Cauca levels in the La Mojana sector on the date between September 29, 2011 and September 1, 2017. With this amount of data 6 subgroups were built, the first sub-group with 64 data falling between 29 September 2011 until 16 February 2014, the second subgroup with 112 data falling between 29 September and 23 May 2012, the third subgroup with 128 data with dates of 29 September until 8 August 2012, the fourth subgroup with 224 data with dates of 29 September until 3 November 2013, the Fifth 488 subgroup data and dates between 29 September and 19 February 2015, and the last Subgroup with the totality of the data. This division is accumulated to observe the self-similarity in the behaviour of the data as you increase the amount of the same. The number of subgroups that have been generated for this series is related to the dividers on the total number of data, i.e. with the number 897. This is possible because no matter the scale at which the analysis is carried out, the behaviour must be the same based on the fractality of the series. In each of this series is calculated as the standard deviation, the average and range- scaled by the difference of the maximum and minimum value of the cumulative sum of deviations. Table 1 describes the data obtained in each subgroup.

| Subgroup | Call forwarding stand | Number of released | Rank | Rescaled | Max  | Min  | Average |
|----------|-----------------------|--------------------|------|----------|------|------|--------|
| 1        | 0.59                  | 64                 | 14.96| 1.77     | -13.20 | 5.26 |
| 2        | 1.03                  | 112                | 42.14| 42.14    | 0     | 4.60 |
| 3        | 0.98                  | 128                | 46.05| 42.47    | -3.57 | 4.60 |
| 4        | 0.93                  | 224                | 53.02| 34.54    | -18.49 | 4.72 |
| 5        | 1.26                  | 488                | 190.79| 170.42   | -20.36 | 3.96 |
| 6        | 1.30                  | 897                | 289.76| 277.23   | -12.53 | 3.64 |

In the Table 2 shows the data needed for the calculation of the coefficient of hurts for this time-series, as well as compare the natural logarithm of the data of each of the subgroups generated versus the natural logarithm of the ratio between the rescaled range and standard deviation. With this data we make a new graph in which we compare the natural logarithm of the data of each of the subgroups generated versus the natural logarithm of the ratio between the rescaled range and standard deviation, these data are related in Figure 4.
Table 2. Related data from the four groups depending on the range and Standard deviation. Calculation of the natural logarithm of the data number and the natural logarithm of the quotient between the rescaled range and the standard deviation.

| Subgroup | Number data | Rescaled range | Standard Deviation | Ln(Num) | Ln(R/S) |
|----------|-------------|----------------|--------------------|---------|---------|
| 1        | 64          | 14.96          | 0.58793287         | 4.158883083 | 3.236835771 |
| 2        | 112         | 42.14          | 1.03               | 4.718498871 | 3.712918796 |
| 3        | 128         | 46.05          | 0.98113714         | 4.852030264 | 3.848672389 |
| 4        | 224         | 53.02          | 0.93105402         | 5.411646052 | 4.042147592 |
| 5        | 488         | 190.79         | 1.25977654         | 6.190315406 | 5.020233087 |
| 6        | 897         | 289.76         | 1.3015687          | 6.799055862 | 5.405807777 |

Figure 2. Logarithmic linear regression of the data supplied in Table 2.

6. Conclusions
Of the 897 data recorded by the governmental entity Cormagdalena between 129 September 2009 June 1993 and 1 September 2017, in the area of La Majana-Las Rods with respect to the levels of fluctuation of the Magdalena River, the coefficient of Hurts associated was of $H = 0.832 > 0.5$, which induces that the fractal dimension is $D = 1.168$, according to [14,15], which indicates that the time series of low fractal dimension and for that reason these series have persistence associated with a volatility of 58.4%. In addition, it was found that this basin meets the above mentioned by Mandelbrot Set with respect to the fractal dimension of the basins with these characteristics must be close to the value of 1.2. These series were not gaussian distribution due to the fact that the Jarque-Bera test was greater than 6 for each [23].

References
[1] Braun E 2011 *Caos, fractales y cosas raras* (Mexico: Fondo de Cultura Económica) p 92
[2] Rodriguez E 2012 *Lat. Amer. Not. Adm.* 3 42-50
[3] Ocampo W, Osorio C, Schuhmacher M and Domingo J 2012 *Env. Int.* 52 17-28
[4] Puerta O, Carvajal Y and Quintero M 2011 *Dyna* 169 112-120
[5] Mesa O, Poveda G and Carvajal L 1997 *Introducción al clima de Colombia* (Colombia: Universidad Nacional de Colombia)
[6] Escuela Superior de Administración Pública ESAP. Consulted on http://cdim.esap.edu.co/Combosdependientes.asp?PnDepartamentos=70&Pnmuni=70771
[7] Ramirez R 2017 *Min. and Geo.* 33 279-292
[8] Ramirez R, Rodriguez A and Ordaz A 2017 *Min. and Geo.* 33 pp 36-49
[9] Barbera P and Rosso R 1989 *Wat. Res. Res.* 25 735-741
[10] Tarboton D 1996 *J. Hyd.* **187** 105-117  
[11] Mandelbrot B 2006 *Los objetos fractales: forma, azar y dimensión* (España: Metatemas)  
[12] Falconer K 2006 *Fractal geometry. Mathematical foundations and applications, second edition* (India: Wiley)  
[13] Barnsley M 2012 *Fractals Everywhere* (United States: Dover Publications)  
[14] Erdf J 1988 *Fractals* (New York: Plenum Press)  
[15] Chen C, Lee Y and Chang Y 2008 *Phys. A* **387** 4643-4648  
[16] Mandelbrot B and Wallis J 1968 *Wat. Res. Res.* **4** 909-918  
[17] Hurst H 1951 *Trans. Am. Soc. Civ. Eng.* **116** 770-808  
[18] Mandelbrot B and Wallis J 1969 *Wat. Res. Res.* **5** 321-340  
[19] Mandelbrot B and Wallis J 1969 *Wat. Res. Res.* **5** 967-988  
[20] Sanchez M, Trinidad J and Garcia J 2008 *Phys. A* **387** pp 5543 -5551  
[21] Jiménez A 2011 *Phys A* **390** 2146-2154  
[22] Martin L, Aranda N and Quimbaya C 2015 *Phys. A* **421** 124 -133  
[23] Gujarati D and Porter D 2010 *Econometría* (Mexico: Mc Graw Hill)