Fractal analysis of the time series of particulate material

D A Prada¹, D Parra¹, J D Tarazona¹, M F Silva¹, P Vera¹, S Montoya¹, A Acevedo¹, and J Gomez²

¹ Universidad Pontificia Bolivariana, Bucaramanga, Colombia
² Servicio Nacional de Aprendizaje, Bogotá, Colombia

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

Abstract. The air quality index shows the pollution of the air we breathe, a very low index indicates that we can have respiratory, eye or heart conditions. An index below the minimum levels declared, requires generating strategies such as days without the presence of private vehicles, among others, to reduce pollution. The city of Medellín, Colombia, has shown very low levels in the air quality index due to several factors, mainly the emission produced by vehicles and companies. Due to this, activities have been initiated to reduce the emission of 2.5 micron and 10 microns particulate material, classified as highly harmful to health. The city installed 12 air quality monitoring stations and designed a decontamination proposal for the year 2030. Conducting a study on the persistence of particulate matter in the air according to the data recorded by the stations, will strengthen the proposed strategy or modify it to achieve the goal. In this work, the data collected by a station from July to December 2018 in which a high persistence in the presence of particulate material was analyzed.

1. Introduction

The quality of life of people can be evidenced from different aspects such as emotional, social, labor, among others, however, an important factor at the physical level is environmental. Due to the contamination, we observe that there are places on the planet where people should use capholes, they cannot perform outdoor exercises and the probability of respiratory-associated diseases is high, which evidently shows a low quality of life in that appearance. Our country is no exception, as there are cities that show alarming levels of pollution or what is similar, the air quality index is very low. Several factors are the causes of such a low index, such as vehicles with more than 20 years of circulation, companies with high emissions, a poor culture regarding the conservation of the environment, among others. Under the notorious presence of air pollution, the city of Medellín, Colombia, has made proposals for the purpose of decontamination by 2030, however it is important to carry out ongoing studies to the respective to show if the proposal framed in days without car, peak and plate between others, it is generating the intended impact.

The generates a compromise on the part of the academic community in order to monitor the behavior of said phenomenon based on different methods and techniques that allow or consolidate the strategy proposed by the city. Through the registration of levels of 2.5 micron and 10 micron particulate material, generated by a station located in the city center and with the support of the technique for calculating the fractal dimension associated with said time series, persistence and volatility are analyzed in order to show the community in general the relevance of the strategy. To calculate the persistence and volatility of the time series, 4416 data recorded from July to December 2018 were analyzed using the rescaled
range method in four time sub-series in order to find the Hurst coefficient by means of a logarithmic regression to calculate the dimension fractal and with it determine persistence and volatility.

2. Theoretical framework

Pure air is composed of a balanced mixture of 78% nitrogen, 21% oxygen and less than 1% amounts of carbon dioxide, argon and other gases. Therefore, air pollution is defined as the change in the balance between these substances [1]. At the moment the atmospheric contamination in conjunction with other factors is producing innumerable risks in the health, biodiversity and alimentary security of the alive beings, given the dependency of diverse substances necessary for the development of processes like the photosynthesis and the breathing, offered by the function of the air in ecosystems [2]. That is, the activities carried out by humans over the years, have generated an alteration in the composition of the air, affecting its quality, which has been evidenced in different studies that indicate that 92% of the world population lives in places lacking air quality standards, generating three million people suffering from premature deaths annually, the death of 1.7 million children under five years of age and 4.9 million adults between 50 and 75 years old [3]. It is also alarming that exposure to air pollutants, especially fine particles, which generates one of the main risks of non-communicable diseases, such as stroke, asthma, among others. Studies claim that air pollution is responsible for half of the deaths from acute respiratory infections in children under five years of age, which makes it one of the main factors of infant mortality worldwide [4].

Mexico, is one of the countries with the highest rates of severe pollution (environmental contingency), which has carried out several studies, among which is a research carried out by students of the “Universidad Autónoma de Mexico (UNAM), México”, in which the behavior was evaluated dynamic, in terms of persistence and self-correlation of the particulate matter (PM) of 10 microns pollutant through the implementation of the Hurst coefficient, which allowed the establishment of pre-environmental event previews, thus providing useful information to monitor the behavior of the levels of such particles in the metropolitan area of the Valley of Mexico, evidencing the importance and impact of such project [5].

Colombia, is considered as one of the most biodiverse countries in the world, has been degraded, like most Latin American countries, due to different factors such as the population growth of the city, agricultural, mining and industrial exploitation, generating the incorporation of foreign substances affecting the health and welfare of society and a large number of ecosystems that are part of the national heritage. Some of the cities in the country have higher pollution levels than others, depending on the geomorphological, orographical and climatological conditions of the regions, given that these are directly related to the processes of dilution, concentration and transport of substances into the air [6]. In particular, the cities of Bogotá and Medellín have been ranked for years as the cities with the worst air quality status in Colombia.

The city of Medellín, Colombia, with an average of 2.5 million inhabitants, known as "the city of eternal spring" due to its temperate climate of between 18 °C and 22 °C, and its famous fair of flowers, presents in its environment PM of 2.5 microns generated by 69% by trucks, dump trucks and industry, which alter the daily conditions of its metropolitan area throughout the day [7]; together with the pollutant PM10 have become the alarming substances in the last 10 years, since they exceed by 400% the national levels and those established by the who, by the growing soot particles intensified by the unfavorable topographic conditions of the city, given the lower layers of the atmosphere, which generate the phenomenon of thermal inversion, which causes a plug that contains pollutants for several days preventing their regular circulation, increasing the risks in people against respiratory diseases [8]. Studies have shown that the air pollution of the city presents a particular behavior for two periods, the first corresponds to the months of March and April, in which there is a passage from the dry season to the rainy season, there is evidence of winds low, which is why the highest indicators for the year are shown, likewise between the months of October and November the same phenomenon occurs, but with less intensity [9]. The city has implemented air quality monitoring systems through 12 monitoring stations and designed a decontamination proposal for the city by 2030, implementing various follow-up
plans and modern ones with automatic technology to carry out a constant evaluation and control of air pollutants [10].

The Hurst coefficient or exponent is a measure of the independence of time series that was initially studied by British scientist Harold Edwin Hurst (1880-1978), as an element to distinguish fractal series. Hurst discovered that many natural phenomena exhibit a behavior that it can be characterized by a biased random process, in which there is “long-term memory” among the observations, that is, that the events of one period influence all of the following [11]. Hurst’s ideas were retaken by Benoit B. Mandelbrot who generalized his work and what called scaled range analysis (R/S), defined as a statistical method used to evaluate the occurrence of unusual events, giving rise to an ideal tool for physical and financial processes, although not limits only to these types of events, so you can be used in any time series [12]. This type of property based on self-similarity is an alternative tool to methodologies used in different types of studies. Because natural phenomena have a type of persistence over time, this technique is easily related to the data collected. On the other hand, proposing studies with techniques other than the conventional ones generates a critical stance that generally allows researchers to consolidate the findings given from classical theorists. In this study regarding air quality, observing the persistence in the data allows generating new ideas regarding the interventions necessary to reduce the emission of particulate material.

3. Mathematical method

The concept of fractal dimension is introduced given the fact that the objects of nature in their great majority do not have an entire dimension, that is, they are in a fractional dimension (of course some objects can have an entire dimension); likewise, this dimension must be greater than its topological dimension [11]. Regarding the Brownian movement, result established by Albert Einstein [13], Hurst managed to develop a formula applicable to time series not necessarily of Brownian movements [14]. Equation (1) shows under the power law, the value of the Hurst coefficient as exponent.

\[(R/S)_{m} = aM^{H},\]

where \((R/S)\) is the notation used for the rescaled range statistic, at a constant of proportionality, \(S\) is the standard deviation of the data, \(M\) is the number of data per interval and \(H\) is the Hurst coefficient. The Hurst exponent can be determined from the fractal dimension or vice versa, using the equation \(D = 2 - H\), where \(D\) is the fractal dimension [11].

The “Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM)” recorded 4416 data, taken between July 1 and December 31, 2018 with intervals of one hour between them, at station 12 located in the downtown area of the city, corresponding to the number of particles per millions of PM2.5 and PM10. With the recorded data, a time series is generated for each of the two sizes of the particulate material. Using the rescaling range method described in [11] it is possible to find the value of the Hurst coefficient associated with this time series. For values of \(H = 0.5\) considered white noise, the process has to be completely random [12], which indicates that there is no relationship between each consecutive data. For values of \(H\) greater than 0.5, we fear that the time series has persistence. A persistent series is characterized by long-term memory effects. Theoretically what happens today will impact the future forever [12]. For the values of \(H\) between zero and less than 0.5, the time series is considered to have antipersistence or average reversion, that is, if the series has been above a certain value that serves as a long-term average in the previous period, it is more likely to be down in the following period and vice versa [15].

4. Results

Due that the series shows fractality, the series of 4416 data is partitioned into 4 subgroups [16]. In Table 1 and Table 2 shows the 4 subgroups of the 4416 data series. In each of them, the rescaled range, the standard deviation, the natural logarithm of the amount of data per series, the natural logarithm of the
quotient between the rescaled range and the standard deviation and in the end the average in each subgroup, and in each group of PM10 and PM2.5 are shown.

**Table 1.** Data of the standard deviation, released number and rescaled range of PM10.

| Subgroup | Number of released | Rank rescaled | Standard deviation | Ln (Num) | Ln(R/S) | Average |
|----------|--------------------|---------------|--------------------|----------|---------|---------|
| 1        | 1104               | 2219.88       | 23.43              | 7.00     | 4.68    | 40.68   |
| 2        | 2208               | 3589.26       | 20.48              | 7.69     | 5.16    | 42.63   |
| 3        | 3312               | 5038.09       | 20.32              | 8.10     | 5.51    | 43.69   |
| 4        | 4416               | 7841.44       | 20.23              | 8.39     | 5.95    | 45.35   |

**Table 2.** Data of the standard deviation, released number and rescaled range of PM2.5.

| Subgroup | Number of released | Rank rescaled | Standard deviation | Ln (Num) | Ln(R/S) | Average |
|----------|--------------------|---------------|--------------------|----------|---------|---------|
| 1        | 1104               | 1461.05       | 11.00              | 7.00     | 4.88    | 19.32   |
| 2        | 2208               | 2106.15       | 12.19              | 7.69     | 5.15    | 20.74   |
| 3        | 3312               | 4250.67       | 12.67              | 8.10     | 5.81    | 21.77   |
| 4        | 4416               | 5926.52       | 12.74              | 8.39     | 6.14    | 22.84   |

With this data, a graph of the natural logarithm of the data of each of the generated subgroups is made versus the natural logarithm of the relationship between the rescaled range and the standard deviation, these data are related in Figure 1 and Figure 2.

**Figure 1.** Logarithmic linear regression of the data supplied in Table 1.

Figure 1 shows a less slope of the regression line than the slope of the line shown in Figure 2, however the correlation coefficient in Figure 2 is smaller than in Figure 1.

**Figure 2.** Logarithmic linear regression of the data supplied in Table 2.
From the logarithmic linear regression of the data of the natural logarithm of the number of data versus the logarithm of the rescaled range between the standard deviation generates an equation whose slope coincides with the Hurst coefficient $H$, which coincides with the value of $H$ that can be calculated using Equation (1). In the case of Figure 1, $H = 0.8808$ and in the case of Figure 2, $H = 0.9193$.

5. Conclusion
From the data collected by one of the stations that measures the amount of particulate material in the city of Medellin, which are socialized to the community through IDEAM, between July 1 and December 31, 2018 it was observed that respect to the time series of 2.5 and 10 micron particulate material has persistence, that is, a long-term memory, because the levels were below the average, this indicates that a period of non-growth will be followed by another analogue, previous because the values of $H$ for each of the time series was greater than 0.5. The fractal dimension of the 2.5-micron particulate material series is $D = 1.0807$ for a mean relative volatility of 54.03%, and for the 10-micron particulate material series the fractal dimension was $D = 1.1192$ for a relative volatility the average of 55.96%. With the evidence in this work, it is possible to comment that the proposal for decontamination purposes in 2030 in the city of Medellin, Colombia, is in an acceptable stage, once the decrease in the amount of particulate material registered by IDEAM is observed and that the long-term tendency is to decrease as analyzed by calculating the fractal dimension.

Acknowledgment
The authors of this article are grateful for the hotbed of mathematical modeling research at the Universidad Pontificia Bolivariana, Bucaramanga, Colombia, and to the scientific and academic committee of the VI International Conference Days of Applied Mathematics (ICDAM).

References
[1] Daniels F, Martínez E, Quinchía R, Morales C, Romero A, Marin A, Arbeláez M 2007 Contaminación atmosférica y efectos sobre la salud de Medellin (Medellín: Universidad de Antioquia)
[2] Querol X 2018 La Calidad del Aire en las Ciudades. Un Reto Mundial (Madrid: Fundación Gas Natural Fenosa)
[3] World Health Organization 2016 Healthier Environment: Healthier People (Narobi: Second Session of the United Nations Environment Assembly of the United Nations Environment Programme)
[4] World Health Organization 2014 Salud y Medio Ambiente: Impacto Sanitario de la Contaminación del Aire (Switzerland: World Health Organization)
[5] Vucovich A, Várquez C, Baristain F 2015 Tendencia de la dinámica no-lineal en una precontingencia ambiental causada por partículas en suspensión Inf. Tec. 26(4) 21
[6] Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) 2017 Informe del Estado de la Calidad del Aire en Colombia (Colombia: Instituto de Hidrología, Meteorología y Estudios Ambientales)
[7] Medellín Cómo Vamos (MCV) 2018 Problemáticas de la Contaminación del Aire en Medellín (Colombia: Medellín Cómo Vamos)
[8] Bedoya J, Martínez E 2009 Dyna 76(158) 7
[9] Alcaldía de Medellín 2018 Pacto por la Calidad del Aire (Colombia: Alcaldía de Medellín)
[10] Gómez C 2017 Contaminación del Aire en Medellín por PM10 y PM2.5 y sus Efectos en la Salud (Medellín: Universidad Militar Nueva Granada)
[11] Delgado O, Delgado J 2011 Estimación del exponente de Hurst y la dimensión fractal de una superficie topográfica a través de la extracción de perfiles $UD$ y La Geomática 5 84
[12] Gutiérrez H 2008 Estudio de geometría fractal en roca fracturada y series de tiempo (Santiago de Chile: Universidad de Chile)
[13] Einstein A 1956 Investigations on the Theory of Brownian Movement (New York: Dover Publications)
[14] Hurst H 1951 The Long-term storage capacity of reservoirs Transactions of the American Society of Civil Engineers 116 770
[15] Luengas D, Ardila E & Moreno J 2010 Metodología e interpretación del coeficiente de Hurst ODEON (5) 265
[16] Braun E 2011 Caos, Fractales y Cosas Raras (México: Fondo de Cultura Económica) p 92