THEORETICAL MODEL FOR ESTIMATING ATMOSPHERIC POLLUTION BASED ON CAR TRAFFIC OBSERVATION AND MONITORING STATION

STOICA[1] Mircea-Edwin[1], MĂNTESCU[2] Yasmina-Andreea[2], DIACONU[3] Diana-Elena[3]

Faculty: Industrial Engineering and Robotics, [1]Specialization: Machine Construction Technology, [2]Specialization: Security Engineering in Industry, Year of studies: II, email: yasminamantescu@yahoo.com

Scientific leaders: S.l.dr.ing. Ștefan Constantin PETRICEANU, S.l.dr.ing. Ana Maria BOGATU

ABSTRACT: In the context of increasing concerns about air pollution and its impact on public health and the environment, this research article proposes a theoretical model to estimate air pollution based on vehicle traffic observation and air quality monitoring station records. The study took place over five days at a chosen location near a monitoring station, and data on car traffic and vehicle emissions were collected and analysed. Using the numerical simulation program MATLAB, based on the data from the specialized literature, the empirical formula was developed to estimate the atmospheric pollution at time t0, taking into account the previous pollution, in 10 minutes increments, in the last 60 minutes before the moment of calculating its value. The results of the theoretical model are compared with the values recorded by the air quality monitoring station, and the subsequent analysis and discussion, based on appropriate statistical tools, address the performance of the model and the factors that may influence its accuracy. The conclusions drawn from this study provide valuable information about the effectiveness of the theoretical model in the estimation of atmospheric pollution and the possibility of its application in the management and monitoring of atmospheric pollution in the urban environment. Prospects and directions for further research are related to the possibility of using in the future, instead of the very expensive monitoring station, a car traffic surveillance camera together with a suitable software that would allow the estimation of air pollution in many more points in the inhabited areas and/or with heavy car traffic.

Keywords: air pollution, air pollution monitoring, Matlab numerical interpolation

1. INTRODUCTION

Air pollution is a global problem with a significant impact on human health, quality of life and the environment [1], [2]. The phenomenon has been associated with a number of conditions, such as cardiovascular, respiratory and even cancer [3], [2].

In this context, air pollution monitoring is essential to understand the origin and distribution of pollutants, to assess the impact on human health and the environment, and to develop pollution reduction policies and strategies [4], [5].

The current study aims to develop a theoretical model for air pollution estimation based on the observation of car traffic and air quality monitoring station records, addressing the need for effective tools for monitoring and managing air pollution in the urban environment [6], [7]. The main goal is to evaluate the performance of the proposed theoretical model in the estimation of atmospheric pollution and to explore the possibility of its application in the management and monitoring of atmospheric pollution in inhabited areas and/or with intense car transit [8], [9].

Also, this study aims to strengthen and expand the existing knowledge in the field of air pollution monitoring [10], [11] and provide new perspectives for the development of effective monitoring solutions, such as the use of traffic surveillance cameras and software specialized for the estimation of air pollution at several points in inhabited areas and/or with intense car transit [12], [13].

2. MATERIALS AND METHODS

The study methodology involved the completion of the following stages:

- The choice of the study location: the area near an air quality monitoring station, which has heavy traffic and where it is possible to clearly observe moving vehicles, was selected as an observation point, namely the Drumul Taberei neighborhood, on Brașov street, on both sides direction of travel, towards Lujerului Street and Ghencea Boulevard, and the air monitoring station where the reference values were taken was B5.
- Observation period: a time period of one hour (between 7 and 8 AM) was selected, as this represents a period of intense traffic, associated with the daily commute, during a week (between 04.04.2023 and 11.04. 2023).
- Allocation of observers: Human observers were placed on each direction of travel of the vehicles. This enabled effective traffic monitoring and identification of the vehicle type and associated pollution rate.
- Registration of vehicle type and pollution standard: observers recorded each vehicle that passed by, identifying its type (car, bus, truck, etc.) and classifying it in an European Union pollution standard (Euro 1, Euro 2 , Euro 3, etc.), based on the year of manufacture and the type of engine (petrol or diesel). A predefined list of pollution norms or a mobile application can be used to facilitate this process.
- Data collection and recording: performed at 10-minute intervals. The observers marked the number of vehicles in order to subsequently sum up the emissions for each type of emission (CO, NOx, HC, PM, SOx, etc.), for all vehicles that traveled in the monitored area. The recording of these data was done tabularly.
- Data analysis and interpretation: at the end of the observation period, the collected data were analyzed and the total emissions for one hour were made, using a calculation formula obtained with the Matlab numerical simulation program, based on information from the specialized literature [12]. The results were compared with those recorded by the air quality monitoring station to assess the compliance of the proposed calculation model.

During this research approach, several problems had to be solved related to the formulated working hypothesis - the estimation of urban pollution at time t0 based on the study of car traffic in the last hour:

- the first of these referred to the way in which some numerical data with different emission reports could be correlated: the technical norms of passenger cars provide for emissions expressed in [g/km] and the values recorded by the monitoring stations presented at time t0 in [g]. The correlation of these values was made considering that the values that arrive at the station are those in its immediate vicinity, i.e. the value given by the manufacturer per kilometer, divided by one thousand.
- the second problem that had to be analyzed was that of the different weight that each car passing through the registration area has when estimating the pollution at time t0, read at the monitoring station. Data from the specialized literature were taken into account and by numerical interpolation of the graph drawn based on them, the numerical weighting coefficients of the emission values every 10 minutes were identified. For confirmation, a study on the evolution of emissions can be carried out at exactly the same monitoring station in Romania, Bucharest, Cartierul Drumul Taberei, str. Brașov, station B5, by representing the emission values for one hour, after car traffic has been restricted in the area. The real curve of the evolution of emissions over time can thus be obtained, through its numerical treatment (symmetrization) and graphic interpolation, thus being able to calculate the real coefficients of importance for each moment of time taken into account by the present study.

The formula for calculating the pollution estimate of vehicle emissions, which circulated for one hour in the area of the monitoring station, is:

$$E_P = 0.0859X_1 \quad 0.0900X_2 \quad 0.1006\,X_3 \quad 0.1196\,X_4 \quad 0.5270X_5 \quad 0.0769X_6 \tag{1}$$

where:

- X_i represent values for emissions recorded at time points 7:10, 7:20, 7:30, 7:40, 7:50,8:00;
- the coefficients of the polynomial function were calculated using the Matlab numerical simulation program based on experimental data extracted from the specialized literature (see figure 1).

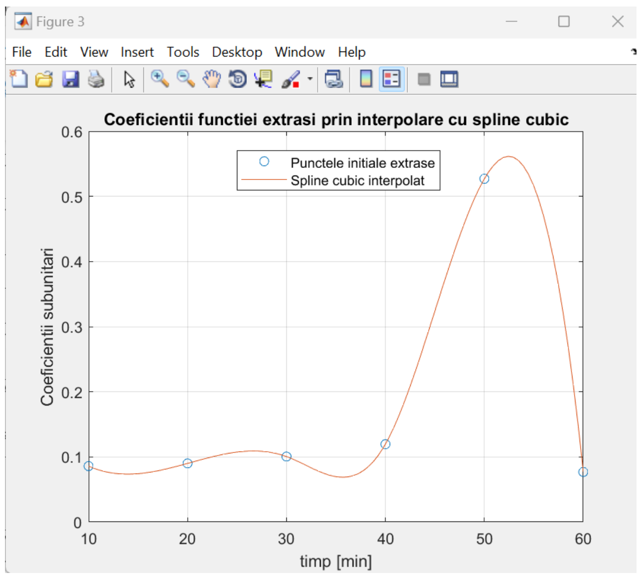

Fig. 1. Determining the coefficients of the polynomial function by numerical interpolation in Matlab

In order to be able to extract significant information about the data strings regarding the analyzed emissions, different statistics were made for their characterization (mean square deviation, median, coefficient of variation):

The standard deviation calculates the degree of dispersion of the values in a data set relative to the arithmetic mean. A large standard deviation indicates that the values are widely dispersed, while a small standard deviation suggests that the values are closer to the arithmetic mean:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{2}$$

The median of a data series gives an indication of the concentration of the data set around the mean value. A large difference between the mean and the median may suggest that there are outliers in the data set that significantly affect the mean value. These extreme values may be higher or lower than the majority of values in the data set and may be significant deviations from the general trend of the data. A large difference may also indicate a skewed distribution, where most values are concentrated in one area of the data set and extreme values are present in another area. On the other hand, a small difference between the mean and the median suggests that the data set is relatively symmetrical and that there are no significant outliers affecting the mean. This may indicate a relatively even distribution of values or a relative concentration of them around a central .

$$Me = \begin{cases} the\ middle\ value\ of\ the\ string, if\ n\ is\ odd \\ the\ sum\ of\ the\ values\ in\ the\ middle/2, if\ n\ \underline{is}\ even \end{cases} \quad (3)$$

The coefficient of variation shows the degree of variation of the data in relation to the arithmetic mean/median taken as a percentage. A large coefficient of variation suggests that the data are highly variable relative to the arithmetic mean, while a small coefficient of variation indicates that the data are relatively constant relative to the arithmetic mean.

$$Cv = \frac{\sigma}{\underline{x}} 100\ [\%] \quad (4)$$

- The Pearson correlation coefficient was calculated between the experimental data and those measured by the monitoring station:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (5)$$

where:

r - is the Pearson correlation coefficient,

n - is the number of pairs of values in the data series,

x_i, y_i - are the individual values in the two data series,

$\bar{x}$, $\bar{y}$ - are the arithmetic means of the x and y data series.

When the correlation coefficient is close to 1 or -1, this indicates a strong linear relationship between the two measurement methods, meaning that both methods provide consistent results and can be used interchangeably. A high correlation coefficient (close to 1) indicates that the two measurement methods can detect emission values in a similar way, which helps to confirm the accuracy and reliability of the experimental data. When the correlation coefficient is close to 0, this may indicate major differences between the two measurement methods and may suggest the need for further analysis to determine their causes, which may include measurement errors, incorrect calibration of measuring instruments or the influence of external factors.

3. Results and Discussion

The mean square deviation for the estimated and read emissions at the monitoring station in the reference interval are presented in table no. 1, on meaning 1:

Table 1. Mean square deviation on mean 1

Chemical elements	Estimated CO	CO read	Estimated HC+Nox	HC+Nox read	Estimated PM	PM read
Sigma	0.024129	0.073824	2.946806	7.923606	0.334557	3.991927

The mean square deviation for the estimated and read emissions at the monitoring station in the reference interval are presented in table no. 2, on meaning 2:

Table 2. Mean square deviation on mean 2

Chemical elements	Estimated CO	CO read	Estimated HC+Nox	HC+Nox read	Estimated PM	PM read
Sigma	0.012771	0.073824	3.057685	7.923606	0.378609	3.991927

It is observed that the mean square deviation for carbon monoxide estimated at 8.00 am is lower than the one calculated based on the reading from the monitoring station, respectively the values 0.0241 and 0.07382. It follows that the estimated daily values are closer to the mean than the read ones, so the data string most likely does not contain outliers caused by recording errors.

The absolute values of the arithmetic mean of the read and estimated values of pollution on the 5 days of monitoring are 0.2629 and 0.72, respectively. A difference of 0.45 g is observed, a result that may come from the pollution estimation methodology used in the calculations.

Regarding the HC and NOx emissions, a rather large difference is observed between the values of the average deviations of the two data sets, 2.9468 estimated and 7.9236 monitored. The higher value in the case of the monitoring station may come from the fact that these emissions are much more sensitive to the weather conditions of the day, wind, fog, cloud cover, etc.

For suspended dust the estimated average value is 0.0045 while the read one is 25.586, therefore it follows that practically suspended dust has other causes than the emissions from internal combustion engines and their source must be looked for in another direction.

Raporting results:

Through random measurements, it was found that the emissions of a motor vehicle remain in the monitoring area defined as a circle with a diameter of 250m, during peak traffic hours, for around 5 minutes. The measurements assumed the summation of the number of cars according to the criteria of years, the value obtained being multiplied by the emissions for each pollution standard.

The first range of the CO correlation coefficient, on the 1st sense, has a weight of -0.67, and on the 2nd sense of -0.81. The 2nd interval of the HC+Nox correlation coefficient, on direction 1, has a weight of -0.58, and on direction 2 of -0.06. On the last interval of the PM correlation coefficient, on the 1st direction, it has a weight of - 0.14, and on the 2nd one of 0.18.

4. Conclusions

The values resulting from the calculation were compared with those measured by the monitoring station over a period of 5 days. From the measurements, the recorded values are higher than the read values, most of the Pearson correlation coefficients are negative, resulting in large differences between the values. So, following this study, it was found that the visual assessment could not cope with the hundreds of cars with the specific characteristics of each brand. That is why it is recommended to mount a camera and create a software for automatic measurement of traffic and noxes in real time, which can be positioned in several locations, being a flexible, practical and cheap alternative to the actual pollution station, which comes to almost 10 thousand dollars per station, excluding the installation and its maintenance.

The mean square deviation in the two directions evolves sinusoidally, for the estimated and read chemical element CO, the values being approximately similar. In the case of estimated and read HC+Nox chemical elements, respectively estimated and read PM, the values differ significantly.

Bibliography

[1] Cohen, A.J., Brauer, M., Burnett, R., et al., Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. The Lancet, 389(10082), 1907-1918, 2017.

[2] WHO, Ambient air pollution: Health impacts. World Health Organization, 2018.

[3] Loomis, D., Grosse, Y., Lauby-Secretan, B., et al., The carcinogenicity of outdoor air pollution. The Lancet Oncology, 14(13), 1262-1263, 2013.

[4] Chen, J., Li, C., Ristovski, Z., et al, A review of biomass burning: Emissions and impacts on air quality, health and climate in China. Science of the Total Environment, 579, 1000-1034, 2013.

[5] Gariazzo, C., Silibello, C., Finardi, S., et al., A gas/aerosol air pollutants study over the urban area of Rome using a comprehensive chemical transport model. Atmospheric Environment, 61, 89-102 2016.

[6] Wang, R., Liu, Y., Xue, T., et al., Traffic-related air pollution modeling during the 2008 Beijing Olympic Games: The effects of an odd-even day traffic restriction scheme. Science of the Total Environment, 563-564, 39-50, 2016.

[7] Zheng, T., Bergin, M.H., Johnson, K.K., et al., Field evaluation of low-cost particulate matter sensors in high- and low-concentration environments. Atmospheric Measurement Techniques, 12(8), 3673-3686, 2019.

[8] Zhang, Y., Li, M., Bravo, M.A., et al., Estimating population exposure to ambient polycyclic aromatic hydrocarbon in the United States—Part I: Model development and evaluation. Environment International, 117, 145-153, 2018.

[9] Chen, H., Bai, S., Eisinger, D., et al., Review of traffic-related air pollution exposure assessment studies. Journal of Environmental Sciences, 89, 116-129, 2020.

[10] Kumar, P., & Morawska, L., Energy-pollution nexus for urban buildings. Environmental Science & Technology, 49(13), 7599-7600, 2015.

[11] Beelen, R., Raaschou-Nielsen, O., Stafoggia, M., et al., Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. The Lancet, 383(9919), 785-795, 2014.

[12] Hasenfratz, D., Saukh, O., Walser, C., et al., Pushing the spatio-temporal resolution limit of urban air pollution maps. IEEE Pervasive Computing, 11(1), 22-32, 2012.

[13] Snyder, E.G., Watkins, T.H., Solomon, P.A., et al., The changing paradigm of air pollution monitoring. Environmental Science & Technology, 47(20), 11369-11377, 2013.