ESTIMATING CONCENTRATIONS OF SUSPENDED PARTICULATE MATTER OVER THE METROPOLITAN AREA OF MEXICO CITY USING SATELLITE AND GEOSPATIAL IMAGERY: PRELIMINARY RESULTS

RODRIGO T. SEPÚLVEDA-HIROSE, ANA B. CARRERA-AGUILAR,
MAGNOLIA G. MARTINEZ-RIVERA, PABLO DE J. ANGELES-SALTO & CARLOS HERRERA-VENTOSA
Faculty of Engineering, Universidad Nacional Autónoma de México, Mexico

ABSTRACT

In order to diminish health risks, it is of paramount importance to monitor air quality; however, this process is accompanied by high costs of physical and human resources. This research is carried out in this context with the main objective of developing a predictive model for concentrations of inhalable particles (PM_{10} and $PM_{2.5}$) using remote sensing. To develop the model, satellite images, mainly from Landsat 8, of the Mexico City's metropolitan area, were used. Using historical PM_{10} and $PM_{2.5}$ measurements of the RAMA (Environmental Monitoring Network of Mexico City) and processing the available satellite images, a preliminary model was generated in which it was possible to observe critical opportunity areas that would allow for the generation of a robust model. After applying this preliminary model to the scenes of Mexico City, three areas of great interest were identified due to their presumed high concentration of PM. The zones of interest all presented high plant density, bodies of water and bare soil without buildings or construction, or vegetation. To date, work continues along these lines to improve the proposed preliminary model. In addition, a brief analysis of six distinct models was made and presented in articles developed in different parts of the world in order to visualize the optimal bands for the generation of a suitable model for Mexico City. It was found that infrared bands have helped modelling in other cities, but the effectiveness of these bands for the specific geographic and climatic conditions of Mexico City is still in need of evaluation.

Keywords: air quality, modelling pollution, particulate matter, remote sensing.

1 INTRODUCTION

Particulate matter (PM) has a variety of sources and compositions. It appears naturally in the atmosphere due to volcanic activity and fires; however, human activity also contributes to high levels of PM pollution. Particles form in the atmosphere as a result of complex reactions of chemicals such as sulphur dioxide and nitrogen oxides. Particle pollution is composed of: PM_{10}, particles with a diameter of 10 micrometers and less, and $PM_{2.5}$ with a diameter of 2.5 micrometers and less. Due to their small size, particles can be easily inhaled causing serious health problems [1]. PM_{10} particles, after inhaled, can travel to the lungs and can even reach the bloodstream, while $PM_{2.5}$, being even smaller, pose a greater health risk.

Access to air pollution monitoring equipment is restricted due to its elevated cost; this explains the use of remote sensing techniques as an alternative method for measurement. Remote sensing provides advantages such as global and exhaustive coverage of land surface, multi-scale and non-destructive observations and repetitive coverage.

Several investigations that have taken place, mainly in Asia, show the feasibility of using satellite images for the estimations of particulate material concentrations [1]–[10]. According to Chu et al. [4], the four most used models for the prediction of particulate material are: Lineal Multiple Regression, Mixed Effects Model (MEM), Model of Chemical Transport and Pondered Geographic Regression.

The project's main objective is to develop air quality models for the estimation of both PM_{10} and $PM_{2.5}$ from both locally measured data and remote sensing.

WIT Transactions on Ecology and the Environment, Vol 236, © 2019 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/AIR190121

2 STUDY AREA AND DATA SETS

2.1 Study area

Mexico City's metropolitan area comprises an area of 2,370 km^2 and has a population of 20,565,000 inhabitants which generates a density of 8,700 people per square kilometer [11]. The city, the capital of Mexico, is located at 19 29 52 N and 99 7 37 O in the central region of Mexico. Dominant wind flow in Mexico City moves from north to south with a confluence line in the southeastern corner of the city. This wind behavior indicates a flow of atmospheric pollution from the north and east to the south and southwest part of the metropolitan area. As a consequence of these orographic conditions in the metropolitan area, it is subject to air stagnation and high concentrations of pollutants, especially in the southern region of the city. Average wind speeds oscillate between 0.5 and 1.5 m/s, with maximum speeds reaching to 2.5 m/s [12].

Mexico City has an automatic air quality monitoring system comprised of 29 stations with a capability of measuring PM_{10} and $PM_{2.5}$ hourly year-round. The coverage of these stations will define the area of study for the project. According to Mexican regulation, the allowed limits for PM pollution are shown in Table 1.

Table 1: Mexican regulations allowed limits for particulate matter pollution.

Pollutant	**Period of measurement**	**Limit (µg/m^3)**
$PM_{2.5}$	Annual average	12
$PM_{2.5}$	24 hour average	45
PM_{10}	Annual average	40
PM_{10}	24 hour average	75

2.2 Data sets

A statistical analysis was carried out to determine the spatial and temporal distribution of the concentrations of particulate matter ($PM_{2.5}$ and PM_{10} particles) present in Mexico City's metropolitan area. Precipitation was also considered as a variable to determine a possible relationship between the behavior of both variables (precipitation and particulate matter concentrations) in a certain time span.

A selection of dates when the satellite Landsat 8 passed over the ZMCM was made. Said satellite made contact with the ZMCM for the first time on May 20, 2013 at 5:02 pm (GMT+0), 11:00 am local time. From that date on, all later dates were considered in this study up until the month of May 2018, accounting the satellite's temporal resolution of 16 days for a total of 115 dates during the 5 years in which the analysis took place.

Regarding the pollutants PM_{10} and $PM_{2.5}$, the data was extracted from the database generated by the measurements of all Automatic Network of Environmental Monitoring stations (Fig. 1) under the Ministry of the Environment of Mexico City.

Both sets of data (PM_{10} and $PM_{2.5}$) were processed. Fig. 2 shows the boxplot graphic for PM_{10}, at 11:00 am, on the dates compatible with the passing of the Landsat 8 platform. As can be observed, the concentrations have an annual cyclical behavior related to the rainy season (May to November) and dry months (November to May) in Mexico City.

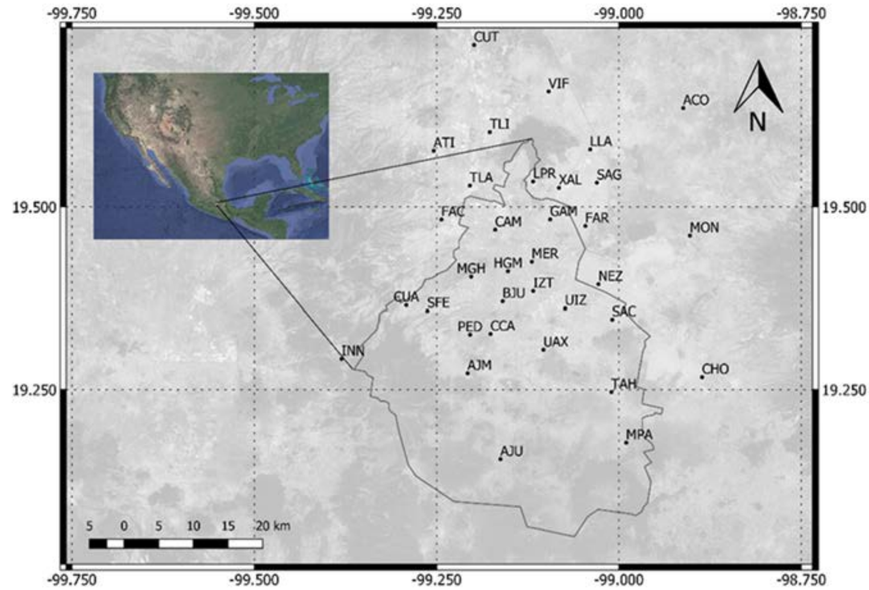

Figure 1: Location of the stations of the Automatic Network of Environmental Monitoring. *(Source: Ministry of the Environment of Mexico City.)*

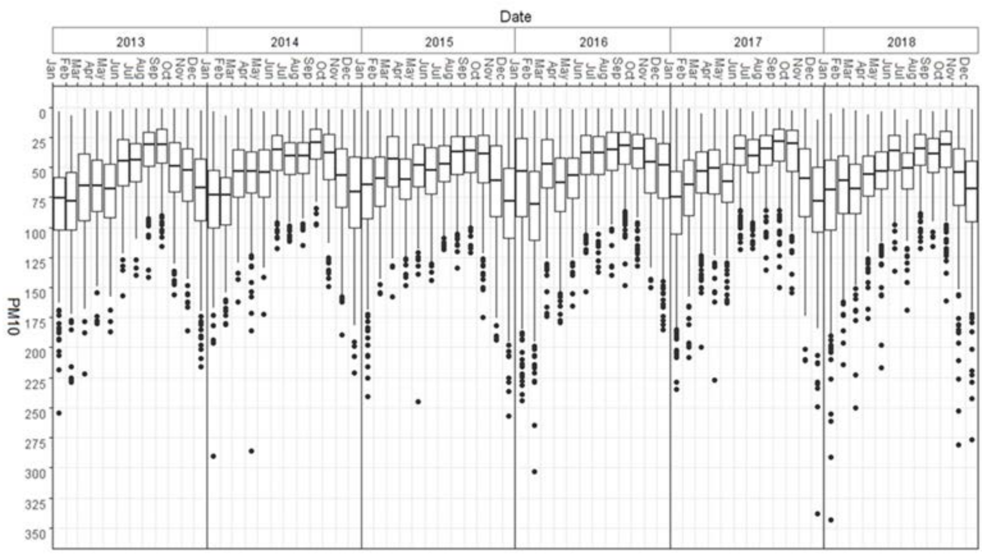

Figure 2: Boxplot graph showing PM_{10} concentrations (µg/m3) in Mexico City on the dates in which satellite Landsat 8 passed over the city in the period between May 2013 and May 2018. The median of measured data is represented by the line in the box, the interquartile range box represents the middle 50% of the data, and the whiskers represent the ranges for the top and bottom 25% of the data. The dots represent outliers.

In addition to the statistical analysis per season, a spatial analysis was also elaborated upon, first based on zoning by quadrants and then, later, spatial interpolation of the data, using the monthly average values as a reference and thus obtaining the pollutants $PM_{2.5}$ and PM_{10} and interpolating the data gathered by the entirety of the stations located within the study area. The evaluation period lasted from the year 2013 until 2018, as it tied in with the information provided by the Landsat 8 platform. Figs 3 and 4 show the mentioned interpolations for $PM_{2.5}$ and PM_{10} respectively.

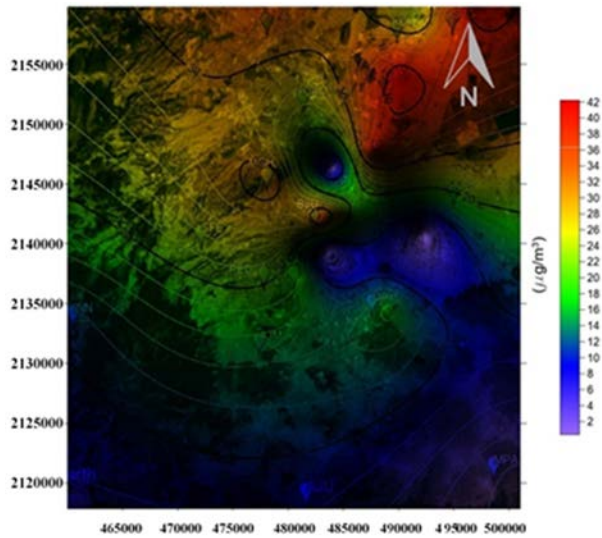

Figure 3: $PM_{2.5}$ concentration, February 2018.

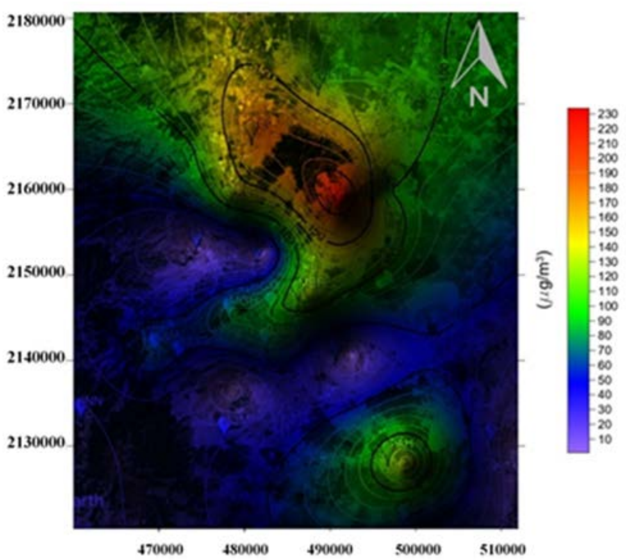

Figure 4: PM_{10} concentration, January 2018.

3 METHODOLOGY

The project methodology begins with the collection of historical data from the SEDEMA databases regarding PM_{10} and $PM_{2.5}$. Once the historical data is collected, a statistical analysis is carried out. The analysis is comprised of one study on the temporal relation and another on the spatial relation to the system. For the temporal study, the monthly averages of particulate matter PM and precipitation are plotted against time and thus the relationship between both variables analyzed. With respect to the spatial study, a spatial representation of PM pollution is made in each of the 4 zones of the Mexico City's metropolitan area and the results are analyzed in conjunction.

Together with the collection of historical data, the satellite image platform is selected. Once a platform is selected, the satellite image bank is made. These images have to be subjected to pre-processing which allows for the retrieval of the corrected reflectance data. The next step is to analyze and classify viable images for the study. In order to do so, the values of cloud percentage in the images were obtained and a visual inspection was made by previewing the images through the digital platform Earth Explorer.

Based on this analysis it was found that, of the 119 scenes available to date, only 44 can be used. Of these 44 scenes, the products of digital numbers (Level 1) are not corrected for atmospheric conditions, however reflectance TOA (Level 2) and surface reflectance (Level 2) were downloaded, as shown in Table 2.

Table 2: Summary Landsat 8 scenes acquired.

Type of images	Number	Percentage
Useful	44	35.77%
Not useful	74	60.16%
Not available	5	4.07%
Total	123	100.00%

3.1 Processing and digital analysis of images

In order to validate the products of Level 2, pre-processing was performed on the May 20, 2013 scene with the help of the software QGIS (Quantum GIS). The following eqn was applied for the calculation of reflectance on the top of the atmosphere (TOA)

$$\rho'_\lambda = M_p Q_{cal} + A_p, \quad (1)$$

where ρ'_λ is top of atmosphere (TOA) reflectance without solar angle correction, M_p is specific rescaling factor by band obtained in metadata, A_p is specific resizing factor by band obtained in metadata, and Q_{cal} is pixel values of standard products quantified and calibrated (DN).

In order to obtain the TOA reflectance corrected for solar angle in the center of the scene, it must be divided by the sine of that angle, information provided in the metadata of the scene. Once this pre-processing was concluded, the reflectance values were compared with the respective Landsat 8 product, and differences of the order of 3% were found. This is due to the Landsat 8 algorithm which performs the correction of the solar angle per pixel in the scene, offering a more precise result. Atmospheric corrections were made via different methods, to verify the surface reflectance. First, the subtraction of dark objects (DOS) correction was made, which identifies the objects that "should be dark" in some bands, having as a consequence the supposition that the reflectance measured by the sensor is only that

www.witpress.com, ISSN 1743-3541 (on-line)

caused by the atmosphere. Another correction also took place with the help of the module FLAASH (fast line-of-sight atmospheric analysis of hypercubes) of ENVI; the problem with these methods is that their correction is constant and when applied to the whole image, they do not give precise information of the atmospheric reflectance.

As a conclusion of these analyses, the use of the Level 2 product of Landsat 8 was chosen. This decision was made based on the fact that it offers a more precise atmospheric correction due to its better calibration. Due to the location of monitoring stations, a vector layer is generated and associated to each image in order to begin processing data. The atmospheric reflectance data is extracted for each location and associated with the historical field data obtained from the monitoring stations.

In order to determine the optimal bands to use in the generation of the model for Mexico City, an analysis of reported models was made. The scene was analyzed using 6 reported models (Table 3). Table 4 shows the mean squared error obtained from each of them.

Table 3: Proposed equations to model PM_{10} concentrations.

Reference	Proposed equation	Study area
Torres and Vivanco [5]	$\mathrm{PM_{10}} = -31.56 + 111.4\mathrm{NDVI} - .03\mathrm{RA.B1} + 0.089\mathrm{RA.B3} - .019\mathrm{RA.B4} - 5.42\mathrm{Season}$	Quito, Ecuador
Abad and Mejía-Coronel [6]	$\mathrm{PM_{10}} = -126.9 + 0.005\mathrm{ND.B11} + 582.7\mathrm{TOA.B2} - 207.1\mathrm{TOA.B5}$	Cuenca, Ecuador
Saraswat et al. [7]	$\mathrm{PM_{10}} = -0.8689 - 94.22\mathrm{RA.B1} + 166.48\mathrm{RA.B2} + 21.01\mathrm{RA.B3} - 78.98\mathrm{RA.B4}$	Delhi, India
Shaheen et al. [8]	$\mathrm{PM_{10}} = 10008\mathrm{RA.B2} - 21356\mathrm{RA.B3} + 10965\mathrm{RA.B4}$	Gaza Strip, Palestine
Nadzri et al. [9]	$\mathrm{PM_{10}} = 396\mathrm{RA.B2} + 253\mathrm{RA.B3} - 194\mathrm{RA.B4}$	Mecca, Saudi Arabia
Ozelkan et al. [10]	$\mathrm{PM_{10}} = 232.66\left(\frac{\mathrm{RA.B6}}{\mathrm{RA.B7}}\right) - 78.673$	Izmir, Turkey

Table 4: Mean squared errors reported after modelling PM_{10} concentrations.

Reference	Reported mean squared error	
	Authors	**Mexico City's Metropolitan Area**
Torres and Vivanco [5]	NA	54.71
Abad and Mejía-Coronel [6]	77.70	45.90
Saraswat et al. [7]	18.98	75.5
Shaheen et al. [8]	9.71	74.33
Nadzri et al. [9]	NA	55.02
Ozelkan et al. [10]	NA	49.59

After processing, classifying and analyzing field measurements as well as the extracted data from remote sensing, best fitted models for air quality parameters in Mexico City can be propounded. Significant variations in air quality across space and time are observed.

4 RESULTS

A multiple linear regression was made in order to model the dependency of PM_{10} on reflectance values obtained from remote sensing. The model that was generated is preliminary and used to observe the performance by means of the parameters R^2, with an adjusted R and root mean squared error.

After applying the preliminary model to the scene for the 44 available dates, it was observed that the model presents temporal variations in the concentration of PM_{10}, which was an expected result.

Model for June 5, 2013:

$$PM10 = 0.553RA.B1 - 1.764RA.B2 + 0.524RA.B3 + 0.432RA.B4 + 3.7RA.B5 - 0.757RA.B6 + 0.3017RA.B7 + 31.646TOA.B10 - 41.306TOA.B11 + 3310.923$$

$$R^2 = 0.93.\ R^2 - adjusted = 0.73.\ RSME = 4.84\ \mu g/m^3. \qquad (2)$$

Model for December 1, 2014:

$$PM10 = -1.835RA.B1 + 29358RA.B2 - 0.279RA.B3 + 0.319RA.B4 - 1.034RA.B5 + 1.796RA.B6 - 0.587RA.B7 - 197.454TOA.B10 + 247.547TOA.B11 - 1537.326$$

$$R^2 = 0.86.\ R^2 - adjusted = 0.45.\ RSME = 10.10\ \mu g/m^3. \qquad (3)$$

Figs 5 and 6 show the PM_{10} concentrations for June 5, 2013 and December 1, 2014. On these dates, low and high concentration values, respectively, were observed. The areas circled in red reflect the zones in which the model did not work properly.

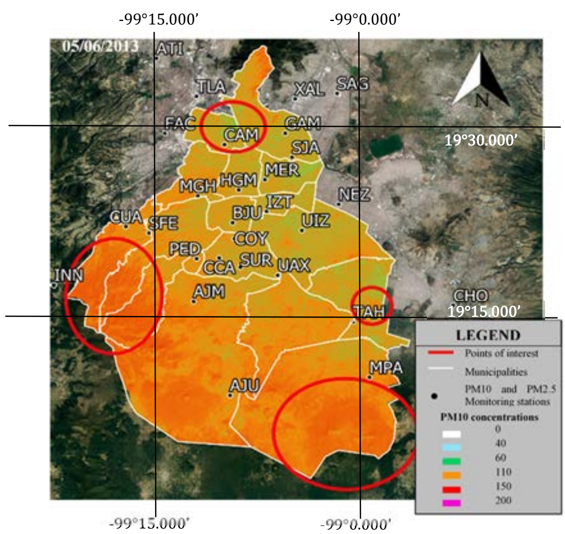

Figure 5: PM_{10} modelled concentrations in Mexico City on June 5, 2013.

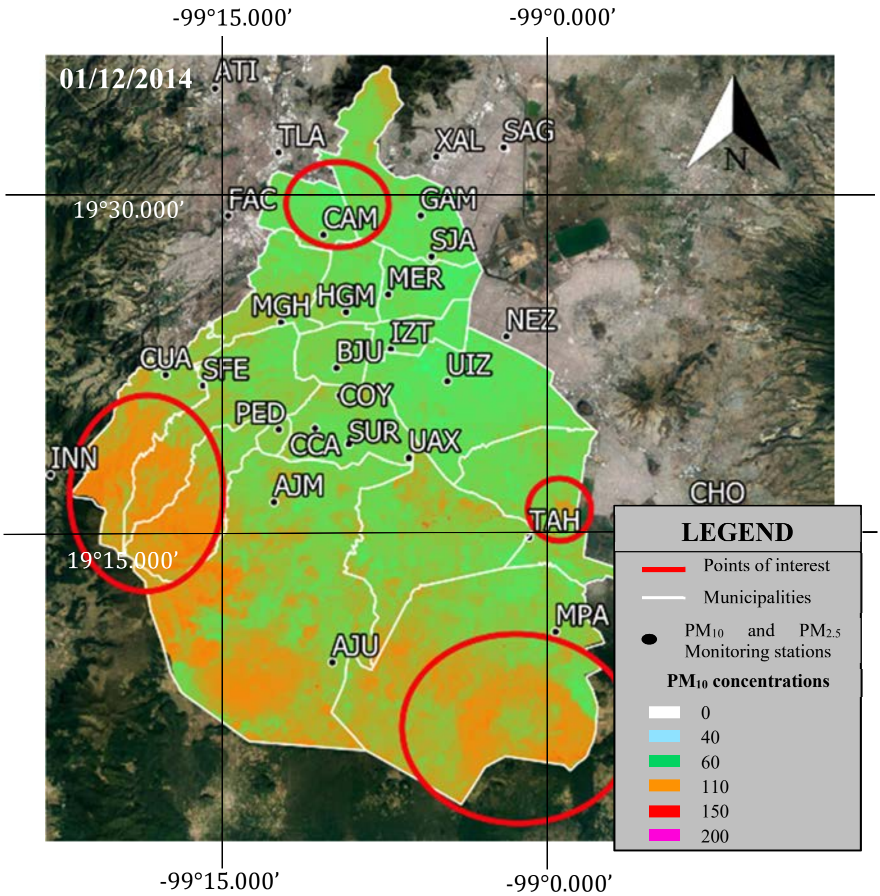

Figure 6: PM_{10} modelled concentrations in Mexico City, December 1, 2014.

5 DISCUSSION

The model for June 5, 2013 corresponds to the rainy season (May to November), whereas the model for December 1, 2014 corresponds to the dry season (November to May). The determined relationship between correlation coefficients were quite strong suggesting a high correlation. Figs 7 and 8 show the observed vs. predicted PM_{10} concentrations for June 5, 2013 and December 1, 2014, respectively. The performance of both models was very good, with a mean R^2 for June of 0.93 and for December of 0.86. As discussed above, the concentration of PM varies considerably across space and time. Thus, it is essential to match the spatial–temporal resolutions of AOD and PM data as closely as possible. Fortunately, for Mexico City, the monitoring system provides hourly air quality data and the acquisition time of Landsat 8 satellite is at 11:00 AM local time.

Other alternatives for remote sensing atmospheric modelling have been tested by several authors, including AOT retrieved from the operational MODIS algorithm. Although kilometric resolution works well in regional scales, the detailed aerosol spatial distribution at city scale may be missing. The develop of the predicting models for PM_{10} concentrations at a 30 m resolution across Mexico City will be a very useful tool for more accurate long-term estimations of PM exposure in this urban area. Although the determined correlation values are good respective to linear regressions, our future research will be geared towards improving algorithms for modelling PM including artificial neural networks.

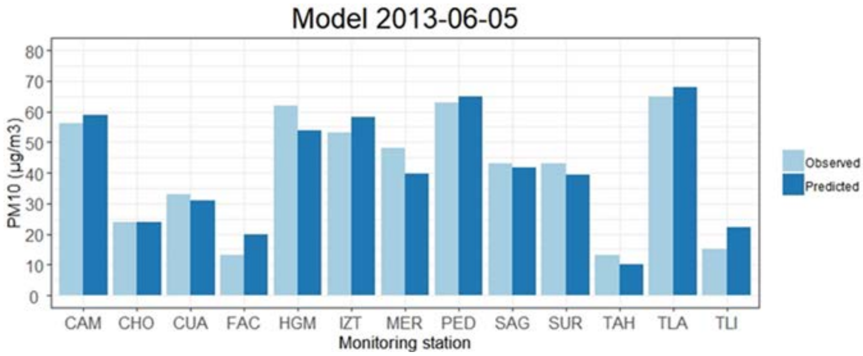

Figure 7: Comparison between observed and predicted values, June 5, 2013.

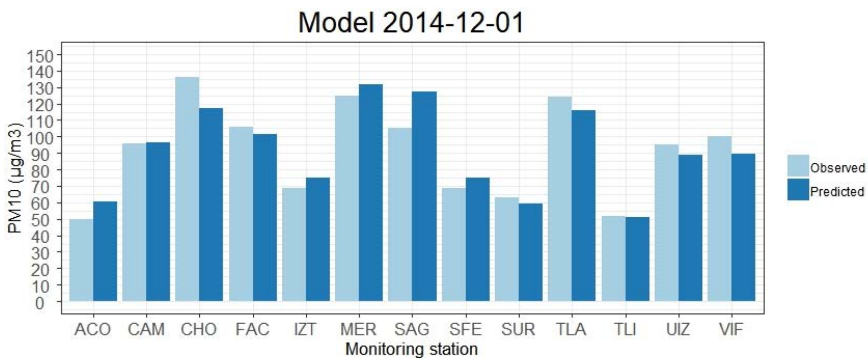

Figure 8: Comparison between observed and predicted values, December 1, 2014.

6 CONCLUSION

One of the main aspects in air pollution modelling is the quality and quantity of data used. An enormous amount of existing information regarding the parameters of this study has been made available thanks to the constant monitoring of SEDEMA during the past decades. Transversal and active cooperation of different governmental instances in the development of this work allowed for the access to high quality data. On the other hand, satellite-based measures of aerosol optical depth (AOD) as compared to fixed ground monitors allows for greater coverage of PM_{10} estimations. In this paper, a method for retrieving PM_{10} from Landsat 8 Operational Land Imager (OLI) images over urban areas is proposed. Important variations in the air quality within space and time can be observed, which will be incorporated into the second phase of this work.

The models presented in this work can be used in various disciplines for estimating PM and may be particularly useful for environmental epidemiology studies. The obtained PM_{10} distribution maps can be integrated into a geographical information system and analyzed with relational data to obtain more reliable interpretations that act against air pollution effects.

ACKNOWLEDGEMENTS

The authors thank the Ministry of Education, Science, Technology and Innovation of Mexico City for the funds, advice and support given during the development of the Project SECITI/089/2017 "Development of Models for the Assessment of Suspended Particulate Material in the Metropolitan Area of Mexico City, with the Utilization of Satellite Images and Geospatial Information." We also thank all the students and professors on our team for their enthusiasm which motivated the success of this project.

REFERENCES

[1] Sun, L., Wei, J., Bilal, M., Tian, X., Jia, C., Guo, Y. & Mi, X., Aerosol optical depth retrieval over bright areas using Landsat 8 OLI images. *Remote Sensing*, **8**(23), pp. 1–14, 2015.

[2] Hamzelo, M., Gharagozlou, A., Sadeghian, S., Baikpour, S.H. & Rajabi, A., Modelling of carbon monoxide air pollution in large cities by evaluation of spectral LANDSAT 8 images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives*, **40**(1W5), pp. 281–285, 2015.

[3] Nguyen, N.H. & Tran, V.A., Estimation of PM_{10} from AOT of satellite Landsat 8 image over Hanoi City. Presented at *International Symposium on Geoinformation for Spatial Infrastructure Development in Earth and Allied Sciences*, Danang, Vietnam, 2014.

[4] Chu, Y. et al., A review on predicting ground $PM_{2.5}$ concentration using satellite aerosol optical depth. *Atmosphere*, **7**(10), p. 129, 2016.

[5] Torres, N. & Vivanco, V. Comparación en la estimación de material particulado PM_{10} usando imágenes satelitales Landsat 7, Landsat 8 y Modis en Quito. (Undergraduate thesis). *Atmospheric Measurement Techniques Discussions*, pp. 1–47, 2018.

[6] Abad, L. & Mejía-Coronel, D., Estimación de la concentración de material particulado menor a 10 micras a través de sensores remotos en el área urbana de la Ciudad de Cuenca. *XVI Conferencia de Sistemas de Información Geográfica*, pp. 381–390, 2017.

[7] Saraswat, I., Mishra, R.K. & Kumar, A., Estimation of PM_{10} concentration from Landsat 8 OLI satellite imagery over Delhi, India. *Remote Sensing Applications: Society and Environment*, **8**(April), pp. 251–257, 2017.

[8] Shaheen, A., Kidwai, A.A., Ain, N.U., Aldabash, M. & Zeeshan, A., Estimating air particulate matter 10 using Landsat multi-temporal data and analyzing its annual temporal pattern over Gaza Strip, Palestine. *Journal of Asian Scientific Research*, **7**(2), pp. 22–37, 2017.

[9] Nadzri, O., Zubir Mat Jafri, M., San Lim, H., Othman, N. & Hwee San, L., Estimating particulate matter concentration over arid region using satellite remote sensing: A case study in Makkah, Saudi Arabia. *Modern Applied Science*, **4**(11), 2010.

[10] Ozelkan, E., Karaman, M., Mostamandy, S., Avci, Z.D.U. & Toros, H., Derivation of PM_{10} levels using obra on Landsat 5 TM images: A case study in Izmir, Turkey. *Fresenius Environmental Bulletin*, **24**(4B), pp. 1585–1596, 2015.

[11] Demographia World Urban Areas, Built up urban areas or world agglometarions, 14th annual ed. http://demographia.com/db-worldua.pdf. Accessed on: 14 Mar. 2019.

[12] Secretaría de Medio Ambiente, Informe climatológico ambiental del Valle de México. Environmental climatological report on the Valley of Mexico, 2006. www.aire.cdmx.gob.mx/descargas/publicaciones/flippingbook/informe_anual_climatologico_2006/#p=1. Accessed on: 12 Mar. 2019.

www.witpress.com, ISSN 1743-3541 (on-line)