ANALYSIS OF MEASURED PM2.5 CONCENTRATIONS AND METHOD FOR DETERMINING THEIR ORIGIN

Nina Ilieva, Vasil Metodiev

University of Chemical Technology and Metallurgy
8 Kliment Ohridski blvd., Sofia 1756, Bulgaria
e-mail: nilieva@uctm.edu

Received 21 September 2022
Accepted 26 May 2023

ABSTRACT

The significance of atmospheric Particulate Matter (PM) size as a determining factor for their source is of utmost importance. The analysis of PM2.5/PM10 ratios serves as a crucial indicator for particle origin. This study utilizes data obtained from measurements of PM2.5 and PM10 concentrations at the "Kamenitsa" automatic measuring station, located in the Kamenitsa district of Plovdiv, Bulgaria. The data covers a period of 6 years. A statistical approach for identifying the source of particulate matter in the air has been investigated and implemented. The findings of the applied method indicate that during winter, the primary source of particles is anthropogenic in nature.

Keywords: fine particulate matter, PM_{10}, $PM_{2.5}$, $PM_{2.5}/PM_{10}$ ratio, $PM_{2.5}$ origin, air pollution.

INTRODUCTION

One of the most common pollutants of ambient air is fine Particulate Matter (PM), which is emitted by variety of anthropogenic and natural sources. PM might be a primary pollutant, which is directly emitted by the emission sources, or secondary which is formed by atmospheric chemistry processes. PM is divided into several fractions, and for two of them - with an equivalent aerodynamic diameter for the first one up to 2.5 µm and for the second up to 10 micrometers. For those two fractions there are ambient air quality and human health protection standards set.

The size of the particles is important, as it determines how long and the distance they travel before being deposited. It also defines their entry into the human body - particles with a smaller aerodynamic diameter enter deeper into the respiratory tract. A growing body of research studies link a number of respiratory and cardiovascular diseases to prolonged exposure to high concentrations of fine fraction ($PM_{2.5}$) [1]. Another important characteristic of particulate matter as an air pollutant in terms of effects on human health is its chemical composition, which is a direct function of its origin process or activity.

Concentrations of PM are measured by Automatic Monitoring Stations (AMS) or by manual sampling. The collection, processing and analysis of such data shall determine the measures and actions of the competent authorities to improve air quality, if needed. The tasks of the analyses include answers to questions if they are of anthropogenic, natural or mixed origins, as well as the relationship with other pollutants.

The determination of particles origins is extremely important and can be done by direct chemical analyses of the samples or indirectly by statistical methods for data analysis. However, to conduct regular chemical analyses, it is necessary to analyze every sample collected, which is related to the following significant problems. According to the standard for the determination of PM_{10} and $PM_{2.5}$ by gravimetric method EN 12341:2014, each sample is the result of a 24-hour continuous sampling, resulting in a very low temporal resolution. Not all the samples collected can be used, as some of them are necessarily used to determine the content of other ambient air pollutants in the solid phase. Chemical analyses are time-consuming and expensive, especially when they have to be conducted continuously for long periods of time.

Undoubtedly, as a first step for the application of a particular type of statistical analysis, it is important to assess if the availability and completeness of the data used is adequate.

The goal of this study is to propose an indirect statistical method for determining the origin of particles in low time-resolution data from a manual sampling set with 24-hour averages of the $PM_{2.5}$ and PM_{10} concentrations measured in the territory of Plovdiv. The time period covered by the study is 01.01.2016 - 31.12.2021.

EXPERIMENTAL

On the territory of Plovdiv, ambient air quality is monitored by two monitoring stations - AMS "Trakia" and AMS "Kamenitsa". The concentration of $PM_{2.5}$ is measured by a gravimetric method by manual 24-hour sampling only in AMS "Kamenitsa", and the PM_{10} concentration is measured automatically by beta-absorption method at both locations.

A Python-developed algorithm was used to represent the entire $PM_{2.5}$ dataset in time-sequence concentration heatmap. Standard algorithms and tests were used to estimate the descriptive statistics parameters and decision tree method is applied to $PM_{2.5}$ origin determination using $PM_{2.5}/PM_{10}$ ratio as more robust to parameter.

Regarding fine particulate matter with aerodynamic diameter up to 2.5µm ($PM_{2.5}$), a reduction standard has been defined in Ordinance No12/30.07.2010 of Bulgarian Ministry of Environment and Water, which is 25 µg m^{-3} effective from 01.01.2015 and 20 µg m^{-3} effective from 01.01.2020. For the period from 01.01.2016 to 31.12.2021, the subject of the present study, hourly average data for city of Plovdiv are available in the Bulgarian National Automated Air Quality Control System (NAAQCS) only until 07:00 on 1 April 2016, and daily average data from manual sampling from the automated monitoring station AMS-Plovdiv-BG0051A-Kamenitsa site - for the entire study period. The lack of hourly average measurements for this pollutant and the availability of only daily average concentration data significantly reduces its usefulness as a carrier of air quality information. The hourly average fractional ratio of $PM_{2.5}/PM_{10}$ is a key indicator for indirectly determining the predominant type of emission sources (solid fuel domestic heating, transport) [2]. Unfortunately, this analysis cannot be performed with the available daily average data on measured concentrations.

RESULTS AND DISCUSSION

The values of the annual mean concentration of $PM_{2.5}$ are presented in Table 1 and Fig. 1. The data indicates a steady improvement trend in air quality, and the impact of pandemic situation in 2019, 2020 and 2021 should be considered.

A heatmap of the values of the daily mean concentration of $PM_{2.5}$ µg m^{-3} is presented in Fig. 2. It clearly shows that high concentrations of $PM_{2.5}$ are observed during the winter period and during the

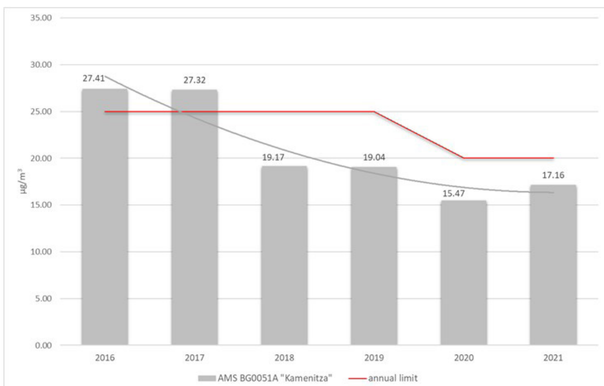

Fig. 1. Annual mean concentration of PM2.5 µg m^{-3} measured in AMS "Kamenitsa" for the period 2016 - 2021.

Table 1. Annual mean concentration of $PM_{2.5}$ in Plovdiv, µg/m³.

Data source	AMS	2016	2017	2018	2019	2020	2021
	limit	25	25	25	25	20	20
NAAQCS	BG0051A-Kamenitsa	27.41	27.32	19.17	19.04	15.47	17.16

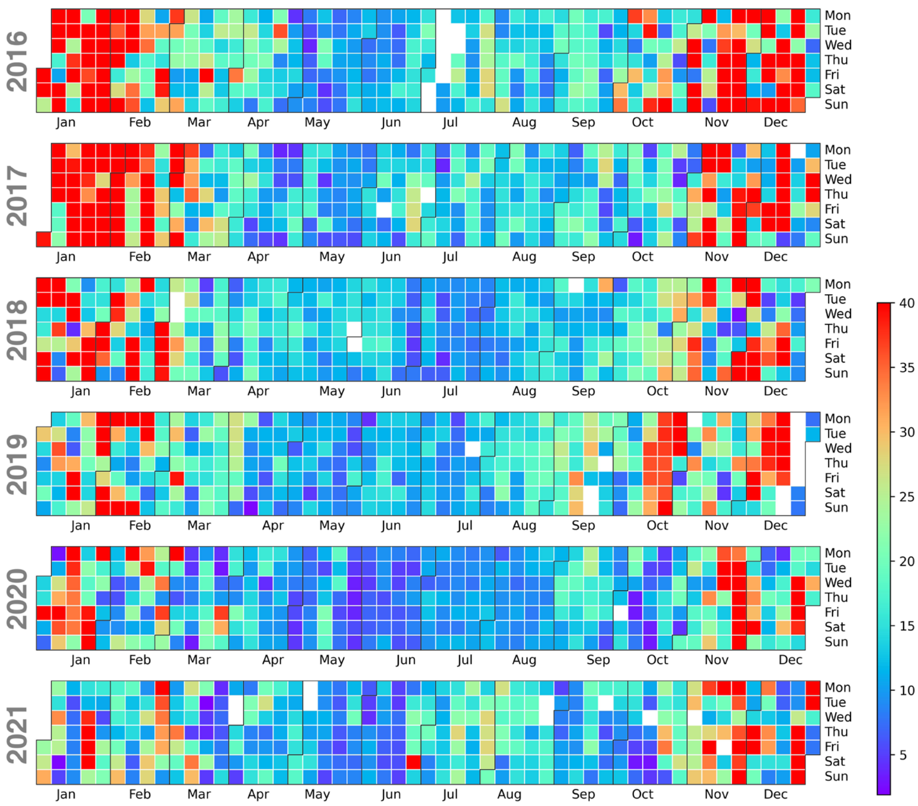

Fig. 2. Heatmap of the measured daily average concentration of PM2.5 μg m^{-3} in AMS "Kamenitsa" for the period 2016 - 2021.

months of January, February, November, and December. This suggests an anthropogenic origin of the particles because of combustion processes for domestic heating. The origin of fine particulate matter could be determined with relatively high probability by chemical analysis of air samples collected on filters. The disadvantage of this method is that it is an expensive and slow process that requires preservation of the samples and analysis of the particles collected on the filters in an analytical laboratory. This means that this method cannot be used to determine the origin of particles over time if the samples are not available and they cannot be stored for too long.

The $PM_{2.5}/PM_{10}$ ratio can provide information on the particle size distribution of the aerosol, which could be used to classify aerosol types by setting threshold values. One of the variables that finds application is the probability distribution function of the $PM_{2.5}/PM_{10}$ ratios [3, 4].

This ratio is a particularly important indicator because the finer fraction is usually due to combustion processes, so values tending towards 1 indicate a predominant influence of anthropogenic and natural (forest fires for example) combustion processes. This ratio is therefore very often used in determining the origin of particles without chemical analysis and supports decision making.

In addition to the distribution function (unimodal, bimodal, or multimodal), two other key parameters are the skewness and the kurtosis coefficient. The skewness coefficient is the third standardized moment defined as:

$$k_{\text{skewness}} = \tilde{\mu}_3 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} =$$

$$= \frac{E[(X-\mu)^3]}{(E[X-\mu^2])^{3/2}} = \frac{k_3}{k_2^{3/2}} \quad (1)$$

Skewness, denoted by $k_{skewness}$, quantifies the extent and direction of deviation from symmetry in a statistical data distribution. It is calculated using various statistical parameters, including the mean (μ), standard deviation (σ), expectation operator (E), third central moment (μ_3), and t-th cumulants (k_t). Skewness provides information about the asymmetry of the distribution and the degree of distortion in the data's spread.

The coefficient of kurtosis ($k_{kurtosis}$) represents the maximum value of the distribution function at the mean. In general, $k_{kurtosis} > 0$ represents a distribution with a distinct peak, and $k_{kurtosis} < 0$ indicates a flat distribution without distinct outliers.

$$k_{kurtosis} = E\left[\left(\frac{x-\mu}{\sigma}\right)^4\right] = \frac{E\left[\left(\frac{x-\mu}{\sigma}\right)^4\right]}{(E[(x-\mu)^2])^2} = \frac{\mu_4}{\sigma^4} \qquad (2)$$

where μ_4 is the fourth central moment and σ is the standard deviation.

Aerosol Type Classification Method

In [5] a novel method for classifying aerosol types based on spatiotemporal analysis of PM2.5/PM10 ratios is proposed. The classification process utilizes the probability density function (PDF) of PM2.5/PM10 ratios, as well as the skewness and kurtosis of related variables. Skewness is a statistical measure that indicates the direction and degree of asymmetry in the distribution of data, providing a numerical representation of data distribution asymmetry. Skewness can be categorized into three types: normal distribution ($k_{skewness} = 0$), right-skewed distribution ($k_{skewness} > 0$), and left-skewed distribution ($k_{skewness} < 0$). On the other hand, kurtosis is a descriptive statistic that represents the peak value of the PDF at the mean. A positive kurtosis value ($k_{kurtosis} > 0$) signifies a peak distribution, while a negative kurtosis value ($k_{kurtosis} < 0$) indicates a flatter distribution.

Fig. 3 depicts the specific classification methodologies. Utilizing the PM2.5/PM10 [5] ratio analysis approach, a comprehensive examination of the statistical distribution parameters enables the identification of the source of PM2.5-related ambient air pollution, even in scenarios involving incomplete or inadequately temporally resolved data (e.g., daily average data as opposed to hourly average data).

Fig. 4 shows a histogram of the distribution of the $PM_{2.5}/PM_{10}$ ratio calculated by the 24-hour concentrations for the whole study period (2152 pairs of data for the period 2016 - 2021). The distribution parameters are as follows: mean 0.53 with 95 % confidence interval (0.526;0537), coefficient of skewness ($k_{skewness}$) 0.767 and

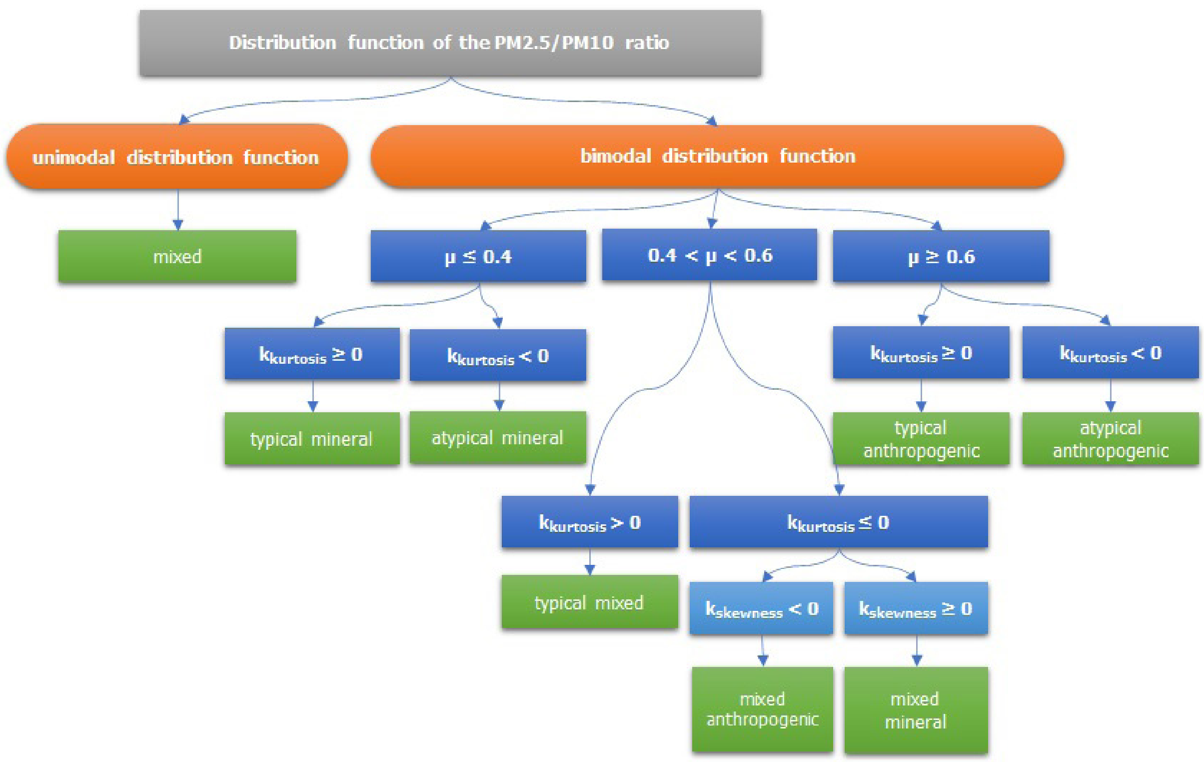

Fig. 3. Determination method for fine particulate matter composition by the ratio of $PM_{2.5}$/ PM_{10}.

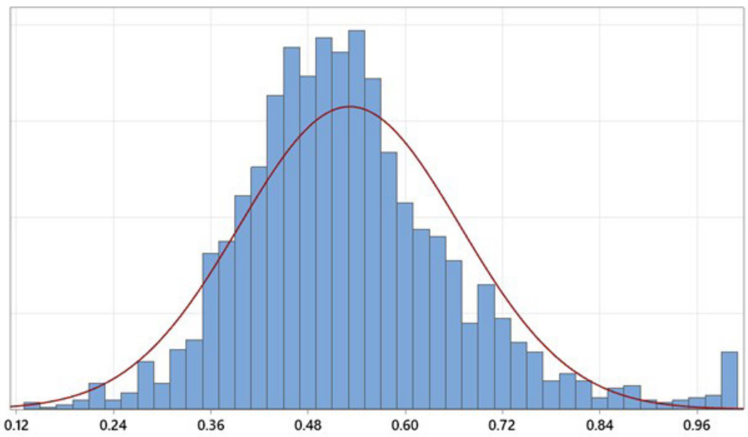

Fig. 4. Distribution histogram of PM2.5/PM10 ratio measured in AMS Kamenitza for the period 2016 - 2021.

Table 2. Descriptive statistics for used daily averages PM2.5, PM10, PM2.5/PM10 data set from AMS BG0051A "Kamenitsa" for period from 01.01.2016 to 31.12.2021.

Statistical parameter	$PM_{2.5}$	PM_{10}	$PM_{2.5}/PM_{10}$
N	2152	2152	2152
Mean	21.03	37.55	0.53
Median	15.00	29.14	0.52
StDev	21.29	28.82	0.14
Min	1.90	6.23	0.13
Max	253.00	323.94	1.00
Skewness	4.12	3.08	0.77
Kurtosis	25.05	14.59	1.45
AD Normal distribution $P < 0.005$	222.95	159.90	18.68
AD Box-Cox Transformation $P < 0.005$	3.397 $\lambda = -0.271$	1.314 $\lambda = -0.338$	6.73 $\lambda = 0.345$

coefficient of excess ($k_{kurtosis}$) 1.44. To apply the method mentioned above, it is necessary to initially determine the type of distribution function. Some statistical parameters of descriptive statistics are calculated and presented in Table 2. As can be seen from the Anderson-Darling test results, shown in Table 2, the probability distribution of $PM_{2.5}$ and PM_{10} is not normal and is highly right (positive) skewed. This is shown in the study of Munir [6], while $PM_{2.5}/PM_{10}$ ratio distribution seems to be normal or close to it.

For the whole period studied, it can be assumed to be bimodal due to the superimposition of the influence of alternating winter and summer periods, as well as the mixing of effects in different years. It should be noted here that the two fractions are measured by different methods (gravimetric and beta-absorption), which leads to an accumulation of combined measurement errors from the two methods, and this in turn leads to values of the ratio of $PM_{2.5}/PM_{10}$ above 1. Since this is physically impossible, all values in the sample above 1 are equated to the maximum value, which appears as a second hump in the distribution function.

Taking further into account the mathematical expectation and the coefficients of excess and asymmetry, the method assumes a typical mixed composition of fine particulate matter. This result is not unexpected given the location and classification of the AMS "Kamenitsa" site as urban background and confirms the expectation that it reflects the influences of all sources of fine particulate matter - including those of natural and anthropogenic origin.

The variation of the parameters of the distribution function by month, presented in Table 3 and Fig. 5, is a matter of interest. Here the particle composition is determined for each month by the proposed method depending on the parameters obtained. As can be seen in Fig. 5, this distribution is quite different for the months of January and June, aggregated for the years 2016 to 2021.

As can be seen from the last column of the table, the anthropogenic influence dominates during the month of January, while during the other months of the year this composition is classified as typical mixed, except for the August when the particles are predominantly of mixed mineral origin. These observations are consistent with the fine particulate matter generating processes within the city, as during the coldest periods of the year, of the anthropogenic sources, the most major influence is the combustion of domestic heating fuels [7].

Table 3. Variation of the distribution function parameters of the PM2.5/ PM10 ratio measured in the AMS "Kamenitsa" for the period 2016 - 2021.

Month	I. PM2.5/PM10 mean value	II. Coefficient of excess $k_{kurtosis}$	III. Coefficient of skewness $k_{skewness}$	Suggested origin of the ambient air pollution with $PM_{2.5}$	
January	0.61	0.3	0.6		typical anthropogenic
February	0.59	0.7	0.6		typical mixed
March	0.53	1.1	0.3		typical mixed
April	0.49	2.9	1.0		typical mixed
May	0.48	3.6	0.9		typical mixed
June	0.48	3.5	1.4		typical mixed
July	0.55	1.7	1.3		typical mixed
August	0.52	0.0	0.4		mixed mineral
September	0.52	0.7	0.6		typical mixed
October	0.49	1.7	0.3		typical mixed
November	0.57	0.9	0.1		typical mixed
December	0.57	0.5	0.5		typical mixed

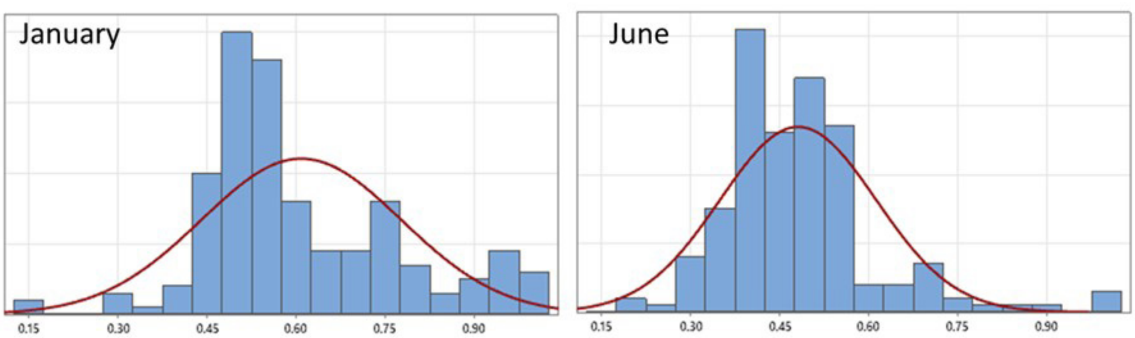

Fig. 5. Histograms of the distribution of the measured $PM_{2.5}/PM_{10}$ ratio for January and June in the AMS "Kamenitsa" for the period 2016 - 2021.

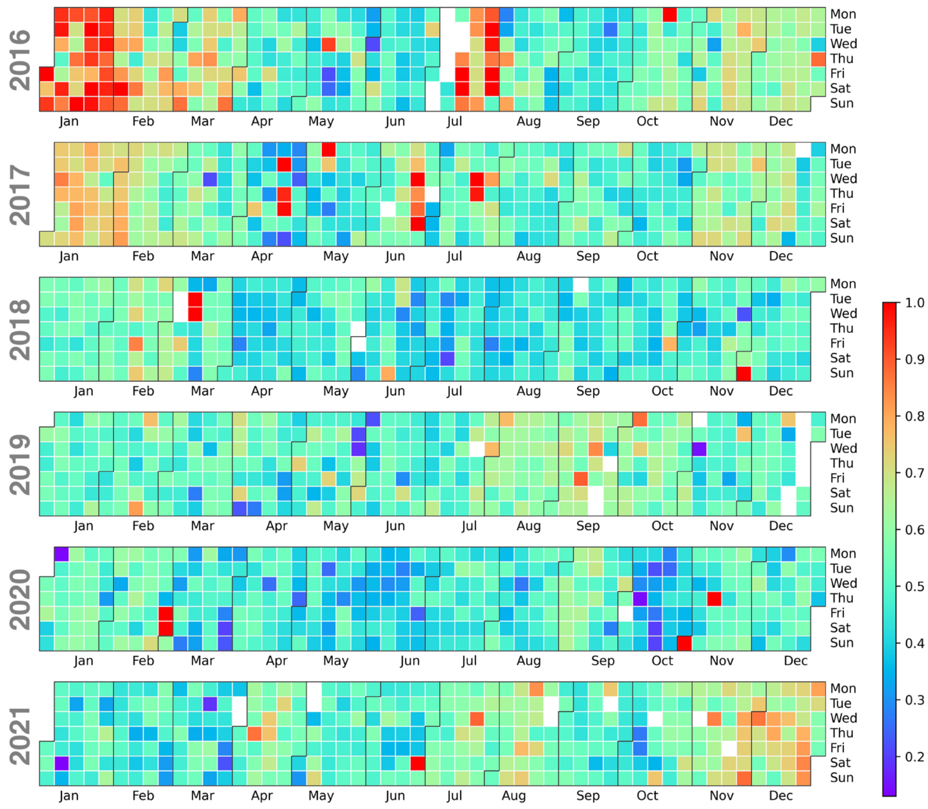

Fig. 6. Map of the measured $PM_{2.5}$/ PM_{10} ratio of the daily average concentrations in AMS "Kamenitsa" for the 2016 - 2021 period.

Fig. 6 presents a map of the measured ratio of $PM_{2.5}$/ PM_{10} to the daily mean concentrations at AMS "Kamenitsa" for the 2016 - 2021 period. When comparing the graphs in Figs. 2 and 6, there is a correspondence of the periods with high values for the concentration of $PM_{2.5}$ and the ratio of $PM_{2.5}$/ PM_{10}.

CONCLUSIONS

Although the average annual concentration of $PM_{2.5}$ is above the norm only the first two years of the study period, careful analysis and monitoring of the levels of this pollutant is necessary and recommended. High concentration values are systematically observed in the months of January, November and December. A statistical method for determining the origin of particulate matter in the air has been studied and applied. It was found to be suitable for analyzing data from a 24-hour manual sample set - low temporal resolution data. According to the method used, the particles generated in winter have predominant anthropogenic origin.

REFERENCES

1. R.M. Harrison, The health effects of air pollution, in Pollution: Causes, Effects and Control, 4 ed., Royal Society of Chemistry, 2001.
2. A.L. de Jesus, M. Rahman, M. Mazaheri, H. Thompson, L.D. Knibbs, C. Jeong, G. Evans, W. Nei, A. Ding, L. Qiao, L. Li, H. Portin, J.V. Niemi, H. Timonen, K. Luoma, T. Petäjä, M. Kulmala, M. Kowalski, A. Peters, J. Cyrys, L. Ferrero, M. Manigrasso, P. Avino, G. Buonano, C. Reche, X. Querol, D. Beddows, R.M. Harrison, M.H. Sowlat, C. Sioutas, L. Morawska, Ultrafine particles and PM2.5 in the air of cities around the world: Are they representative of each other? Environment International, 129, 2019, 118-135.
3. G. Mishra, K. Ghosh, A. K. Dwivedi, M. Kumar, S. Kumar, S. Chintalapati, S. Tripathi, An application of probability density function for the analysis of PM2.5 concentration during the COVID-19 lockdown period, Science of The Total Environment, 782, 2021.

4. X. Wang, R.J. Chen, B.H. Chen, H.D. Kan, Application of Statistical Distribution of PM10 Concentration in Air Quality Management in 5 Representative Cities of China, Biomedical and Environmental Sciences, 26, 8, 2013, 638-646.
5. H. Fan, C. Zhao, Y. Yang, X. Yang, Spatio-Temporal Variations of the PM2.5/PM10 Ratios and Its Application to Air Pollution Type Classification in China, Frontiers in Environmental Science, 2021.
6. S. Munir, Analysing Temporal Trends in the Ratios of PM2.5/PM10 in the UK, Aerosol and Air Quality Research, 17, 1, 2017, 34-48.
7. S.J. Jędruszkiewicz, B. Czernecki, M. Marosz, The variability of PM10 and PM2.5 concentrations in selected Polish agglomerations: the role of meteorological conditions, International Journal of Environmental Health Research, 27, 6, 2017, 441-462.