Journal of Fundamental and Applied Sciences

ISSN 1112-9867

Available online at http://www.jfas.info

Research Article

Special Issue

SELECTED MALAYSIA AIR QUALITY POLLUTANTS ASSESSMENT USING CHEMOMETRICS TECHNIQUES

N. L. A. Rani[1], A. Azid[1,2,*], S. I. Khalit[1,2], M. B. Gasim[1,3] and H. Juahir[1,3]

[1]Faculty of Bioresources and Food Industry, Universiti Sultan Zainal Abidin, 22200 Besut, Terengganu, Malaysia

[2]UniSZA Science and Medicine Foundation Centre, Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Kuala Nerus, Terengganu, Malaysia

[3]East Coast Environmental Research Institute (ESERI), Universiti Sultan Zainal Abidin, Gong Badak Kampus, 21300 Kuala Nerus, Terengganu, Malaysia

Published online: 08 August 2017

ABSTRACT

Air quality played an important role as polluted air quality could harm human health, environment as well as property. Thus, a study of air quality pollutants assessment using chemometrics was performed with the objectives to ensure the air quality data analysis is valid, acceptable and interpreted well. Analysis of PCA, FA, KMO and Bartlett's test were done on five main air quality pollutants (O_3, NO_2, SO_2, CO and PM_{10}) from all around Malaysia. From the data analysis obtained, the concentrations of air quality pollutants all around Malaysia starting from 2008 to 2011 were acceptable and the most dominant major pollutants had been highlighted. KMO obtained in this study is 0.7760, which show that the results are factor well. While, Bartlett's test shows that the variables correlated to each other's. From these tests, air quality data were acceptable for factor analysis.

Keywords: air pollution; chemometrics; PCA; FA.

Author Correspondence, e-mail: azman.azid@gmail.com

doi: http://dx.doi.org/10.4314/jfas.v9i2s.23

1. INTRODUCTION

Principal component analysis (PCA) is used to form the most significant parameter with least loss of the original variables by excluding the less significant parameter [1] and this allow identification of the pollution source [2]. Diverse possible pollution sources according to the activities in the air quality monitoring environment can be identified by applying Factor Analysis (FA) where usually carried out after successfully applied the PCA [3]. According to [4], FA suggests important variants to explain the observed variances in the data and it is one of the data reduction technique while PCA is for different factors extraction. PCA is considered as one of the useful statistical methods for the potential structure of a set variables and one of the most prevalent. Besides, it can be used to cut down a lot of data set. Most dominant major pollutant in this study were determined by PCA where subgroups were pool based on the correlation patterns between two or more air pollutants. Analysis from the whole data does not include less significant variables with less original data loss. The first factor in PCA explains the major variables amount within the original data. While, the second factor explained by the factor that has not been explained by the first factor and subsequently [5].

It is advisable to rotate the (Principle Components) PCs by varimax rotation with eigenvalues equal to or greater than 1, as PCs produced by PCA without rotations are at times not readily presented for interpretation [6]. Furthermore, according to [7], most general factor with negative coefficients and similar size coefficients on all variables disappears and loadings structure been simplified. Besides, simpler structure of coefficients from rotated PCs make them easier to interpret. Moreover, varimax factors's number gained by varimax rotations usually equivalent to the variables number of the unobservable, hypothetical and hidden variables [8]. Better relationship between the PCs and the original variables can be achieved when PCs been rotated by PC varimax rotation. It is vital to rotate the PCs since the factor loadings after rotation reveal to what amount one variable is similar to the other and how much the variable contributes to that particular PC [3, 9]. Varimax factors (VFs) are new variables group obtained from the varimax rotation [10], while variables amount contributes to that particular factor and its similarity to the other also can be reflected by the factor loadings after rotation [3, 5, 9]. According to [11], PCA and varimax rotation were methods used for factor extraction and matrix rotation respectively. Thus, factor analysis will be discussed as principal factor analysis (PFA) since the factor extraction method employed is PCA [12].

According to [13], rotated factors with two or less variables should be interpreted with attention. Highly correlated variables ($r > 0.70$) only considered for the factor with two variables. However, it seems fairly uncorrelated with other variables. Data error will be lessen for a larger sample size. Furthermore, according to [14], at least three variables needed for something to be labelled as a factor although this depends on the study design.

Kaiser Meyer Olkin (KMO) and Bartlett's test were tested for data suitability prior to applying of factor analysis [15]. KMO and Barlett's tests were implemented in the Principal factor analysis (PFA) commencement where the KMO test forecasts whether data of interest are factor well or not. While, Sphericity Bartlett's test used to confirm that there are correlated variables used in PFA from the rejected results from the hypothesis used. Samples adequacy had been tested by applying the KMO of sampling adequacy (MSA) [16] before extracting the factors in the PCA, and MSA is acceptable if the value of KMO is ranging between 0.60 to 1.00 [12] (see Table 1).

Table 1. Guiding rules for interpretation of KMO test results [17]

KMO Value	Interpretation
0.90-1.00	Marvelous
0.80-0.89	Meritorious
0.70-0.79	Middling
0.60-0.69	Mediocre
0.50-0.59	Miserable
0.00-0.49	Unacceptable

Variable with high factor loading shows many variables contributes to the variation of that factor [18]. Correlation coefficient matrix between the variables was used for the ranking of factor loadings [16]. This support by the research done by [19], which states that variables with high loadings were grouped in the same factors whereas association between variable and factor shown by the larger the factor loading. From the factor scores obtained, each component can be connected with a kind of source [20]. The quantity of variable contributes to the factor can be measured by variable of a factor loading. Therefore, factors size are well accounted for by the variables [13].

Scree plot graph for PCA loadings indicates the cut-off point where strong factors are selected for interpretation [21]. Interpretation is based on significance factors [13] and used to clarify the extraction method of different factors [4]. Besides, it is also used to decide how many factors to retain. According to [13], the cut-off selection may be determine from the ease of interpretation together with how complex variables are being handled.

It is considered strong if factor loadings are greater than 0.75, considered moderate is the range between 0.50 to 0.75 and weak if range of factor loadings between 0.30 and 0.49 [3, 22]. However, according to [3], for the principal component analysis, only factor loadings with absolute values greater than 0.5 are selected in practice. In contrast, in [23] classified range of factor loadings quite different from [3, 22] whereby they classified the factor loadings as excellent, very good, good, fair and poor if the loadings are in the values of 0.71, 0.63, 0.55, 0.45 and 0.32 respectively. However, these loadings still in range with the study applied by [3, 22].

While, the range of factor loadings referred by [24-25] in their studied are taken from the range of factor loadings applied by [11, 26] respectively. Whereby, they classified the factor loadings as strong, moderate and weak if the loadings are in the values of greater than 0.75, 0.75-0.50 and 0.50-0.30 respectively. These range of factor loadings applied almost the same been applied by the [3, 22].

2. RESULTS AND DISCUSSION

In this study, four-years (2008-2011) daily average data of five major air pollutants variables which are ozone (O_3), nitrogen dioxide (NO_2), carbon monoxide (CO), particulate matter (PM_{10}) and sulphur dioxide (SO_2) were studied. Interrelated variables were interpreted using PCA where new variables known as principle components (PC) were created. Thus, source of emission can be identified [27] through analysis of PCA from factor analysis.

Before extracting factors in PCA, KMO measure of sampling adequacy (MSA) was applied to test samples adequacy which might be caused by underlying factors [11]. Table 2 shows result of KMO measure of sampling adequacy. From the result obtained, measure of sampling adequacy (MSA) was acceptable as the value obtained is greater than 0.5. All variables are considered adequate and can implement for further analysis. Correlations between variables of air quality and the extracted factors can be assess from applying factor loadings [28]. Besides, according to [11], principal component/factor analysis may be convenient if high value which is close to 1 obtained. In this study, the value of KMO obtained is 0.7660. Thus, the principal/factor analysis may be considered convenient. Besides, based on Guiding rules for KMO test results interpretation [17] (Table 1), the air quality data is middling which is in the range from 0.70 to 0.79. This shows that the data would factor well and air quality data is agreeable to PFA [12].

Table 2. Kaiser-Meyer-Olkin measure of sampling adequacy

O_3	0.8265
NO_2	0.7110
CO	0.7362
PM_{10}	0.8444
SO_2	0.8906
KMO	0.7660

Observed chi-square value obtained from this analysis is 111132.783 ($p < 0.0001$, df = 10) (see Table 3). Null hypothesis (Ho) was rejected and alternative hypothesis (Ha) was accepted as the computed p-value is lower than the significance level alpha = 0.05. This Bartlett’s test of sphericity shown that the air quality variables were correlated and not orthogonal [12] and gives correlated and unbiased scores with their own factor. Thus, factor analysis obtained from principle component analysis (PCA) will agree for the data variability interpretation with less than the original number [13]. From these types of test, it is confirmed that the adequate factors had been extracted in PCA were adequate, factor well and correlated with each other’s.

Table 3. Kaiser-Meyer-Olkin measure of sampling adequacy (observed values)

Chi-square (Observed value)	111132.7833
Chi-square (Critical value)	18.3070
DF	10
p-value	< 0.0001
alpha	0.05

Table 4 shows the loading factors of selected air pollutants parameters for Malaysia air monitoring stations (2008-2011). Based on the resulted obtained as depicted in Table 4, it can be inferred that correlations between air pollutants parameters can be carried out from the factor loadings. The Pearson's correlation between selected air pollutants parameters were also applied for the data interpretation (Table 5).

Table 4. Loading factors of selected air pollutants parameters for Malaysia air monitoring stations (2008-2011)

Parameters	F1	F2	F3
O_3	**0.7705**	-0.2736	-0.2331
NO_2	**0.8803**	-0.1305	-0.1078
CO	**0.8600**	-0.0635	-0.1626
PM_{10}	0.3768	**0.9079**	-0.1544
SO_2	0.5897	0.0649	**0.8012**
Eigenvalue	2.5980	0.9245	0.7581
% of variance	51.9598	18.4899	15.1621

Table 4 and Figure 1 show the highlights of selected factors with strong positive loadings (> 0.75), eigenvalues greater than 0.70 and percentage of variance. There are three factors represent 85.61% percentage of variability after varimax rotation. Factor 1 (F1) consist of O_3, NO_2 and CO, Factor 2 (F2) consist of PM_{10} and Factor 3 (F3) consist of SO_2. Fig. 2 shows scree plots of PCA with five PCs. Generally, to clarify the origin of variation, principal components were removed based on an eigenvalue greater than 1 [29]. Although the scree plot shows only F1 has an eigenvalue greater than 1, it is still acceptable for F2 and F2 which has an eigenvalue less than 1 as based on Jolliffe's criterion, it suggested retaining factors with eigenvalue above 0.70 [13].

For the selection of factor loadings, previous study done by [3, 11, 23, 26] does mentioned that loading factors with values greater than 0.75 are considered strong. These air pollutants parameters (O_3, NO_2, CO, PM_{10} and SO_2) which has loading factors greater than 0.75 then categorized as potential air pollutants contributor but coming from different kind of sources, as they were divided into three factors (F1, F2, F3) after extracts the factors in the PCA.

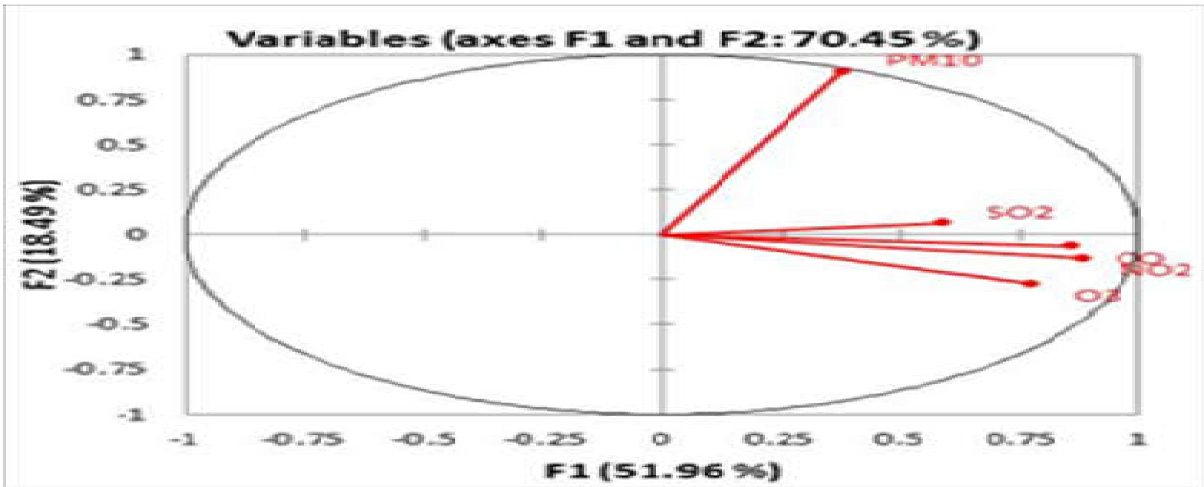

Fig.1. Factor loading plot after varimax rotation

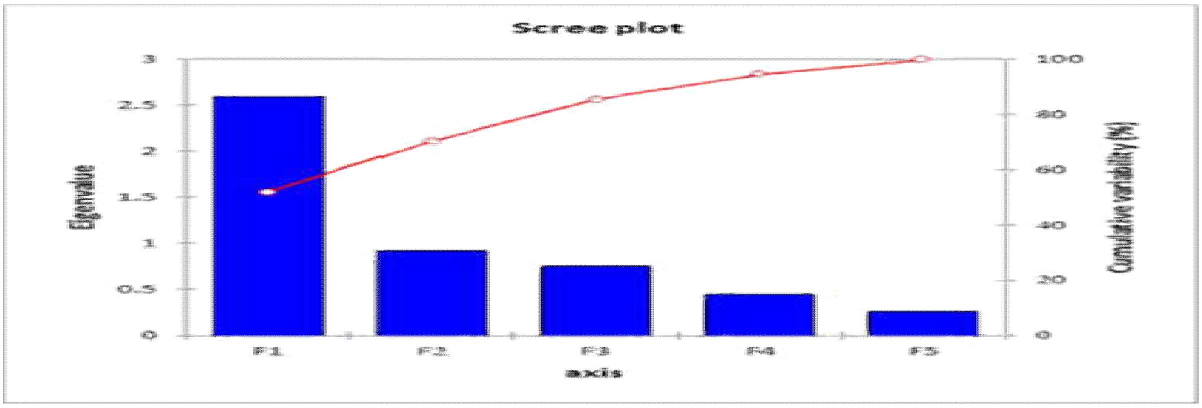

Fig.2. Scree plots for PCA

Factor 1 with higher factor include O_3, NO_2 and CO signified that the source is coming from diesel fuel. The Pearson's correlation coefficient (see Table 5) is showing higher correlation between O_3, NO_2 and CO. All pollutants shows positive correlation between each other. There are strong correlation between O_3, NO_2 and CO with correlation between O_3 and NO_2 (r = 0.6055), O_3 and CO (r = 0.5573) and NO_2 and CO (r = 0.7260). Among these three strong correlation, air pollutants correlation between NO_2 and CO shows the highest correlation (r = 0.7260). This confirmed the results of PCA analysis. In [30] also shows the same finding where there is strong correlation between NO_2 and CO with r = 0.805 and r = 0.901 in 2007 and 2012 respectively. As according to them, CO and NO_2 are the main pollutants from diesel fuel vehicles. This supported by the research done by [31] found out that CO and NO_2 pollution in ambient air are coming out from the mobile source which are also contributed to the formed of secondary pollutant namely ozone [30]. Thus, present of CO and NO_2 indirectly contributed to the present of O_3 from the same sources. These results lend support to the suggestion by [21], whereas O_3 concentration is largely dependent on its precursors (NO_x, and CO) availability.

In [32] mentioned that NO_2 is one of the traffic air pollutant while according to [33], local anthopogenic activities such as traffic, industries and agriculture generated O_3 pollutant. Other authors, in [34] have observed the aspect of O_3 where O_3 production consist of top five compounds come mainly from road traffic.

Factor 2 and Factor 3 show highest factor loading of PM_{10} and SO_2 respectively. Factor 2 has high factor loading of PM_{10} (r = 0.9079) signified that the source is coming from dust fall which possibly comes from the construction sites, industrial activities, soil dust and the transportation exhaust emission [27]. Besides, present of PM_{10} was shown to be contributed from the forest burning [35]. Anthropogenic activities such as wood burning, vehicles combustion activities and power plant also contribute to this most harmful pollutants [36].

While, Factor 3 with highest factor loading of SO_2 (r = 0.8012). According to [37], industrial activities one of the contributor to the high concentration of SO_2 present. The results of Pearson's correlation (Table 5) shows other major pollutants namely PM_{10} and SO_2 show low correlation with all of the pollutants and the lowest correlation shown by the correlation between O_3 and PM_{10} (r = 0.1281). Sites proximities of industrial plants may experiences high concentration of SO_2 compared to remote sites [38].

Table 5. Pearson's correlation between selected air pollutants

Air Pollutants	O_3	NO_2	CO	PM_{10}	SO_2
O_3	1				
NO_2	0.6055	1			
CO	0.5573	0.7260	1		
PM_{10}	0.1281	0.2157	0.2521	1	
SO_2	0..2896	0.3976	0.3575	0.1640	1

3. EXPERIMENTAL

Air pollution data were obtained from 50 continuous air monitoring stations around Malaysia (see Figure 3). In order to standardize all the data from 2010 to 2015, continuous air monitoring stations at locations number 13 and 34 which represent continuous air monitoring stations at ILP, Miri and Taman Semarak, Nilai were removed from the analysis due to lack of data from 2010 to 2011 and 2015 respectively.

All the data were collected and gathered from the Air Quality Division (DOE) Malaysia (2015). The data starts from 1st January 2008 to 31st December 2011. In this study, 367,080 data points were analyzed and used which compromises of five pollutants (73,416 data for each pollutants). All the data were interpreted in daily average. The locations of continuous air monitoring stations are stated in Table 6 (a) to Table 6 (n). Main air pollutant parameters namely ozone (O_3), nitrogen dioxide (NO_2), sulphur dioxide (SO_2), carbon monoxide (CO) and particulate matter (PM_{10}) were studied in this study.

BAM-1020 Beta Attenuation Mass Monitor from Met One Instrument, Inc. USA was used to monitor PM_{10}. While, SO_2, NO_2, CO and O_3 were monitored using the Teledyne API Model 100A/100E, Teledyne API Model 200A/200E, Teledyne API Model 300/300E and Teledyne API Model 400/400E respectively. Because of their accuracy, robustness and reliability, these instruments were chosen.

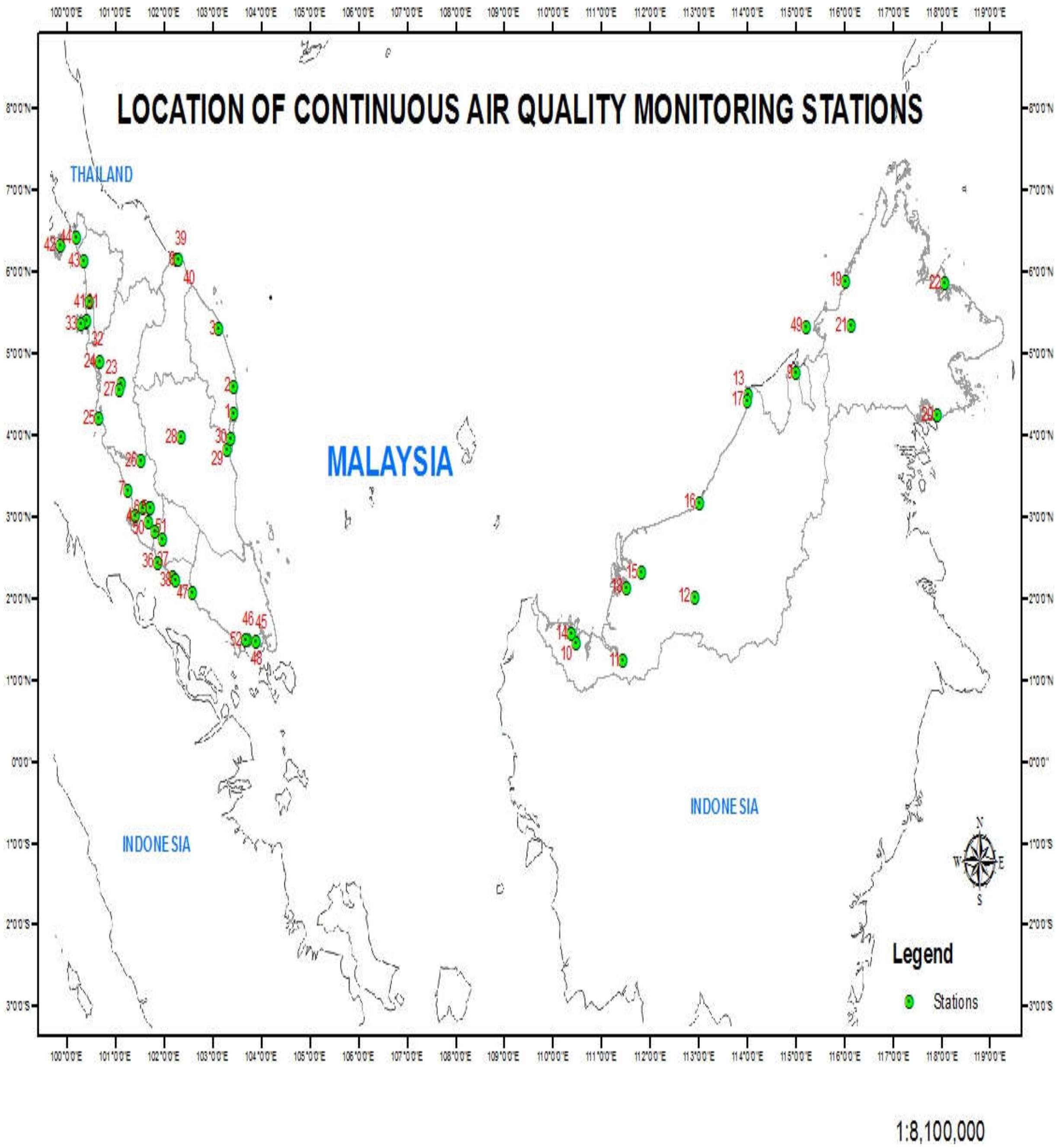

Fig.3. Air monitoring stations throughout Malaysia

Table 6 (a). Sampling point for air quality monitoring stations in Selangor

	Locations	Lat. (N)	Long. (E)
4	SM (P) Raja Zarina, Klang	N03°00.602	E101°24.484
5	SK Bandar Utama, Petaling Jaya	N03°06.612	E101°42.274
6	SK TTDI Jaya, Shah Alam	N03°06.286	E101°33.367
7	SM Sains, Kuala Selangor	N03°19.592	E101°15.532
8	Kolej MARA Banting	N02°49.001	E101°37.381

Table 6 (b). Sampling point for air quality monitoring stations in Johor

	Locations	Lat. (N)	Long. (E)
45	SM Pasir Gudang 2	N01°28.225	E103°53.637
46	Institut Perguruan, Larkin	N01°28.225	E103°53.637
47	SM Teknik, Muar	N02°03.715	E102°35.587
48	SMA Bandar Penawar	N01°33.500	E104°13.310

Table 6 (c). Sampling point for air quality monitoring stations in Kedah

	Locations	Lat. (N)	Long. (E)
41	SK Bakar Arang, Sg Petani	N05°37.886	E100°28.189
42	Kompleks Sukan Langkawi	N06°19.903	E099°51.517
43	SM Agama Mergong	N06°08.218	E100°20.880

Table 6 (d). Sampling point for air quality monitoring stations in Kelantan

	Locations	Lat. (N)	Long. (E)
39	SMK Tanjung Chat, Kota Bharu	N06°09.520	E102°15.059
40	SMK Tanah Merah	N05°48.671	E102°08.000

Table 6(e). Sampling point for air quality monitoring stations in Melaka

	Locations	Lat. (N)	Long. (E)
37	SMK Bukit Rambai, Melaka	N02°15.510	E102°10.364
38	SM. Tinggi	N02°12.789	E102°14.055

Table 6 (f). Sampling point for air quality monitoring stations in Negeri Sembilan

	Locations	Lat. (N)	Long. (E)
35	SM. Teknik Tuanku Jaafar,	N02°43.418	E101°58.105
36	Pusat Sumber Pendidikan	N02°26.458	E101°51.956

Table 6 (g). Sampling point for air quality monitoring stations in Pahang

	Locations	Lat. (N)	Long. (E)
28	Pej. Kajicuaca Batu Embun	N03°58.238	E102°20.863
29	SK Indera Mahkota, Kuantan	N03°49.138	E103°17.817
30	SK Balok Baru, Kuantan	N03°57.726	E103°22.955

Table 6 (h). Sampling point for air quality monitoring stations in Perak

	Locations	Lat. (N)	Long. (E)
23	SM Jalan Tasek, Ipoh	N04°37.781	E101°06.964
24	SK. Air Puteh, Taiping	N04°53.940	E100°40.782
25	Pej. Daerah Manjung	N04°12.038	E100°39.841
26	UPSI, Tanjung Malim	N03°41.267	E101°31.466
27	SM. Pagoh, Ipoh 2, Perak	N04°33.155	E101°04.856

Table 6 (i). Sampling point for air quality monitoring stations in Perlis

	Locations	Lat. (N)	Long. (E)
44	ILP, Kangar	N06°25.424	E100°11.046

Table 6 (j). Sampling point for air quality monitoring stations in Pulau Pinang

	Locations	Lat. (N)	Long. (E)
31	SK Cenderawasih	N05°23.470	E100°23.213
32	SK. Sebarang Jaya II, Perai	N05°23.890	E100°24.194
34	USM, Pulau Pinang	N05°21.528	E100°17.864

Table 6 (k). Sampling point for air quality monitoring stations in Sabah

	Locations	Lat. (N)	Long. (E)
19	SMK Putatan, Tg Aru	N05°53.623	E116°02.596
20	Pejabat JKR, Tawau, Sabah	N04°15.016	E117°56.166
21	SMK. Gunsanad, Keningau	N05°20.313	E116°09.769
22	Pej JKR Sandakan	N05°51.865	E118°05.479

Table 6 (l). Sampling point for air quality monitoring stations in Sarawak

	Locations	Lat. (N)	Long. (E)
14	Medical Store, Kuching	N01°33.734	E110°23.329
15	Ibu Pej. Polis Sibu, Sarawak	N02°18.856	E111°49.906
16	Balai Polis Pusat Bintulu	N03°10.587	E113°02.433
17	SM Dato' Permaisuri Miri,	N04°25.456	E114°00.731
18	Balai Polis Pusat Sarikei	N02°07.992	E111°31.351
9	Dewan Suarah, Limbang	N04°45.529	E115°00.813

10	Pej. Daerah, Kota Samarahan	N01°27.308	E110°29.498
12	Pej. Perumahan, Sri Aman	N01°14.425	E111°27.629
11	Stadium Tertutup, Kapit	N02°00.875	E112°55.640

Table 6 (m). Sampling point for air quality monitoring stations in Terengganu

	Locations	Lat. (N)	Long. (E)
1	SK. Bukit Kuang	N04°16.260	E103°25.826
2	Kuarters TNB, Paka-Kertih	N04°35.880	E103°26.096
3	Sek. Keb.Chabang Tiga,	N05°18.455	E103°07.213

Table 6 (n). Sampling for air quality monitoring stations in Wilayah Persekutuan

	Locations	Lat. (N)	Long. (E)
50	SK. Putrajaya 8(2), Jln P8/E2	N02°55.915	E101°40.909
51	SMK. Seri Permaisuri, Cheras	N03°06.376	E101°43.072
52	SK. Batu Muda, Batu Muda	N03°12.748	E101°40.929
49	Taman Perumahan MPL	N05°19.980	E115°14.315

4. CONCLUSION

Analysis of data using chemometrics are reliable where concentrations of five main air quality pollutants consist of ozone (O_3), nitrogen dioxide (NO_2), sulphur dioxide (SO_2), carbon monoxide (CO) and particulate matter (PM_{10}) from 2008 to 2011 were acceptable as KMO and Bartlett's test obtained in this study is factor well and the variables correlated to each other's respectively. Generally, air quality data were acceptable for factor analysis.

5. ACKNOWLEDGEMENTS

The authors acknowledge the Air Quality Division of the Department of Environment (DOE) under the Ministry of Natural Resource and Environment, Malaysia for giving us permission to utilize air quality data, advice, guidance and support for this study.

6. REFERENCES

[1] Al-Odaini N A, Zakaria M P, Zali M A, Juahir H, Yaziz M I, Surif S. Application of chemometrics in understanding the spatial distribution of human pharmaceuticals in surface water. Environmental Monitoring and Assessment, 2012, 184(11):6735-6748

[2] Zhang X, Jiang H, Zhang Y. Spatial distribution and source identification of persistent pollutants in marine sediments of Hong Kong. Environmental Monitoring and Assessment, 2013, 185(6):4693-4704

[3] Mutalib S N, 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 and Impacts, 2013, 15(9):1717-1728

[4] Alkarkhi AF, Ahmad A, Ismail N, mat Easa A, Omar K. Assessment of surface water through multivariate analysis. Journal of Sustainable Development, 2009, 1(3): 27-33

[5] Rahman S R, Ismail S N, Raml M F, Latif M T, Abidin E Z, Praveena S M. The assessment of ambient air pollution trend in Klang Valley, Malaysia. World Environment, 2015, 5(1):1-1

[6] Juahir H, Zain S M, Yusoff M K, Hanidza T T, Armi A M, Toriman M E, Mokhtar M. Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques. Environmental Monitoring and Assessment, 2011, 173(1-4):625-641

[7] Jolliffe I. T. Principal component analysis. New York: Springer-Verlag, 2002

[8] Azid A, Juahir H, Toriman ME, Endut A, Rahman A, Nordin M, Kamarudin MK, Umar R. Identification source of variation on regional impact of air quality pattern using chemometric. Aerosol and Air Quality Research, 2015, 15: 1545-1558

[9] 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, 2012, 60:172-181

[10] Brūmelis G, Lapiņa L, Nikodemus O, Tabors G. Use of an artificial model of monitoring data to aid interpretation of principal component analysis. Environmental Modelling and Software, 2000, 15(8):755-763

[11] Shrestha S, Kazama F. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environmental Modelling and

Software, 2007, 22(4):464-475

[12] Gazzaz N M, Yusoff M K, Ramli M F, Aris A Z, Juahir H. Characterization of spatial patterns in river water quality using chemometric pattern recognition techniques. Marine Pollution Bulletin, 2012, 64(4):688-698.

[13] Yong A G, Pearce S. A beginner's guide to factor analysis: Focusing on exploratory factor analysis. Tutorials in Quantitative Methods for Psychology, 2013, 9(2):79-94

[14] Tabachnick B. G., Fidell L. S. Using multivariate statistics. Massachusetts: Allyn and Bacon, 2007

[15] Tesfazghi E S, Martinez J A, Verplanke J J. Variability of quality of life at small scales: Addis Ababa, Kirkos Sub-City. Social Indicators Research, 2010, 98(1):73-88

[16] Mustapha A, Aris A Z, Juahir H, Ramli M F. Surface water quality contamination source apportionment and physicochemical characterization at the upper section of the Jakara Basin, Nigeria. Arabian Journal of Geosciences, 2013, 6(12):4903-4915

[17] Kaiser H F. An index of factorial simplicity. Psychometrika, 1974, 39(1):31-36

[18] Jolliffe I. T. Principal component analysis. New York: Springer, 1986

[19] Zhou H, Kusaka Y, Tamura T, Suganuma N, Subhannachart P, Siriruttanapruk S, Dumavibhat N, Zhang X, Sishodiya P K, Van Duy K, Hering K G. Proficiency in reading pneumoconiosis radiographs examined by the 60-film set with 4-factor structuring 8-index. Industrial Health, 50(2):142-146

[20] Charron A, Harrison R M. Primary particle formation from vehicle emissions during exhaust dilution in the roadside atmosphere. Atmospheric Environment, 2003, 37(29):4109-4119

[21] Isiyaka H A, Azid A. Air quality pattern assessment in Malaysia using multivariate techniques. Malaysian Journal of Analytical Sciences, 2015, 19(5):966-978

[22] Liu C W, Lin K H, Kuo Y M. Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan. Science of the Total Environment, 2003, 313(1):77-89

[23] Li G, Weng Q. Measuring the quality of life in city of Indianapolis by integration of remote sensing and census data. International Journal of Remote Sensing, 2007,

28(2):249-267

[24] Samsudin M S, Juahir H, Zain S M, Adnan N H. Surface river water quality interpretation using environmetric techniques: Case study at Perlis River Basin, Malaysia. International Journal of Environmental Protection, 2011, 1(5):1-8

[25] Baharuddin N, Nor'ashikin S A, Zain S M. Characterization of spatial patterns in river water quality using chemometric techniques. Sains Malaysiana, 2014, 43(9):1355-1362

[26] Reghunath R, Murthy T S, Raghavan B R. The utility of multivariate statistical techniques in hydrogeochemical studies: An example from Karnataka, India. Water Research, 2002, 36(10):2437-2442

[27] Azid A, Juahir H, Toriman M E, Endut A, Kamarudin M K, Rahman M N, Hasnam C N, Saudi A S, Yunus K. Source apportionment of air pollution: A case study in Malaysia. Jurnal Teknologi, 2015, 72(1):83-88

[28] Azid A, Juahir H, Toriman M E, Kamarudin M K, Saudi A S, Hasnam C N, Aziz N A, Azaman F, Latif M T, Zainuddin S F, Osman M R. Prediction of the level of air pollution using principal component analysis and artificial neural network techniques: A case study in Malaysia, Water, Air, and Soil Pollution, 2014, 225(8):1-14

[29] Ibrahim A, Juahir H, Toriman M E, Kamarudin M K, Isiyaka H A. Surface water quality assessment of Terengganu River Basin using multivariate techniques. Advances in Environmental Biology, 2014, 8(24):48-58

[30] Kovač-Andrić E, Radanović T, Topalović I, Marković B, Sakač N. Temporal variations in concentrations of ozone, nitrogen dioxide, and carbon monoxide at Osijek, Croatia. Advances in Meteorology, 2013, 2013:1-7

[31] Chelani A B. Study of extreme CO, NO_2 and O_3 concentrations at a traffic site in Delhi: Statistical persistence analysis and source identification. Aerosol and Air Quality Research, 2013, 13(1):377-384

[32] Al-Anzi B, Abusam A, Khan A R. Evaluation of temporal variations in ambient air quality at Jahra using multivariate techniques. Environmental Technology and Innovation. 2016, 5:225-232

[33] Durao R M, Mendes M T, Pereira M J. Forecasting O_3 levels in industrial area

surroundings up to 24 h in advance, combining classification trees and MLP models. Atmospheric Pollution Research, 2016, 7(6):961-970

[34] Franco J F, Pacheco J, Behrentz E, Belalcázar L C. Characterization and source identification of VOC species in Bogotá, Colombia. Atmósfera, 2015, 28(1):1-11

[35] Shafie S H M, Mahmud M. Pencemaran habuk di Lembah Klang melalui analisis statistik boxplot. Malaysian Journal of Society and Space, 2015, 11(11):144-155

[36] Mishra D, Goyal P, Upadhyay A. Artificial intelligence based approach to forecast PM 2.5 during haze episodes: A case study of Delhi, India. Atmospheric Environment, 2015, 102:239-248

[37] Mapoma H W, Tenthani C, Tsakama M, Kosamu I B. Air quality assessment of carbon monoxide, nitrogen dioxide and sulfur dioxide levels in Blantyre, Malawi: A statistical approach to a stationary environmental monitoring station. African Journal of Environmental Science and Technology, 2014, 8(6):330-343

[38] Petracchini F, Paciucci L, Vichi F, D'Angelo B, Aihaiti A, Liotta F, Paolini V, Cecinato A. Gaseous pollutants in the city of Urumqi, Xinjiang: Spatial and temporal trends, sources and implications. Atmospheric Pollution Research, 2016, 7(5):925-934

How to cite this article:
Abd Rani NL, Azid A, Khalit SI, Gasim MB, Juahir H. Selected Malaysia air quality pollutants assessment using chemometrics techniques. J. Fundam. Appl. Sci., 2017, *9(2S), 335-351*.