

*Original Research*

# Assessing Indoor Air Quality Using Chemometric Models

**Azman Azid<sup>1\*</sup>, Mohammad Azizi Amran<sup>1\*</sup>, Mohd Saiful Samsudin<sup>1</sup>,  
Nurul Latiffah Abd Rani<sup>1</sup>, Saiful Iskandar Khalit<sup>1</sup>, Muhammad Barzani Gasim<sup>1,2</sup>,  
Kamaruzzaman Yunus<sup>3</sup>, Ahmad Shakir Mohd Saudi<sup>4</sup>,  
Siti Noor Syuhada Muhammad Amin<sup>5</sup>, Ku Mohd Kalkausar Ku Yusof<sup>1</sup>**

<sup>1</sup>Faculty Bioresources and Food Industry, Universiti Sultan Zainal Abidin, Besut Campus, Terengganu, Malaysia

<sup>2</sup>East Coast Environmental Research Institute, Universiti Sultan Zainal Abidin, Gong Badak Campus, Terengganu, Malaysia

<sup>3</sup>Kulliyah of Science, International Islamic University Malaysia, Pahang, Malaysia

<sup>4</sup>Institute of Medical Science and Technology, University of Kuala Lumpur, Selangor, Malaysia

<sup>5</sup>Faculty of Health Science, Universiti Sultan Zainal Abidin, Gong Badak Campus, Terengganu, Malaysia

*Received: 6 April 2017*

*Accepted: 1 October 2017*

## Abstract

The objectives of this study are to identify the significant variables and to verify the best statistical method for determining the effect of indoor air quality (IAQ) at 7 different locations in Universiti Sultan Zainal Abidin, Terengganu, Malaysia. The IAQ data were collected using in-situ measurement. Principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA), linear discrimination analysis (LDA), and agglomerative hierarchical clustering (AHC) were used to classify the significant variables as well as to compare the best method for determining IAQ levels. PCA verifies only 4 out of 9 parameters (PM<sub>10</sub>, PM<sub>2.5</sub>, PM<sub>1.0</sub>, and O<sub>3</sub>) and is the significant variable in IAQ. The PLS-DA model classifies 89.05% correct of the IAQ variables in each station compared to LDA with only 66.67% correct. AHC identifies three cluster groups, which are highly polluted concentration (HPC), moderately polluted concentration (MPC), and low-polluted concentration (LPC) area. PLS-DA verifies the groups produced by AHC by identifying the variables that affect the quality at each station without being affected by redundancy. In conclusion, PLS-DA is a promising procedure for differentiating the group classes and determining the correct percentage of variables for IAQ.

**Keywords:** indoor air quality (IAQ), pattern recognition, PCA, PLS-DA, LDA

## Introduction

Humans need clean, comfortable, and safe indoor air environments [1]. People working in poor indoor air quality (IAQ) conditions may face higher health risks [2] such as sick building syndrome (SBS) and other occupational hazards [3].

Particulate matter (PM) is one of the contributed pollutants in indoor air. The common types of PM measured in indoor air are a diameter less than 10  $\mu\text{m}$  ( $\text{PM}_{10}$ ) and 2.5  $\mu\text{m}$  ( $\text{PM}_{2.5}$ ). From both these types of particles,  $\text{PM}_{2.5}$  can penetrate deep into the lung tissue and may result in reduced lung function in children and adults. As a result, diseases such as pneumonia, shortness of breath, poor cardiovascular effects and elevated chronic obstructive pulmonary disease (COPD) [4-5] can occur. Exposure to this pollutant can give oxidative stress to human DNA [5-6].

The sources of IAQ pollutants are caused by humans entering the indoors and outdoors of buildings. These activities can disturb the concentration of PM in the building [6]. The most important sources that affecting IAQ are particulate matter (PM), carbon monoxide (CO), carbon dioxide ( $\text{CO}_2$ ), temperature, and relative humidity (RH) [7-8]. However, the major contributor to SBS risk is CO and PM [2, 9-12].

IAQ depends on the efficiency of air conditioning and mechanical ventilation (ACMV) systems that distribute air to people in the rooms. Previous researchers have proven that the distribution of indoor air pollutants will vary from one building to another depending on the ventilation system [4, 13-14].

The objectives of this study are: 1) to identify the significance of variable values influencing the impact of IAQ in selected offices at the Universiti Sultan Zainal Abidin, Terengganu, Malaysia and 2) to compare the best methods in chemometrics for evaluating the above objective.

## Experimental

### Study Location

Seven IAQ sampling points (Table 1) at Universiti Sultan Zainal Abidin, Terengganu, Malaysia were chosen for this study. The different indoor environmental conditions and operations such as ventilation systems, space size, and building age were considered when measuring IAQ levels.

### *Sampling and Data Collection of Indoor Air Quality*

Sample data for the IAQ study were gathered from *in-situ* measurements. The standard breathable level (1.0-1.5 m) of 8 hours of sampling time [4] was used for monitoring all locations. The sampling was measured using such equipment as a portable Turnkey Instruments

Table 1. Sampling locations of indoor offices.

| No. | Station | Location                                  |
|-----|---------|---|
| 1   | S1      | Vice Chancellor's Office, Canselor        |
| 2   | S2      | Deputy Vice Chancellor's Office, Canselor |
| 3   | S3      | Office Room Level 1, ESERI                |
| 4   | S4      | Meeting Room, ESERI                       |
| 5   | S5      | Office Room Level 2, ESERI                |
| 6   | S6      | Computer Room, Library                    |
| 7   | S7      | Reading Corner, Library                   |

dust mate detector (for total suspended particulate (TSP),  $\text{PM}_{10}$ ,  $\text{PM}_{2.5}$ , and  $\text{PM}_{1.0}$ ) and a 3M Quest EVM series gas meter for carbon monoxide (CO), carbon dioxide ( $\text{CO}_2$ ), relative humidity (RH), and temperature (T). The ability to monitor in real time is the reason why these instruments were selected. Each reading of IAQ parameter was taken every one minute.

For statistical analysis, chemometric techniques such as linear discriminant analysis (LDA), principal component analysis (PCA), and partial least square – discriminant analysis (PLS-DA) were applied. These techniques were used for comparing the best model to find the most variables that contribute in IAQ readings in the study locations.

### *Using PCA to Identify the most Significant Variables Influencing Indoor Air Quality*

PCA was developed by Pearson [15], where each principal component (PCs) is an evaluation of a linear combination of original multiple responses [16]. Every component in PCA is orthogonal between each other. The variance of a large set of interrelated variables can be transformed into a new (smaller set of uncorrelated (independent) variables when applying this method [16-17]. In this study, the most significant parameter contributing to IAQ can be recognized by this method. PCA was calculated based on Equation 1:

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

...where  $z$  is the component score,  $a$  is component loading,  $x$  is the measure of the IAQ variable,  $i$  is the component number, and  $m$  is the total of IAQ variables. PCA will produce the list of IAQ variables that are dominant from less significant ones until the most significant variables that influence positive maximum and negative minimum peak of the spatial variation pattern. The finding of PCs will provide the eigenvectors with the factors given by the eigenvalues to identify the number of rotations in varimax-rotations. The eigenvalues in PCA were calculated based on Equation 2:

$$V^{-1}AVe_i = K_{e_i} = \lambda_i e_i \quad (2)$$

...which shows that  $V$  is the inversely varimax of total computed values, and that a matrix  $A$  with a linearly independent system of eigenvectors is similar to a diagonal matrix formed from its eigenvalues. The number of rotations depends on the eigenvector values of 1.0 and above. This value will deliver the detection values for variables either related to or dismissed between each other.

#### *PLS-DA for Classifying Groups and Discrimination Problems*

PLS-DA is a statistical approach that allows for a certain relationship to principal components regression, which discovers the variance that responds to independent variables. PLS-DA was used to reduce the dimension, redundancy, and discrimination problems [18-19]. PLS-DA was created when  $X$  (dimensional  $I \times J$ ) is present.  $X$  is a center matrix of IAQ predictive variables of the calibration set. Then  $g$  can be a vector of  $I$  integer values coding the qualitative groups to deliver the group number associated with observation  $I$ , and  $g$  denotes the total number of qualitative groups. The general underlying model of multivariate PLS mentioned in Equations 3 and 4:

$$X = TP^T + E \quad (3)$$

$$Y = UQ^T + F \quad (4)$$

...where  $X$  is  $n \times p$  matrix of responses while  $Y$  is matrix of qualitative, including the IAQ station and classified groups. PLS-DA will show a summary of the IAQ variables which indicators of the groups are. Higher prediction of IAQ data can be achieved by simply classifying the observations in the group process operated by PLS-DA. Reducing redundancy disturbs and influences the non-precise values shown by the percentage of correct values examined by PLS-DA.

#### *LDA and Agglomerative Hierarchical Clustering (AHC) for Comparison of PLS-DA*

LDA was used to identify the cluster classification accuracy of IAQ parameters that are available in the office building [20]. LDA is a supervised classification method that maximizes differences between sampling locations by generating a weighed linear combination of the discriminating variables [20-22] where  $k$  is the number of sub-locations. LDA comprises the calculation of  $k-1$  discriminant functions such as in Equation 5:

$$D = b_1X_1 + b_2X_2 + \dots + b_nX_n + c \quad (5)$$

... where  $D$  is discriminant function,  $b$  is the discriminant coefficient for that variable, and  $X$  discriminates IAQ variables, i.e., IAQ values collection for variables  $c$ , a constant, and  $n$ , the number of discriminant variables,

and finally the confusion matrix values in percentage [20]. Then, LDA values were compared with the PLS-DA confusion value instead of redundancy by modification of data variables.

The other methods to compare PLS-DA were AHC. AHC is an unsupervised pattern recognition approach to separate a large group of datasets into smaller or simpler ones [16, 23]. AHC were performed based on the normal distribution of datasets through Ward's method by means of Euclidean distances, as a measure of the connection between the datasets or variables [24]. In this study, AHC was used to classify the cluster groups from the selected IAQ office building.

## **Results and Discussion**

### **Principal Component Analysis**

Pattern recognition provides by using PCA to give the correlation value of each IAQ variable without any disturbance from the persistence of redundancy in the analysis. PCA performed by XLSTAT 2014 software gives a better and more accurate value of correlation by many types of testing hypothesis such as Bartlett's test, the Kaiser-Meyer-Olkin measure (KMO), and Pearson's correlation matrix. Findings from Bartlett's test (Table 2a) shows that the alpha value was less than 0.0001. This means that the observation data has a correlation with each other. Further correlation testing methods were needed for detecting the critical IAQ variables that call for more concern and control. Detecting critical IAQ variables can be detected by correlation of KMO (Table 2b), which shows the specific correlation value at the specific variables. From the KMO results, the most important pollutant has a higher tendency to cause harm to the respiratory system and is a variable with a value greater than 0.75. IAQ variables that respond with this value were  $PM_{10}$ ,  $PM_{2.5}$ ,  $PM_{1.0}$ , CO, and  $CO_2$ , consistent with the KMO values of 0.751, 0.752, 0.798, 0.781, and 0.892, respectively. Based on the European Union, air pollution from PM groups is predictable as causing an average decline in life anticipation of around 8.6 months per person [25-27]. The PM characteristics are invisible to human eyes, odourless, and colourless as a main subject in risk assessment of public health exposure [28]. Other characteristics of PM were based on chemical composition such as carbon organic compounds and heavy metals; while the biological composition such as bacteria and fungi can make the human defence system damaged and infected [29-30]. There are several health implications from PM such as lung inflammation, respiratory symptoms, adverse cardiovascular effects, and an increase in the prevalence of chronic obstructive pulmonary disease (COPD) [4,5]. Meanwhile, exposure to CO at a higher level will affect cardiovascular and neurobehavioral functions. Even a small concentration of CO can cause unconsciousness and even death [31].

Table 2. Correlation hypothesis testing: a) Bartlett's sphericity test and b) Kaiser-Meyer-Olkin measure.

| a)                          |                   |
|-----------------------------|-------------------|
| Chi-square (observed value) | 495.617           |
| Chi-square (critical value) | 50.998            |
| DF                          | 36                |
| <b>p-value</b>              | <b>&lt;0.0001</b> |
| alpha                       | 0.05              |

| b)   |              |
|--|--------------|
| <b>Temperature (°C)</b>                    | 0.651        |
| <b>RH (%)</b>                              | 0.619        |
| <b>TSP (µg/m<sup>3</sup>)</b>              | 0.715        |
| <b>PM<sub>10</sub> (µg/m<sup>3</sup>)</b>  | <b>0.751</b> |
| <b>PM<sub>2.5</sub> (µg/m<sup>3</sup>)</b> | <b>0.752</b> |
| <b>PM<sub>1.0</sub> (µg/m<sup>3</sup>)</b> | <b>0.798</b> |
| <b>CO (ppm)</b>                            | <b>0.781</b> |
| <b>CO<sub>2</sub> (ppm)</b>                | <b>0.892</b> |
| <b>O<sub>3</sub> (ppm)</b>                 | 0.467        |
| <b>KMO</b>                                 | <b>0.714</b> |

Then the Pearson's correlation matrix was tested to show the findings of the relationship between each variable in the IAQ study. In the Pearson testing hypothesis, the variables included with the selected value of alpha should be equal to or less than 0.05 ( $p < 0.05$ ). When this happen, it means that the data have a relationship with each other. Table 3 shows the finding of Pearson's correlation matrix test. The bold values are different from 0 with a significance level of alpha ( $\alpha \leq 0.05$ ). This hypothesis testing shows that the temperature variable has the greatest influence and relationship with RH, PM<sub>10</sub>,

Table 4. Eigenvalue to determine rotation number.

|                 | <b>F1</b>     | <b>F2</b>     | F3     | F4     |
|-----------------|---------------|---------------|--------|--------|
| Eigenvalue      | <b>3.931</b>  | <b>2.208</b>  | 0.908  | 0.868  |
| Variability (%) | <b>43.681</b> | <b>24.529</b> | 10.088 | 9.649  |
| Cumulative (%)  | <b>43.681</b> | <b>68.210</b> | 78.298 | 87.947 |

PM<sub>2.5</sub>, PM<sub>1.0</sub>, CO, and O<sub>3</sub>. This is because the temperature influences the distribution and dispersion of the selected variable concentration into indoor environments.

In Malaysia, the Industrial Code of Practice on Indoor Air Quality (ICOP) act was applied as a standard guideline. ICOP lists the acceptable range for specific physical parameters by the Department of Safety and Health (DOSH) in 2010 [32]. Based on ICOP, the acceptable range for indoor temperature is 23-26°C, which is suited to climatic conditions in Malaysia. Besides, the bacteria will grow at an optimum temperature of around 35°C [32].

PCA also provides the eigenvector value to determine the number of rotations that can be run by this method to give an understanding of reliable variables contributing to the concentration of other variables. Eigenvector values greater than 1 (Table 4) can be accepted for varimax-rotated factors (VFs) stepwise.

Based on Table 4, only two rotations in VFs can occur. These two VFs were conducted to determine the predictive variation pattern between variables (Table 5). Table 5 concludes the significant prediction and contribution of variables in the IAQ study. These values demonstrate the predictive value of the variables that need to be monitored and controlled in order to find out the cause and dispersing agent for the purposes of IAQ management in the future.

This forecast reflects the detection value adjustment. The control measures and frequent maintenance of the air-conditioning system needs to be considered with appropriate filters such as UV filters in air conditioning systems, so that the hazardous pollutants

Table 3. Pearson (n) correlation matrix.

| Variables                                 | Temp.        | RH           | TSP          | PM <sub>10</sub> | PM <sub>2.5</sub> | PM <sub>1.0</sub> | CO           | CO <sub>2</sub> | O <sub>3</sub> |
|---|--------------|--------------|--------------|------------------|-------------------|-------------------|--------------|-----------------|----------------|
| Temp. (°C)                                | 1            |              |              |                  |                   |                   |              |                 |                |
| RH (%)                                    | 0.473        | 1            |              |                  |                   |                   |              |                 |                |
| <b>TSP (µg/m<sup>3</sup>)</b>             | 0.187        | -0.011       | <b>1</b>     |                  |                   |                   |              |                 |                |
| <b>PM<sub>10</sub> (µg/m<sup>3</sup>)</b> | <b>0.247</b> | 0.045        | <b>0.790</b> | <b>1</b>         |                   |                   |              |                 |                |
| PM <sub>2.5</sub> (µg/m <sup>3</sup> )    | 0.390        | 0.487        | 0.368        | 0.622            | 1                 |                   |              |                 |                |
| PM <sub>1.0</sub> (µg/m <sup>3</sup> )    | 0.578        | 0.429        | 0.456        | 0.642            | 0.804             | 1                 |              |                 |                |
| <b>CO (ppm)</b>                           | <b>0.544</b> | 0.186        | 0.167        | <b>0.240</b>     | <b>0.366</b>      | <b>0.560</b>      | <b>1</b>     |                 |                |
| <b>CO<sub>2</sub> (ppm)</b>               | 0.154        | -0.125       | <b>0.553</b> | <b>0.649</b>     | <b>0.435</b>      | <b>0.499</b>      | <b>0.360</b> | <b>1</b>        |                |
| <b>O<sub>3</sub> (ppm)</b>                | <b>0.607</b> | <b>0.464</b> | -0.118       | -0.209           | -0.100            | -0.019            | 0.088        | <b>-0.311</b>   | <b>1</b>       |

\*Values in bold are different from 0 with a significance level alpha ( $\alpha = 0.05$ )

Table 5. Varimax rotation factors of each IAQ variable.

| IAQ Variable              | VF1          | VF2          |
|---------------------------|--------------|--------------|
| Temperature (°C)          | 0.613        | 0.637        |
| RH (%)                    | 0.391        | <b>0.691</b> |
| TSP (µg/m³)               | <b>0.660</b> | -0.404       |
| PM <sub>10</sub> (µg/m³)  | <b>0.806</b> | -0.406       |
| PM <sub>2.5</sub> (µg/m³) | <b>0.824</b> | 0.034        |
| PM <sub>1.0</sub> (µg/m³) | <b>0.912</b> | 0.092        |
| CO (ppm)                  | <b>0.602</b> | 0.224        |
| CO <sub>2</sub> (ppm)     | <b>0.668</b> | -0.492       |
| O <sub>3</sub> (ppm)      | 0.018        | <b>0.834</b> |

\*Values in bold are selected value (>0.6) or whichever is greater in any rotation numbers.

in indoor air can be removed from the building. This process is necessary in order to improve air quality in the workplace.

*Partial Least Square Discriminant Analysis*

PLS-DA was tested for classifying IAQ variables into main groups and to compare with findings from AHC. Group results from PLS-DA show the concentration distribution of pollutants clearer than AHC, where it only provides information without any kind of pollutants contained in the resulting group. Fig. 1 shows that the IAQ variables have been classified into three main groups. In the groups, they consist of the information about stations or sampling buildings that are influenced by the concentration of every IAQ variable. The three main groups shown by Fig. 1 are low-polluted concentration (LPC), moderately polluted concentration (MPC), and highly polluted concentration (HPC).

Based on Fig. 1a) using PLS-DA, the group of LPC shows no IAQ variables that are highly influenced to this group and the stations involved in this group are S2, S4, and S6. In the MPC group, there are several variables (O<sub>3</sub>, RH, and CO) that highly influence station S7. Meanwhile, the HPC group shows that many numbers of IAQ variables are highly correlated to cause serious respiratory damage: TSP, PM<sub>10</sub>, PM<sub>2.5</sub>, PM<sub>1.0</sub>, and CO<sub>2</sub> and station involved there. The stations in HPC that require precaution and preventive measures were S1, S3, and S5.

Fig. 1b) shows the different cluster groups that have occurred where S3 has been clustered with S7 in MPC. However, there is a dissimilarity between these methods as it still results in highly clustered groups of S3. There are some weaknesses with the AHC method because it cannot illustrate the IAQ variables that influence each cluster group.

IAQ variables that show a correlation value >0.75 are considered the most significant value of IAQ. Through the PLS-DA, the most important of IAQ variables have been determined, which are PM<sub>2.5</sub> (0.843) and PM<sub>1.0</sub> (0.916). These two variables were categorized as of fine particulate size. People who are exposed to the highest concentrations of these variables can cause harmful effects, especially in the respiratory system. The source of these ultrafine particulates inside buildings comes from many possible mediums. One of the examples comes from carpeting materials, which can disperse via air. Higher concentrations of these variables can be achieved by human movement in and out of a building on a regular basis. Table 6 shows the other highest correlation values, which were PM<sub>10</sub> (0.750) and O<sub>3</sub> (0.807).

Soil and dust particles were brought into the building through human movement in and out of and released into the indoor environment by being supported with air-conditioning and airflow. People who spend more time indoors with this serious condition may experience respiratory disease caused by exposure duration and concentration levels. By demonstrating

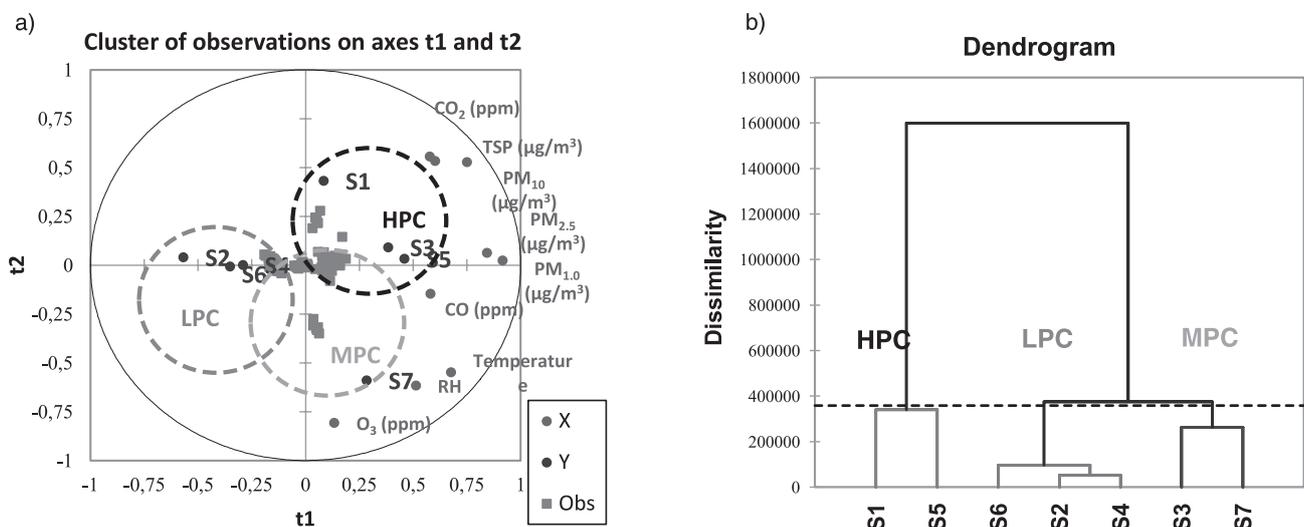


Fig. 1. Cluster of station influenced by IAQ variables using a) PLS-DA and b) AHC.

Table 6. Correlation values of each IAQ variable.

| Parameter                                      | Correlation Value |
|--|-------------------|
| Temperature (°C)                               | 0.676             |
| RH (%)   | 0.616             |
| TSP ( $\mu\text{g}/\text{m}^3$ )               | 0.602             |
| PM <sub>10</sub> ( $\mu\text{g}/\text{m}^3$ )  | 0.750             |
| PM <sub>2.5</sub> ( $\mu\text{g}/\text{m}^3$ ) | 0.843             |
| PM <sub>1.0</sub> ( $\mu\text{g}/\text{m}^3$ ) | 0.916             |
| CO (ppm)                                       | 0.581             |
| CO <sub>2</sub> (ppm)                          | 0.577             |
| O <sub>3</sub> (ppm)                           | 0.807             |

the value of this correlation, it is evident that these variables must be monitored more closely to ensure the cleanliness of the indoor building and that indoor air pollution can be controlled efficiently.

In another perspective, the findings of PLS-DA in the context of a confusion matrix percentage to determine the reliability and significance of collected IAQ data were slightly different from LDA. In LDA, the percentage of

the confusion matrix value is much lower than the value in PLS-DA because of the redundancy that disturbed the findings and discrimination problems compared with PLS-DA, which is less of redundancy and discrimination problems. Results derived from PLS-DA in Table 7a) give an 89.05% value for the percentage of the correct data. This value was classified as a significant value when approaching 75% and above. While in LDA (Table 7b) with discrimination problems resulting of 66.67%, those not strong enough values to be significant impact IAQ variables in this study. In PLS-DA, each station of IAQ has a different value to identify the variation of data collected. Based on that table, the suitable stations for detailed evaluation in the future are S2 and S3.

More collection of IAQ data may increase the percentage of confusion matrix and suitable of data analysis. Each IAQ variable in overall 7 stations provides much difference of the percentage of correctness. Determining IAQ data collection and analysis can be described as a strongly reliable data to use for an analysis and modelling procedure. A confusion matrix of PLS-DA provides the actual significance percentage of IAQ data to perform other reliable analysis, modelling, and prediction.

Table 7. Confusion matrix for estimating IAQ variable samples in each station for methods a) PLS-DA and b) LDA.

## a) PLS-DA

| from \ to    | S1        | S2        | S3        | S4       | S5        | S6        | S7       | Total     | Percentage correct |
|--------------|-----------|-----------|-----------|----------|-----------|-----------|----------|-----------|--------------------|
| S1           | 9         | 0         | 0         | 2        | 0         | 1         | 0        | 12        | 75.00              |
| S2           | 0         | 12        | 0         | 0        | 0         | 0         | 0        | 12        | 100.00             |
| S3           | 0         | 0         | 10        | 0        | 2         | 0         | 0        | 12        | 83.33              |
| S4           | 2         | 1         | 0         | 6        | 0         | 3         | 0        | 12        | 50.00              |
| S5           | 0         | 0         | 3         | 0        | 9         | 0         | 0        | 12        | 75.00              |
| S6           | 0         | 6         | 0         | 0        | 0         | 6         | 0        | 12        | 60.00              |
| S7           | 0         | 0         | 4         | 0        | 2         | 0         | 6        | 12        | 60.00              |
| <b>Total</b> | <b>11</b> | <b>19</b> | <b>17</b> | <b>8</b> | <b>13</b> | <b>10</b> | <b>6</b> | <b>84</b> | <b>89.05</b>       |

## b) LDA

| from \ to    | S1       | S2        | S3        | S4        | S5        | S6       | S7       | Total     | Percentage correct |
|--------------|----------|-----------|-----------|-----------|-----------|----------|----------|-----------|--------------------|
| S1           | 6        | 0         | 0         | 6         | 0         | 0        | 0        | 12        | 50.00              |
| S2           | 0        | 12        | 0         | 0         | 0         | 0        | 0        | 12        | 100.00             |
| S3           | 0        | 0         | 8         | 0         | 4         | 0        | 0        | 12        | 66.67              |
| S4           | 1        | 1         | 0         | 9         | 0         | 1        | 0        | 12        | 75.00              |
| S5           | 0        | 0         | 3         | 0         | 8         | 0        | 1        | 12        | 66.67              |
| S6           | 0        | 5         | 0         | 2         | 0         | 5        | 0        | 12        | 41.67              |
| S7           | 0        | 0         | 1         | 0         | 3         | 0        | 8        | 12        | 66.67              |
| <b>Total</b> | <b>7</b> | <b>18</b> | <b>12</b> | <b>17</b> | <b>15</b> | <b>6</b> | <b>9</b> | <b>84</b> | <b>66.67</b>       |

### Mitigation and Controlling Actions for PM Group Reductions

Based on this study, S1, S3, and S5 were the stations detected as having higher concentrations of the PM groups CO and CO<sub>2</sub>. Central air conditioning can effectively remove the airborne pollutants compared to non-central air conditioning. This study suggests that improvement or filtered air conditioning, airflow types, and management of human activities in different work areas can improve IAQ in the location at any clustered group [2]. The application of modern ultraviolet (UV) can be used to reduce the growth or killing of bacteria and fungi contained inside airborne particulate for healthy IAQ. Certain costs might be needed for doing this recommended mitigation action, but the benefits of occupant health and good IAQ can be achieved efficiently.

### Conclusions

IAQ control should be implemented by the employer to ensure the comfort of the occupants as long as they are at work. Pollutants such as PM<sub>10</sub>, PM<sub>2.5</sub>, PM<sub>1.0</sub>, and O<sub>3</sub> need to be given attention by all people because they are the largest contributors to the significant variable in IAQ as shown in the PCA model. The PLS-DA method has proven that the percentage correlation (89.05%) is higher than for the LDA (66.67%), and this proves that this method provides a high percentage of accuracy to the pollutants involved in the IAQ. The PLS-DA method can also provide the same number of groups with AHC (HPC, MPC, and LPC), but it is better than AHC because it can provide the types of pollutants involved in each group.

In conclusion, PLS-DA is a good method of analysis for identifying the significance, correlation, location, and percentage of correct variables of IAQ. From these findings, the IAQ monitoring assessment can be done efficiently due to extensive air-quality monitoring that requires higher cost, labour, and time.

### Acknowledgements

The authors acknowledge the East Coast Environmental Research Institute (ESERI), UniSZA for supplying the sufficient IAQ equipment during sampling and monitoring activities in this study. This manuscript was funded through the university research fund (grant No. UniSZA/2016/GOT/01).

### References

- ZHU C., LI N., RE D., GUAN J. Uncertainty in indoor air quality and grey system method. *Building and Environment*. **42**, 1711, **2007**.
- JUNG CC., WU PC. TSENG CH., SU HJ. Indoor air quality varies with ventilation types and working areas in hospitals. *Building and Environment*. **85**, 190, **2015**.

- WANG X. Hospital indoor PM<sub>10</sub>/PM<sub>2.5</sub> and associated trace elements in Guangzhou, China. *Sci Total Environ*. **366** (1), 124, **2006**.
- MENDES A., BONASSI S., ANGUIAR L., PEREIRA C., NEVES P., SILVA S., MENDES D., GUIMARAES L., MORONI R., TEIXEIRA J.P. Indoor air quality and thermal comfort in elderly care centers. *Urban Climate*. **2014**.
- SORENSEN M. Personal exposure to PM<sub>2.5</sub> and biomarkers of DNA damage. *Cancer Epidemiol*. **2003**.
- ZHANG X.Y., AHMADI G., QIAN J., FERRO A. Particle detachment, resuspension and transport due to human walking in indoor environments. *Journal Adhes Sci Technol*. **25**, 591, **2008**.
- WAN G.H., CHUNG F.F., TANG C.S., SHAN K., SINJHUNA G. Long-term surveillance of air quality in medical center operating rooms. *Am J Infect Control*. **39**, 302, **2011**.
- SAAD S.G. Integrated environmental management for hospital. *Indoor Built Environ*. **12**, 93, **2013**.
- REYNOLDS S.J., B D.W., BORIN S.S., BREUER G., BURMEISTER L.F., FUORTES L.J. Indoor environmental quality in six commercial office buildings in the midwest United States. *Appl Occup Environ Hyg*. **16**, 1065, **2011**.
- APTE M.G., BUCHANAN I.S.H., MENDELL M.J. Outdoor ozone and building-related symptoms in the BASE study. *Indoor Air*. **18**, 156, **2008**.
- BELL M.L., EBISU K., PENG R.D., SAMET J.M., DOMINICI F. Hospital admissions and chemical composition of fine particle air pollution. *Am J Resp Crit Care*. **179**, 1115, **2009**.
- BARAZZA-VILLARREAL A., ESCAMILLA-NUNEZ M.C., HERNANDEZ-CADENA L., TEXCALAC-SANGRADOR J.L., SIENRA-MONGE J.J., DEL RIO-NAVARRO B.E. Elemental carbon exposure and lung function in school children from Mexico City. *Eur Respir Journal*. **38**, 548, **2011**.
- KEDJARUNE U., KUKIATTRAKOON B., YAPONG B., CHOWANADISAI S., LEGGAT P. Bacterial aerosols in the dental clinic: effect of time, position and type of treatment. *Int Dent J*. **50**, 103, **2000**.
- HELMIS C.G., TZOUTZAS J., FLOCAS H.A., HALIOUS C.H., STATHOPOUIOU O.I., ASSIMAKOPOULOS V.D. Indoor air quality in a dentistry clinic. *Sci Total Environ*. *Medical Dictionary by Farlex*. Excess on 5<sup>th</sup> March 2015. **377**, 349, **2007**.
- PEARSON K. On lines and planes of closest fit to systems of points in space. *Philos Mag Series Series* **2** (11), 559, **1901**.
- AZID A., JUAHIR H., TORIMAN M.E., ENDUT A., KAMARUDIN M.K.A., RAHMAN M.N.A., HASNAM C.N.C., SAUDI A.S.M., YUNUS K. Source Apportionment of Air Pollution: A Case Study in Malaysia. *Jurnal Teknologi (Sciences & Engineering)*. **72** (1), 83, **2015**.
- AZID A., JUAHIR H., TORIMAN M.E., KAMARUDIN M.K.A., SAUDI A.S.M., HASNAM C.N.C., AZIZ N.A.A., AZAMAN F., LATIF M.T., ZAINUDDIN S.F.M., OSMAN M.R., YAMIN M. Prediction of the Level of Air Pollution Using Principal Component Analysis and Artificial Neural Network Techniques: A Case Study in Malaysia. *Water Air Soil Pollut*. **225**, 2063, doi: 10.1007/s11270-014-2063-1. **2014**.
- BAKER M., RAYENS W. Partial Least Square for Discrimination. *Journal of Chemometrics*. **17**, 166, **2003**.
- CHEVALLIER S., BERTRAND D., KOHLER A., COUCOUX P. Application of PLS-DA in multivariate image analysis. *J. Chemometrics*. **20**, 1, **2006**.

20. DOLPHENS M., BARBARA C., COOREVITS P., VLEEMING A., PLAMAN T., DANNEELS L. Posture class prediction of pre-peak height velocity subjects according to gross body segment orientations using linear discriminant analysis. *Eur Spine J.* **23**, 530, **2014**.
21. KLEINBAUM D.G., KUPPER L.L., MULLER K.E. Applied regression analysis and other multivariable methods. PWS-Kent Publishing, Boston. **1988**.
22. VANDEGINSTE B.G.M., MASSART D.L., BUYDENS L.M.C., DE JONG S., LEWI P.J., SMEYERS-VERBEKE J. Data handling in science and technology. Handbook of chemometrics and qualimetrics: Part B. Elsevier B.V., Amsterdam. **1998**.
23. AMRAN M.A., AZID A., JUAHIR H., TORIMAN M.E., MUSTAFA A.D., HASNAM C.N.C., AZAMAN F., KAMARUDIN M.K.A., SAUDI A.S.M., YUNUS K. Spatial Analysis of the Certain Air Pollutants Using Environmetric Techniques. *Jurnal Teknologi.* **75** (1), 241, **2015**.
24. AZID A., JUAHIR H., EZANI E., TORIMAN M.E., ENDUT A., RAHMAN M.N.A., YUNUS K., KAMARUDIN M.K.A., 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**, 1545, **2015**.
25. EEA. Environmental Signals. Improving air quality in Europe. Copenhagen, Denmark. **2013**.
26. WHO. Review of evidence on health aspects of air pollution – REVIHAAP project. Technical Report. Bonn, Germany: World Health Organisation European Centre for Environment and Health. **2013**.
27. CHALLONER A., GILL L. Indoor/outdoor air pollution relationships in ten commercial buildings: PM<sub>2.5</sub> and NO<sub>2</sub>. *Building and Environment.* **80**, 159, **2014**.
28. CAO A.G., YU G., CHEN Y.S., CAO Q.M., FIEDLER H., DENG S.B. Particle size: a missing factor in risk assessment of human exposure to toxic chemicals in settled indoor dust. *Environ Int.* **49**, 24, **2012**.
29. KOPANAKISI, ELEFThERiADiSK., MIHALOPOULOS N., SIMANTRIS-LYDAKIS N., KATSIVELA E., PANTARI D. Physico-chemical characteristics of particulate matter in the Eastern Mediterranean. *Atmos Res.* **106**, 93, **2012**.
30. BZDEK B.R., PENNINGTON R., JOHNSTON M.V. Single particle chemical analysis of ambient ultrafine aerosol: a review. *Journal of Aerosol Science.* **52**, 109, **2012**.
31. RAUB J.A., MATHIEU-NOIF M., HAMPSON N.B., THOM S.R. Review of Carbon monoxide poisoning – a public health perspective. *Toxicology.* **145** (1), 1, **2000**.
32. DEPARTMENT OF SAFETY AND HEALTH. The Industrial Code of Practice Indoor Air Quality (ICOP) published on 2010. **2010**.