https://doi.org/10.25100/iyc.v18i2.2155

GEOMÁTICA

Análisis de componentes principales funcionales para el estudio de material particulado (pm10) en la ciudad de Bogotá

GEOMATICS

Functional principal component analyses for study of particulate matter (pm10) in the city of Bogotá

Miguel A. Ramírez*§, Jánel Arias*, Luis F. Santa*

**Ingeniería Catastral y Geodesia, Universidad Distrital Francisco José de Caldas. Bogotá D.C., Colombia.*
§miaramirezg@correo.udistrital.edu.co, jaariasc@correo.udistrital.edu.co, lfsantag@unal.edu.co

(Recibido: Agosto 21 de 2015 – Aceptado: Diciembre 03 de 2015)

Resumen

La contaminación del aire de la ciudad de Bogotá es un problema que se ha venido incrementando paralelo al crecimiento de la ciudad. Las mediciones de la contaminación en Bogotá requieren metodologías que permitan su análisis, para que a su vez se emitan diagnósticos. A continuación se realiza el análisis de Componentes Principales Funcionales a reportes de material particulado PM10, resultando esta una metodología apropiada para tratar grandes cantidades de datos, pues resume estadísticos funcionales tales como la media y la covarianza, obteniendo como resultados diagnósticos del comportamiento de las mediciones de forma sintetizada y concisa.

Palabras Clave: *Componentes principales funcionales, datos funcionales, PM10.*

Abstract

Bogotá city air pollution is a problem that has been increasing parallel to the growth of the city. The measurements of pollution in Bogotá require methodologies that allow its analysis; therefore, at the same time, that diagnoses are issued. Below is the functional principal component analysis to reporting material particulate matter PM10. As a result, an appropriate methodology for treating large amounts of data is developed because it summarizes functional statistics such as the average and the covariance obtaining a diagnostic of the behavior of the medic results.

Keywords: *Functional data analysis, functional principal components, PM10.*

1. Introduction

Knowing the importance of atmospheric concentrations, it is significant to start mentioning that these concentrations of particulate matter with diameter less than 10 μm (PM10) in Bogotá city tend to exceed the limits set by the standards of air quality as Gaitán et al. (2007) and Bautista (2010) have noted. Likewise, this situation is reflected in the incidence of respiratory and heart problems in the population; hence, it has called for some real alerts in the environmental control bodies as Rodríguez-Moreno et al. (2013), García-Ubaque et al. (2011), Otman et al., (2010), Blanco-Becerra et al., (2014) and Franco et al., (2009) have pointed this out. Also, Preciado (2008) and García-Ubaque et al. (2011) suggested that the continued growth in economic terms in Bogotá has caused that the control measures which already exist for the proper management of PM10 are upstaged by the increase in the presence of particulate material. For this reason, it is necessary to develop new methodologies that allow studying its spatio-temporal behavior in order to facilitate environmental authorities and the pollution diagnoses for PM10 in Bogotá; consequently, some management strategies which are conducive to the well-being of the population and can be suggested.

Seen in this light, the majority of studies have been conducted to determine the influence of the presence of particulate matter on health. The results obtained it is concluded in all cases that the increase in the presence of particulate matter PM10 is directly proportional to the increase in respiratory diseases as Cattani et al. (2010), Zarkey (2008) and Aldunate et al. (2006) have pointed.

In Colombia, several air quality studies have been done. From one of the most prominent studies in this country, a result of the most appropriate statistic method to study the pollution data has been obtained. This research was mentioned by Reina & Olaya (2012) and as a main conclusion the authors indicated that the Spline method is the most appropriate to model the pollution data. In another research, García et al. (2006) found that other statistical tools such as analysis of variance and multiple range test were used and it was determined that the emission of particulates comes from anthropogenic sources.

Taking into account the air quality research that has taken place in Bogotá, Gaitán et al. (2007) conducted a descriptive analysis of the PM10 concentration at different times of the year; moreover, a relationship with other environmental factors was established such as the speed of the wind and carbon dioxide. At the same time, Bautista (2010) and Vargas & Rojas (2009) made a description of the chemical and physical characteristics of the PM10 and PM2.5 concluding that they are a health risk because of the high presence of metals in them. This conclusion is also supported by the high correlation that exists between these two pollutants as Galvis & Rojas (2005) and Blanco-Becerra et al. (2015) pointed this out.

Similarly, it is important to note that Blanco-Becerra et. al (2014) used classical time series along with simple delay and distributed lag models. Then, it was found that mortality from respiratory causes is strongly influenced by the increase in the concentration of the pollutant; likewise, this is also supported by Rodríguez-Moreno et al. (2013) who indicated that exposure to this contaminant significantly increases the risk to any respiratory illness.

On the other hand, Bohórquez (2010) conducted and analysis for predictive purposes the presence of PM10 in Bogotá by using a non-separable covariance function, resulting in a predictive model of distribution space in the pre-

sence of PM10. However, Herrera (2013) noted that using functional data (FDA) next to the interpolator space cokriging analysis, it was proposed a functional model. This model is based on the wind speed and the behavior of the pollutant would be able to predict the maximum average hours of the PM10.

Seen in this light, this research proposed the use of the statistical technique of functional data analysis which has been used for the analysis of temperature, measurement of depth and risk of occurrence of an accident as Giraldo (2007), Levitin et al. (2007) and Segovia et al. (2005) have found this out. This technique provides an alternative in which it is possible to distinguish the main characteristics of the data such as functions of mean and covariance without any analysis separated in time intervals. Additionally, one of the main functional components was the suggested technique. According to Ramsay & Silverman (2005) this technique summarizes the variability of the data in the few principal functions understanding this as the weighting function of the variability of the data which is more easily manipulated. In order to get the main conclusions, it is important to see that Chávez et. al (2015) mentioned that a conceptual extension

of the classic principal components where instead of treating vectors, functions is used, being somewhat improved characterization of the nature of the data.

This, in terms of pollution would allow the improvement of curricula and existing environmental management policies that seek to improve the quality of Bogotá's inhabitants.

2. Methodology

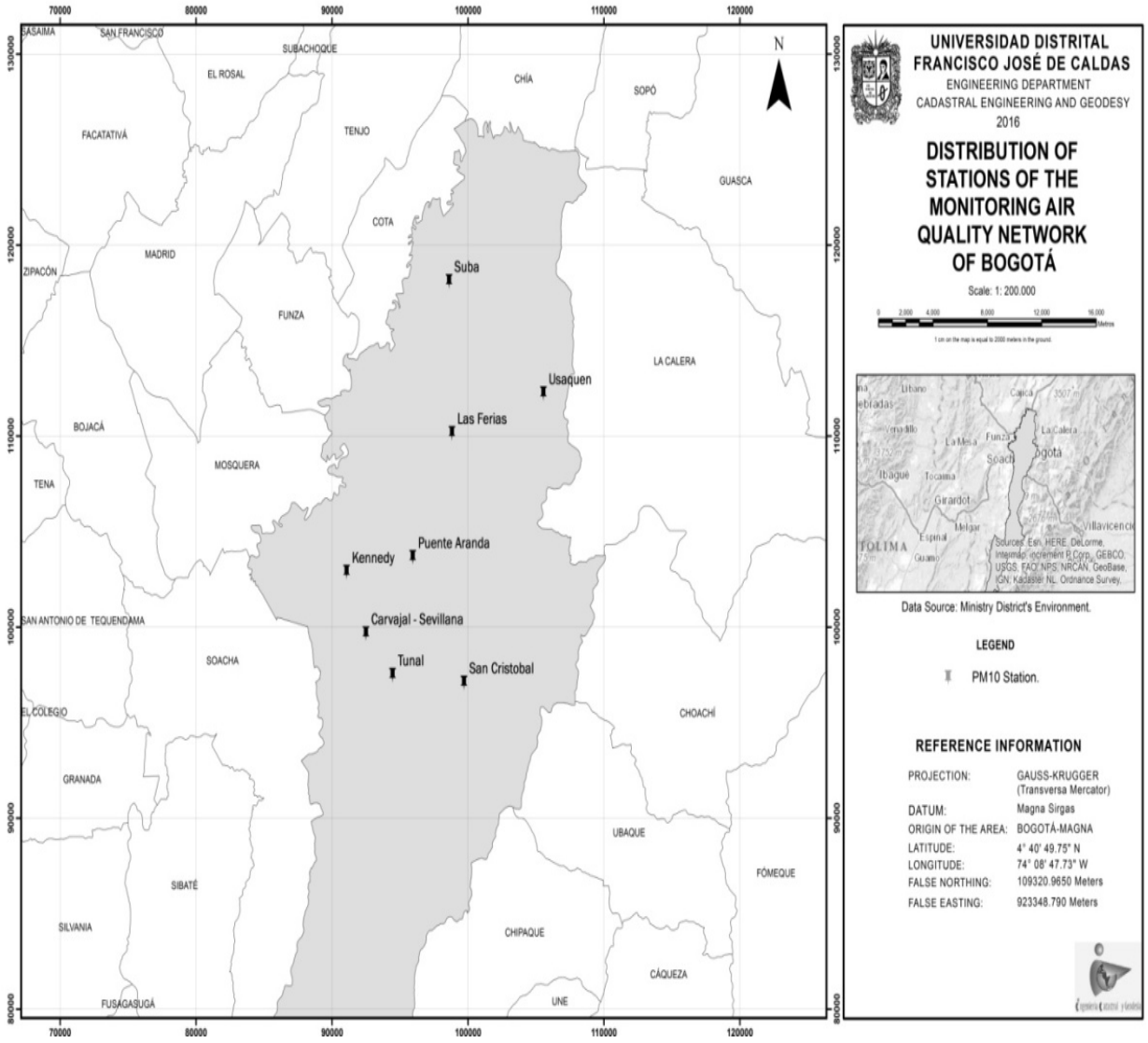
2.1 Study Area

The city of Bogota has a network of stations which are distributed throughout the city, it is important to note that not all stations that are willing to measure concentrations of PM10 are not available; therefore, only nine out of fifteen stations were taking into account for this research; additionally, eight out of those nine stations registered the required data for this research (see table 1 next to figure 1).

Table 1. *Stations used in environmental monitoring network with their respective coordinates datum MAGNA-SIRGAS*

Station	North (m)	East (m)	Height (m)
Carvajal-Sevillana	99637.709	92495.434	2559.138
San Cristóbal	97060.376	99705.904	2548.624
Usaquén	112204.733	105556.503	2775.947
Kennedy	102856.982	91074.557	2554.873
Puente Aranda	103628.018	95934.516	2557.045
Las Ferias	110125.45	99820.121	2550.817
Suba	118080.677	98613.068	2554.268
Tunal	97468.198	94451.448	2564.361

Figure 1. *Location of stations used for the Environmental Monitoring Network.*



2.2 Materials

Knowing the importance of measuring the concentrations of PM10 at Bogotá city, it is important to mention that in order to implement this study; firstly, it was necessary to achieve a construction of a database that took as input particulate material PM10 which was supplied by the Network for Monitoring air quality in Bogota (Red de Monitoreo de Calidad de Aire).

These data were obtained through the website of the Ministry district's environment. Given the availability of information, the database contained reports of PM10 measurements corresponding to the days Monday and Sunday with one hour temporal resolution during May and August from 2011 to 2013 for each selected station. For each concentration of PM10 the data were included: date, time, PM10, season, day of the week, north coordinate, east coordinate

and height above sea, specifying that the coordinates are in the spatial reference system MAGNA-SIRGAS.

Clearly, for the proper analysis of the data and to facilitate processes related to the treatment of the same were used the following software: R Project 3.1.2 for the statistical analysis of the database, Postgis 1.18.1 as the database management system and QGis 2.4 for spatial analyses that were necessary. For the integration of the base data with R, the use of the package (version 0.4) RPostgreSOL was required. Given the functional data methodology used in the study, it was necessary to package fda (version 2.4.4) thus, this package provides functions and methods that allowed the analysis of functional data as well as it allowed getting the principal functional components. Moreover, it was also necessary to use the package gstat (version 1.0.25) to perform the spatial interpolation of the principal functional components obtained for each station.

2.3 Methods

Besides addressing the materials used for this research, it is important to see that the absence of 134 pollution concentrations was detected before the final structuring of the database. In order to not be more robust, it was necessary to find an appropriate method to complete missing data; therefore, this procedure was performed using the software remove measurements so that the analysis would R using package Amelia (version 1.7.3) using the time series module as proposed by Honake et to the. (2014).

By the nature of the data, it was necessary to create and implement functions in R that facilitates handling the same data and that were also consistent with the parameters of the functions of the fda library. With the purpose of facilitating the processes carried out in relation to functional data analysis and taking into account the issues raised by Ramsay & Silverman (2005) about the definition of functional data on measurements repeated in a continuous system of reference being a curve through space. Mathematically it is defined as shown:

$$Y_i(t) = \sum_{j=1}^{k} c_j \, \Phi_i(t) \quad (1)$$

Where, Yi (t) denotes the functional data, cj are the weighting coefficients and ϕj (t) are the basis functions of smoothing; therefore, in order to represent hourly concentrations of PM10 for each date was selected as a base B-Spline by its computational flexibility and adequate description of local effects. As Carreño (2013) pointed this out, these characteristics are conducive for the pollution data as Montero & Fernández-Avilés (2015) also noted. Therefore, for its use, it was necessary to define the number of polynomials and the grade used with each sample. For this reason, it was created and implemented a function on R which made the validation procedure cross to obtain these parameters, Aristizabal (2011) mentioned that this method is the most suitable to define the parameters of this base; hence, the definition and characterization of each functional data was accurate.

Later, exploratory analysis was subsequently performed functional data, this corresponded to the first diagnosis made to data reported in each month and year and particulate material which consisted of obtaining together with the analy sis of the statistical average $\bar{x}(t)$ and bivariate covariance-centric v(s,t) functions corresponding to each period, that are defined as:

$$\acute{x}(t) = \frac{1}{n}\sum_{i=1}^{n} x_i(t) \quad (2)$$

$$v(s,t) = \frac{1}{n-1}\sum_{i=1}^{n}[x_i(s) - \bar{x}_i(s)][x_i(t) - \bar{x}_i(t)] \quad (3)$$

Where xi (t) is the ith functional data while s and t are times in which measurement was made. This

analysis was conducted with the aim to have firsthand a robust examination of the temporal behavior of the contaminant.

As the exploratory analysis complement, two principal functional components were found from the functional data of each month studied. Taking into account that this method is intended to calculate a specific number of ξ (t) from the eigenvalues μ weighting functions (these determine the portion of the explained variability by each weighting function in ξ (t)) bivariate covariance-centric function (these are defined in equation 4 which corresponds to the functional equation of itself), with the aim of reducing the dimensionality of the whole functional data. The method must accomplish the following mathematical characteristics, as Ramsay & Silverman (2002) pointed out:

$$\int \left(v(s,t)\xi(t)_j\right)dt = \mu_j \xi(s)_j \quad (4)$$

$$Z_i = \int \xi(t) Y_i(t) dt \quad (5)$$

$$\int \xi(t)^2 dt = 1 \quad (6)$$

$$\int \xi(t)_j \xi(t)_{j-1} dt = 0 \quad (7)$$

In the functional context, it is important that these equations are valid for all weighting functions in, due that in the case of equation 5 ensures the generation of a new space functional Z, while the In the functional context, it is important that thee quation 6 designates the condition of normality which allows comparisons between the generated functions too; likewise, the equation 7 designates the aspect of orthogonality which is appropriate since it ensures that the weighting functions are independent. However, it was necessary to use the VARIMAX rotation algorithm with them. According to Ramsay & Silverman (2005) components are affected by the systematic patterns of orthogonal functions; therefore; two interpretations were only appropriate in the sense of identifying particular characteristics of the PM10 behavior that explain each eigenfuncion and it retains about 80% of the variability of the data.

In addition, the functional principal components for each station were found because of the purpose which was based on that visualization of the spatial behavior of variability explained by each eigenfuncion of the pollution in the hours in which the functional average was maximum and minimum; therefore, the spatial interpolation model deterministic IDW was employed because the calculation with this method is much simpler. As Villatoro et al. (2008) noted its use is preferable at sampled points that are separated in large distances being the case of the spatial arrangement of stations. To use this interpolation, it was necessary to define the power of weighting, making this process by cross-validation as says Mehrjardi et to the. (2008). In Figure 2, you can see synthetically methods that were employed.

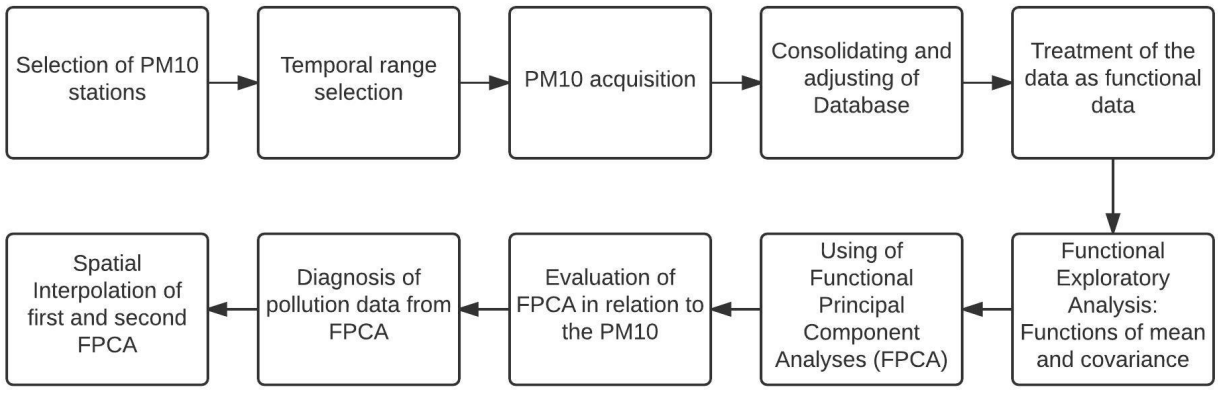


Figure 2. *Diagram of the methodology used for the analysis.*

3. Results and discussion

3.1 Results

As the first result was obtained with all full records database, this was finally consolidated with a total of 7752 records.

The following result corresponded to the functions of media for the months and years chosen for the study. In figures 3 and 4, the functions of media in which the axis of ordinates in the values of the PM10 concentrations average are observed; likewise, the hours of the day where it is easy to establish that time for greater concentration of the contaminant is 10 am while the minor takes place at 4 am. It is important to see that all dates of study are represented on the horizontal axis.

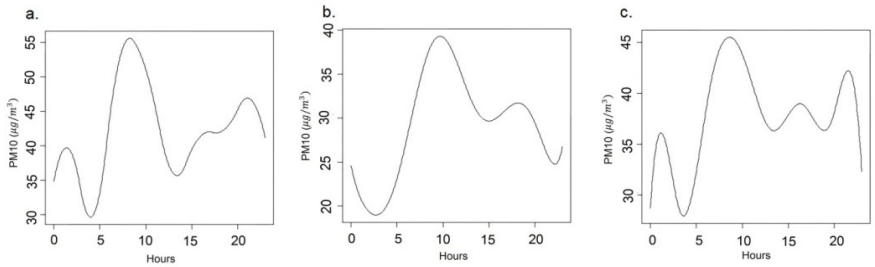


Figure 3. *Mean functions for August. a). 2011 b). 2012 c). 2013.*

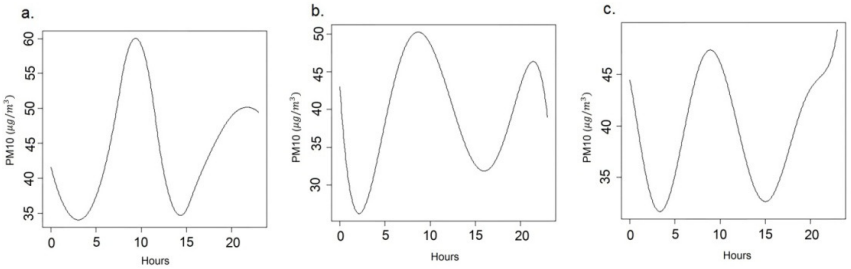


Figure 4. *Mean functions for May. a). 2011 b). 2012 c). 2013.*

Another of the results corresponding to the exploratory analysis of functional data was the function of bivariate covariance-centric; this was evidenced by the variability of the data in the time at the zone used. In figures 5 and 6, the covariance functions obtained are observed. For this result interpretation, it is observed in the graph that in the ordinates of the day are shown while and in the abscissa the values of variability are seen. It is also observed that most of the times the more dispersed data in the majority of the functions are those corresponding to the slot in the morning especially between 5 am and 12 pm while there are more homogeneity in them in the close to 7 pm and 10 pm hours.

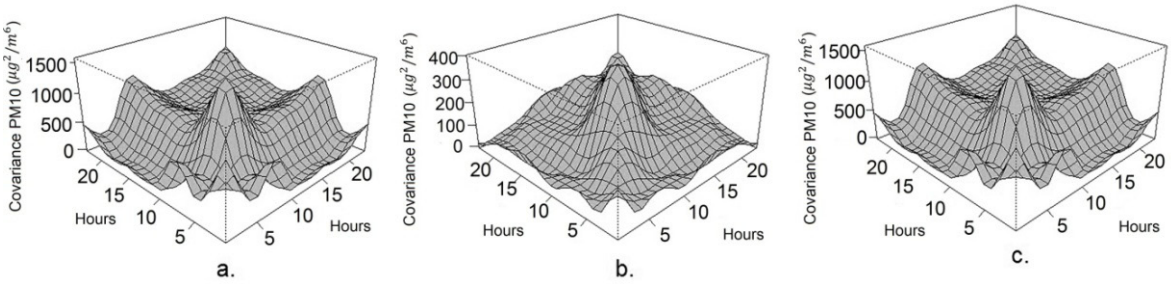


Figure 5. *Covariance functions for August. a). 2011 b). 2012 c). 2013.*

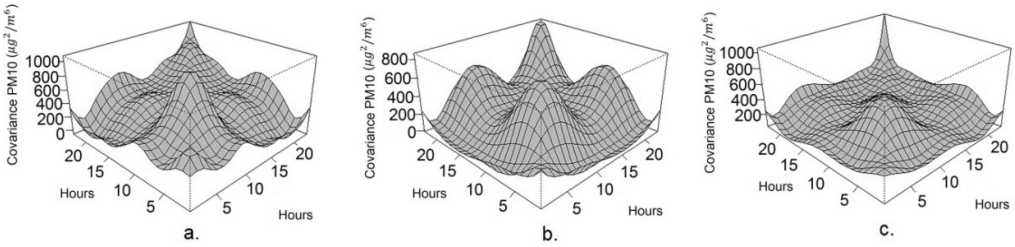


Figure 6. *Covariance functions for May. a). 2011 b). 2012 c). 2013.*

From the functional data of each date of study, it was possible to obtain two functional principal components for each; however, it was necessary to rotate them using the VARIMAX algorithm (see figures 7 and 8 for this process results).

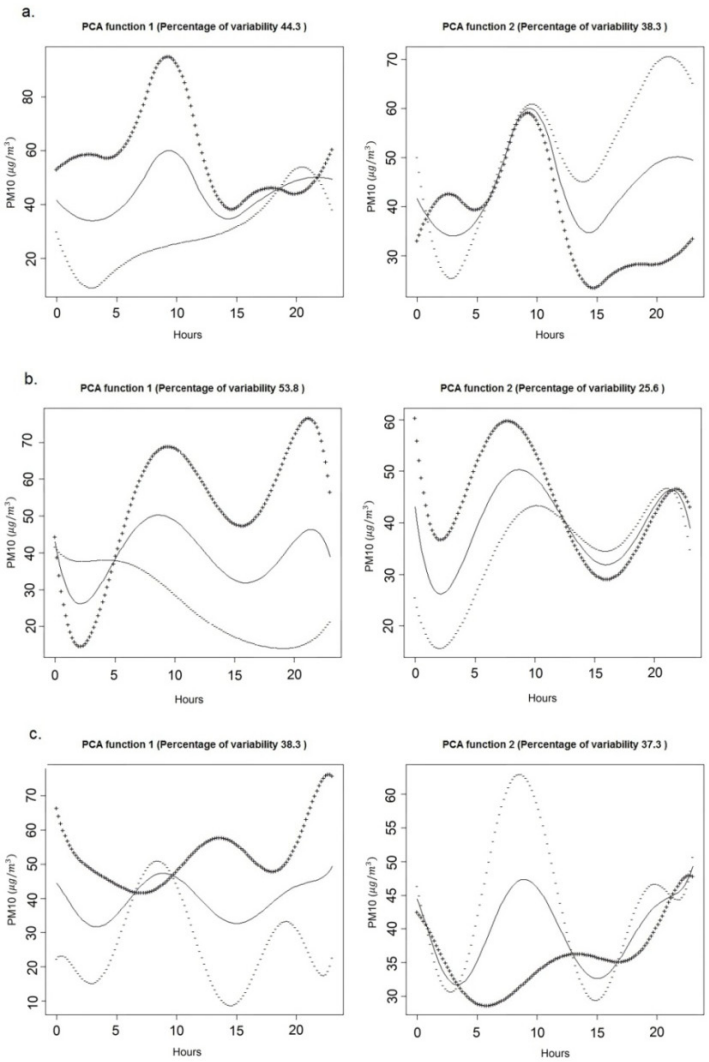


Figure 7. *Functional Principal Component rotated with the VARIMAX algorithm, plotted as positive perturbations (dashed lines '+') and negative (dashed lines '-') of the functional mean (thick solid line). to). May 2011. b). May 2012. c). May 2013.*

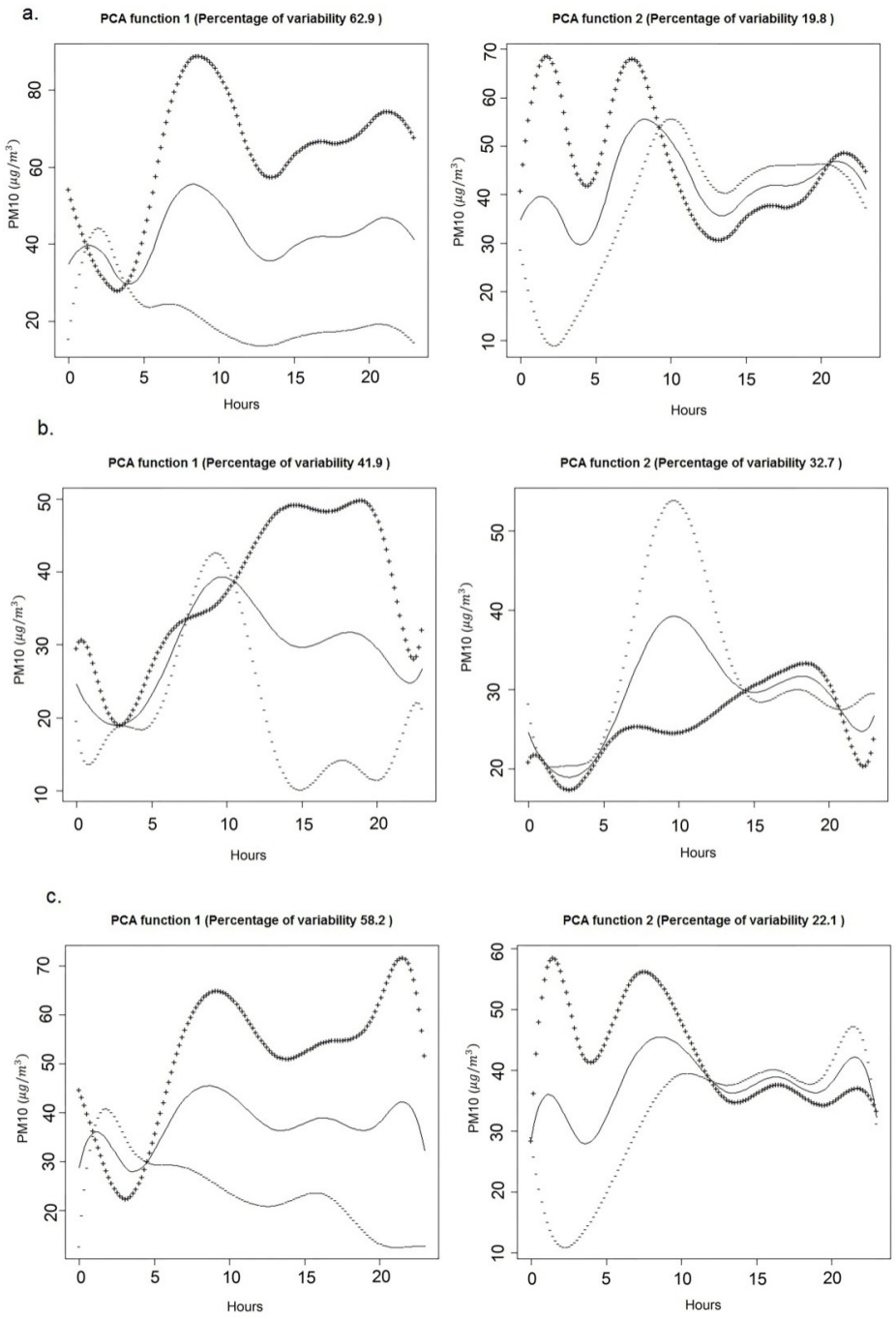


Figure 8. *Functional Principal Component rotated with VARIMAX algorithm, plotted as positive perturbations (dashed lines '+') and negative (dashed lines '-') of the functional mean (thick solid line). to). August 2011. b). August 2012. c). August 2013.*

Finally, the model deterministic IDW was used for the spatial interpolation and the functional principal components of each season; as a result, four maps that correspond to the first and second eigenfuncion from 10 o'clock and 4 o'clock in the morning were obtained (see figures 9 and 10).

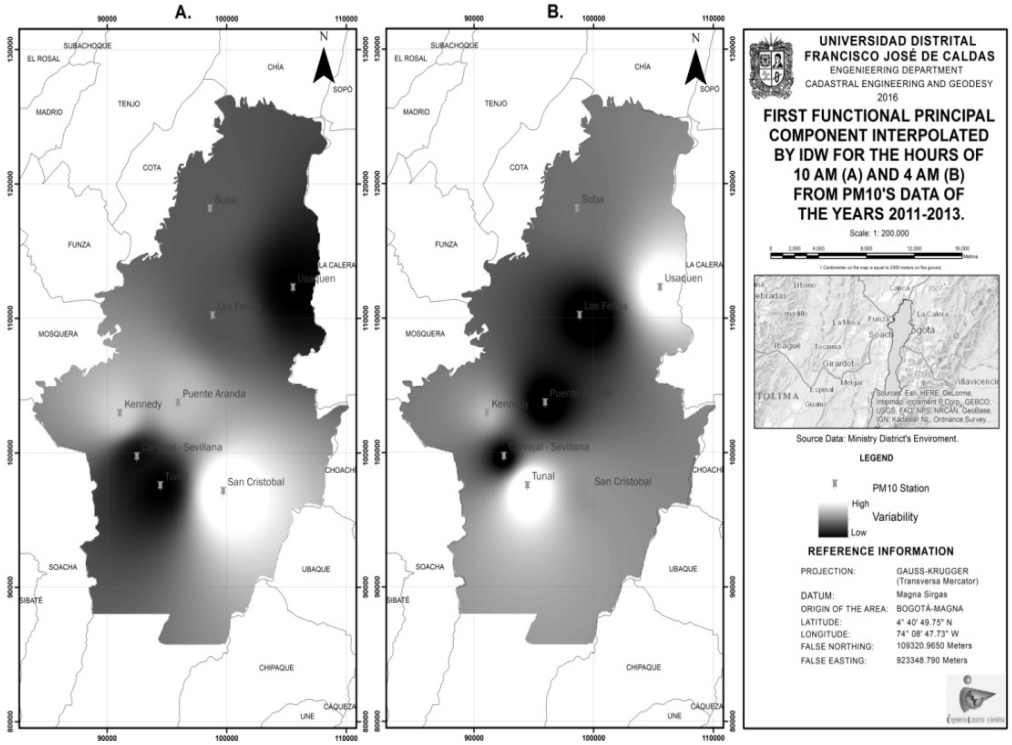


Figure 9. *Spatial interpolation of the first functional principal component by IDW. to). For 10 am b). For 4am.*

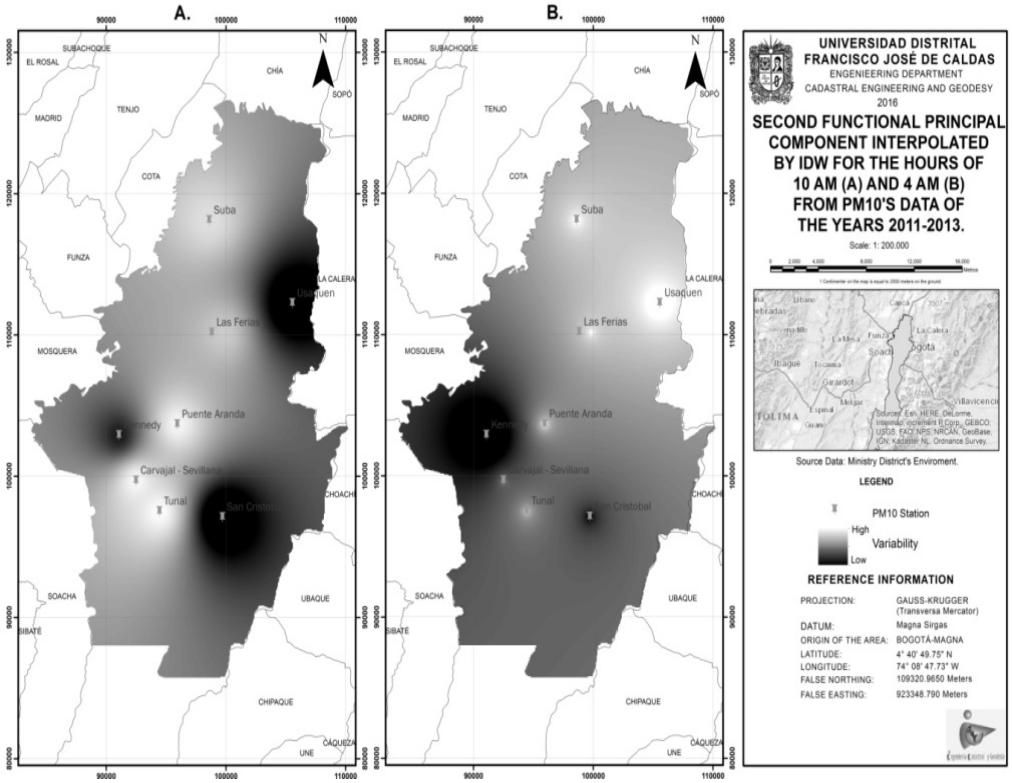


Figure 10. *Spatial interpolation of the second functional principal component by IDW. to). For 10 am b). For 4am*

3.2 Discussion

Regarding the exploratory analysis of functional data, media functions and bivariate covariance functions were found. These functions allowed a robust diagnosis of data which in the case of the media function allowed observing that in general. In August (see figure 3) the highest values of PM10 concentrations are concentrated between the 5 and 10 o'clock in the morning. Similar results were observed in the diagnosis of the data in May (see figure 4). In the graphics, it was observed: the values of the average in each hour of the day allowing a general diagnosis of the behavior of the contaminant. It is important to highlight that as it was stipulated in resolution 0610 2010 about the limits of exposure to PM10 in terms of 24 hours (100 μg/m3) the means of measurements functions showed that the records did not exceeded this limit. As it shown in figures 5 and 6, the diagnosis of variability is visible and easily identified and the times where there is more dispersion among the data noting that in the case of August, specifically for the year 2011 (Figure 5a) and 2013 (figure 5b) the dispersion is concentrated in the strip from 5 am to 10 am while in 2012 I group in the 8 am to 1 pm. Similarly, in May (see figure 6), it was observed that in all cases, the more dispersed data were located in the strip from 5 am to 12 pm; additionally, in 2012 (see figure 6b) another scattered strip was from 4 pm to 8 pm. It was also noted that the data in 2013 (see figure 6) are homogeneous due to the behavior of its covariance function. This means that in some way, the data are reliable because the variability was constant in the majority of the hours of the day.

Concerning the interpretation of the functional principal components rotated for each date, this was performed for each one because it depended exclusively on behavior of the bivariate covariance-centric function and which possess weighting function compared with the average function (as shown in figure 7 and 8). Hence, in the case on May in 2011 (figure 7a), the first component retained was 44.3% of the data variability; additionally, the first nineteen hours a day were described. The second component has the 38.3% of the variability, but the first hours of the day (from 1 am to 6 am) were only described; it means that the two weighting functions functions were not able to explain the behavior of the whole time slot. The opposite happened with data from 2012 (see figure 7b), since the first component has 53.8% of the variability of the data, that it described much of the day except the time slot from 12 am to 5 am, while the second retained the 25.6% of the variability, successfully synthesizing only the first twelve hours of the day. Finally, in 2013 (Figure 7 c), the first component has about the 38.3% of the data variability; moreover, part of the day was summed up. The only exception was made from 6 am to 10 am in the morning while the second retained the 37.3% of the variability explaining only the afternoon hours from 1 pm to 5 pm.

In the case of August, in 2011 rotated functional principal components (see figure 8a), presented a supplementary behavior. Since the first component retained around the 62.9% of the data variability of describing much of the day except for the hours early morning hours like from 1 am to 4 am, while the second retained 19.8% of the variability, explaining only the first nine hours of the same second functional principal component . Meanwhile, in 2012 (figure 8b), weighting functions were not adequate, since the first component explained about the 41.9% of the data variability, adequately describing only the slot from 11 am until the end of the day, while the second retained the 32.7% of the variability explaining properly the slot from 3:00pm to 8.00pm. Finally, in 2013 the eigenfuncion (see figure 8c), it was noted that the first component explained about the 58.2% of the data variability describing properly most of the day except its first five hours, while the second retained the 22.1% of the variability explaining properly only eleven hours of the day; consequently, they are not suitable because they do not perform a correct explanation of all day long.

As a result of spatial diagnostic from functional components as it can be seen in figures 9 and 10,

the maps which correspond to the interpolation of the eigenfuncion, noting that the first component retained in most cases 60% of the data variability, while the second shows 20%. Figure 9a shows that the stations of surrounding areas: Suba, Usaquén, Las Ferias, Carvajal-Sevillana and Tunal have low variability in the data, while stations of Kennedy, Puente Aranda and San Cristóbal have high variability in the data. As to the second functional principal component for the hour of highest contamination (figure 10a) it was noted that the stations of Tunal, Carvajal-Sevillana, Puente Aranda, Las Ferias and Suba have a high variability, and the stations of Kennedy, Usaquén and San Cristóbal have low variability. Regarding to the first and second principal component, it is observed that the highest time of contamination shows that in general the area that is not adequately explained by the major functional components of Usaquén, Tunal and Kennedy present high of variability levels. In terms when it comes to less pollution (figures 9b and 10b), it is observed in the case of figure 9b that stations of Usaquén, Tunal and Kennedy present high levels of variability, meanwhile Las Ferias, Puente Aranda and Carvajal Sevillana have low levels of variability. In figure 10b it is seen that there is a high variability in the North, in the stations Suba, Usaquén and Las Ferias; the stations Kennedy and San Cristobal have a low variability and other stations showed average variability. The first and second main functional component for reducing pollution time evidenced that the area better explained by the FPCA was the area of Usaquén followed by climb, while the Carvajal zone do not contribute information relevant to these eigenfuncions.

From the spatial interpretation, applicability which would have this type of study for the characterization of the material data can be observed PM10 particulate for the dates of study, due to possessing with only two maps is around 80% of the variability of the data (equivalent to own this proportion of the data). In addition, the functional statistical characteristics of pollution data, could be diagnoses of the average behavior of contaminant in the regions of greater reliability without manipulating the data or statistics functions exposed previously, being faster process, depending on their reliability only the correct characterization data and the proper interpretation of the principal components.

4. Conclusions

The methodology used in the collection and analysis of the principal functional components rotated for each study date casts as a result that the first and the second eigenfuncion characterize synthetically the temporal behavior of the contaminant, in the majority of cases; consequently, this feature in more appropriate in the sense that the functions are more precise and robust in the description of the pollutant compared with functional exploratory analysis. Therefore, the interpretation of these is more appropriate since the explanation is made across the range of temporary using only the two main components without needing to use statistical functions and direct manipulation of the entire database as a whole.

Seen in this light, it is important to note that the diagnosis made from the spatial interpolation of the main functional components, it was evident that regardless of the time of reporting, the eigenfuncion were able to explain the behavior of the pollutant in most of the study area, as it was the North-East and South-West of Bogotá city. However, there were areas that provided no variability due to special characteristics of these regions; however, this generality is ideal because it would possess spatial synthetic report of the contaminant behavior at that moment of time having no need to manipulate and compare the records of all the dates required by the study, meeting fully the purpose of the method since it will significantly decrease time that would make a reliable diagnosis of concentration of PM10, and thus government entities could make decisions regarding their control.

5. Acknowledgements

A special thanks to *Secretaria Distrital del Medio Ambiente* from Bogotá for providing some data and to *Universidad Distrital Francisco José de Caldas* for supporting this study.

6. References

Aldunate, P., Paz, O. & Halvorsen, K. (2011). Los efectos de la contaminación atmosférica por PM10 sobre la salud ciudad de La Paz–Bolivia (3650 msnm). *Acta Nova* 3 (2), 422-442.

Aristizabal, J. (2011). *Metodología estadística para el análisis de datos funcionales cerebrales: Una aproximación con potenciales evocados*. Master thesis, Facultad de Estadística, Universidad Nacional de Colombia, Bogotá D.C, Colombia.

Bautista, L. (2010). *Evaluación de las características físicas y químicas del material particulado en Bogotá*. Master thesis, Facultad de Ingeniería, Universidad de los Andes, Bogotá D.C, Colombia.

Blanco-Becerra, L., Roa, R. & Rojas, A. (2015). Influence of precipitation scavenging on the PM2. 5/PM10 ratio at the Kennedy locality of Bogota, Colombia. *Revista Facultad de Ingeniería* (76), 58-65.

Blanco-Becerra, L., Miranda-Soberanis, V., Hernández-Cadena, L., Barraza-Villarreal, A., Junger, W., Hurtado-Díaz, M. & Romieu, I. (2014). Effect of particulate matter less than 10μm (PM10) on mortality in Bogota, Colombia: a time-series analysis, 1998-2006. *Salud pública de méxico* 56 (4), 363-370.

Bohórquez, M. (2010). *Diferenciabilidad de funciones de covarianza espacio temporal no separables*. Master thesis, Facultad de Estadística, Universidad Nacional de Colombia, Bogotá D.C, Colombia.

Carreño, S. (2013). *Aplicaciones en meteorología de modelos funcionales de predicción en componentes principales*. Master thesis, Departamento de Estadística, Universidad de Granada, Granada, España.

Cattani, G., Bucchianico, D., Menno, A., Dina, D., Inglessis, M., Notaro, C., Settimo, G., Viviano, G. & Marconi, A. (2010). Evaluation of the temporal variation of air quality in Rome, Italy from 1999 to 2008. *Annali dell'Istituto Superiore di Sanità* 46 (3), 242–253.

Chávez, C., Sánchez, J. & DelaCerda, J. (2015). Análisis de componentes principales funcionales en series de tiempo económicas. GECONTEC: *Revista Internacional de Gestión del Conocimiento y la Tecnología* 3 (2), 13-25.

Franco, J., Rojas, N., Sarmiento, O., Hernández, L., Zapata, E., Maldonado, A., Matiz, L. & Behrent, E. (2013). Niveles de material particulado en colegios distritales ubicados en vías con alto tráfico vehicular en la ciudad de Bogotá: estudio piloto. *Revista Facultad de Ingeniería* (49), 101-111.

Gaitán, M., Cancino, J. & Behrentz, E. (2007). Análisis del estado de la calidad del aire en Bogotá. *Revista de Ingeniería* 26 (11), 81–92.

Galvis, B. & Rojas, N. (2005). Relación entre PM 2,5 y PM10 en la ciudad de Bogotá. *Revista Acta Nova* 3 (2), 54-60.

García, F., Agudelo, R. & Jiménez, K. (2008). Distribución espacial y temporal de la concentración de material particulado en Santa Marta, Colombia. *Facultad Nacional de Salud Pública* 24 (2), 73-82.

García-Ubaque, J., Garcia-Ubaque, C. & Vaca-Bohórquez, M. (2011). Medical consultation in productive age population related with air pollution levels in Bogota city. *Procedia Environmental Sciences* 4, 165-169.

Giraldo, R. (2007). Análisis exploratorio de variables regionalizadas con métodos funcionales. *Revista Colombiana de Estadística* 30 (1), 115–127.

González, M., Casas, F. & Peinado, P. (2005). *Estudio del riesgo de ocurrencia de un siniestro utilizando el análisis de componentes principales funcionales*. In 13h Annual ASEPUMA Conference, Myrtle Beach, South Carolina, p. 1-11.

Herrera, L. (2012). *Una metodología para la predicción espacial de una variable escalar georeferenciada basada en su información y la de una variable funcional*. Master thesis, Facultad de Estadística, Universidad Nacional de Colombia, Bogotá D.C, Colombia.

Honaker, J., King, G. & Blackwell, M. (2014). *Amelia: Amelia II: A Program for Missing Data (Version 1.7.3)*. http://gking.harvard.edu/amelia. [Last visited: 9 December 2015]

Levitin, D., Nuzzo, R., Vines, B. & Ramsay J. (2007). Introduction to functional data analysis. *Canadian Psychology/Psychologie canadienne* 48 (3), 135–155.

Mehrjardi, R., Jahromi, M, Mahmod, S. & Heidari, A. (2008). Spatial Distribution of Groundwater Quality with Geostatistics (Case Study: Yazd-Ardakan Plain). *World Applied Sciences Journal* 4 (1), 9-17.

Montero, J. & Fernández-Avilés, G. (2015). Functional Kriging Prediction of Pollution Series: The Geostatistical Alternative for Spatially-Fixed Data. *Estudios de Economía Aplicada* 33 (1), 145-174.

Ramsay, J. & Silverman, B. (2005). *Functional Data Analysis*, 2nd edn, New York: Springer.

Othman, N., Jafri, M. Z. M., & San, L. H. (2010). Estimating particulate matter concentration over arid region using satellite remote sensing: A case study in Makkah, Saudi Arabia. *Modern Applied Science* 4 (11), p131.

Preciado, J. (2008). *Crecimiento Urbano, Pobreza y Medio Ambiente en Bogotá: Los efectos Socio Ambientales en Tres Humedales*. En VII Seminario Nacional de Investigación Urbano-Regional. Universidad Nacional de Colombia, Bogotá, Colombia, p. 1-5.

Reina, J. & Olaya, J. (2012). Curve fitting nonparametric methods for studying behavior from air pollution pm10. *Revista EIA* (18), 19–31.

Rodríguez-Moreno, N., Sarmiento-Suarez, V., Medina-Palacios, K. & Hernández, L. (2013). Factores de riesgo para enfermedad respiratoria en población de 5 a 14 años de una Localidad de Bogotá, 2012-2013. *Rev. Salud pública* 15 (3), 408-420.

Vargas, F. & Rojas, N. (2009). Source apportionment of PM10 in Bogotá, Colombia. *In 8th Annual CMAS Conference, Bogotá, Colombia* p. 1-4.

Villatoro, M., Henríquez, C. & Sancho, F. (2007). Comparación de los interpoladores IDW y Kriging en la variación espacial de pH, CA, CICE y P del suelo. *Agronomía Costarricense* 32 (1), 95-105.

Zakey, A., Abdel-Wahab, M., Pettersson, J., Gatari, M. & Hallquist, M. (2008). Seasonal and spatial variation of atmospheric particulate matter in a developing megacity, the Greater Cairo, Egypt. *Atmósfera* 21 (2), 171–189.

Revista Ingeniería y Competitividad por Universidad del Valle se encuentra bajo una licencia Creative Commons Reconocimiento - Debe reconocer adecuadamente la autoría, proporcionar un enlace a la licencia e indicar si se han realizado cambios. Puede hacerlo de cualquier manera razonable, pero no de una manera que sugiera que tiene el apoyo del licenciador o lo recibe por el uso que hace.