Pol. J. Environ. Stud. Vol. 31, No. 3 (2022), 2223-2235
DOI: *10.15244/pjoes/143295* ONLINE PUBLICATION DATE: 2022-03-23

Original Research

Identification of Air Pollution Sources and Temporal Assessment of Air Quality at a Sector in Mosul City Using Principal Component Analysis

Abdulmuhsin S. Shihab*

Mosul University, Environmental Research Centre, Mosul, IRAQ

Received: 26 July 2021
Accepted: 23 October 2021

Abstract

This study was carried out to apply principal component analysis (PCA) as a tool to identify the major sources responsible stand behind air pollution variation in a sector of Mosul city for the first time. In addition, Besides, PCA was used to construct a temporal overall air quality assessment index to find the period of best air quality along the year. The data was collected through a monitoring station located in the public library on a side of a very crowded highway and near a traffic light intersection in Mosul city. It The data involves the measurements of O_3, NO, NO_2, NO_x, SO_2, CO, CO_2, CH_4, TH, NMHC and PM_{10} for a year. Air quality parameters were analyzed using PCA seasonally and yearly. The study found that the pollutants produced by vehicular traffic exhibited more variation with a percentage of 56.91 to 73.75%. The results showed that traffic pollution is the main contributor to air quality variation with 56.91 to 73.75%. It is verified by the gases CO, NO, NO_x, O_3, THC and CO_2. The temporal assessment of monthly air quality showed that the Best air quality was recorded in March followed by April. Whereas, the worst air quality was observed in January. The study concluded that the application of PCA to air quality data had drawn the parameters responsible for air quality variation and detect the sources of air pollution efficiently. In addition, The results of PCA can be helpful in the design of the program of measurements in the monitoring station.

Keywords: principal component analysis, traffic pollution, ambient air quality, Mosul city, pollution sources, temporal assessment of air quality, principal component scores

*e-mail: mss_qzz@uomosul.edu.iq

Introduction

Clean air became an essential requirement for perfect human life. One of the major threats facing large cities is air pollution. Urbanization and industrialization with the increase in population and fossil fuels using use have had made the atmosphere more polluted. In other words, the amounts of air pollutants released in air became not compatible with normal mechanism.

Air pollutants consist of primary gases and particles emitted from stationary and mobile sources. In addition to secondary pollutants are formed from aerosols by chemical reactions. Air is considered polluted when pollutants accumulate to a level that causes a harmful effect to humans, animals and plants or exceed the standards. Other gases may contribute to global warming or climate change like greenhouse gases. Air pollution can cause harm or discomfort to humans. It is associated with respiratory and cardiovascular diseases [1-2]. Thus, air pollution control is necessary to avoid the harmful situation of air pollution in the future. Air quality monitoring is the first step in air pollution control. Additionally, the interpretation of the variability of air quality parameters requires the monitoring of their concentration and statistical analysis. The availability of air quality data and the results of the analysis can be very helpful in air quality management [3].

Multivariate techniques is verified as one of the efficient tools used to analyze air quality data as it can identify the sources of pollution [4]. Of these multivariate techniques, principal component analysis (PCA) has been widely applied to model the dynamic characteristics of air pollution [5-10]. Also, it can identify the pattern of air quality data, revealing the redundant measurements, reduce the dimensionality of the data and demonstrate the contribution percentage of pollutants in air quality variation [5].

Air pollution has had been studied in Mosul city. The earlier studies focused on dustfall and suspended particulate distribution [11, 12]. After that, Al-Jarrah [13] conducted a study to identify the levels of pollutants solely within the largest population community in Northern Iraq, Mosul city. Furthermore, Shihab and Al-Jarrah [14] investigated the levels of ozone and nitrogen oxides including their relationships with metrological factors, moreover, Shihab [15] use air quality index to assess air quality in Mosul city. This research tries to use multivariate techniques (also known as chemometric techniques) in analyzing air quality parameters in Mosul city for the first time.

Hence, no study has been conducted to identify the sources of air pollution in Mosul city, the current study aims to apply principal component analysis to recognize the possible sources of air pollution. In addition, to determine the contribution of air pollutants in air quality variation in Mosul city. Furthermore, the study intends to identify the weekly and seasonally temporal variation in the pattern of air quality within a selected sector in Mosul city, Iraq.

The objectives of this study are to recognize the possible sources of air pollution and to determine the contribution of each pollutant in air quality variation. In addition to identify the weekly and seasonally temporal variation in the pattern of air quality within a selected sector in Mosul city, Iraq.

Materials and Methods

Study Site

The present study was conducted in Mosul city, northern Iraq. The monitoring station was named the public library as it was placed in the public library of the city. It lies near a traffic light intersection on the side of a very crowded two-way highway in the left bank of the city as shown in Fig. 1. Also it is surrounded by the buildings of Iraqi Engineers Union and the courthouse of the city in addition to a residential area from the North.

Air Quality Monitoring Station

The monitoring site includes a stationary monitoring station type Horiba (German made). It can measure the air quality parameters: O_3, NO, NO_2, NO_x, SO_2, CO, CO_2, CH_4, total hydrocarbons (TH), non-methane hydrocarbon (NMHC) and PM_{10}. The devices in the station are calibrated automatically using span gases and zero gas. The measurements were conducted every three minutes and then the average of 30 minutes was calculated. This station belongs to the Ninevah Environment Directorate. The surveillance operation was continued from Feb 2013 till Jan 2014.

Mosul city has a semi-arid climate with extremely hot, prolonged, dry summer, mild autumn and spring and moderately wet cool winter. The dominant wind direction in the study area is NW with 17.2% calm conditions (Fig. 1). Dominant wind speed ranged between more than 0 to 2 km/hr.

Statistical Analysis

Descriptive Statistics

Mean, median, standard deviation, minimum and maximum were calculated to describe air quality parameters for the study area using IBM SPSS statistics 26 software.

Principal Component Analysis

Principal component analysis is a multivariate technique used to treat a set of variables together. This type of analysis has the ability to convert the actual

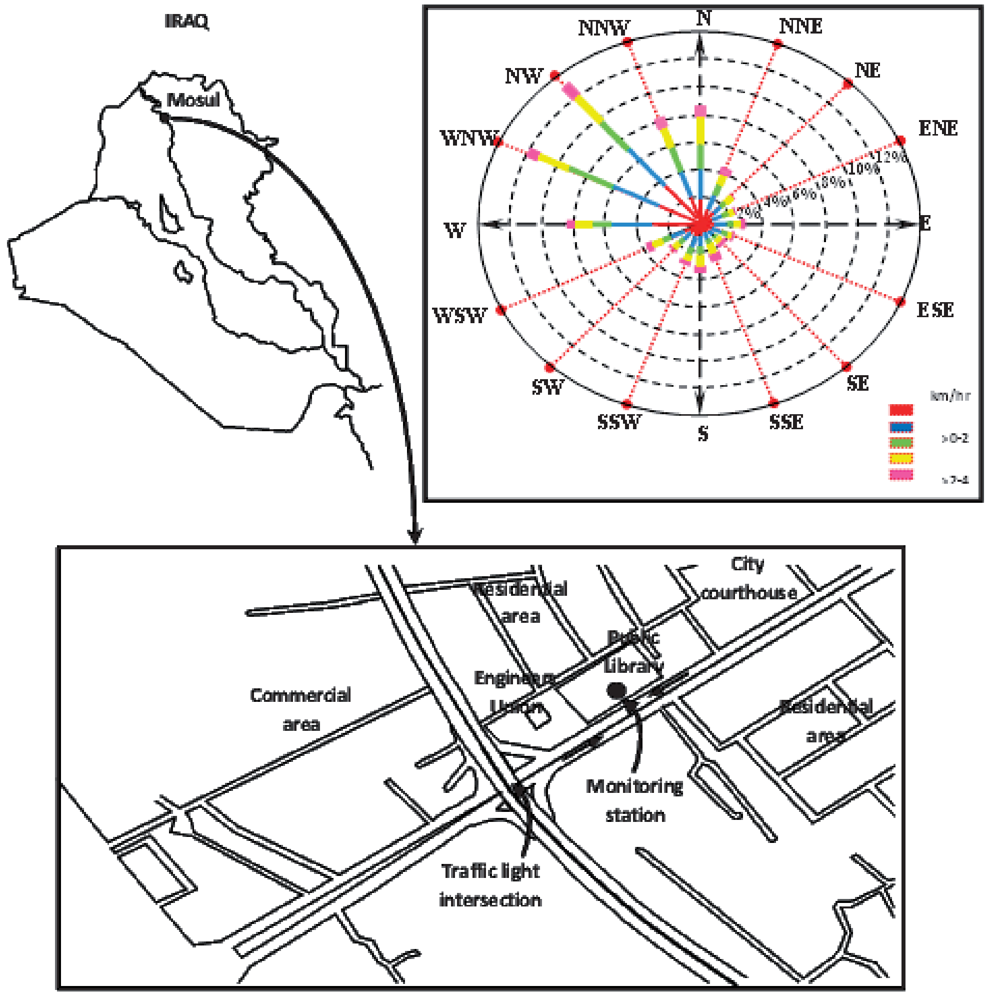

Fig. 1. Location of the study site with monitoring station and wind rose diagram.

correlated air quality parameters into a new set of orthogonal and uncorrelated composite components [16-18]. These components are linear combinations of the original data. It can be expressed as in Equation (1):

$$Zij = a_{i1}x_{1j} + a_{i2}x_{2j} + a_{i3}x_{3j} + ... + a_{im}x_{mj} \quad (1)$$

where Z is the component score, a is the component loading, x is the measured value of the variable, i is the component number, j is the sample number, and m is the total number of variables [3].

The raw data need to be standardized as air quality parameters have different magnitudes and scales of measurements according to Z-scale to a mean of 0.0 and variance of 1.0 by using equation (2), [19, 20]:

$$Z_{ij} = (X_{ij} - \mu)/\ \sigma \quad (2)$$

Where Z_{ij} is the standard score of jth value of the measured variable i; μ is the variable mean value and σ is the standard deviation. This standardization will give equal weights to air quality variables in the statistical analysis process. Besides, this process will homogenize the variance of the distribution [21].

Principal component analysis used the correlation matrix of observations (X) to estimate a sorted matrix of eigenvalues (λ) and corresponding eigenvectors (component loading V). The characteristic equation is $[X-\lambda I]\ V = 0$, where each eigenvalue λ is associated with an eigenvector V. The components with eigenvalues equal to or greater than one are retained using Kaiser criterion [22].

The amount of variance extracted in each variable is named as communality. It represents the sum of squared loadings (V) for the extracted components of a specified air quality parameter. The high communalities indicate that the extracted components represent the variables well. To exclude measures that are not useful in the interpretation, a value of 0.5 was selected as a cut-off communality.

The sampling adequacy of the data for analysis was tested using Kaiser-Meyer-Olkin (KMO) and Bartlett's test for sphericity. KMO values of 0.5 and more with significant Bartlett's test ($p \leq 0.05$) indicate that sampling is adequate for analysis [23, 24].

To facilitate the interpretation of air quality data, rotation may be applied to the components axis to provide a simple structure. The components loadings will be rescaled back to the proper size after rotation. The total variance outlined by all components remains the same, while the total variance defined by each component will be different. The most common orthogonal rotation is Varimax. It minimizes the number of variables that have high loadings on each component and simplifies the interpretation of the components. Varimax rotation confirms that each variable is maximally correlated with only one component and a near zero association with the other components [25].

The direct relationships between air quality parameters were obtained by drawing the loadings for PC1 versus PC2 as arrows. Each arrow represents an air quality parameter. The angle between the arrows denoted the relationship between them. The angle closer to zero denoted a strong relationship versus weak for an angle closer to 90 degree. On the other hand, an angle closer to 180 degree denoted a strong inverse relationship.

For weekly and monthly assessment of air quality, PC scores were utilized. According to the component score coefficient matrix, the scores of the components were calculated for each week and month of the study period. From the results of the variance contribution rate, the score of the components extracted were converted to a score for each week and month. The week with lower score represents the period with best air quality and the same way for the months [10].

Results and Discussion

Principal component analysis was conducted on air quality parameters for the monitoring station at the public library, Mosul city for each season and for the whole year.

The descriptive statistics for air pollutants including mean, median, minimum and maximum in the monitoring site are shown in Table 1.

Kaiser-Meyer-Olkin (KMO) values were more than 0.5 and ranged between 0.661-0.839 for seasonally and yearly analysis. In addition, Bartlett's test for sphericity shows significant results ($p \leq 0.001$). This indicates that sampling adequacy is highly suitable for PCA [24].

For air quality in winter, PCA generated two components of eigenvalues greater than one after varimax rotation. These components explained 72.88% of air quality variation in the study site at Mosul city (Table 2).

PC1 explained 56.99% of the variance in air quality versus 15.89% for PC2. NO_x has the highest loading on this PC with 0.96 followed by NO with 0.95. NO is produced from the reaction of nitrogen with oxygen in the air as they are in touch with the very high temperature vehicles engines. Consequently, its concentrations will vary with traffic volume and the number of vehicles in the traffic light intersection, especially traffic congestion during rush hours as stated by Lee et al. [26]. Additionally, its concentration undergoes more variation as solar radiation increases due to its relationship with ozone [27, 28]. CO is loaded on this PC with a value of 0.89. It is the result of incomplete combustion of fuel in vehicle engine due to acceleration, load and poor maintenance. Its concentration varied with traffic circulation level and air/fuel ratio [29]. Furthermore, the vehicles at the nearby traffic light intersection are in inactive position with engine exuded more CO compared to free flow condition [30]. The removal of CO as it reacts with OH radicals produced by the photolysis of O_3 contributes in the variation of its concentration [28]. Additional gases emitted from the vehicles are loaded also on this PC like NMHC, THC, SO_2 and NO_2 with 0.84, 0.83, 0.73 and 0.68 loadings respectively. NMHC and THC come from unburned vehicle fuel especially at their stopping in the nearby traffic light intersection as the urban area around the monitoring station does not include any industrial activities. The variation in these two parameters

Table 1. Descriptive statistics for air pollutants.

Parameters	Mean	Median	SD	Min.	Max.
CO (ppm)	1.24	0.99	0.73	0.39	3.56
NO (ppb)	35.88	20.55	35.26	3.55	184.41
NO_2 (ppb)	26.20	25.08	8.93	8.93	55.32
NO_x (ppb)	62.03	48.04	40.76	14.13	217.11
O_3 (ppb)	27.56	28.45	15.72	0.79	69.38
SO_2 (ppb)	15.58	13.10	8.33	3.44	41.57
CO_2 (ppm)	54.62	52.45	14.85	28.68	104.89
CH_4 (ppm)	1.68	1.82	0.53	0.02	2.26
NMHC (ppm)	0.76	0.66	0.45	0.02	2.55
THC (ppm)	2.42	2.48	0.84	0.04	4.04
PM10 ($\mu g/m^3$)	145.05	143.63	69.25	2.33	348.00

Table 2. Principal components loadings with communalities, eigenvalues and variance after varimax rotation of air quality parameters at winter season.

Parameter	PC1	PC2	Communality
NO_x	**0.96**	0.07	0.921
NO	**0.95**	0.00	0.904
CO	**0.89**	0.30	0.876
NMHC	**0.84**	-0.07	0.712
THC	**0.83**	0.25	0.754
SO_2	**0.73**	0.34	0.647
O_3	**-0.71**	0.53	0.778
NO_2	**0.68**	0.44	0.664
PM10	**0.68**	0.30	0.549
CO_2	0.17	**0.79**	0.665
CH_4	0.09	**0.74**	0.554
Eigen values	6.27	1.75	-
% Variance	56.99	15.89	-
% Cumulative	56.99	72.88	-

KMO = 0.770, p<0.001

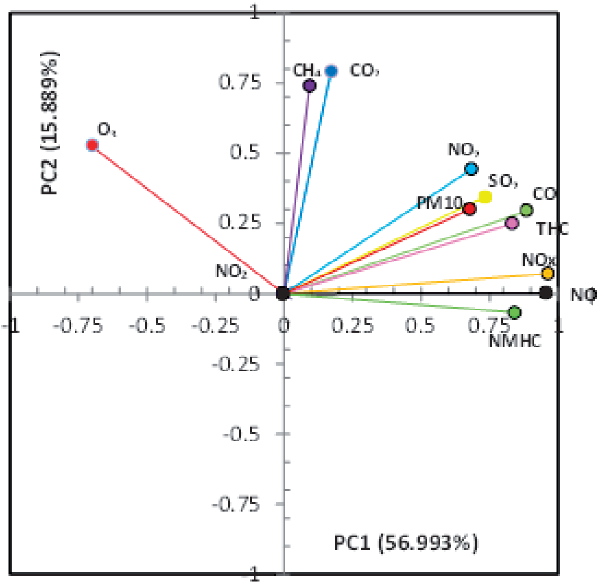

Fig. 2. PC1 loading versus PC2 for winter season.

may be attributed to the variation in the traffic volume and the type of vehicles. The main source of SO_2 is the combustion of fossil fuels containing sulfur [28]. SO_2 is produced by heavy diesel engines, buses and lorries [31]. Also, using kerosene in heating contribute in SO_2 variation in the nearby residential area at winter. Its variation is associated with the number of vehicles of diesel engines passing the monitoring station road in addition to its reaction with OH radicals produced by ozone photolysis. For NO_2, the emissions from motor vehicles represent the main source in high traffic area [32]. Its concentration varies with the load and speed of the vehicle in addition to traffic volume [33]. The concentration of NO_2 increased in rush hours, while its reduction can be explained in terms of traffic volume in addition to its relationship with ozone [34]. Ozone is loaded negatively on this first component. It means that ozone is inversely correlated with other pollutants loaded on this component which coincided with results of Kovac-Andric et al. [27]. NO_2 photolysis is stimulated with solar radiation increase to generate ozone and decrease NO_2 concentration. This trend is inversed during the day. Additionally, ozone is consumed in the oxidation of traffic originated NO [35]. PM_{10} is also loaded on this PC with a value of 0.68. Its concentration in this season is related with the rain, wind speed and direction and its role in re-suspension of soil dust with scarcity of green areas in the study region. In addition, traffic volume contributes to its variation as vehicular smoke emission from diesel engines is another source of PM_{10} [11, 36]. These results indicate that this PC represents the traffic pollution.

PC2 represents global warming gases as CO_2 and methane were loaded on it with values 0.79 and 0.74 respectively. It represents about quarter the variance in air quality of PC1 with a percentage of 15.89. The main source of CO_2 is the fuel combustion in vehicles and in heating at homes during the cold season winter. On the other hand, decaying of organic domestic waste in the nearby residential area and wastewater are the main sources of methane.

Direct relationships among some the variables which is denoted by low angle between their vectors like NO and NO_x, except for ozone (Fig. 2). Ozone has an angle near 180 degree with more of the pollutants, which indicates a reverse relationship with them, which coincided with the results of Lee et al. [26]. On the other hand, the angle near 90 degree of CO_2 and CH_4 with other parameters indicates their weak relationship with the parameters loaded on PC1 as illustrated in Fig. 2.

For spring season, PCA extracted two components to represent 88.11% of air quality variation in the study site (Table 3). Six air pollutants are loaded on PC1 versus three on PC2. NO_x (0.98), NO (0.98), NO_2 (0.89), CO (0.94), and SO_2 (0.94) are loaded on PC1 positively except ozone which is loaded negatively (-0.65). Vehicles are the main source of these pollutants as explained in the case of the analysis of winter results. Therefore, this PC may be called traffic pollution or fossil fuel combustion component. Additionally, the slight angles among the arrows of these variables in Fig. 2 exhibit the high relationships among them. On the other hand, ozone is inversely correlates with these variables according to the angles with its arrow of about 180 degree.

Table 3. Principal components loadings with communalities, eigenvalues and variance after varimax rotation of air quality parameters at spring season.

Parameter	PC1	PC2	Communality
NO_x	**0.98**	0.13	0.939
NO	**0.98**	0.08	0.911
CO	**0.94**	0.12	0.983
SO_2	**0.94**	-0.01	0.907
NO_2	**0.89**	0.21	0.967
O_3	**-0.65**	0.33	0.834
THC	0.02	**0.99**	0.979
CH_4	-0.10	**0.96**	0.532
NMHC	0.31	**0.90**	0.877
Eigen values	5.12	2.81	-
% Variance	56.91	31.20	-
% Cumulative	56.91	88.11	-

KMO = 0.661, $p<0.001$

Table 4. Principal components loadings with communalities, eigenvalues and variance after varimax rotation of air quality parameters at summer season.

Parameter	PC1	PC2	PC3	Communality
NMHC	**0.90**	0.27	-0.10	0.938
THC	**0.88**	0.28	0.22	0.893
CO	**0.76**	0.54	0.21	0.906
NO_2	**0.72**	0.55	0.35	0.911
CO_2	**0.72**	0.00	-0.02	0.889
O_3	-0.200	**-0.81**	0.25	0.931
NO	0.494	**0.80**	0.06	0.960
SO_2	0.533	**0.75**	0.25	0.757
PM10	-0.033	**0.75**	0.17	0.905
NO_x	0.611	**0.74**	0.19	0.509
CH_4	0.088	0.07	**0.96**	0.585
Eigen values	6.82	1.24	1.13	-
% Variance	62.00	11.24	10.25	
% Cumulative	62.00	73.24	83.50	-

KMO = 0.715, $p<0.001$

The hydrocarbonate gases are loaded positively on PC2. It exhibits high loading from THC (0.99), CH_4 (0.96) and NMHC (0.90). Also, it represents more than half (31.20%) of the variation in air quality characterized by PC1 (56.91%) as shown in Table 2. Vehicle traffic volume contribute to the variation in THC and NMHC as vehicle is the main source of them. While, the variation in CH_4 may be attributed to the decomposition of organic materials of solid waste placed in open containers in the nearby residential area. Carbon dioxide does not contribute in this variation since its saturation level or its communality is less than 0.50. Furthermore, the variables loaded on this PC show weak correlation with those of component 1 as their arrows are approximately perpendicular to them (Fig. 3). This PC may be called as the global warming component as the hydrocarbons loaded on it contribute to the global warming.

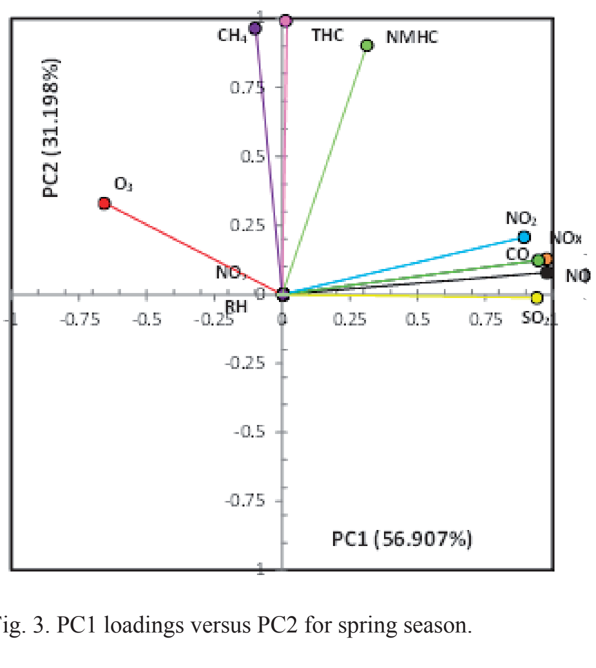

Fig. 3. PC1 loadings versus PC2 for spring season.

During summer season, PCA extracted three components to represent 83.50% of the cumulative variance of air quality (Table 4). PC1 demonstrates 62.0% of the variance in air quality data. It exhibits high loading from NMHC (0.90), THC (0.88) and CO (0.76). NO_2 (0.72) and CO_2 (0.71) are also loaded on this PC. Vehicles are the main source of these gases during their inactive position at the nearby traffic light intersection including the evaporation of fuel with temperature increase in this season. Therefore, these gases vary with vehicle traffic volume, smooth traffic condition and percentage of heavy vehicles and buses. In addition, the levels of hydrocarbons elevated at the high traffic sits located at intersections [37]. Additionally, high solar radiation and temperature during this season stimulates the photolysis of NO_2 and the generation of OH radicals. Furthermore, it can be seen that NMHC and THC moved from PC2 to PC1 of more variance in this season compared with spring (Table 4). This can be attributed to the increase in photochemical removal by hydroxyl radical in summer [38].

PC2 represents 11.24% of the variance in air quality data. It exhibits high loading from O_3 (-0.81), NO (0.80), SO_2 (0.75), PM10 (0.75) and NO_x (0.74). Ozone is negatively loaded on this PC as it is inversely correlated with parameters loaded on this PC (Fig. 4). Traffic pollution, photochemical oxidation and soil re-suspension by wind contribute in the variation of these pollutants. PC1 and PC2 combine the effect of traffic emissions, traffic light intersection and local climate condition.

PC3 represent 10.25% of the variation in air quality data. It exhibits high loading from CH_4 (0.96) only. It demonstrates the decomposition of organic materials in wastewater and domestic waste in the nearby residential area [39]. This PC represents the decaying of organic materials.

In autumn, PCA shows different results as temperature decreases. The analysis extracted two PCs to demonstrate 86.73% of the cumulative variance in air quality data, 73.75% for component 1 and 12.98% for component 2 (Table 5). Seven of air quality parameters are loaded on PC1 versus three for PC2. PC1 exhibits its loadings from CO_2 (0.88), O_3 (-0.84), NO (0.81), THC (0.80), NMHC (0.79), CO (0.77) and NO_x (0.74). Most of these parameters are highly correlated with each other as shown in Fig. 4. The main source of these gases is vehicle fuel combustion and the photochemical reactions [40]. In addition, gases may be produced by the reactions between air constituents near the hot vehicle machine like nitrogen and oxygen. This component represents the traffic pollution.

PC2 exhibits its loading from PM_{10} (0.91), SO_2 (0.85) and NO_2 (0.81). The source of PM_{10} varied among wind re-suspension of soil dust, vehicle exhaust emission and domestic solid waste burning. In addition, the environmental conditions during the study period may also contribute to the variation in PM10 [41-42]. On the other hand, motor vehicle emission is the source of NO_2, while SO_2 is emitted from heavy diesel engines, buses and lorries. This PC represents the effects of local climate condition and the emissions of diesel engines.

Table 5. Principal components loadings with communalities, eigenvalues and variance after varimax rotation of air quality parameters at autumn season.

Parameter	PC1	PC2	Communality
CO_2	**0.88**	-0.05	0.930
O_3	**-0.84**	-0.22	0.942
NO	0.81	0.52	0.899
THC	**0.80**	0.55	0.916
NMHC	**0.79**	0.55	0.784
CO	**0.77**	0.56	0.961
NO_x	**0.74**	0.64	0.755
PM10	0.03	**0.91**	0.887
SO_2	0.41	**0.85**	0.771
NO_2	0.35	**0.81**	0.827
Eigen values	7.38	1.30	
% Variance	73.75	12.98	
% Cumulative	73.75	86.73	

KMO = 0.839, p<0.001

Fig. 5 shows inverse relationship between O_3 and more of the parameters (an angle close to 180 degree) versus a weak one with PM10, NO_2 and SO_2 of PC2 (an angle close to 90). Additionally, there is a direct relationship between THC and CO in winter, which coincided with the results of Eslami et al. [43]

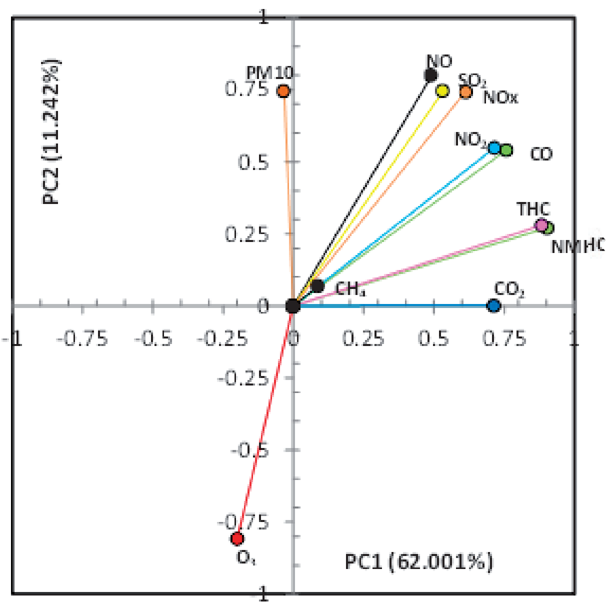

Fig. 4. PC1 loadings versus PC2 for summer season.

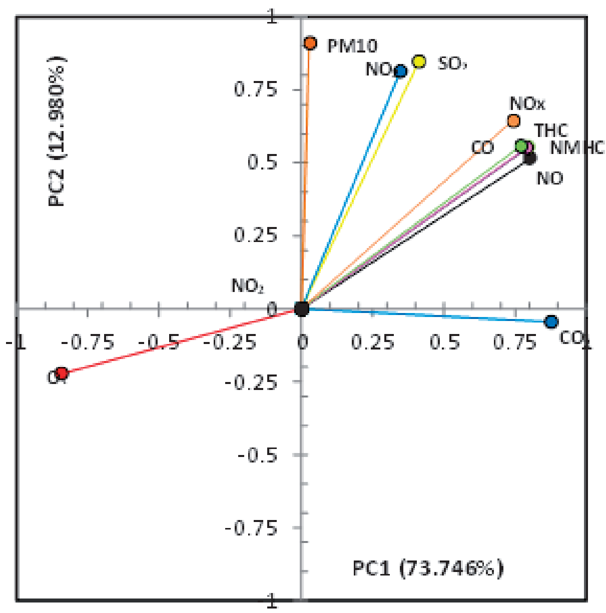

Fig. 5. PC1 loadings versus PC2 for Autumn season.

Table 6. Principal components loadings with communalities, eigenvalues and variance after varimax rotation of air quality parameters for the study period along a year.

Parameter		PC1	PC2	PC3	Communality
NO		**0.91**	0.24	0.15	0.932
CO		**0.88**	0.21	0.32	0.871
O_3		**-0.87**	-0.10	-0.08	0.984
CO_2		**0.81**	-0.11	-0.21	0.918
SO_2		**0.69**	0.12	0.57	0.901
CH_4		-0.07	**0.96**	0.04	0.721
THC		0.26	**0.95**	0.16	0.765
NMHC		0.61	**0.67**	0.25	0.814
NO_2		-0.03	-0.03	**0.85**	0.721
PM10		0.06	0.40	**0.72**	0.684
Eigen values		4.90	2.11	1.30	
% of Variance		49.00	21.08	13.03	
% Cumulative		49.00	70.08	83.11	
Rotation sum of square loading	Total	3.92	2.56	1.83	
	% of Variance	39.23	25.57	18.30	
	Cumulative %	39.23	64.80	83.11	

KMO = 0.727, p<0.001

For the yearly variation, the analysis of all data extracted three PCs which demonstrate 83.11% of the cumulative variance in air quality (Table 6). PC1 exhibits high loading from NO (0.91), CO (0.88), O_3 (-0.87), CO_2 (0.81) and SO_2 (0.69). These parameters undergo more variation along the year as they represent the highest part of the cumulative variance with 49.00%. It represents the traffic pollution and photochemical reactions. The hydrocarbonate gases CH_4 (0.96), THC (0.95) and NMHC (0.67) are loaded on component 2. They have lesser contribution in air quality variation with 21.08%, which represent less than half that of PC1., as This can be attributed to the fact that PC1 indicates the principal source of air pollution [36], besides,

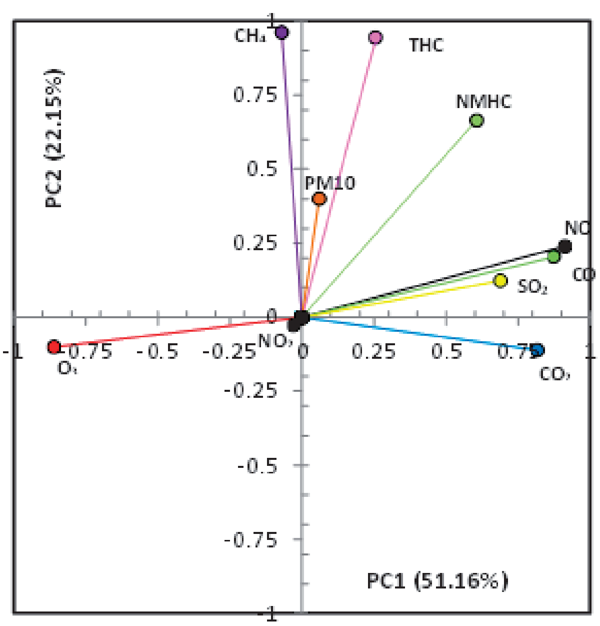

Fig. 6. PC1 loadings versus PC2 for the study period.

Table 7. Component score coefficient matrix.

Parameters	PC1	PC2	PC3
NO	0.235	0.001	-0.020
CO	0.216	-0.044	0.103
O_3	-0.255	0.024	0.140
CO_2	0.272	-0.099	-0.188
SO_2	0.146	-0.108	0.298
CH_4	-0.121	0.469	-0.132
THC	-0.031	0.408	-0.081
NMHC	0.094	0.221	-0.001
NO_2	-0.075	-0.158	0.565
PM10	-0.082	0.064	0.400

Table 8. Month factor score and ranking.

Period (month)	PC1	Rank	PC2	Rank	PC3	Rank	PC	Rank
Feb-13	0.82	10	0.08	3	-0.25	5	0.36	9
Mar-13	0.42	8	-2.15	1	-0.29	4	-0.53	1
Apr-13	-0.19	6	-0.91	2	-0.38	2	-0.45	2
May-13	-0.39	5	0.17	5	-0.13	6	-0.16	7
Jun-13	-0.89	2	0.22	6	0.32	9	-0.28	4
Jul-13	-1.06	1	0.25	7	0.26	8	-0.37	3
Aug-13	-0.86	3	0.34	10	0.53	10	-0.19	6
Sep-13	-0.65	4	0.16	4	0.17	7	-0.22	5
Oct-13	-0.10	7	0.33	9	0.80	11	0.23	8
Nov-13	0.69	9	0.48	11	0.53	10	0.59	11
Dec-13	1.02	11	0.32	8	-0.32	3	0.51	10
Jan-14	1.64	12	0.82	12	-1.28	1	0.75	12

PC2 are less affected by seasonal variation. This PC2 represents the decomposition of organic materials and unburned vehicle fuel especially at their stopping in the nearby traffic light intersection. PC3 shows the parameters of least seasonal variation, NO_2 (0.85) and PM10 (0.72). They represent 13.03% of the variation in air quality during the study period. The main source of NO_2 is mainly from fuel burning in vehicles. As the PM10 is emitted from diesel by motor vehicle exhaust, Also it can be produced by wind gusts. Ozone shows reverse correlation with air quality parameters loaded on PC1. On the other hand, it correlates weakly with parameters loaded on PC2 and PC3 (Fig. 6).

From the PCA for all seasons, the parameters loaded more on PC1 are CO, NO, NO_x, O_3, THC and CO_2 which constitute the traffic pollution gases and those of photochemical reactions. This means that these parameters are subjected to more variation than others as PC1 covers the higher percentage of variance in the data. This can be attributed to the location of the monitoring station which lies on a side of main highway and near traffic light intersection. In addition to the variation in traffic volume and the effect of weather condition on the emitted gases.

PCA was used to study the monthly and weekly temporal assessment of air quality in the study station. Component score coefficient matrix was used to calculate the components scores for each month and week for the study period (Table 7). The scores of the components were merged using variance contribution rate (percentage of variance for each component) listed in Table 6 and converted to Equation (3) which was used

Table 9. Week factor score and ranking.

Period (week)	PC1	Rank	PC2	Rank	PC3	Rank	PC	Rank
1-7 Feb13	1.59	47	0.24	27	-0.50	12	0.71	45
15-21 Feb13	0.91	42	0.10	13	-0.61	8	0.33	38
22-28 Feb13	0.60	38	0.03	8	0.08	30	0.31	37
1-7 Mar13	0.65	40	-0.45	6	-0.17	24	0.13	32
8-14 Mar13	0.48	36	-3.08	1	-0.52	11	**-0.83**	1
15-21 Mar13	0.33	33	-2.43	4	-0.34	17	-0.66	6
22-28 Mar13	0.37	35	-3.05	2	0.05	28	**-0.75**	2
29 Mar-4 Apr13	0.13	30	-2.50	3	0.17	33	-0.67	5
5-11 Apr13	-0.33	24	-1.53	5	-0.47	13	**-0.73**	3
12-18 Apr13	-0.39	21	0.16	19	-0.74	7	-0.30	12
19-25 Apr13	0.06	29	-0.19	7	-0.99	4	-0.25	16

Table 9. Continued.

26 April–2May13	-0.41	20	0.20	21	-0.11	25	-0.15	26
3-9 May13	-0.09	27	0.21	22	-0.17	23	-0.01	29
10-16 May13	-0.36	23	0.15	17	-0.60	9	-0.26	15
17-23 May13	-0.53	17	0.09	12	-0.07	26	-0.24	18
24-30 May13	-0.66	16	0.25	31	0.21	34	-0.19	24
31 May-6 Jun13	-0.95	9	0.27	34	-0.29	18	-0.43	9
7-13 Jun13	-0.80	11	0.21	23	0.49	39	-0.20	21
14-20 Jun13	-0.68	14	0.09	11	0.66	42	-0.15	27
21-27 Jun13	-0.96	8	0.23	26	0.67	43	-0.23	19
28 Jun-4 Jul13	-0.96	6	0.22	24	0.15	32	-0.35	10
5-11 Jul13	-1.02	5	0.27	33	0.40	38	-0.31	11
12-18 Jul13	-1.44	1	0.23	25	-0.34	16	-0.68	4
19-25 Jul13	-1.12	3	0.25	29	0.07	29	-0.44	8
26 Jul- 1 Aug13	-0.96	7	0.27	32	0.76	44	-0.20	22
2-8 Aug13	-1.07	4	0.28	35	0.60	41	-0.29	13
9-15 Aug13	-1.17	2	0.36	38	0.03	27	-0.44	7
16-22 Aug13	-0.52	18	0.41	43	0.87	48	0.07	30
23-29 Aug13	-0.79	12	0.28	36	0.55	40	-0.17	25
30 Aug- 6 Sep13	-0.69	13	0.15	16	0.82	46	-0.10	28
7-13 Sep13	-0.81	10	0.10	14	0.30	36	-0.28	14
14-20 Sep13	-0.68	15	0.16	18	0.13	31	-0.24	17
21-27 Sep13	-0.37	22	0.20	20	-0.37	14	-0.20	23
28 Sep- 5 Oct13	-0.18	26	0.33	37	0.82	47	0.20	35
6-12 Oct13	-0.49	19	0.25	30	-0.28	19	-0.22	20
13-19 Oct13	0.34	34	0.40	41	1.95	51	0.72	46
20-26 Oct13	0.04	28	0.07	10	0.34	37	0.12	31
27 Oct- 2 Nov13	-0.20	25	0.54	45	1.30	49	0.36	40
3-9 Nov13	0.32	32	0.43	44	1.66	50	0.65	43
10-16 Nov13	0.31	31	0.40	40	-0.58	10	0.14	34
17-23 Nov13	0.55	37	0.38	39	0.27	35	0.44	42
24-30 Nov13	1.43	46	0.68	48	0.77	45	1.05	51
1-7 Dec13	1.06	45	0.82	49	-0.37	15	0.67	44
8-14 Dec13	0.94	44	0.07	9	-0.24	21	0.41	41
15-21 Dec13	0.67	41	0.25	28	-0.26	20	0.34	39
22-28 Dec 13	0.62	39	0.11	15	-0.88	6	0.13	33
29 Dec13- 4 Jan14	2.00	51	0.40	42	-0.20	22	1.02	49
5-11 Jan14	0.93	43	0.67	47	-1.69	1	0.27	36
12-18 Jan14	1.71	48	0.94	50	-1.44	2	0.78	48
19-25 Jan14	1.97	50	1.03	51	-0.91	5	1.05	50
26-31 Jan14	1.79	49	0.64	46	-1.22	3	0.78	47

to calculate the overall score for each month (Table 8) and each week (Table 9). The results show that the best air quality was recorded in March with lowest score, then April (Table 8). On the other hand, the highest score or the worst air quality is noticed in January. The parameters loaded on PC2 in March contributed in its lower score. In contrast, the parameters loaded on PC1 and PC2 in January contributed in its highest score.

$$PC\ score = (39.234PC1 + 25.569PC2 + 18.304PC3)/83.107 \quad (3)$$

For weekly scores, the best air quality was found in spring season (Table 9). The week 8-14 March in rank 1 with lowest score, then the week 22-28 March (rank 2) and 5-11 April (rank 3). On the other hand, the worst air quality week was recorded in autumn at the week 24-30 November (rank 51).

Conclusions

Principal component analysis declares the degree of variation in the concentration of air pollutants according to seasons. The pollutants NO_x, NO and SO_2 exhibited more variation in Winter and Spring versus Winter and Summer for NO_2. CO shows more variation in all seasons versus lower variation for CH_4. Additionally, O_3, NMHC and TH exhibits more variation along the year except in Summer for O_3 and Spring for NMHC and TH. The higher variation in CO_2 concentration was recorded in Summer and Autumn versus Winter for PM_{10}. It is worth mentioning that air pollutants of low variation may be at high or low level, but they do not exhibit a significant change along the study period. It was found that the local traffic emissions are responsible for 56.91% in winter to 73.75% in autumn of the variation in air quality in the monitoring site. The results of PCA can be useful for the design of measurement intervals for air pollutants. The pollutants included in PC1 (NO, CO, O_3, CO_2, SO_2) exhibit more variation and must have the intense measurements of low intervals. On the other hand, the parameters included in the other components must have longer interval of measurements as they exhibit lower variation. The temporal assessment of air quality shows that the best air quality occurs in March, while the worst one occurs in January. Principal component analysis has been proved as a useful tool to reduce the dimensionality of a large amount of data to identify the sources of pollution and to find the scores for temporal assessment.

Acknowledgment

Funding: The author is very grateful to Mosul University/ Environmental Research Center for their funding this work and providing the facilities which helped to improve the quality of this work.

Conflict of Interest

The author declares no conflict of interest.

References

1. MANAN N.A., HOD R., SAHANI M., YUSOFF H.M., ISMAIL R., MAHIYUDDIN W.R.W. The impact of air pollution on hospital admissions for cardiovascular and respiratory disease. International Journal of Public Health Research, **6** (1), 707, **2016**.
2. KHAMUTIAN R., NAJAFI F., SOLTANIAN M., SHOKOOHIZADEH M.J., POORHAGHIGHAT S., DARGAHI A., SHARAFI K., AFSHARI A. The association between air pollution and weather conditions with increase in the number of admissions of asthmatic patients in emergency wards: a case study in Kermanshah. Medical Journal of the Islamic Republic of Iran, **29**, 229, **2015**.
3. ISIYAKA H.A., AZID A. Air quality pattern assessment in Malaysia using multivariate techniques. Malaysian Journal of Analytical Sciences, **19** (5), 966, **2015**.
4. AZID A., JUAHIR H., EZANI E., TORIMAN M. E., ENDUT A., RAHMAN M.N.A., YUNUS K., KAMARUDIN M.K., HASNAM C.N.C., SAUDI A.S.M., UMAR R. Identification source of variation on regional impact of air quality pattern using chemometric. Aerosol and Air Quality Research, **15** (4), 1545, **2015**.
5. PIRES J.C.M., PEREIRA M.C., ALVIM-FERRAZ M.C.M., MARTINS F.G. Identification of redundant air quality measurements through the use of principal component analysis. Atmospheric Environment, **43** (25), 3837, **2009**.
6. KUMAR A., GOYAL P. Forecasting of air quality in Delhi using principal component Regression technique. Atmospheric Pollution Research, **2** (4), 436, **2011**.
7. DOMINICK D., JUAHIR H., LATIF M.T., ZAIN S.M., ARIS A.Z. Spatial assessment of air quality patterns in Malaysia using multivariate analysis. Atmospheric Environment, **60**, 172, **2012**.
8. DOGRUPARMAK S.C., KESKIN G.A., YAMAN S., ALKAN A. Using principal component analysis and fuzzy c–means clustering for the assessment of air quality monitoring. Atmospheric Pollution Research, **5** (4), 656, **2014**.
9. PANDEY B., AGRAWAL M., SINGH S. Assessment of air pollution around coal mining area: Emphasizing on spatial distributions, seasonal variations and heavy metals, using cluster and principal component analysis. Atmospheric Pollution Research, **5** (1), 79, **2014**.
10. DE SUN. Air quality assessment in various regions of china based on factor analysis and cluster analysis. Statistics and Application, **8** (2), 234, **2019** [In Chinese].
11. SHIHAB A.S., FADHIL M.N., JUMAA O.KH. Distribution of dustfall in selected sites in Mosul city and around it/ Iraq. Tishreen University Journal for Scientific Researches and Studies, Engineering Series, **32** (6), 50, **2010**.

12. SHIHAB A.S., TAHA A.T. Suspended particulate levels in the left bank residential areas of Mosul city and its relation with some metrological factors. Al-Rafidain Engineering Journal, **22** (3), 158, **2014**.
13. AL-JARRAH O.A.I.. The levels of some air pollutants in selected sites of Mosul city, and its relationship with the traffic volume and meteorological factors. M.Sc. thesis, Mosul University, **2015** [In Arabic].
14. SHIHAB A.S., AL-JARRAH O.A.I. The levels of ozone and nitrogen oxides and its relationship with metrological factors. Al-Rafidain Engineering Journal, **23** (4), 98, **2015** [In Arabic].
15. SHIHAB A.S. Assessment of ambient air quality of Mosul city/Iraq via Air Quality Index. Journal of Ecological Engineering, **22** (10), 241, **2021**.
16. LAU J., HUNG W.T., CHEUNG C.S. Interpretation of air quality in relation to monitoring station's surroundings. Atmospheric Environment, **43**, 769, **2009**.
17. LU W.Z., HE H.D., DONG L.Y. Performance assessment of air quality monitoring networks using principal component analysis and cluster analysis. Building and Environment, **46**, 577, **2011**.
18. OZBAY B. Modeling the effects of meteorological factors on SO_2 and PM10 concentrations with statistical approaches. Clean-Soil, Air, Water **40** (6), 571, **2012**.
19. ÖZDEMIR Ö. Application of multivariate statistical methods for water quality assessment of Karasu-Sarmisakli creeks and Kizilirmak river in Kayseri, Turkey. Polish Journal of Environmental Studies, **25** (3), 1149, **2016**.
20. ANGELLO Z.A.,TRÄNCKNER J., BEHAILU B.M. Spatio-Temporal Evaluation and Quantification of Pollutant Source Contribution in Little Akaki River, Ethiopia: Conjunctive Application of Factor Analysis and Multivariate Receptor Model. Polish Journal of Environmental Studies, **30** (1), 23, 2**021**.
21. SIMEONOV V., EINAX J.W., STANIMIROVA I., KRAFT J. Environmetric modeling and interpretation of river water monitoring data. Analytical and Bioanalytical Chemistry, **374** (5), 898, **2002**.
22. DAVIS J.G. Statistics and data analysis in geology. John Wiley and Sons, Inc. New York, 473, **1973**.
23. UNCUMUSAO ĞLU A.A., MUTLU E. Evaluating spatial and temporal variation in Tuzaklı pond water using multivariate statistical analysis. Polish Journal of Environmental Studies, **28** (5), 3861, **2019**.
24. FIELD A.P. Discovering statistics using IBM SPSS Statistics. North American Edition, 5th ed. SAGE publishing, University of Sussex, UK, **2017**.
25. KHAN R.A., ZAIN SH.M., JUAHIR H., YUSOFF M.K. AND TG HANIDZA T.I. Using Principal Component Scores and Artificial Neural Networks in Predicting Water Quality Index; Chemometrics in Practical Applications, Edited by Kurt Varmuza, Chapter **12**, 271, **2012**.
26. LEE D.H., AN S.S., SONG H.M., PARK O.H., PARK K.S., SEO G.Y., CHO Y.G., KIM E.S. The Effect of traffic volume on the air quality at monitoring sites in Gwangju. Journal Environmental Health Science, **40** (3), 204, **2014** [In Korean].
27. KOVAC-ANDRIC E., RADANOVIC T., TOPALOVIC I., MARKOVIC B., SAKAC N. Temporal variations in concentrations of ozone, nitrogen dioxide, and carbon monoxide at Osijek, Croatia. Advances in Meteorology, **2013**, 1, **2013**.
28. KIM S.U., KIM K.Y. Physical and chemical mechanisms of the daily-to-seasonal variation of PM10 in Korea. Science of the Total Environment **712**, **2020**.
29. DENG B., LI Q., CHEN Y ., LI M., LIU A., RAN J., XU Y., LIU X., FU J., FENG R. The effect of air/fuel ratio on the CO and NOx emissions for a twin-spark motorcycle gasoline engine under wide range of operating conditions. Energy, **169** (C), 1202, **2019**.
30. ZAKRI N.L., SAUDI A.S.M., JUAHIR H., TORIMAN M.E., ABU I.F., MAHMUD M.M., KHAN M.F. Identification source of variation on regional impact of air quality pattern using chemometric techniques in Kuching, Sarawak. International Journal of Engineering & Technology, **7** (3.14) 49, **2018**.
31. MUTALIB S.N.S.A., JUAHIR H., AZID A., SHARIF S.M., LATIF M.T., ARIS A.Z., ZAIN S.M., DOMINICK D. Spatial and temporal air quality pattern recognition using environmetric techniques: A case study in Malaysia. Environmental Science: Processes & Impacts **15** (9), 1717, **2013**.
32. RAHMAN S.R.A., ISMAIL S.N.S., RAMLI M.F., LATIF M.T., ABIDIN E.Z., PRAVEENA S.M. The assessment of ambient air pollution trend in Klang Valley, Malaysia. World Environment, **5** (1), 1, **2015**.
33. COOPER C.D., ALLEY F.C. Air Pollution Control: A Design Approach, 4th ed.; Waveland Press Inc., Long Grove, Illinois, USA, 573, **2014**.
34. JENKIN M.E., CLEMITSHAW K.C. Ozone and other secondary photochemical pollutant: chemical processes governing their formation in the planetary boundary layer. Atmospheric Environment, **34** (16), 2499, **2000**.
35. VOUKANTSIS D., KARATZAS K., KUKKONEN J., RÄSÄNEN T., ARPPINEN A., KOLEHMAINEN M. Intercomparison of air quality data using principal component analysis, and forecasting of PM_{10} and $PM_{2.5}$ concentrations using artificial neural networks, in Thessaloniki and Helsinki. Science of the Total Environment, **409**, 1266, **2011**.
36. LEVY J.I., BENNETT D.H., MELLY S.J., SPENGLER J.D. Influence of traffic patterns on particulate matter and polycyclic aromatic hydrocarbon concentrations in Roxbury, Massachusetts. Journal of Exposure Analysis and Environmental Epidemiology **13** (5), 364, **2003**.
37. KIM S.J., PARK M.K., LEE S.E., GO H.J., CHO B.C., LEE Y.S., CHOI, S.D. Impact of traffic volumes on levels, patterns, and toxicity of polycyclic aromatic hydrocarbons in roadside soils. Environmental Science: Processes & Impacts, **21**, 174, **2019**.
38. AL-JEELANI H.A. Diurnal and Seasonal Variations of Surface Ozone and Its Precursors in the Atmosphere of Yanbu, Saudi Arabia. Journal of Environmental Protection, **5** (5), 408, **2014**.
39. JIMENEZ J., LEI H., STEYER J.P., HOUOT S., PATUREAU D. Methane production and fertilizing value of organic waste: organic matter characterization for a better prediction of valorization pathways. Bioresource Technology, **241**, 1012, **2017**.
40. SUBRAMANI T., VAN S.A., GANESAN S.K. Analysis of highway air pollution. International Journal of Engineering Research and Applications. **4** (6), 173, **2014**.
41. DARGAHI A., DEHGHANZADEH R., FAHIMINIA V., JABBARI Y., AZIZI, F. Studying Air Quality Changes in Tabriz in Terms of PM10 Pollutant Density Using AQI Index and its Relation with Drop of Water Level in Uremia Lake during 2008-2011. Journal of Environmental Science Technology, **18** (2), 55, **2016** [In Persian].
42. PIRSAHEB M., DARGAHI A., ASASDI F., FAHIMINIA V., AZIZI F., NAVAZESHKHAH F. The quality of air

pollutant concentrations in Kermanshah emphasizing PM10 due to the occurrence of dust (2011-2013). Journal of Kermanshah University of Medical Sciences, **18** (11), e74002, **2015** [In Persian].

43. ESLAMI F., SALARI M., DEHGHANI M.H., DARGAHI A., NAZMARA S.H., YAZDANI M., BEHESHTI A. Relationship of Formaldehyde Concentration in Ambient Air with CO, NO_2, O_3, Temperature and Humidity: Modeling by Response Surface Model. Arch Hygiene Science **8** (1), 9, **2019**.