Jurnal Teknologi

Full paper

Source Apportionment of Air Pollution: A Case Study In Malaysia

Azman Azid[a,*], Hafizan Juahir[a], Mohd Ekhwan Toriman[a], Azizah Endut[a], Mohd Khairul Amri Kamarudin[a], Mohd Nordin Abdul Rahman[b], Che Noraini Che Hasnam[a], Ahmad Shakir Mohd Saudi[a], Kamaruzzaman Yunus[c]

[a]*East Coast Environmental Research Institute (ESERI), Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Kuala Terengganu, Terengganu, Malaysia*
[b]*Centre for Research & Innovation Management (CRIM), Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Kuala Terengganu, Terengganu, Malaysia*
[c]*Kulliyyah of Science, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia*

*Corresponding authors: azmanazid@unisza.edu.my

Article history

Received: 9 May 2014
Received in revised form:
27 October 2014
Accepted: 1 December 2014

Abstract

Air pollution is becoming a major environmental issue in Malaysia. This study focused on the identification of potential sources of variations in air quality around the study area based on the data obtained from the Malaysian Department of Environment (DOE). Eight air quality parameters in ten monitoring stations for seven years (2006 – 2012) were gathered. The Principal Component Analysis (PCA) method from chemometric technique was applied to identify the source identification of pollution around the study area. The PCA method has identified methane (CH_4), non-methane hydrocarbon (NmHC), total hydrocarbon (THC), ozone (O_3) and particulate matter under 10 microns (PM_{10}) are the most significant parameters around the study area. From the study, it can be concluded that the application of the PCA method in chemometric techniques can be applied for the source apportionment purpose. Hence, this study indicated that for the future and effective management of the Malaysian air quality, an effort should be placed as a priority in controlling point and non-point pollution sources.

Keywords: Air pollution; principal component analysis; chemometric; source apportionment

1.0 INTRODUCTION

Air pollution is a serious issue that needs to be given immediate and serious attention by all relevant authorities around the globe, as it is one of the most important factors that contributes to the quality of life and living. Air pollution is becoming a major environmental issue in Malaysia due to the increasing number of transportations (mobile sources), trans-boundary pollution from neighbouring countries and the industrial activities (stationary sources), and they are the main sources of air pollution in Malaysia.[1] The effect of air pollution may cause acute and chronic to humans or other living organisms, and cause damage to the natural environment or built environment, when enter into the atmosphere.[2] Symptoms such as nose, throat, eye and skin irritation, headache, fatigue, dizziness, and difficulty in breathing are general health effect experienced by human due to poor of air quality.[3]

Controlling the source of air pollutants is one of the major challenges in the world. In Malaysia, the Malaysian Department of Environment (DOE) has been consistently monitored air quality status and collecting data in order to inform people about major pollutant concentrations in real time.[4] Once the lack of compliance is determined, the data can be used to advise or caution the decision makers or planners in lieu of health effects.[1,5] Two major of air pollutants are PM_{10} and surface ozone (O_3), particularly in the urban and suburban areas in Malaysia[1,6,7] and has been recognised as one of the major concerns that have high potential for deleterious effects on human health.[8,9,10,11]

The chemometric techniques (also known as multivariate techniques) believed as a better tool for analysing air quality. Chemometric in the environmental field is verified to be a functional tool to identify the sources of pollution.[1,12] Chemometric methods also offer the recognition of the potential sources that are accountable for variations in air quality and manipulate the air quality. Therefore, the methods have been proven as priceless tools for developing suitable plans for efficient management of the air monitoring network.[13] Purposely, this study is to identify the potential sources of variations in air quality around the study area.

2.0 EXPERIMENTAL

Study Area

Ten continuous air monitoring stations were selected. The stations are Pasir Gudang (ST01: N01° 28.225, E103° 53.637), Kuching (ST02: N01° 33.734, E110° 23.329), Bukit Rambai (ST03: N02° 15.510, E102° 10.364), Tasek (ST04: N04° 37.781, E101° 06.964), Nilai (ST05: N02° 49.246, E101° 48.877), Klang (ST06: N03° 00.620, E101° 24.484), Balok Baru (ST07: N03° 57.726, E103° 22.955), Pengkalan Chepa (ST08: N06° 09.520, E102° 17.262), Paka (ST09: N04° 35.880, E103° 26.096), and Labuan (ST10: N05° 19.980, E115° 14.315). The locations of the air quality monitoring stations are shown in **Figure 1**. Eight stations are located in the Peninsular Malaysia and another two are in East Malaysia. These stations were selected due to their location differences, which lies in the heavily industrial areas, residential areas and surrounded by congested main roads. Based on the DOE report, the overall status of air quality in Malaysia within good and moderate levels most of the time.[14] There are no major natural disaster (such as typhoon, volcanic eruption and earthquake) occurrences in these areas. The value of the air pollution index (API) in Malaysia is usually influenced by the concentration of suspended particulate matter (PM_{10})[15] because of the concentration value of PM_{10} is always higher than other pollutants.[16]

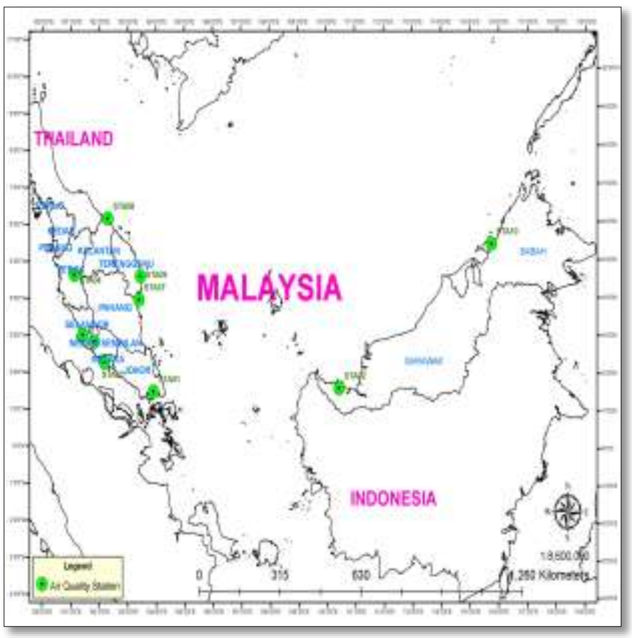

Fig 1 Location of the ten selected air quality monitoring stations in Malaysia

Data collection

The air quality data were gathered from the Air Quality Division in the Department of Environment (DOE) Malaysia. The data were collected and monitored by Alam Sekitar Malaysia Sdn. Bhd. (ASMA), the authorized agency for DOE. All stations were identified based on the availability of data start from January 1, 2006 to December 31, 2012. The air quality variables used in this study are carbon monoxide (CO), ozone (O_3), particulate matter under 10 microns (PM_{10}), sulphur dioxide (SO_2), nitrogen dioxide (NO_2), methane (CH_4), non-methane hydrocarbon (NmHC) and total hydrocarbon (THC). The measurements recorded for these variables are hourly. The equipment used by ASMA to monitor the air quality data are from Teledyne Technologies Inc. USA, and Met One Instrument Inc. USA. Based on the Standard Operating Procedures for Continuous Air Quality Monitoring (2007),[17] the analyzers used by ASMA to monitor PM_{10} using a BAM-1020 Beta Attenuation Mas Monitor from Met One Instrument, Inc. USA. This instrument has a fairly high resolution of 0.1 μg m^{-3} at a 16.7 L min^{-1} flow rate, with lower detection limits of <4.8 μg m^{-3} and <1.0 μg m^{-3} for 1 h and 24 h, respectively.. The instruments used by ASMA to monitor SO_2, CO and O_3 were the Teledyne API Model 100A/100E, Teledyne API Model 200A/200E, Teledyne API Model 300/300E and Teledyne API Model 400/ 400E, respectively from Teledyne Technologies Inc., USA. SO_2 measurement was based on the UV fluorescence method, where the lowest level of detection is at 0.4 ppb. CO was measured using the non-dispersive, infrared absorption (Beer Lambert) method with 0.5% precision and the lowest detection of 0.04 ppm. While, O_3 was measured through the UV absorption (Beer Lambert) method with a detection limit of 0.4 ppb. The measurements of SO_2, CO and O_3 were at a precision level of 0.5%. For THC, CH_4 and NmHC, the analyzer used by ASMA were measured using a Teledyne API M4020 from Teledyne Technologies Inc., USA, which equipped with aflame-ionization detector (FID) and a measurement accuracy of 1%. These instruments were used due to well-proven accuracy, reliability, and robustness.

Data pre-treatment

A total of 202,080 data points (8 variables x 25,260 data set) was utilized in this analysis. The total number of missing data in the data points was very small (~3%) from the overall data. Based on data sets provided by DOE, there are some data, such as O_3 and CO in a certain stations are not available. In order to facilitate the unavailable or missing data, the nearest neighbour method [18,19] was applied and computations performed using XLSTAT 2014 add-in software. The nearest neighbour method was based on the endpoints of the gaps using **Equation 1**:

$$y = y_1 \text{ if } x \le x_1 + [(x_2 - x_1)/2] \text{ or}$$
$$y = y_2 \text{ if } x > x_1 + [(x_2 - x_1)/2] \quad (1)$$

where; y is the interpolant, x is the time point of the interpolant, y_1 and x_1 are the coordinates of the starting point of the gap, and y_2 and x_2 are the endpoints of the gap.

Principal component Analysis

Dimension of a huge data set can be trimmed down by using principal component analysis (PCA), which it is considered as one of the most prevalent and useful statistical methods for uncovering the potential structure of a set of variables. This method used for explaining the variance of a large set of interrelated variables by transforming them into new, smaller set of uncorrelated (independent) variables, namely as principal components (PCs).[1,6,20] PCs are orthogonal and uncorrelated to each other and have linear combinations of the original variables.[21,22,23,24] PCA has the ability to show the most significant variables which can indicate the source of the pollutants. This is because, in the analysis process the variables that are less significant are omitted from the data set with a minimal loss of original data.[13,25,26] A total of 25,260 data sets and 8 air quality variables were used in this study. The raw air quality variables were standardized through Z-scale transformation to a mean of 0.0 and variance of 1.0 by applying the **Equation 2**:

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

where Z_{ij} is the *j*th value of the standard score of the measured variable *i*; X_{ij} is the *j*th observation of variable *i*; μ is the variable's mean value; and σ is the standard deviation. *Z*-scale transformation method was used to ensure the different air quality variables had equal weights in the statistical analysis process. Besides, these transformations will homogenize the variance of the distribution and prevent any classification errors that may occur from groups described by variables of completely different sizes.[12] Then, the data matrix was decomposed into scores or components and loadings (correlations between the original variables and the PCs extracted by the analysis) for the variables.

The Barlett's test of sphericity was performed at the beginning of the PCA in order to examine the correlation of the variables used in the PCA.[27] The null hypothesis, H_0 of this test states that there is no correlation significant difference from 0 between the variables. While the alternative hypothesis, H_a state that at least one of the correlations between the variables is significantly different from 0. As the computed *p*-value is lower than the significance level alpha=0.05, one should reject the null hypothesis H_0, and accept the alternative hypothesis H_a. The risk to reject the null hypothesis, H_0 while it is true is lower than 0.01%. When the null hypothesis, H_0 result is rejected, then it is confirmed that the variables used in the PCA are correlated.[27]

The PCs generated by PCA sometimes are not readily interpreted and should be rotated using any of a number of applicable methods, e.g.; varimax rotation. The varimax rotation goal is to minimize the complexity of the components by making the large loadings larger and the small loadings smaller within each component. The varimax rotation method was applied because this method simplifies the factor structure and therefore makes its interpretation easier and more reliable. In the varimax rotation method, only the PCs with eigenvalues of more than one (>1.0) are used and considered significant[28] in order to obtain the new variables, known as varifactors (VFs) or factor loadings. This approach is known as Kaiser Criterion. Kaiser Criterion is used to solve the problem of the number of components to be retained.[29] The numbers of VFs obtained by varimax rotations are equal to the number of variables in accordance with common features and can include unobservable, hypothetical and latent variables. The VFs are values that use to measure the correlation between variables. VFs values which are greater than 0.75 (> 0.75) is considered as "strong", the values range from 0.50-0.75 ($0.50 \geq$ factor loading ≥ 0.75) is considered as "moderate" and the values range from 0.30-0.49 ($0.30 \geq$ factor loading ≥ 0.49) is considered as "weak" factor loadings.[30] In this study, the VFs with absolute values greater than 0.75 was set as the selection threshold. Then, the results of factor scores after varimax rotation was used for artificial intelligence modelling. The PCA was examined using XLSTAT 2014 add-in software.

3.0 RESULTS AND DISCUSSION

A seven-year daily average secondary data was used in this study. The database consists of eight air pollutant variables and the air pollutant index (API). The overall descriptive statistics of the air pollutant variables and the API are summarized in **Table 1**. The average values of all the parameters are within the value of the Recommended Malaysian Air Quality Guideline (RMAQG). This means that the air quality in Malaysia is still in controlled conditions. Although the average values for all parameters at each station is under RMAQG permitted levels, but there are some stations such as Pasir Gudang (STA01), Bukit Rambai (STA03), Tasek (STA04), Klang (STA06), Balok Baru (STA07), and Paka (STA09) were exposed to the maximum amount of O_3 values higher than the level allowed by RMAQG. This condition is caused by these stations are located in the urban and industrial area. Besides, all the stations are also facing a number of the maximum PM_{10} values which exceeded permitted level by RMAQG (> 150). The highest value for PM_{10} was recorded in STA01 (Pasir Gudang), STA06 (Klang), and STA07 (Balok Baru), which are the value of 780, 732, and 760 respectively. The other stations such as Kuching (STA02), Bukit Rambai (STA03), Tasek (STA04), Nilai (STA05), Pengkalan Chepa (STA08), Paka (STA09), and Labuan (STA10) also recorded the highest PM_{10} values of 384 μg m^{-3}, 427 μg m^{-3}, 202 μg m^{-3}, 385 μg m^{-3}, 430 μg m^{-3}, 470 μg m^{-3}, and 357 μg m^{-3}, respectively.. The results show that STA06 (Klang) recorded the highest concentration of PM_{10}, NmHC, CO, THC, and NO_2. STA04 (Tasek) recorded the highest concentration of O_3, STA05 (Nilai) recorded the highest concentration values for CH_4.and the highest concentration of SO_2 was recorded at STA09 (Paka). The mean range for PM_{10} was 53.70μg $m^{-3} \leq PM_{10} \leq$ 98.15μg m^{-3} and CO mean range was 0.57ppm $\leq$ CO $\leq$ 1.24ppm. For O_3 and SO_2, the mean range was 0.02ppm $\leq O_3 \leq$ 0.03ppm and 0.00ppm $\leq SO_2 \leq$ 0.37ppm, respectively. For NmHC and THC, the mean range was 0.40ppm $\leq$ NmHC $\leq$ 0.80ppm and 2.54ppm $\leq$ THC $\leq$ 3.53ppm, respectively. The NO_2 mean range recorded as 0.01ppm $\leq NO_2 \leq$ 0.03ppm, while CH_4 mean range recorded as 2.20ppm $\leq CH_4 \leq$ 2.97ppm.

The Bartlett's test of sphericity revealed that the air quality data met the sphericity assumption since it had an observed chi-square value of 185152.046 ($p < 0.05$, df = 28), therefore confirming that the air quality variables were correlated and not orthogonal. This suggests that PCA will allow for interpretation of the variability in the data with less than the original number of variables.[31]

The estimation of the factor loadings was carried out for assessing the correlations between air quality variables and the extracted factors. After varimax rotation, from eight PCs, there are only two VFs which represent 64.24% of the variance of the data were selected due to the eigenvalues larger than one (> 1.0). Despite of the cumulative variance is less than 70%, the cut-off point of the factors were determined using scree plot graph (**Figure 2**). The eigenvalues with lower than one (<1.0) are neglected because of redundant with more important factors. It means that multicollearity was present among original variables. In this study, the VFs with absolute values greater than 0.75 was set as the selection threshold, because these values are solid and stable, which exhibit moderate to strong loadings on the extracted factors. **Table 2** and **Figure 3** highlights that five (5) out of eight (8) air quality variables used in this study satisfy the 0.75 factor loadings threshold. These variables are CH_4, NmHC, THC, O_3 and PM_{10}. These pollutants are then classified as the potential contributor pollutants in the selected monitoring stations in Malaysia.

Table 1 Overall descriptive statistics of daily average air quality and air pollutant index (API) in the study areas, 2006-2012

Station	Statistic	Parameter								
		CO (ppm)	O_3 (ppm)	PM_{10} (µg m^{-3})	SO_2 (ppm)	NO_2 (ppm)	CH_4 (ppm)	NmHC (ppm)	THC (ppm)	API
STA01	No. of observations	2526	2526	2526	2526	2526	2526	2526	2526	2526
	Minimum	0.00	0.00	19.86	0.00	0.00	0.00	0.00	0.00	2.13
	Maximum	4.85	0.12	780.00	0.13	0.06	9.75	5.15	10.50	125.88
	Mean	**1.24**	**0.03**	**81.24**	**0.01**	**0.02**	**2.49**	**0.55**	**2.96**	**57.84**
STA02	No. of observations	2526	2526	2526	2526	2526	2526	2526	2526	2526
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	14.00
	Maximum	3.79	0.08	384.00	0.10	0.12	8.32	4.54	8.90	173.13
	Mean	**0.74**	**0.02**	**57.76**	**0.01**	**0.01**	**2.32**	**0.45**	**2.69**	**40.57**
STA03	No. of observations	2526	2526	2526	2526	2526	2526	2526	2526	2526
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.04
	Maximum	3.77	0.12	427.00	0.10	0.05	7.53	8.01	9.80	125.88
	Mean	**1.10**	**0.03**	**95.02**	**0.01**	**0.02**	**2.15**	**1.10**	**3.14**	**57.87**
STA04	No. of observations	2526	2526	2526	2526	2526	2526	2526	2526	2526
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Maximum	2.84	0.16	202.00	0.10	0.06	9.33	4.81	9.60	158.00
	Mean	**0.86**	**0.04**	**58.70**	**0.01**	**0.02**	**2.91**	**0.41**	**3.24**	**50.14**
STA05	No. of observations	2526	2526	2526	2526	2526	2526	2526	2526	2526
	Minimum	0.00	0.00	19.29	0.00	0.00	0.00	0.00	0.00	21.96
	Maximum	3.11	0.12	385.00	0.09	0.08	8.03	4.80	9.40	176.42
	Mean	**0.83**	**0.03**	**89.56**	**0.01**	**0.02**	**2.97**	**0.48**	**3.36**	**57.48**
STA06	No. of observations	2526	2526	2526	2526	2526	2526	2526	2526	2526
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Maximum	10.52	0.12	732.00	0.11	0.13	9.75	3.97	10.60	494.88
	Mean	**1.75**	**0.03**	**98.15**	**0.01**	**0.03**	**2.85**	**0.80**	**3.53**	**60.97**
STA07	No. of observations	2526	2526	2526	2526	2526	2526	2526	2526	2526
	Minimum	0.00	0.00	27.00	0.00	0.00	0.00	0.00	0.00	3.63
	Maximum	3.82	0.12	760.00	0.06	0.06	6.40	6.17	8.20	151.00
	Mean	**0.99**	**0.02**	**94.66**	**0.01**	**0.01**	**2.24**	**0.58**	**2.75**	**57.32**
STA08	No. of observations	2526	2526	2526	2526	2526	2526	2526	2526	2526
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Maximum	4.65	0.07	430.00	0.10	0.06	6.89	8.01	9.80	77.83
	Mean	**0.93**	**0.02**	**68.90**	**0.01**	**0.01**	**2.38**	**0.56**	**2.85**	**46.00**
STA09	No. of observations	2526	2526	2526	2526	2526	2526	2526	2526	2526
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Maximum	4.17	0.12	470.00	47.96	0.06	9.39	4.54	10.10	94.88
	Mean	**0.79**	**0.02**	**53.70**	**0.37**	**0.01**	**2.33**	**0.46**	**2.72**	**37.70**
STA10	No. of observations	2526	2526	2526	2526	2526	2526	2526	2526	2526
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.63
	Maximum	3.32	0.06	357.00	0.04	0.04	6.64	4.54	7.60	97.00
	Median	0.52	0.02	51.00	0.00	0.01	2.03	0.32	2.33	37.00
	Mean	**0.57**	**0.02**	**55.56**	**0.00**	**0.01**	**2.20**	**0.40**	**2.54**	**38.41**
	Standard deviation (n-1)	0.33	0.01	29.86	0.01	0.01	0.57	0.35	0.70	11.50
	Averaging time	1hr	1hr	24hrs	1hr	1hr	1hr	1hr	1hr	
	RMAQG	30.00	0.10	150.00	0.13	0.17				

The VF1 contributes 32.329% of the variation in the air quality data. It has high loadings from three variables, which are CH_4 (0.949), NmHC (0.787) and THC (0.988). This factor can be interpreted as a potential of gaseous pollutants. Considering the nature of these three air quality variables, this factor is mostly probably related to the processes of petrochemical production from petrochemical industries and the fuel combustion from transportation activities.[32] Besides, it is also probably related to the process of biomass burning, grazing and residual of agricultural product from agricultural activities.[33]

The VF2 demonstrates 31.912% of the variance in the data. It exhibits high loading from O_3 (0.753) and PM_{10} (0.838). The concentration of these pollutants is potentially related to the secondary pollutant (O_3) and non-gas pollutant (PM_{10}). O_3 released into the atmosphere as a result of photochemical oxidation and the main component of smog.[34] The concentration of O_3, especially in urban and suburban is probably contributed by the mono-nitrogen oxide (NO_x)[35] and the downwind plume of O_3 precursors from the industrial activities.[36,37] PM_{10} is the main component of dust fall, which it potentially comes from the industrial activities and construction sites,[38] the transportation exhaust emission and soil dust[39] and also open burning activity around the study area. According to the Malaysian Ministry of Transport (MOT),[40] the total amount of new registered motor vehicles in Malaysia was increased 4.42% from 934,367 in 2004 to 1,160,082 in 2010. Based on this information, motor vehicles in Malaysia are one of the major factors that contribute to the deterioration of atmospheric conditions.

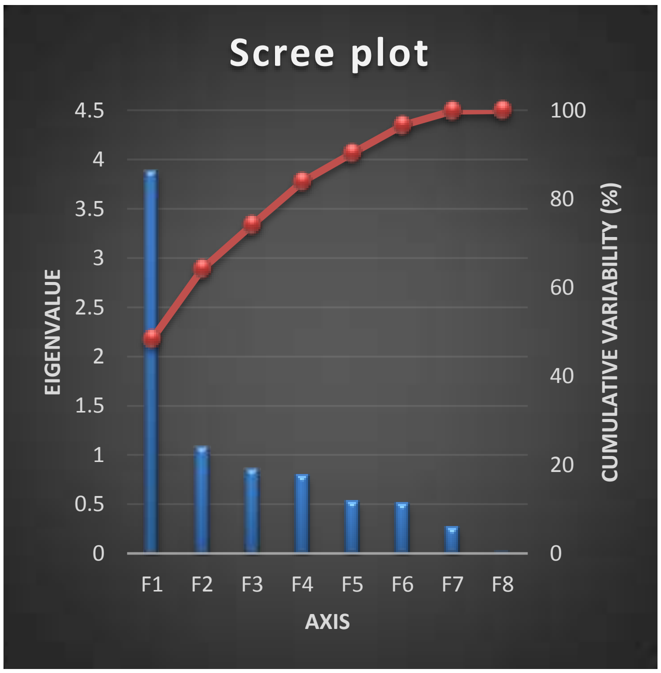

Fig. 2 Screen plots for PCA

Table 2 Varifactors after varimax rotation and the possible source category in the study area

Variable	VF1	VF2
CO	0.217	0.687
O_3	0.152	**0.753**
PM_{10}	-0.123	**0.838**
SO_2	-0.010	-0.007
NO_2	0.062	0.664
CH_4	**0.949**	0.015
NmHC	**0.787**	0.261
THC	**0.988**	0.063
Eigenvalue	3.118	2.021
Variability (%)	32.329	31.912
Cumulative %	32.329	64.241

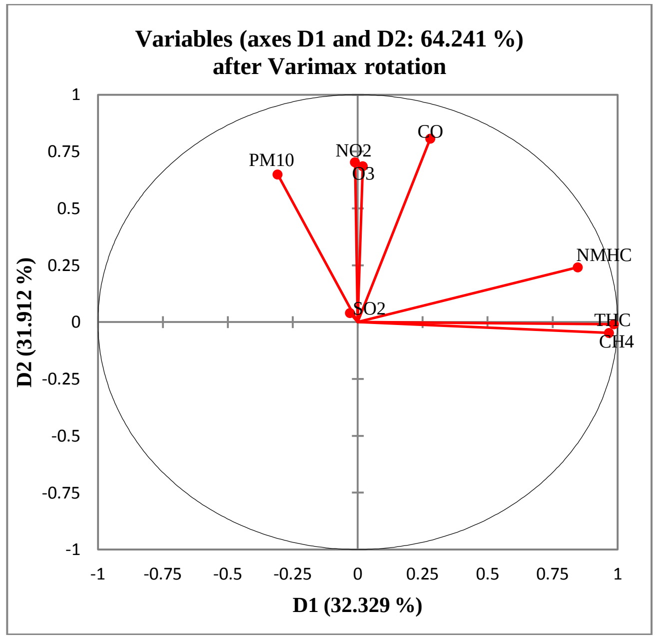

Fig 3 Factor loading plot after varimax rotation

4.0 CONCLUSION

Air quality monitoring programs have generated huge, multidimensional and complex data set, which require chemometric techniques for data analysis and interpretation of the underlying information. In this study, we applied the method of PCA to identify the pollution sources for air quality variation in a certain area in Malaysia even without field visit. The two VFs generated by rotated PCA indicate that the parameters such as CH_4, NmHC, THC, O_3 and PM_{10} are responsible for air quality variations in the study area. Based on the Malaysian Ministry of Transport data, it is believed that motor vehicles are one of the major factors that contribute to the formation of these pollutants. Thus, this study indicated that for the future and effective management of the Malaysian air quality, an effort should be placed as a priority in controlling point and non-point pollution sources.

Acknowledgement

The authors acknowledge the Air Quality Division from Department of Environment (DOE), Malaysian Ministry of Natural Resource and Environment; Professor Dr. Mohd Talib Latif from Faculty of Science and Technology, Universiti Kebangsaan Malaysia; and Professor Dr. Sharifuddin Mohd Zain from Chemistry Department, Universiti Malaya for their permission to utilize air quality data, advice, guidance and support for this study.

References

[1] S.N.S.A. Mutalib, H. Juahir, A. Azid, S.M. Sharif, M.T. Latif, A.Z. Aris, S.M. Zain, D. Dominick. 2013. *Sci.: Processes & Impacts.* 15: 1717–1728. DOI: 10.1039/c3em00161j.

[2] K.P. Moustris, I.C. Ziomas, A.G. Paliatsos. 2010. *Water, Air & Soil Pollution.* 209(1): 29–43. DOI: 10.1007/s11270-009-0179-5.

[3] H. Xie, F. Ma, Q. Bai. 2009. *Fifth International Conference on Natural Computation*. 412–418.

[4] E.G. Dragomir. 2010. *Series Mathemathics, Informatics, Physics*. 1: 103–108.

[5] M.M Kamal, R. Jailani, R.L.A. Shauri. 2006. *4th Student Conference on Research and Development*. 115–119. DOI: 10.1109/SCORED.2006.4339321.

[6] D. Dominick, H. Juahir, M.T. Latif, S.M. Zain, A.Z. Aris. 2012. *Atmospheric Environment*. 60: 172–181. DOI: 10.1016/j.atmosenv.2012.06.021.

[7] M.T. Latif, L.S. Hey, L. Juneng. 2012. *Atmospheric Environment.* 61 : 434–445. DOI: 10.1016/j.atmosenv.2012.07.062.

[8] W.R.W Mahiyudin, M. Sahani, R. Aripin, M.T. Latif, T.Q. Thach, C.M. Wong. 2013. *Atmospheric Environment*. 65: 69–79. DOI: 10.1016/j.atmosenv.2012.10.019.

[9] S. Mazrura, W.R. Wan Mahiyuddin, R. Aripin, M.T. Latif, TQ. Thach, C-W. Wong. 2011. *Epidemiology*. 22(1): 159.

[10] L.Y. Siew, L.Y. Chin, P.M.J. Wee. 2008. *The Malaysian Journal of Analytical Sciences*. 12(1): 257–263.

[11] R. Afroz, M.N. Hassan, N.A. Ibrahim. 2003. *Environmental Research.* 92: 71–77. DOI: 10.1016/S0013-9351(02)00059-2.

[12] V. Simeonov, J.W. Einax, I. Stanimirova, J. Kraft. 2002. *Analytical and Bioanalytical* Chemistry. 374: 898–905.

[13] K.P. Singh, A. Malik, S. Sinha. 2005. *Analytica Chimica Acta.* 538: 355–374. DOI: 10.1016/j.aca.2005.02.006.

[14] Department of Environment Malaysia (DOE). 2010. *Malaysia Environmental Quality Report, 2009*. Kuala Lumpur: Ministry of Science, Technology and Environment.

[15] M.B. Awang, A.B. Jaafar, A.M. Abdullah, M.B. Ismail, M.N. Hassan, R. Abdullah, S. Johan, H. Noor. 2000. *Respirology*. 5: 183–196. PMID: 10894109 [PubMed – indexed for MEDLINE].

[16] N. Othman, M.Z. Mat Jafri, L.H. San. 2010. *Morden Applied Science*. 4(11): 131–142.

[17] Alam Sekitar Malaysia Sdn Bhd (ASMA). 2007. *Standard Operating Procedure for Continuous Air Quality Monitoring.* Shah Alam, Selangor Malaysia.

[18] H. Junninen, H. Niska, K. Tuppurainen, J. Ruuskanen, M. Kolehmainen. 2004. *Atmospheric Environment*. 38: 2895–2907. DOI: 10.1016/j.atmosenv.2004.02.026.

[19] A. Azid, H. Juahir, M.T. Latif, S.M. Zain, M.R. Osman. 2013. *J. Env. Prot.* 4(12A): 1–10. doi: 10.4236/jep.2013.412A001.

[20] V. Simeonov, J.A. Stratis, C. Samara, G. Zachariadis, D. Voutsa, A. Anthemidis, M. Sofoniou, T. Kouimtzis. 2003. *Water Research*. 37: 4119-4224.

[21] S.A. Abdul-Wahab, S.M. Alawi. 2002. *Environmental Modelling and Software*. 17: 219–228.

[22] M. Viana, X. Querol, A. Alastuey, J.I. Gil, M. Menéndez. 2006. *Chemosphere*. 65: 2411–2418.

[23] C.B. Skrbi, S.I. Duri, C.N. C-Mladenovi. 2007. *Chemosphere*. 68: 2144–2152.

[24] L. Juneng, M.T. Latif, F.T Tangang, H. Mansor. 2009. *Atmospheric Environment.* 45: 4370–4378.

[25] K.P. Singh, A. Malik, D. Mohan, S. Sinha. 2004. *Water Research*. 38: 3980–3992. DOI: 10.1016/j.watres.2004.06,011.

[26] S. Shrestha, F. Kazama. 2007. *Environmental Modelling and Software*. 22: 464–475. DOI: 10.1016/j.envsoft.2006.02.001.

[27] B.G. Tabachnick, L.S. Fidell. 2013. *Using Multivariate Statistics*, 6th ed., Pearson, Boston.

[28] J.O. Kim, C.W. Mueller. 1987. *Introduction to factor analysis: What it is and how to do it.* Quantitative applications in the social science series. Sage University Press, Newbury Park.

[29] H.F. Kaiser. 1958. *Psychometrika*. 23(3): 187–200. DOI: 10.1007/BF02289233.

[30] C.W. Liu, K.H. Lin, Y.M. Kuo. 2003. *Sci. Tot. Environ.* 313: 77–89. DOI: 10.1016/S0048-9697(02)00683-6.

[31] V.H. McNeil, M.E. Cox, M. Preda. 2005. *Journal of Hydrology*. 310 (1–4): 181–200.

[32] R. Koppmann. 2007. *Volatile organic compounds in the atmosphere.* Blackwell Publishing Ltd, ISBN 978-1-4051-3115-5, Singapore.

[33] I. Haiduc, M.S. Beldean-Gale. 2011. http://www.intechopen.com/books/air-qualitymodels-and-applications/variation-of-greenhouse-gases-in-urban-areas-case-study-co2-co-and-ch4-in-threeromanian-cities. Accessed April 2, 2014.

[34] N. Banan, M.T. Latif, L. Juneng, F. Ahamad. 2013. *Aerosol and Air Quality Research.* 13: 1090–1106. DOI: 10.4209/aaqr.2012.09.0259.

[35] Y. Sadanaga, M. Sengen, N. Takenaka, H. Bandow. 2012. *Aerosol Air Qual. Res.* 12: 161–168. *Building and Environment.* 46: 577-583.

[36] X. Wei, Q. Liu, K.S. Lam, T. Wang. 2012. *Adv. Atmos. Sci.* 29: 635–645.

[37] A. Monteiro, A. Strunk, A. Carvalho, O. Tchepel, A.I. Miranda, C. Borrego, S. Saavedra, A. Rodríguez, J. Souto, J. Casares, E. Friese, H. Elbern. 2012. *Environ. Pollut.* 162: 176–189.

[38] S. Pandey, B. Tripathi, V. Mishra. 2008. *J. Environ. Manage.* 86: 132–138.

[39] C. Arsene, R.I. Olariu, N. Mihalopoulos. 2007. *Atmos. Environ.* 41: 9452–9467.

[40] Ministry of Transport Malaysia (MOT). 2010. *Malaysia Transport Statistics Report, 2010*. Putrajaya, Malaysia.