Computer vision and file transfer protocol to identify and transmit evolutionary curves of PM10

Viviana Bernal-Benítez
Electronic Engineer
Institution: Universidad Santo Tomás
Full address: Av. Universitaria No. 45 – 202 Tunja - Colombia
E-mail: viviana.bernal@usantoto.edu.co

Juan Gómez-Malagón
Electronic Engineer
Institution: Universidad Santo Tomás
Full address: Av. Universitaria No. 45 – 202 Tunja - Colombia
E-mail: juan.gomezm@usantoto.edu.co

Camilo Pardo-Beainy
Master's Degree in Electronic Engineering
Institution: Universidad Santo Tomás
Full address: Av. Universitaria No. 45 – 202 Tunja - Colombia
E-mail: camilo.pardo@usantoto.edu.co

ABSTRACT
The following article presents the techniques developed for the treatment of images obtained from the modeling process with AERMOD of data on the immission concentration of PM10 particulate matter. This data conditioning was carried out in order to generate, through the analysis of dispersion images, isolines that identify and quantify the areas of the Sogamoso Valley where PM10 concentrations occur from emissions from the limestone firing process in Nobsa, Boyacá. It is a first approach to the prediction of the dispersion phenomenon and the spatio-temporal determination of the influence that this immission has on the air quality conditions of a region.

Keywords: AERMOD, Computer Vision, Evolutionary Curves, FTP.

1 INTRODUCTION

In Latin America there is the difficulty of acquiring or implementing stations for measuring and monitoring environmental conditions; that equipment's are expensive and must meet EPA standards that add complexity to its handling. In 2015, only 17 of the 33 countries that make up the Latin American region measured air quality (Chile, 2017). But, modeling software such as CALPUFF, ATSTEP and AERMOD allow the visualization of the dispersion phenomenon of different pollutants (Oliveri Conti, Heibati, Kloog, Fiore, & Ferrante, 2017) in an integrated way with orography and atmospheric conditions,

being a process that reduces costs, time and does not necessarily require information from field measurements. This is advantageous, but even so, these computational aids have limitations.

Nobsa is characterized for being a center of artisan work and being the predominant sector of limestone extraction and transformation ("Contaminación por producción de cal en Nobsa, Colombia | EJAtlas", s/f). Because of this last activity, it presents high rates of environmental contamination; that has had an impact on the quality of life of its inhabitants and surrounding cities.

Developing a system that generates the evolution curves of the concentration levels of a pollutant over time, offers communities a low-cost alternative that indicates how exposed their inhabitants are to certain concentrations of particulate matter; in addition to synthesizing and supporting the isolinear modeling generated by the AERMOD software that identifies and quantifies the areas of the Sogamoso Valley where PM10 (Quijano, Quijano, Quijano, & Manzano, 2015) concentrations occur from emissions from the limestone firing process in Nobsa.

2 RELATED WORKS

The serious consequences of climate change have been evidenced by the work of researchers around the world. (Potter, Emberlin, & Mullins, 2000) were one of the first researchers to study the size, composition and concentration of particulate material using image analysis techniques to extract quantitative, geometric and densitometry data from PM10.

(Liu, Zhou, & Xiao, 2013) investigated a new image threshold segmentation method based on the genetic algorithm and the Otsu method to identify the main morphological parameters of PM10 particulate matter suspended indoors.

(Shu, Chen, & Xiong, 2020) used integrated technology in order to acquire images of the same environment at different instants of time, and then applied quantification models. Furthermore, using the variance of the gradient function, they formed new images with different levels of clarity in order to analyze the correlation of environmental quality.

CASAP is a scientific space in which works oriented to air quality in Colombia are exposed and socialized ("CASAP 2017", s/f) like the one developed by Guevara. M. (Luna, Andrés, Gonzalez, Ceron, & Carlos, 2019), who space-time evaluated the PM10 and PM2.5 (Miskell, Pattinson, Weissert, & Williams, 2019) pollutants in the country through a process generated in Python scripts of satellite information and air quality measurements.

As evidenced in related works, the AERMOD software analyzes the concentration and deposition of air pollution caused by various sources and creates a dispersion model (Software Científico, s/f). Generally, a graphic report is acquired with all the levels of dispersion of the pollutant, but this does not

allow an exhaustive or particular analysis (specific area of a certain concentration level) to be made at each of the concentration levels provided by the software; Furthermore, it does not have the facility to compare the variations in the dispersion of pollutants generated by different reports.

For this reason, the project is oriented towards a process of conditioning the images obtained in order to generate a system that allows the most relevant concentrations of the PM10 pollutant to be identified and quantified in Las Caleras-Nobsa and generates reduction scenarios as a management mechanism for the quality of the air resource.

3 MATERIALS AND METHODS

For the generation of the evolutionary curves of the pollutant, pre-processing and processing techniques are applied to the results obtained from the project "Modeling strategies to reduce particulate matter in the air (Caleras-Nobsa)" (Fig. 1); A file transfer protocol (Buchanan, 2002) is also implemented in order to interactively transfer the reports generated between various systems (Fig. 2). The general stages that describe the development of the project are: acquisition, pre-processing, processing and extraction of characteristics to finish with the presentation of the results.
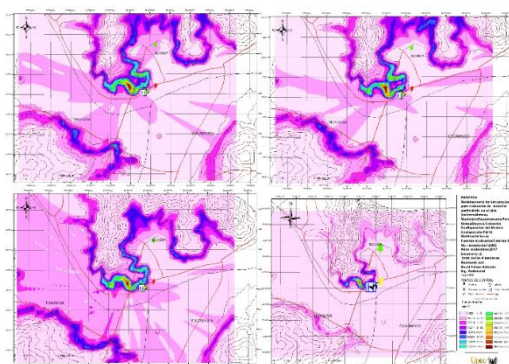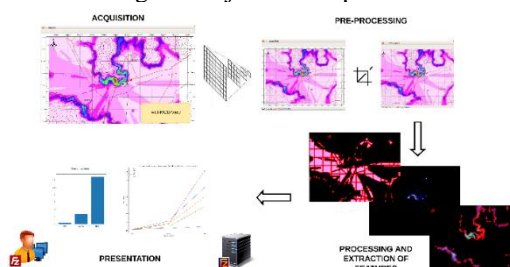
Fig. 1 Scenario modeling.

Fig. 2 Project development.

To start, a scale operator (Adelson, Burt, Anderson, Ogden, & Bergen, 1984) was used because the initial image was oversized and when working with images in this state, a greater processing resource is used that is not considered necessary.

Gaussian reduction: This algorithm is mainly used to reduce the resolution of the images, it can be described as a pyramid, where a level g (i) has a better resolution than its upper level g (i + 1). By default the kernel described below is used: (Publikoa, Gradua, Grado, & Lana, s/f).

$$\frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

As an objective of the project was the identification of three different degrees of concentration of the pollutant in question, logical operators (González, Martínez, & Pernía, 2006) were used in order to create masks that allowed highlighting the desired concentration levels. In order to obtain a better perception of the information, a grouping of pixels is carried out by detecting the most significant changes in intensity in the image, in other words, performing edge detection (Cavanagh & Leclerc, 1989); To achieve this, the Canny edge detector (Canny, 1986) and the morphological erosion, opening and closing operators were used.

Opening: Equation (2) describes the actions of erosion and then expansion to execute the "opening" operation. The result of this operation is the adjustment of the structuring element in the input image. The operator is called "opening top-hat".

$$A \circ B = (A \ominus B) \oplus B$$
$$A \circ^\wedge B = A - (A \circ B)$$

Closing: The inverse operation to opening is "closing" which describes dilation followed by erosion. The operator is called "closing top-hat". Equation (3), if considered as a linear process, can be regarded as a moving average filter.("Hands-on Morphological Image Processing - Edward R. Dougherty, Roberto A. Lotufo - Google Libros", s/f).

$$A \bullet B = (A \oplus B) \ominus B$$
$$A \bullet^\wedge B = A - (A \bullet B)$$

Canny's Algorithm: For edge detection using the algorithm created by John F. Canny in 1986 (Corke, 2011), it is necessary to apply a set of operations (Fig. 3).

Fig. 3 Process for applying Canny's algorithm.
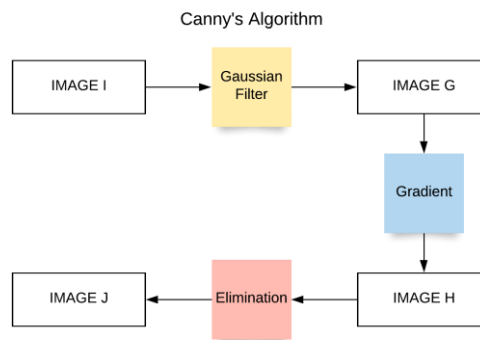
a. Applying a Gaussian filter (Lambraño García, Lázaro Plata, Lázaro Plata, Trigos Quintero, & Trigos Quintero, 2017) the input image I is smoothed, this is described by the famous gauss bell function expressed in (4) where σ takes a default value of 0.5 and (x, y) are the coordinates of a pixel central in the image, generating image G.

$$f(x,y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

b. Gradient calculation, this operator uses the calculation of the partial derivatives (x, y) in each of the RGB planes where g(x, y) is the image and n is the direction of the gradient which determines the angle θ. Image H is obtained.

$$\frac{\partial g(x,y)}{\partial n} = \begin{bmatrix} \nabla R(X,Y) \cdot n \\ \nabla G(X,Y) \cdot n \\ \nabla B(X,Y) \cdot n \end{bmatrix};$$

$$n = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$$

c. Next, the magnitude (m) and the direction of the resulting gradient (θ) of image H are calculated, this in order to determine the areas of the image where this value is maximum or where the image has a large discontinuity.

$$|m(x,y)| = |\lambda_{\max}|$$
$$m^2 = \| Cx \|^2 \cdot \cos(\theta)^2 + 2 \cdot Cx \cdot Cy \cdot \sin(\theta) \cdot \cos(\theta) + \| Cy \|^2 \cdot \sin(\theta)^2 \square$$

d. A non-maximum values elimination is performed, which consists of comparing the pixel with its neighbors in order to see if it is part of an edge. This occurs when the magnitude of this pixel is greater than that of its two neighbors assigned by the direction given by the angle θ. Image J is obtained.

$$|m(x,y)| > |m_2(x_2,y_2)|$$
$$|m(x,y)| > |m_3(x_3,y_3)|$$

e. Finally, equation (8) describes the hysteresis thresholding of image J, which compares the resulting pixels with two threshold values u1 and u2. Two thresholds are considered because if the value is too high, information from the edges can be lost and, conversely, if it is too low, noise could be generated in the image; if the pixels exceed those thresholds they will be considered as edges. (Valverde, 2007)(Burger & Burge, 2016).

$$u_1 < u_2$$
$$J(x,y) > u_2$$
$$J > u_1$$

For the quantification of the affected space, the geometric property of regions is used: Area (9), as it is of interest to calculate the number of pixels in the region delimited by edge detection. Where N is the number of pixels in the region "R".

$$A(R) = |R| = N$$

To complete the extraction stage, a process of pixel equivalence to Km^2 is created using the grid of the initial image and knowing the number of pixels associated with the manipulated image. This process was repeated for five more images, but it was not identical, because the size of some images varied and also the grid of these.

Finally, the information of the High, Medium and Low concentrations for the months of January to May is collected in easy-to-view formats and is transmitted using the file transfer protocol (Loshin, 2000) belonging to the application layer in the TCP/IP model (Alvernia & Bautista, 2017) (Fig. 4). Free software FileZilla is used to create an FTP server and client; An FTP client is nothing more than an application that connects to a server with the same protocol to manage and access files. The server is a service that uses the described protocol to share files with other users (Fig. 5). Access to this FTP server

is done by entering the necessary data for the connection, such as the server address, username and password.

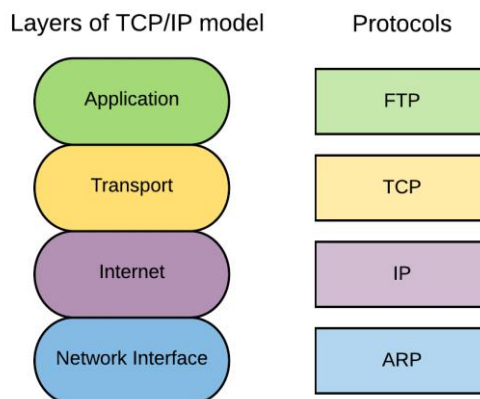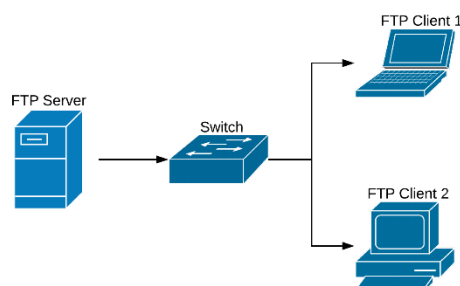Fig. 4 Layers of the TCP/IP model and its main protocols.

Fig. 5 File transmission protocol topology.

4 RESULTS

After using this methodology, it is verified that the image processing techniques are complementary tools to the dispersion models acquired from the AERMOD software.

The images had a rescaling and trimming process in order to use less processing resources (Fig. 6 (a)) to then create ranges of shades to identify the different concentrations; applying this step, the graph b shown in Fig. 6 was obtained. The extracted color is again highlighted when calculating the conjunction of the pre-processed image with the mask (Fig. 6 (c)).

Fig. 6 Processing (a) Pre-processing (b) Mask (c) Low tonality mask.

When the canny algorithm is carried out, the pixels present in the limits of the objects of an image are used, when making the definition of the border, a small loss of information occurs.

To calculate the areas that describe the different PM10 concentrations , morphological operations are carried out on the original image. Erosion can be defined as the elimination of pixels at the limits of an object in an image, and dilation, being the opposite operation, is the addition of pixels in these contours. In order to eliminate noise and unrepresentative data, first a closing operation (dilation and then erosion) is performed and to open up the definition of the point objects in the image, an opening operation is performed (erosion and then dilation), then The canny algorithm is used, which by means of different mathematical operators finds the most abrupt changes in intensity in the image, defining them as edges and allowing the arithmetic calculation of the respective areas to be carried out at each concentration of PM10 (Fig. 7).

Fig. 7 Feature extraction.
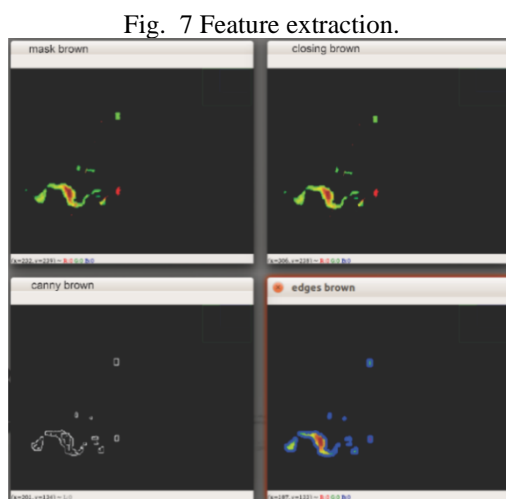
By condensing the model information into evolutionary curves, great support is provided for air quality management. This facilitates the understanding of the scatter diagrams produced by the AERMOD format. The final reports consist of a graph that condenses the information extracted for each month (Fig. 5), and an .xlsx file that specifies in km2 the area affected by the particulate material in the three important levels (Table 1). This was determined by making a relationship between the pixels in the image and the size of the grid.
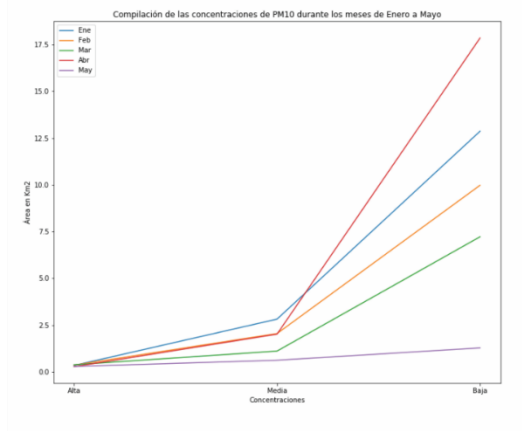
Table. 1 Affected area in km^2 by each concentration level.

	High	Medium	Low
January	0,336024	2,815709	12,87089
February	0,354715	2,043974	9,979796
March	0,375068	1,106512	7,223275
April	0,271228	2,012615	17,85996

| May | 0,282584 | 0,616115 | 1,275899 |

Fig. 8 Compilation of PM10 concentrations during the months of January to May.

For the last stage (Fig. 1) in the development of the project, a PC had to be configured as a server. Initially, the port and address that will be referred to the FileZilla server are assigned. In this case, PC 1 will be the server and its address will be the localhost. Also, some settings should be edited such as the range of ports for users, the welcome message and the creation of the certificate associated with the server. Next, an access profile is created that the user will use to enter the server and the files that he decided to share.

The client must be aware of the server address, the username and password modified on the server and the port assigned to the FTP to achieve the successful connection. By zooming in on Fig. 9, it can be identified that a client with address 19X.6X.1X9.247 is accessing the server with global address 18X.1X1.9X.154 to acquire the reports shared by the server.
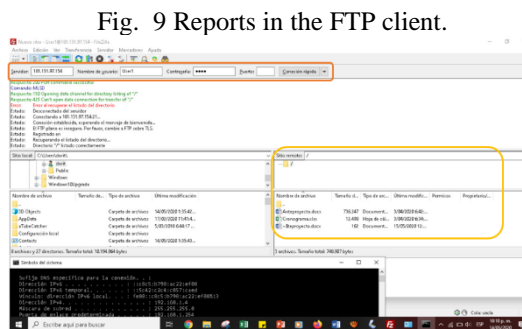
Fig. 9 Reports in the FTP client.

5 CONCLUSIONS

The images corresponding to the three levels of concentration of the contaminant were obtained. (High, Medium, Low), with this information, individual reports were generated for each month and they are compiled into a total graph to make it easy to analyze the degrees of dispersion throughout the months.

The viewing of these reports can be done on the device that performed the analysis or on another by transferring files using a transfer protocol.

As the information that was desired to be obtained did not depend on the resolution of the images obtained in the software, scaling processes were carried out with the intention of reducing the processing requirements. This allowed the scripts to run efficiently.

The application of techniques for the processing and extraction of characteristics in an image, provides a positive impact to the exploration of new methodologies to analyze dispersion reports such as that generated by the AERMOD software, since it gives other study areas greater tools that enhance the capacity of information analysis and allows to study in more detail a specific behavior.

As evidenced in the compilation graph of PM10 concentrations, for May 2019 it presented a drastic variation in the low concentration level. This data is taken as a reference to generate an intensive analysis of the modeling report for that month.

The acquisition of this type of reports made it possible to display in a more efficient way a concentration level dangerous for human health, in the future, it would allow generating real-time alerts on the behavior and degree of concentration of other types of pollutants.

SOUTH FLORIDA
JOURNAL OF DEVELOPMENT

REFERENCES

Adelson, E. H., Burt, P. J., Anderson, C. H., Ogden, J. M., & Bergen, J. R. (1984). Pyramid Methods in Image Processing. RCA engineer, 29(6), 33–41.

Alvernia, S. A., & Bautista, D. R. (2017). ANALYSIS OF A NETWORK IN AN IPV6 ENVIRONMENT: A VIEW FROM NETWORK INTRUSIONS AND THE TCP / IP MODEL. Revista Colombiana De Tecnologias De Avanzada (Rcta), 1(29), 81–91. Recuperado de http://ojs.unipamplona.edu.co/ojs_viceinves/index.php/RCTA/article/view/2490/1246

Buchanan, W. J. (2002). The Complete Handbook of the Internet. En The Complete Handbook of the Internet. https://doi.org/10.1007/978-0-306-48331-8

Burger, W., & Burge, M. J. (2016). Digital Image Processing an Algorithmic Introduction Using Java. En Springer (Vol. 24). https://doi.org/10.1007/978-1-4471-6684-9

Canny, J. (1986). A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6), 679–698. https://doi.org/10.1109/TPAMI.1986.4767851

CASAP 2017. (s/f). Recuperado el 2 de diciembre de 2019, de http://casap.com.co/es/informacion.asp

Cavanagh, P., & Leclerc, Y. G. (1989). Shape from shadows. Journal of Experimental Psychology: Human Perception and Performance, 15(1), 3–27. https://doi.org/10.1037/0096-1523.15.1.3

Chile, B. del C. N. de. (2017). La contaminación del aire en Latinoamérica - Programa Americas.

Contaminación por producción de cal en Nobsa, Colombia | EJAtlas. (s/f). Recuperado el 3 de diciembre de 2019, de https://ejatlas.org/conflict/contaminacion-en-nobsa-colombia

Corke, P. (2011). Image formation. En Springer Tracts in Advanced Robotics (Vol. 73). https://doi.org/10.1007/978-3-642-20144-8_11

González, A., Martínez, F., & Pernía, A. (2006). Técnicas y algoritmos básicos de visión artificial. En U. de la Rioja (Ed.), Universidad de La Rioja (Vol. 1). La Rioja, España.

Hands-on Morphological Image Processing - Edward R. Dougherty, Roberto A. Lotufo - Google Libros. (s/f). Recuperado el 16 de mayo de 2020, de https://books.google.com.co/books?id=-ch3fZTh08EC&printsec=frontcover&hl=es&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Lambraño García, E. D., Lázaro Plata, J. L., Lázaro Plata, J. L., Trigos Quintero, A. E., & Trigos Quintero, A. E. (2017). Revisión De Técnicas De Sistemas De Visión Artificial Para La Inspección De Procesos De Soldadura Tipo Gmaw. Revista Colombiana De Tecnologias De Avanzada (Rcta), 1(29). https://doi.org/10.24054/16927257.v29.n29.2017.2486

Liu, H., Zhou, X., & Xiao, L. (2013). Measurement of concentration and size distribution of indoor PM10. Advances in Intelligent Systems and Computing, 212, 697–705. https://doi.org/10.1007/978-3-642-37502-6_83

Loshin, P. (2000). Big book of Internet file transfer RFCs. Morgan Kaufmann.

Luna, G., Andrés, M., Gonzalez, V., Ceron, B., & Carlos, L. (2019). Evaluación espacial y temporal de PM 10 y PM 2 . 5 en Colombiautilizando información satelital (CAMS , MODIS-AOD) y mediciones de calidad del aire en superficie. 8–12.

Miskell, G., Pattinson, W., Weissert, L., & Williams, D. (2019). Forecasting short-term peak concentrations from a network of air quality instruments measuring PM 2.5 using boosted gradient machine models. Journal of Environmental Management, 242(February), 56–64. https://doi.org/10.1016/j.jenvman.2019.04.010

Oliveri Conti, G., Heibati, B., Kloog, I., Fiore, M., & Ferrante, M. (2017). A review of AirQ Models and their applications for forecasting the air pollution health outcomes. Environmental Science and Pollution Research, 24(7), 6426–6445. https://doi.org/10.1007/s11356-016-8180-1

Potter, C., Emberlin, J., & Mullins, J. (2000). The use of image analysis for PM10 characterisation of volumetric air samples taken in Cardiff since 1956. Aerobiologia, 16(1), 131–135. https://doi.org/10.1023/A:1007610620280

Publikoa, N. U., Gradua, I. I., Grado, T. F. D. E., & Lana, G. B. (s/f). Restauración de imágenes mediante pirámide gaussiana y técnicas de reducción y ampliación.

Quijano, A., Quijano, M. J., Quijano, L. A., & Manzano, S. T. (2015). Air Toxicity of the City of Villa Del Rosario in Samples of. 88.

Shu, Y., Chen, Y. G., & Xiong, C. (2020). Application of image recognition technology based on embedded technology in environmental pollution detection. Microprocessors and Microsystems, 75, 103061. https://doi.org/10.1016/j.micpro.2020.103061

Software Científico, A. (s/f). Software : AERMOD View 9.8. Recuperado el 19 de noviembre de 2019, de https://www.addlink.es/productos/aermod-view

Valverde, J. (2007). Detección de bordes mediante el algoritmo de Canny. ResearchGate, (October 2007), 8.