Environment Protection Engineering

Vol. 40 2014 No. 2

DOI: 10.5277/epe140204

GRZEGORZ MAJEWSKI[1], PIOTR O. CZECHOWSKI[2],
ARTUR BADYDA[3], ANDRZEJ BRANDYK[1]

EFFECT OF AIR POLLUTION ON VISIBILITY IN URBAN CONDITIONS. WARSAW CASE STUDY

The influence of air pollutants on visibility in Warsaw Agglomeration has been investigated. Following pollutants were considered: PM_{10}, SO_2, NO_2 and O_3, while meteorological parameters included: air temperatures (mean, minimum, maximum), solar radiation, relative air humidity, rainfall rates and wind speed. Initial analyses were performed with the use of principal component analysis (PCA). In next stages, the logistic regression (LR), the analysis of variance (ANOVA), one-way classification and a model path of generalized regression models (GRM) were applied. PCA analysis showed that in the cold season the visibility index depends on PM_{10}, SO_2, NO_2 and the temperatures: T, $T_{\min}$ and $T_{\max}$. In the warm season, the index of visibility is mostly shaped by four elements: O_3, T, $T_{\max}$ and solar radiation. Logistic regression model indicated that in the warm season only two variables are significantly related to visibility: PM_{10} and relative humidity of air. Regularities in the cold season shown by the LR correspond with the conclusions from the PCA. Among meteorological conditions, the most important is air temperature, but only $T_{\max}$ preserves the same direction of influence as the one pointed by the PCA model.

1. INTRODUCTION

Effect of air pollution from anthropogenic sources on human health and visibility has been examined for decades. Many studies throughout the world were conducted not only to estimate benefits for human health resulting from reducing emission of air pollutants but also to identify how air pollutants decrease visibility [1–6].

Scientific research reveals that the size, chemical composition and mass concentration of particles suspended in air considerably affect visibility [5, 7, 8]. The distor-

[1]Warsaw University of Life Sciences, Division of Meteorology and Climatology, 02-776 Warsaw, ul. Nowoursynowska 166, Poland, corresponding author G. Majewski, e-mail: grzegorz_majewski@sggw.pl

[2]Gdynia Maritime University, Information Systems Department, 81-225 Gdynia, ul. Morska 83, Poland.

[3]Warsaw University of Technology, Environmental Engineering Faculty, 00-653 Warszawa, ul. Nowowiejska 20, Poland.

tions of visibility occur as a result of such processes as light absorption by water vapor and scattering by aerosol particles [9–11]. A significant role in shaping visibility is played also by meteorological conditions, particularly air relative humidity [12–14]. Some of the polluting particles serve, because of their chemical composition, as a good condensation medium for nuclei. In such cases, condensation may occur even in air which is not saturated with water vapour. Such a process, called aerosol melting, results in reduced visibility in urban areas when the relative humidity of air reaches 70–80% [14].

Air quality monitoring in Warsaw has been carried out for several dozen years, however only as late as in 2004 it was adjusted to comply with the regulations and requirements of the EU. Since the monitoring system has been in operation, air quality can be recorded in order to warn the community of high levels of pollution as well as do research on the influence of air pollutants on human health [15–17]. Because of the lack of scientific research concerning factors influencing visibility in Warsaw, the authors of this paper have sough solutions to the question how air pollutants affect this parameter in various meteorological conditions.

Due to the fact that improving visibility in urban conditions requires understanding which atmospheric pollutants reduce visibility and what is their source, the undertaken research is considered to be important and justifiable.

2. MATERIALS AND METHODS

The study was based on the results of measurement, performed at the air quality monitoring station of MzWarszUrsynów (λ_E = 21°02′, φ_N = 52°09′), located in the southern part of Warsaw (Fig. 1). At the station, concentrations of various pollutants such as SO_2, NO_2, O_3 and PM_{10} – particulate matter are recorded. Measurement data considered in this paper relate to the period 2008–2010. The investigated measurement site of MzWarszUrsynów is representative of the city background and well characterises the imission of pollutants within residential areas exposed to the effect of traffic, municipal and industrial emission. Meteorological data come from the meteorological station of Ursynów-SGGW (λ_E = 21°02′, φ_N = 52°09′), which is a part of the Division of Meteorology and Climatology of WULS. The research material comprised air temperatures: average daily (T), maximum (T_{max}) and minimum (T_{min}), solar radiation intensity (Rad), the relative humidity of air (f), the amount of atmospheric precipitation (Prec) and wind velocity (v). The measurements at that station are taken in compliance with the instructions for the network of state stations of IMGW. The visibility data were collected at the only station which conducts such measurements, i.e. from the station of Okęcie (λ_E = 20°59′, φ_N 52°09′). The distance between these two stations is ca. 6 km.

The aim of the research was to identify the main factors determining visibility in the time domain and the frequency domain. It applies both to imission and meteorological factors and also takes into consideration the seasonal factor from a model perspective through identifying the relationships of the past measurement results of exogenous variables and a series of visibility results as a function of an exogenous variable.

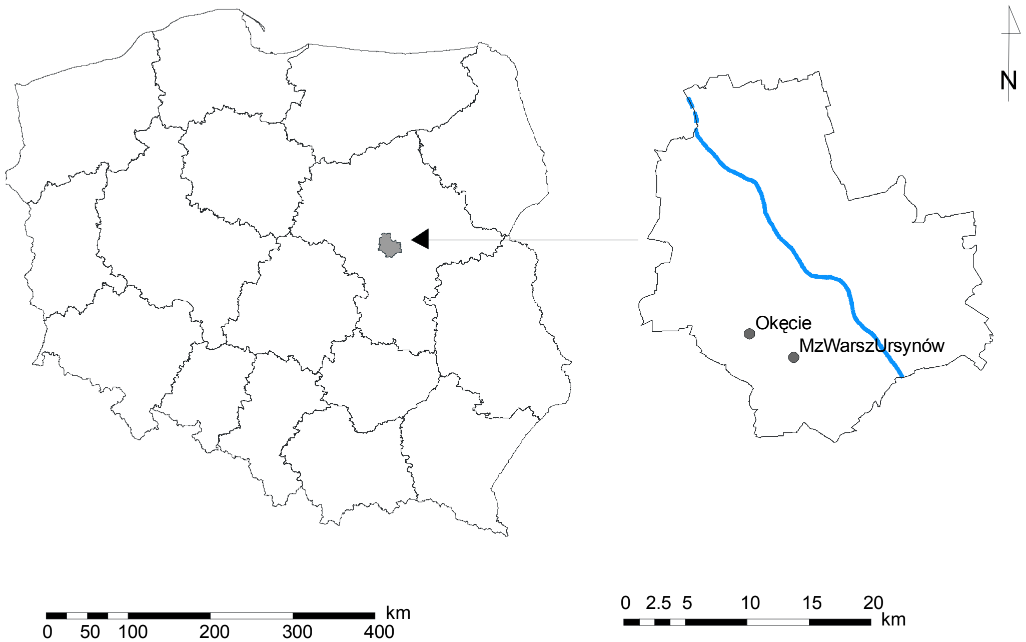

Fig. 1. Location of monitoring stations at Warsaw city, Poland

Achieving the aim of the research was possible after completion of the partial research goals by using appropriate stochastic and exploration models in compliance with a proper order of investigation so as to finally solve the research questions.

The initial stage was to precisely identify mutual relationships of factors in the frequency domain. It was important both to quantify the relationships and to estimate the importance of new synthetic variables, not specifying the direction of a cause-and-effect relationship at this point. Such an approach enables the researcher to temporarily answer the question about the occurrence of strongly nonlinear relationships, including, e.g. seasonality. The study used Hotteling's model of the principal component analysis (PCA), of which first principal component, called the phenomenon index, can be understood as an equivalent to the first harmonic of Fourier's model and the next harmonics, mutually orthogonal, enable addressing the initially proposed research thesis, e.g. on the occurrence of non-linearity and seasonality.

The consideration of the phenomenon index in the Principal Component Analysis denotes the component with the highest eigenvalue of the matrix of correlation, accounting for the largest part of the total variance of the system and related with the highest number of real factors. It is the component which maximizes variance [18–20].

Component loadings in the PCA model are not the same as the regression coefficients and must be understood differently. The model of multiple regression is essentially an internally a linear model which possesses a number of additional assumptions which are difficult to be satisfied in practice such as: stationarity of a random component or homoscedasticity of its variance.

The next, partial research goal was related to the importance of identifying the cause-and-effect relationship from the stochastic perspective, which cannot be overestimated. Out of the family of regression models, this research work employs the logistic regression (LR) which has an advantage over linear and linearised models as it does not need to satisfy any necessary assumptions which result from the parameter estimation method. Moreover, it allows the results to assume values of a given range, assess probability of assuming a high/low visibility value, i.e. assess the risk of a decrease in visibility and chances for "improvement", and additionally enables clear, practical interpretation of the estimated parameters [19].

The models which are the ultimate confirmation of preliminary dependences and mutual relationships are following: analysis of variance (ANOVA) with one-way classification and a generalised regression model path. However, it is worth noting that ANOVA, being at the most a two-dimensional model, in practice is treated as complementary to generalised regression model GRM.

GRM, as opposed to the previously mentioned models, is not a model in the strict sense, but an authentication path which enables conducting step-wise regression comprehensively. It allows using general, linear model methods which enable building models for the systems containing effects with many degrees of freedom for qualitative predictors, as well as for the systems containing effects with single degrees of freedom for continuous predictors [21]. Therefore, GRM enables precise identification of the factors in the time domain, including seasonality from a one-dimensional perspective (e.g. year) as well as interactions between periods (e.g. year–season–month) from a multi-dimensional perspective. It was assumed that this model ultimately confirms the occurrence of dependence between visibility and the considered exogenous factors. The GRM would also prove the earlier identified relationships in the frequency domain (PCA and LR) and the time domain.

3. RESULTS

Deteriorating air quality continues to be one of the most evident impacts of the formation of large city agglomerations. The degradation of visibility is readily evident to the inhabitants, while air pollution episodes involving primary and secondary pollutant species gain national notoriety [15]. For instance the smog events, caused by industrial coal combustion or by automobile emissions combined with photochemical reactions, exert a significant influence on the cleanliness of the sky. Both types of smog cause a significant reduction in visibility, nonetheless, when no such events happen, higher levels of air pollution in large conurbations used to contribute to certain visibility deterioration. It is still required in Warsaw Agglomeration to comply with the EU requirements concerning air pollution, first of all: PM_{10}, NO_x and O_3, which justifies the search of dependences between parameters of air quality and visibility.

3.1. DATA OVERVIEW

The area where the measurement station is located (MzWarszUrsynów) was characterized in many previous studies as having an average degree of air pollution [19, 22, 23]. In the analysed period 2008–2010 at the station of MzWarszUrsynów mean daily values of NO_2 concentration ranged from 21.4 $\mu g \cdot m^{-3}$ (2010) to 23.4 $\mu g \cdot m^{-3}$ (2009), which explains from 53.6% to 58.4% of the binding permissible limit value (40 $\mu g \cdot m^{-3}$, Table 1).

Table 1

Mean concentrations of the analysed pollutants [$\mu g \cdot m^{-3}$] at the MzWarszUrsynów station in 2008–2010

Pollutant	Characteristics	Year		
		2008	2009	2010
NO_2	Yearly mean	23.2	23.4	21.4
	Warm half-year (IV–IX)	21.4		
	Cold half-year (X–III)	24.0		
	S99.8[1] (1-h)	104.5	117.4	91.2
SO_2	Yearly mean	7.6	8.7	8.1
	Warm half-year (IV–IX)	5.8		
	Cold half-year (X–III)	10.4		
O_3	Yearly mean	46.9	43.5	47.1
	Warm half-year (IV–IX)	57.4		
	Cold half-year (X–III)	34.0		
	Number of days with the exceeded target level	22	12	15
PM_{10}	Yearly mean	28.0	33.8	35.2
	Warm half-year (IV–IX)	26.7		
	Cold half-year (X–III)	37.7		
	S90.4[1] (24-h)	46.0	56.4	63.2

[1]Percentile values of 1-hour (24-hour) concentrations).

Mean concentration of NO_2 in all years was slightly higher in the cold half-year than in the warm one which is connected with low emission diversity of this pollutant in the course of the year within the research area. It was also found that a significant share of NO_2 emission resulted from mobile sources [19, 22]. Relating the NO_2 concentration with the values of visibility it can be found that the increase of visibility to the value higher than 10.5 km corresponds with lower average NO_2 concentration record, both in the cold and the warm half-year: 30% and 12%, respectively.

Mean yearly concentration of SO_2 ranged between 7.6 $\mu g\cdot m^{-3}$ (2008) and 8.7 $\mu g\cdot m^{-3}$ (2010) corresponding to 37.8% to 43.4%, respectively, of the binding permissible limit value (20 $\mu g\cdot m^{-3}$) (Table 1). SO_2 is characterised by a more distinct seasonal structure. Big differences in concentrations in the two seasons are caused by the influence of close, local sources of the analysed pollutant around the measurement station in the winter season. As it has been shown in many research studies, within the research area there occur local boiler plants and individual domestic furnaces [19, 24]. Similar to NO_2 concentrations, in all cases when visibility exceeded 10.5 km, the averaged SO_2 concentrations were lower both in the cold and the warm half-year: 35% and 20%, respectively.

Mean yearly values of the concentration of particulate matter PM_{10} ranged from 28.0 $\mu g\cdot m^{-3}$ to 35.2 $\mu g\cdot m^{-3}$, corresponding to 70.0% and 88.1% of the permissible limit value (40 $\mu g\cdot m^{-3}$), respectively (Table 1). Similarly as concentrations of SO_2, much higher concentrations of PM_{10} occurred in the cold half-year. Differences in seasonal concentrations are caused by higher winter emission of suspended dust whose origin is fuel combustion from low emission sources, occurring along with relatively worse conditions of ventilation and self-cleaning of air, which take place in the cold season of the year [24]. Mean daily permissible limit value of PM_{10} was exceeded from 24 (2008) to 55 times (2009), when the frequency of exceeding permissible values amounts to 35 days in a year. Initial analyses of relationships between PM_{10} and visibility prove that the increase of the measured visibility to values higher than 10.5 km is associated with decreased PM_{10} concentrations, both in the cold and the warm half-year by 46% and 22%, respectively.

Mean yearly concentrations of O_3 ranged from 43.5 $\mu g\cdot m^{-3}$ to 47.1 $\mu g\cdot m^{-3}$ (Table 1). The target level of 8-hour concentration for ozone amounts to 120 $\mu g\cdot m^{-3}$ and may be exceeded for not more than 25 days a year. The biggest number of days with the exceeded target level occurred in 2008. In the case of this pollutant, a steady trend of changes within the research area has not been observed. Variability of the concentration of O_3, which is a secondary pollutant, results mostly from differences in weather conditions occurring in subsequent years (solar radiation intensity, air temperature), and also from a share of O_3 precursors of (e.g. nitrogen oxides, hydrocarbons and other pollutants participating in the process of formation of this pollutant) in the atmospheric air [22] Relating ozone concentrations to visibility, a contrary situation can be observed, in comparison to other pollutants. As visibility increases, there occur also higher recorded ozone concentrations. In the cold half-year, the visibility exceeding 10.5 km. is associated with ozone concentrations higher by 31% on the average, while in the warm half-year they are higher by 9% on the average.

Mean daily visibility values in 2008–2010 at the station of Okęcie ranged from 1.6 km to more than 12 km. Visibility of the range of 10–12 km occurred most often, which accounts for approximately 62% of all measurement results. Visibility below 2 km and above 12 km constituted ca. 0.5% of all measurements.

3.2. ANOVA ANALYSIS

The aim of preliminary calculations was to answer the question about the occurrence of seasonality in the results of visibility measurements. The biggest statistically significant differences in the examined years were identified between warm and cold seasons (Fig. 2, Table 2). In the table, StatisticsKW means the Kruskal–Wallis statistics (Kruskal–Walllis test); *p*(F) – *p* level in the Fisher–Snedecor test, *p*(KW) – *p* level in the Kruskal–Wallis test; *p* – significance level.

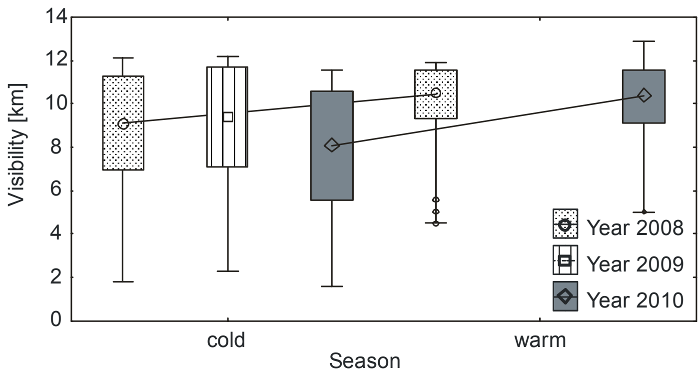

Fig. 2. Presence of seasonality in the measurements of visibility (grouped by Season; categories grouped by Year; points: mean values; boxes: standard deviations; ends: minimum and maximum values)

Table 2

ANOVA for the seasonality factor

Year	Variable	Statistics F(1;363)	*p*(F)	Statistics KW	*p*(KW)
2008	Visibility, km	55.9663	0.0000	48.9995	0.0000
2009		151.0327		99.5671	
2010		121.4233		97.9483	

The differences, with the use of ANOVA, were observed both between the seasons of each single year and between the seasons in all the analysed years. In this model, each of the examined factors proved to be statistically significant, and, in the warm season visibility was generally greater than in the cold one. The seasonality factor strongly influenced the phenomenon, which was considered during further analyses.

3.3. PRINCIPAL COMPONENT ANALYSIS (PCA)

During the PCA, identification the influence of seasonality was taken into account. Model estimation was carried out separately for the warm and the cold season. In the cold season the index of the phenomenon comprised relationships of three pollutants: PM_{10}, SO_2, NO_2 and three temperatures: T, T_{min} and T_{max} (Table 3). With the performed analysis, it was demonstrated that the loadings of temperature and concentrations of above mentioned pollutants on the first principal component were opposite, and the loading of visibility was negative. Thus, it could be concluded that the pollutant concentrations would increase along with a decrease in air temperature, and visibility should decrease along with an increase in pollution and a fall in air temperature.

The first principal component accounts for more than 35% of the total variability of the system. It should be emphasised that the research, due to the lack of data, did not consider mixed layer depth which plays a key role in dispersion of pollutants on a local scale [25]. Taking this variable into consideration could improve the precision of the model and reduce dimensionality.

Table 3

Identification of relationships between various factors (PCA model, R correlations), the cold season

Parameter	Component			
	1	2	3	4
PM_{10}, µg·m^{-3}	0.75	0.00	–0.51	0.09
SO_2, µg·m^{-3}	0.71	0.20	–0.33	0.20
NO_2, µg·m^{-3}	0.70	–0.12	–0.48	0.17
O_3, µg·m^{-3}	–0.41	0.75	0.04	0.21
T, °C	–0.75	–0.48	–0.41	0.06
T_{min}, °C	–0.74	–0.57	–0.25	–0.02
T_{max}, °C	–0.75	–0.38	–0.50	0.07
f, %	0.29	–0.69	0.50	–0.08
Precipitation, mm	–0.14	–0.27	0.41	0.84
Radiation, W·m^{-2}	–0.38	0.46	–0.63	0.14
Visibility, km	–0.62	0.51	–0.04	–0.08
V, m·s^{-1}	–0.44	0.39	0.55	0.03
Output value	4.23	2.48	2.21	0.87
Share	0.35	0.21	0.18	0.07

The second principal component in the cold season was characterized by a high, positive loading of ozone (O_3, $R = 0.75$) which might point at a considerable influence of this pollutant on visibility, explaining 21% of visibility common variance. Atmospheric precipitation displayed a specific character through its significant loading on the fourth component ($R = 0.84$) and most likely decreasing the visibility (–0.1) whose low value points to the absence of direct effect. The first three components, contrary to the fourth one, would considerably shape the visibility in the cold season of the year due to estimated share in the common variance (Table 3). On the other hand, a fully reliable analysis undoubtedly needs further elaboration of scatterplots in new coordinates (principal components) to evidence probably strong non-linearity of their correlation with visibility. In total, the first four components, identified on the basis of Cattel's criterion, accounted for approx. 82% of total variability. The achieved results prove usefulness and correctness of the PCA model for analysing cold season visibility variation in the Warsaw Agglomeration.

Table 4

Identification of relationships between factors (PCA model, R correlations), the warm season

	Component:				
	1	2	3	4	5
PM_{10}, $\mu g \cdot m^{-3}$	–0.09	–0.83	–0.24	0.25	0.06
SO_2, $\mu g \cdot m^{-3}$	–0.03	–0.68	–0.09	0.44	–0.12
NO_2, $\mu g \cdot m^{-3}$	0.09	–0.78	–0.27	–0.04	–0.25
O_3, $\mu g \cdot m^{-3}$	–0.71	–0.05	0.43	0.24	0.24
T, °C	–0.85	0.23	–0.44	0.00	0.03
T_{min}, °C	–0.64	0.44	–0.56	–0.02	0.00
T_{max}, °C	–0.91	0.07	–0.35	0.02	0.01
f, %	0.68	0.40	–0.53	–0.02	–0.05
Precipitation, mm	0.17	0.44	–0.27	0.68	–0.12
Radiation, $W \cdot m^{-2}$	–0.81	–0.18	0.34	–0.06	0.05
Visibility, km	–0.39	0.29	0.39	0.01	–0.76
V, $m \cdot s^{-1}$	0.22	0.50	0.42	0.39	0.15
Output value	3.82	2.73	1.75	0.94	0.77
Share	0.32	0.23	0.15	0.08	0.06

Slightly different relationships occur in the warm period (Table 4). Visibility is most strongly related to the fifth principal component and marks its presence in particular components with various intensity. The fifth principal component is most significantly correlated with O_3 and NO_2, pointing at presumably considerable influence of those pollutants on visibility. The index of the phenomenon, represented by the first principal component, comprises the relationship of four measurements: O_3, T, T_{max} and Rad. They have negative loadings on this component, as well as the visibility. It can be supposed then that visibility increases upon increasing values of those four measurements. The first component accounts for about 32% of variability. The second most important relationship is formed by measurements of PM_{10} and NO_2 which have negative loadings on the second principal component, and might reduce visibility (R = 0.28). In this component, one can observe a significant but lower than 0.7 contribution of SO_2 (–0.68) and O_3 (–0.05) which proves a weaker influence of these pollutants on visibility than in the cold season. The remaining components do not allow simple identification of relationships. However, they may point at further, less significant effect on visibility. The first five components account in total for slightly more variances than in the cold season (about 84% versus 82%), but similarly to the cold season the first three components shape the visibility most significantly due to their share in the common variance (Table 4).

PCA confirmed the negative impact of PM_{10} and NO_2 on visibility and also a significant impact of sulfate from SO_2 emissions. Basing on simulated results from an empirical model, the maximum influence on visibility is exerted by PM_{10}, not excluding the role of ozone and its precursors. However, a 50% reduction in ozone concentration with all other parameters unchanged leads to increase in visibility by just 0.03 km, while a 50% reduction in PM_{10} with all other parameters unchanged leads to increase in visibility by 1.1 km, demonstrating that a reduction in PM_{10} concentration causes a significant improvement in visibility [26], and proving less significant influence of ozone.

It can also be stressed that visibility increased with temperature and also had relatively high correlation with, PM_{10} and NO_2. The literature data also point at wind speed, atmospheric pressure, and CO as exerting a significant influence on visibility [27]. Relative loadings of visibility in PC1 (the first principal component) and PC2 (the second principal component) reach, according to that data, 0.76 and 0.55, respectively [27], indicating that the visibility is primarily affected by SO_2, CO, NO_2 and PM_{10} but also the products of photochemical reactions such as ozone precursors seem to affect visibility. This may explain a high loading of ozone on the second principal component in the cold season and on the first principal component in the warm season.

3.4. LOGISTIC REGRESSION MODEL

Logistic regression is a certain mathematical model which can be employed to describe the influence of independent variables on a dichotomous dependent variable [18]. In order to perform in-depth identification of factors shaping visibility and precisely estimate a degree of the influence on visibility, logistic regression and generalised regression (GRM) models were built. Each of them conveys slightly different information about the phenomenon, which is valuable during the investigation. The conclusions drawn from the three types of models are complementary.

Changeable visibility underwent dichotomisation on the basis of the distribution median (ME = 10.5 km). The measurements not smaller than median were given value 1 and smaller ones were given value 0, with the aim of identifying the logistic model. Final results of the estimation (the tables present only the variables which significantly shape a dependent variable) for all measurements are shown in Table 5. A key parameter for the interpretation is the unit odds ratio. Variables which significantly affect visibility (estimation of model parameters), with the effect which is directly proportional, are: PM_{10} (1.08), f (1.11), SO_2 (1.11) and T (1.47). Interpreting the effect of PM_{10} one could observe that the odds for extended visibility are higher with each 1 μg per m^3 of PM_{10} by approx. 8%, assuming that the remaining variables do not change. Minimum and maximum temperatures have inversely proportional effect on visibility, i.e. they lower the chances for an increase in visibility. The remaining variables can be interpreted in a similar way, and it is worth noting that the influence of temperature is the strongest which confirms the regularities identified in the PCA model.

Table 5

Logistic regression models (cold and warm season jointly)
dependent variable: visibility

Logistic regression model (dependent variable coding; 0 < ME; 1 >= ME)							
Estimation	Constant B_0	PM_{10}	T_{min}	f	T	T_{max}	SO_2
	–11.29	0.08	–0.26	0.10	0.39	–0.17	0.11
t(959)	–10.02	7.92	–3.71	9.48	3.17	–2.23	3.27
Unit odds ratio	0.00	1.08	0.77	1.11	1.47	0.85	1.11

In the warm season the logistic model shows that only two variables play an important role, i.e. PM_{10} and f (Table 6). In the table, CL–/+95% denote 95%-confidence interval for odds ratio, t(959) – the value of the distribution function in the Student's t-distribution of a given number of degrees of freedom, unit odds ratio is a change in % of a dependent variable (in the logistic regression called the odds) with a change of a dependent variable by one unit, assuming that the remaining variables play no significant roles, constant B_0 is a constant of the model. Other comments the same as for Table 5.

Table 6

Logistic regression models (warm season)
dependent variable: visibility

Estimation	Constant B_0	PM_{10}	f
	–6.91	0.07	0.06
t(479)	–5.103	5.17	5.00
Unit odds ratio	0.00	1.07	1.07
–95% CL	0.00	1.04	1.04
+95% CL	0.01	1.10	1.09

The influence of the remaining ones is not statistically significant, but p values, except for SO_2, assume values not higher than 0.12. This discrepancy in conclusions in relation to the PCA model is most likely connected with the loss of information resulting from dichotomisation of visibility measurements and enables clear identification of PM_{10} as the most important factor shaping visibility in the warm season.

Regularities in the cold season are in compliance with the conclusions from the PCA model: variables which mostly shape visibility are: PM_{10} (1.08), SO_2 (1.16) and f (1.21) (Table 7). Out of the meteorological factors the most important is air temperature, but only T_{max} (0.87) has the same direction of influence as the one shown by the PCA model. The differences are connected with big discrepancies in values and dichotomisation of a dependent variable, which in a way forces identification of only the strongest relationships, which is an advantage from the perspective of finding crucial factors. During the identification of values, however, it causes problems concerned with precise estimation of a relationship.

Table 7

Logistic regression models (cold season); dependent variable: visibility

Estimation	Constant B_0	PM_{10}	T_{min}	f	T	SO_2
	–19.30	0.08	–0.44	0.19	0.57	0.15
t(473)	–8.57	5.44	–3.76	8.25	2.79	3.18
Unit odds ratio	0.00	1.08	0.64	1.21	1.76	1.16
–95% CL	0.00	1.05	0.51	1.15	1.18	1.06
+95% CL	0.00	1.12	0.81	1.26	2.62	1.27

3.5. GENERAL REGRESSION MODEL (GRM)

In this work, the GRM was used to verify earlier theses and the results obtained with PCA and logistic models. It constitutes the final stage of identifying factors which shape visibility expressed in their natural measurement scales. As a complement to the PCA model, it enables more precise evaluation of the influence of variables significantly related to visibility. In addition, it shows alternative models and relationships which shape these models, consisting of different systems of variables-predictors as compared to the finally selected model.

The adjusted coefficient of determination R^2 which contains a kind of "punitive element" for too complex models was chosen as a criterion for model selection. The estimation results are presented in Table 8. In the table, R is the multiple correlation coefficient, R^2 – the coefficient of determination; Adj. R^2 – adjusted coefficient of determination; SS – sum of squares, df – degrees of freedom, MS – variance, F – the Fisher–Snedecor statistics, p – significance level. They confirm all the regularities found earlier. The seasonality factor is crucial. In the cold season, visibility is lower.

Table 8

The results of GRM model identification

	Adj. R^2	No. of variables	NO_2	O_3	*T*	T_{min}	T_{max}	*f*	PM_{10}	Season	
1	0.55	9	0.08	0.12	–1.40	0.46	1.02	–0.35	–0.43	–0.10	
2	0.55	8		0.09	–1.47	0.47	1.07	–0.36	–0.39	–0.11	
Visibility [km]	*R*	R^2	Adj. R^2	*SS*	*df*	*MS*	*SS*	*df*	*MS*	*F*	*p*
	0.75	0.56	0.55	2543.0	9	282.5	2021.0	944	2.14	131.98	0.000

The concentration of PM_{10}, air temperature and *f* play a key role. On the other hand, concentrations of NO_2 and O_3 prove to be statistically significant depending on the final structure of a model, i.e. they contain a part of shared information which enables eliminating chosen variables depending on the structure of the entire model. In Step 3, all variables which significantly affect visibility were identified (Fig. 3).

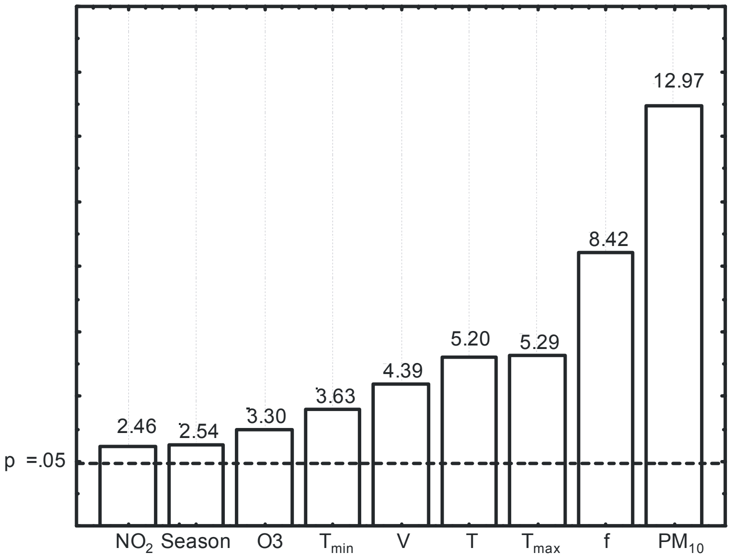

Fig. 3. Variables that significantly influence the visibility based on Student's *t*-distribution

4. DISCUSSION

The results of the study reveal that concentration of an air pollutant significantly influences visibility and by the way it also impacts atmospheric turbidity according to contemporary literature data [26]. Even if it was not for the wind speed in this work to be an important meteorological parameter that affects atmospheric turbidity as well as visibility, in other studies it caused concentrations of air pollutants to remain the same over selected periods of observation [26]. From the observed values, under very turbid conditions when visibility reaches ≤5 km, the wind velocity would be lower than 5 $m \cdot s^{-1}$, regardless of the atmospheric pollutant concentration [27] which is a vital issue and may need a further research to be proved in the Warsaw Agglomeration. However, it was evident in this study that air temperatures and humidity exerted a vital influence on visibility. Owing to air pollutant concentration effects related with air temperatures (the index of phenomenon, determined by PCA, comprising pollutant concentrations and air temperatures), visibility has higher values in the summer and lower values in the winter. The relationships between critical atmospheric visibility and atmospheric air pollutant concentrations have been established, and the general regression model (GRM) proved high correlation between analysed visibility range and known concentrations of air pollutants. Furthermore, a logistic regression model that incorporates the concentrations of atmospheric air pollutants and air temperatures has been developed in this work, and in practice this model can be used to simulate realistic ranges of critical atmospheric visibility. Moreover, the logistic regression model would determine these variables that contain part of shared information about the analysed phenomenon, i.e. visibility. On the other hand, LRM enables eliminating chosen variables from the phenomenon analysis, depending on the adopted structure of the entire statistical model.

In this work, it was estimated that significant visibility differences between clear and episodic air along with the noted difference in PM_{10} show that visibility is strongly negatively correlated with concentration of PM_{10}, of which an abundant evidence is given also by Ying et al. [27]. According to those authors, visibility increased upon increasing temperature and wind speed, decreased upon increasing atmospheric pressure, and also had relatively high correlation with CO ($R = -0.96$), PM_{10} ($R = -0.95$) and NO_2 ($R = -0.94$). The research performed in the Warsaw Agglomeration found as well that higher PM_{10} especially resulted in low visibility and vice versa. A reduction in PM_{10} does not completely improve the visual range [28], owing to the dominant role played by $PM_{2.5}$ in deterioration of visibility. However, there is a strong association between the presence of PM_{10} and the presence of $PM_{2.5}$, not only in Warsaw [15] but elsewhere in the world [29]. Furthermore, the visibility extinction in the present study can be explained well by referring to PM_{10} mass, NO_2 values, and relative humidity, similar to Lai and Sequeira [30], indicating again that long-term visibility trends are highly correlated to PM_{10} concentrations.

All statistical tools used in this work finally proved there were good correlations among SO_2, PM_{10} and NO_2, demonstrating that these pollutants may be similarly sourced. There were four principal components with eigenvalues 4.23 (1), 2.48 (2), 2.21(3) and 0.87 (4), representing a total 82% of the variance. The first PC (PC1) consists of the major species affecting visibility in the Warsaw Agglomeration: PM_{10}, SO_2, NO_2 and a composition of three temperature parameters: T, T_{min} and T_{max}, and accounts for 35% of the explainable variance. Variations in PM_{10}, NO_2 and SO_2 concentration were, however, the main factors affecting visibility, forming a relationship that was stronger also at numerous stations in the world [27]. Burning of heavy fuels, particularly at the industrial part of Warsaw, was the most probable source of SO_2. Photochemical reaction is required to convert this into sulfate, which can influence visibility. Thus, SO_2 and visibility exhibited a moderate correlation. According to Ying et al. [27], relative loadings of visibility in PC1 and PC2 are 0.76 and 0.55, respectively, indicating that the visibility is primarily affected by SO_2, CO, NO_2 and PM_{10}, but that the products of photochemical reactions also affect the visibility. In Warsaw, the PC1 and PC2 loadings are similar: 0.62 and 0.51, however the CO concentration was not taken into consideration due to limitations of the available data but it can be stressed that secondary reaction products would also contribute to visibility deterioration within the area of the Warsaw Agglomeration.

The carried out analysis is not sufficient to fully recognise the effect of air pollutants on visibility within the research area. As shown by scientific research conducted in many cities in the world it is necessary not only to examine temporal and spatial change of visibility in the area of the research but also to identify main causes of visibility reduction which is not possible without performing detailed analyses of chemical composition of aerosols [5, 7]. At present, concentrations of heavy metals in Warsaw are measured only periodically, merely in PM_{10} [15, 31], which undoubtedly makes precise examination of the issue studied in this work impossible.

5. SUMMARY AND CONCLUSIONS

The paper presents results of the analyses concerning the influence of air pollutants on visibility in Warsaw. All the performed analyses and models showed a significant role of seasonality in shaping visibility.

PCA results show that in the summer season the index of the phenomenon consists of the relationships of three pollutants: PM_{10}, SO_2, NO_2 and relationships of three temperature parameters: T, T_{min} and T_{max}. The relationship is inversely proportional in its character, i.e. concentrations of pollutants increase along with a decrease in air temperature. It is quite strongly directly connected with visibility ($R = –0.62$) which decreases along with an increase in pollution and a fall in air temperature. This component accounts for more than 35% of the total variability of the system. The other most important component is a directly proportional relationship of ozone ($R = 0.75$) and visibility, explaining another ca. 21% of variability. Atmospheric precipitation displayed its specific character through its considerable share in the fourth component ($R = 0.84$) and an inversely proportional relationship with visibility, whose low value points to the absence of direct effect. In total, the four first components explain ca. 82% of the total variability, which is a result proving usefulness and correctness of the PCA model.

In the warm season, the index of the phenomenon is to a large extent formed by a directly proportional relationship of four measurements: O_3, T, T_{max} and solar radiation intensity. It is also directly proportional to visibility. The first component explains about 32% of variability. The other most important relationship of the phenomenon is formed by measurements of PM_{10} and NO_2 which are directly proportional to each other and inversely proportional to visibility (R = 0.28). In this component, one can observe a considerable but below 0.7 contribution of SO_2 and O_3, which shows a smaller influence of these pollutants on visibility than in the cold season. Like in the cold season this influence is inversely proportional.

Built as a complement, the model of logistic regression in the warm season shows that only two variables are significantly related to visibility, i.e. PM_{10} and f. The regularities in the cold season are consistent with the PCA model conclusions: variables which shape visibility most are: PM_{10}, SO_2, and the relative humidity of air. Out of the meteorological factors the most important is air temperature but it is worth noting that only the maximum temperature (0.87) maintains the same direction of influence as the one shown by the PCA model.

The analysis with the GRM model enabled verification of the earlier results obtained with PCA and logistic models. A seasonality factor plays an essential role (the season variable statistically significant in the models). In the cold season visibility is lower. The concentration of PM_{10} dust, air temperature and the relative humidity of air play a key role. On the other hand, concentrations of NO_2 and O_3 prove to be statistically significant depending on the final structure of a model, i.e. they contain a part of shared information which allows the elimination of selected variables depending on the structure of the whole model.

REFERENCES

[1] DZUBAY T.G., STEVENS R.K., LEWIS C.W., *Visibility and aerosol composition in Houston, Texas*, Environ. Sci. Technol., 1982, 16, 514.

[2] TSAI Y.I., *Urban visibility and airborne PM_{10} aerosol in southern Taiwan*, J. Aerosol Sci., 2002, 31, 895.

[3] CHEUNG H.C., WANG T., BAUMANN K., GUO H., *Influence of regional pollution outflow on the concentrations of fine particulate matter and visibility in the coastal area of southern China*, Atmos. Environ., 2005, 39, 4779.

[4] KIM Y.J., KIM K.W., KIM S.D. LEE B.K. HAN J.S., *Fine particulate matter characteristics and its impact on visibility impairment at two urban sites in Korea: Seoul and Incheon*, Atmos. Environ., 2006, 40, 593.

[5] YUAN C.S., LEE C.G., LIU S.H., CHANG J.C., YUAN C., YANG H.Y., *Correlation of atmospheric visibility with chemical composition of Kaohsiung aerosols*, Atmos. Res., 2006, 82, 663.

[6] ELIAS T., HAEFFELIN M., DROBINSKI P., GOMES L., RANGOGNIO J., BERGOT T., CHAZETTE P., RAUT J.-C., COLOMB M., *Particulate contribution to extinction of visible radiation: pollution, haze, and fog*, Atmos. Res., 2009, 92, 443.

[7] MALM W.C., PITCHFORD M.L., *Comparison of calculated sulfate scattering efficiencies as estimated from size-resolved particle measurements at three national locations*, Atmos. Environ., 1997, 31 (9), 1315.

[8] YANG L., ZHOU X., WANG Z., ZHOU Y., CHENG S., XU P., GAO X., NIE W., WANG X., WANG W., *Airborne fine particulate pollution in Jinan, China. Concentrations, chemical compositions and influence on visibility impairment*, Atmos. Environ., 2012, 55, 506.

[9] APPEL B.R., TOKIWA Y., HSU J., KOTHNY E.I., HAHN E., *Visibility as related to atmospheric aerosol constituents*, Atmos. Environ., 1985, 19, 1525.

[10] CHAN Y.C., SIMPSON R.W., MCTAINSH G.H., VOWLES P.D., COHEN D.D., BAILEY G.M., *Source apportionment of PM2.5 and PM10 aerosols in Brisbane Australia by receptor modeling*, Atmos. Environ., 1999, 33 (19), 3251.

[11] LATHA K.M., BADARINATH K.V.S., *Black carbon aerosols over tropical urban environment. A case study*, Atmos. Res. 2003, 69, 125.

[12] TANG I.N., WONG W.T, MUNKELWITZ H.R., *The relative importance of atmospheric sulfate and nitrates in visibility reduction*, Atmos. Environ., 1981, 15, 2463.

[13] TSAI Y.I., CHENG M.T., *Visibility and aerosol chemical compositions near the coastal area in central Taiwan*, Sci. Total Environ., 1999, 231, 37.

[14] STULL R.B., *Meteorology for Scientists and Engineers*, Brooks/Cole, Pacific Grove, USA, 2000.

[15] MAJEWSKI G., KLENIEWSKA M., BRANDYK A., *Seasonal variation of particulate matter mass concentration and content of metals*, Pol. J. Environ. Stud., 2011, 20 (2), 417.

[16] BADYDA A., DĄBROWIECKI P., LUBIŃSKI W., CZECHOWSKI P.O., MAJEWSKI G., *Exposure to traffic-related air pollutants as a risk of airway obstruction*, Adv. Exp. Med. Biol., 2013, 755, 35.

[17] POPEK R., GAWROŃSKA H., WROCHNA M., GAWROŃSKI S.W., SÆBØ A., *Particulate matter on foliage of 13 woody species: deposition on surfaces and phytostabilisation in waxes. A 3-year study*, Int. J. Phytorem., 2013, 15, 3.

[18] STANISZ A., *Simple statistics course with the use of Statistica PL software. Examples from medicine*, StatSoft Polska, Cracow 2007 (in Polish).

[19] LAROSE D.T., *Methods and models for data exploration*, PWN, Warsaw 2012 (in Polish).

[20] BALICKI A., *Multidimensional statistical analysis and its application to socio-economical problems*, Gdansk University Press, 2009 (in Polish).

[21] KLEINBAUM D.G., KUPPER L.L., NIZAM A., MULLER K.E., *Applied regression analysis and other multivariable*, Duxbury Press, Belmont 1998.

[22] ROZBICKA K., ROZBICKI T., MAJEWSKI G., *Variability of tropospheric ozone concentration in the southern part of Warsaw Agglomeration*, Air Protection in Theory and Practice, 2010, 2, 293 (in Polish).

[23] KLENIEWSKA M., *Relationship between meteorological elements and Upshur dioxide concentration in Warsaw Agglomeration area*, PhD thesis, Warsaw University of Life Sciences, Warsaw 2005 (in Polish).

[24] MAJEWSKI G., PRZEWOŹNICZUK W., *Study of particulate matter pollution in Warsaw area*, Pol. J. Environ. Stud., 2009, 18 (2), 293.

[25] CZECHOWSKI P.O., *Forecasting system of high concentrations of nitrogen dioxide in the atmosphere (for example Gdansk Agglomeration)*, PhD thesis, Warsaw University of Technology, Warsaw 2004 (in Polish).

[26] WEN C.C., YEH H.H., *Comparative influences of airborne pollutants and meteorological parameters on atmospheric visibility and turbidity*, Atmos. Res., 2010, 96, 496.

[27] YING I., *Atmospheric visibility trends in an urban area in Taiwan 1961–2003*, Atmos. Environ., 2005, 39, 5555.
[28] BAIK N.-J., KIM Y.P., MOON K.C., *Visibility study in Seoul, 1993*, Atmos. Environ., 1996, 30, 2319.
[29] CHEN M.-L., MAO I.-F., LIN I.-K., *The PM2.5 and PM10 particles in urban areas of Taiwan*, Sci. Total Environ., 1999, 226, 227.
[30] LAI L.Y., SEQUEIRA R., *Visibility degradation across Hong Kong: its components and their relative contributions*, Atmos. Environ., 2001, 35, 5861.
[31] MAJEWSKI G., CZECHOWSKI P.O., BADYDA A., KLENIEWSKA M., BRANDYK A., *Estimation of total gaseous mercury (TGM) concentration at the regional background station in Granica KPN (Mazovia Province, Poland), 2010–2011*, An. Set Environ. Prot., 2013, 15, 1302 (in Polish).