The assessment of Danau Kota Lake water quality using chemometrics approach

Hafizan Juahir1,2*, Adiana Ghazali1*, Azimah Ismail1, Mahadzirah Mohamad3, Firdaus Mohamad Hamzah4, Sunardi Sudianto5, Mohd LokmanMohd Lasim6, Mohamed Aidi Shahriz7

1East Coast Environmental Research Institute (ESERI), Universiti Sultan Zainal Abidin (UniSZA), Gong Badak Campus, 21300 Kuala Nerus, Terengganu, MALAYSIA.
2Faculty of Bioresources and Food Industry, Universiti Sultan Zainal Abidin (UniSZA), Besut Campus, 22200 Besut, Terengganu, MALAYSIA.
3Faculty of Business Management and Accountancy, Universiti Sultan Zainal Abidin (UniSZA), Gong Badak Campus, 21300 Kuala Nerus, Terengganu, MALAYSIA.
4Unit of Fundamental Engineering Studies, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, MALAYSIA.
5Department of Biology and Institute of Ecology, Faculty of Mathematic and Natural Sciences, Universitas Padjadjaran, Jalan Raya Bandung-Sumedang Km. 21, Jatinagar, Sumedang 45363, West Java, INDONESIA.
6Kuala Lumpur City Hall, Menara DBKL 1, Jalan Raja Laut, 50350 Kuala Lumpur, MALAYSIA.
7Impian Ekuiti Sdn. Bhd., No.20B, Jalan MJ/1, Taman Maju Jaya, Jalan Klang Lama, 46000 Petaling Jaya, Selangor, MALAYSIA.

*Corresponding authors E-mail: hafizanjuahir@unisza.edu.my, adiana.ghazali@gmail.com

Abstract. Danau Kota Lake is a public attraction area, located in the middle area of Kuala Lumpur. An excessive count of visitors visiting the lake has weakening the lake water quality. The present study was driven to evaluate the lake water quality for environmental conservation. Some water quality and metals data were in-situ measured whereas some were collected for laboratory analysis. BOD, OG, AN and TC are identified as major pollutants towards Danau Kota Lake water quality. The decreasing concentration of DO and increasing of turbidity towards the deeper lake water column has been deteriorated the living water communities and caused hypoxic condition in the bottom water of Danau Kota Lake. Spatial variation of metals in the lake sediment has evidenced the accumulation of metals in the middle part of the lake area due to the sedimentation process. Based on PCA, the possible pollutant sources for this lake are the organic pollutants from the wastewater and industrial pollutants as well as the abundance of suspended algae. Since the water quality of Danau Kota Lake has marked some issues, the present study may assist Kuala Lumpur City Council to prevail the lake water quality challenge promptly by cost saving and in efficient manner.

1. Introduction

A lake usually knows as slow flowing open water with a land surrounding. It is also a group of water bodies that consist of ponds and impoundments. In other term, lakes are absent or lack of exchange with the ocean as they are situated in the inland. They are original from the geographical processes and terminate resulted from the ponding mechanisms with evaporation changing the hydrological balance. Chapman [1] has reviewed numerous result of this mechanism which produces 11 major different lake types and sub-divided into 76 subtypes. These different types of lakes were sharing their ecological and
biogeochemical processes which include biological, physical and chemical properties made the lake a complex ecosystem. Because of this, most organisms are depending on the lake due to output which provides a great ‘ecosystem services’ to the consumer. Nowadays, compilations of studies or research related to the lake are increasing in order to improve a knowledge regarding this subject. Continues gain of knowledge will help to provide a management framework as arise of a problem and issue in the lake dynamic ecosystem [2].

Water catchment development has resulted from a great impact and major transformations of the watershed at the lake have led to the disappearance of some of the ecological component in the lake. Definition of this activities could be defined in two dimensions which through chemical emission and physical degradation [3]. Chemical emission is the product from the pollutant and is solved by the regulation of the system and technology while the physical degradation made up by the environment modification which is needed new and systematic approach. Data collection reflects the change in the surrounding environment which detecting the spatial and temporal changes [4]. This outcome effect of many environment parameters (biotic and abiotic factor) could provide a complete representative change that happens in the lakes Result from the research by Yang et al. [5] reveals that lake natural environment of a lake is a multivariate complex system and difficult to analyze a large number of measured variables and it is required multivariate approaches and metrics for quality assessment. Therefore, a lot of statistical methods like data classification, interpretation and modelling were made to assessing the lake quality parameter [6-8].

Various parameters that present in the water body were detected by many research models and one example is the water quality models (WQMs). This model is valuable tools which supporting the prediction of quality in the water and addressing the characteristic of the waterscape. Most water quality models are created in the best structure for measuring a complex parameter in the water [9]. However, increasing number of parameter that found made research gaps in the water quality models conceded. There is a certain condition that consider will be added some parameter in the water bodies. For example, the effect of rates of flow and wind characteristic could be the effect the water quality parameter. This supported by Jiang et al. [10], simulating of water quality is difficult which only a few studies have recorded that conduct in the different situation. Besides that, most parametric studies need an accurate and high rationality because of the sensitivity of models itself. This change situation could show a full evaluation of the system and this model was expected to produce a high-resolution result when the data collected were applied [11].

Application of environment dataset in the water quality model is important. Eutrophication is the problem that being assessing by the water quality models over 25 years of the research [2]). This becomes a major concern in developing country which gives larges impact to social-economic in quality context of life and human health [12]. An excessive richness of nutrient in the water bodies creates higher biological oxygen demand (BOD) which is indicated a lower level of dissolved oxygen. Increase eutrophication could stress the water quality. This is because; high concentrations of nutrient (nitrogen and phosphorus) in the lake can portray a good water quality. However, with this kind condition will make incapable of producing any useful result in the water quality monitoring. Research is needed to create various techniques to improve water quality as theeutrophication will limit the effectiveness of the models. One of the technique have found are the mixed between various models to produces the best water quality prediction and being revised in the many previous studies [13-14]. overall, water quality needs to be secure in their advantages and should be considered because most parameters in the lake are inherent temporal and spatial variability [15].

The present study was carried out at Danau Kota Lake, Setapak, Kuala Lumpur with the main purpose of investigating the current water quality status of the lake. Besides that, in any research, data analysis is compulsory in order to extract as much information and interpretation for a precise decision making based on monitored data. The collected water quality data from Danau Kota Lake water column has been compared to Malaysian National Water Quality Standard (NWQS) guided by Malaysian Department of Environment. There are six (6) classes (I, IIA, IIB, III, IV and V) which describe the descending order of water quality as Class I is the ‘best’ and Class V is the ‘worst’ [16]. The current
survey has compared the Danau Kota Lake water quality against Class II NWQS (suitable for recreational use body contact) due to the fact that this lake is actively dominated by recreational activities. Additionally, the present study was intentionally applied chemometrics techniques on the collected water quality of Danau Kota Lake and the concentration of heavy metals in the lake sediment.

2. Methodology

2.1. Study area

Danau Kota Lake is a recreational lake which is located in the center city of Kuala Lumpur, Malaysia (Figure 1). This area is considerably as one of the densely populated area and tourist attraction area in Peninsular Malaysia. Additionally, the lake surround area is dominated by the domestic, commercial and industrial land-based activities [17-20]. Danau Kota Lake is a well-known recreational lake as it provides approximately 1.2 km of track length circling around the lake for walking, jogging and cycling. Since this lake is surrounded with various human activities, the lake water is facing a critical pollution issue as the lake sometimes produce unpleasant smell and some irresponsible visitors has been littering around the area which is inconvenient situation for other visitors.

Figure 1. Water quality sampling point at Danau Kota Lake.

2.2. Sample and in-situ data collection

The present study has selected 10 sampling location within Danau Kota Lake area (Figure 1). Samples and in-situ data were collected spatially and temporally on the 27th and 30th of December 2016 as well as on the 3rd to 4th of January 2017. 15 water quality parameters were collected from Danau Kota Lake water during day and night time at all stations. Water quality parameters namely pH, temperature,
dissolved oxygen (DO), total dissolved solid (TDS), ammoniacal nitrogen (AN), total nitrogen (TN) and total phosphate (TP) were in-situ measured using YSI ProMultimeter. In the meantime, turbidity, biological oxygen demand (BOD) and chlorophyll-a (Chl-a) were also in-situ measured, using HACH 2100 Portable Turbidity Meter, Modernwater Portable BOD Meter and Hydrolab DS 5x Sonde respectively. Besides that, some water samples were collected for laboratory analysis in order to define the level of chemical oxygen demand (COD), total suspended solid (TSS), oil and grease (OG), Escherichia coli (E-coli) and total coliform (TC). Analysis for the listed laboratory test water quality parameters were carried out using HACH Method 8000 (COD Reactor Model HACH DRB 200) for COD, gravimetric method (APHA 2540-D) for TSS, USEPA Method 1664b for OG as well as USEPA Method 1604 for E-coli and TC. Meanwhile, the concentration of 22 metals were defined from the lake sediment of five (5) sampling locations. Those parameters are arsenic (As), mercury (Hg), aluminium (Al), cadmium (Cd), chromium (Cr), copper (Cu), lead (Pb), magnesium (Mg), manganese (Mn), selenium (Se), Argentum (Ar), tin (Sn), zinc (Zn), barium (Ba), boron (B), calcium (Ca), nickel (Ni), potassium (K), strontium (Sr), vanadium (V) and cobalt (Co). Danau Kota Lake sediment was digested using mixed concentrated acids and analysed using Inductively Coupled Plasma Mass Spectrometry in order to define the concentration of heavy metals [21-24].

2.3. Statistical analysis

The data matrix in this analysis involves [11 (parameters) x 10 (sampling points)] for surface water quality analysis; [5 (parameters) x 6 (different depths)] for different depth of water quality analysis; as well as [22 (parameters) x 5 (sampling points/sediment samples)] for lake sediment analysis.

2.4. Cluster analysis (CA)

In this study, Hierarchical Aligned Cluster Analysis (HACA) was employed to investigate the grouping of the sampling sites (spatial). HACA is a common method to classify [25] variables or cases (observations/samples) into clusters with high homogeneity level within the class and high heterogeneity level between classes with respect to a predetermined selection criterion [26-27]. Ward’s method using Euclidean distances as a measure of similarity [28-30] within HACA has proved to be a very efficient method. The result is illustrated by a dendogram, presenting the clusters and their proximity [31]. The Euclidean distance (linkage distance) is reported as $D_{\text{link}} / D_{\text{max}}$, which represents the quotient between the linkage distance divided by the maximal distance. The quotient is usually multiplied by 100 as a way to standardize the linkage distance represented by the y-axis [32].

2.5. Discriminant analysis (DA)

Discriminant analysis determines the variables that discriminate between two or more naturally occurring groups/clusters. It constructs a discriminant function (DF) for each group [33-34]. DFs are calculated using equation (1):

$$ f(G_i) = k_i + \sum_{j=1}^{n} w_{ij} P_j $$

where $i$ is the number of groups ($G$), $k_i$ is the constant inherent to each group, $n$ is the number of parameters used to classify a set of data into a given group, and $w_{ij}$ is the weight coefficient assigned by DF analysis (DFA) to a given parameter ($P_j$).

In this study, DA was applied to determinwhether the groups differ with regard to the mean of a variable and to use that variable to predict group membership. Two groups for temporal analysis (two different sampling time are day and night), which standard, forward stepwise and backward stepwise modes. These were used to construct DFs to evaluate temporal variations in the lake water quality. The sampling time (temporal) were the grouping (dependent) variables, while all the measured parameters constitute the independent variables. In the forward stepwise mode, variables are included step by step beginning
with the most significant variable until no significant changes were obtained. In the backward stepwise mode, variables are removed step by step beginning with the less significant variable until no significant changes were obtained.

2.6. Principal Component Analysis (PCA)

The most powerful pattern recognition technique that is usually coupled with HACA is the PCA. It provides information on the most significant parameters due to spatial and temporal variations that describes the whole data set by excluding the less significant parameters with minimum loss of original information [32]. The principle component (PC) can be expressed as:

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + \ldots + a_{im}x_{mj}.$$  

(2)

where $z$ is the component score, $a$ is the component loading, $x$ is the measured value of the variable, $i$ is the component number, $j$ is the sample number, and $m$ is the total number of variables.

The FA is usually applied as a method to interpret a large complex data matrix and offers a powerful means of detecting similarities among variables or samples [35]. The PCs generated by PCA are sometimes not readily interpreted; therefore, it is advisable to rotate the PCs by varimax rotation. Varimax rotations applied on the PCs with eigenvalues more than 1 are considered significant [36-37] in order to obtain new groups of variables called varimax factors (VFs). The number of VFs obtained by varimax rotations is equal to the number of variables in accordance with common features and can include unobservable, hypothetical, are considered as “strong”; 0.75–0.50, as “moderate”; and 0.50–0.30, as “weak” significant factor loadings [38].

Source identification of different pollutants was made on the basis of different activities in the surrounding lake area in light of previous literatures. The basic concept of Factors Analysis (FA) is expressed as:

$$z_{ij} = a_{i1}f_{1j} + a_{i2}f_{2j} + \ldots + a_{im}f_{mj} + e_{ij}.$$  

(3)

where $z$ is the measured value of a variable, $a$ is the factor loading, $f$ is the factor score, $e$ is the residual term accounting for errors or other sources of variation, $i$ is the sample number, $j$ is the variable number, and $m$ is the total number of factors.

In this study, PCA and FA was applied to the data sets of the lake water quality laboratory analysis (11 variables) for the 10 different sampling stations (points). The temporal variation of lake water quality was carried out due to two different sampling time which is day and night. The input data matrices (variables × cases) for PCA and FA were 11x10 for both sampling periods, day and night.

3. Results and discussion

3.1. Danau Kota Lake water quality criteria assessment based on Class IIB Malaysian NWQS

Danau Kota Lake water quality which has been observed spatially at 10 sampling stations were compared to Class IIB NWQS. Table 1 shows that lake water quality data which is comply (below limit) and not-comply (above limit) Class IIB NWQS for both sampling period i.e. during the day and night time. Based on the comparison, it is clearly demonstrating that, BOD, OG, AN and TC are 100% not comply with NWQS during daytime monitoring of Danau Kota Lake water quality samples. DO concentration is not complying with Class IIB NWQS during the day time samples at nine (9) stations except for station SP1. All DO, BOD, OG and AN data collected during the night time period are not complying with Class IIB NWQS. Besides that, 56% (61 count data) and 48% (53 count data) of temperature, pH, TSS and TP are complying with Class IIB NWQS during day and night time period respectively.
### Table 1. The assessment of observed Danau Kota Lake water quality in comparison with Class IIB NWQS.

| Station | DO | pH   | BOD  | COD  | TSS  | OG   | AN  | TN   | TP   | Chl-a  | TC   | E.coli |
|---------|----|------|------|------|------|------|-----|------|------|--------|------|--------|
|          |    |      |      |      |      |      |     |      |      |        |      |        |
| **DAY** |    |      |      |      |      |      |     |      |      |        |      |        |
| SP1     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP2     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP3     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP4     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP5     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP6     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP7     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP8     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP9     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP10    |    |      |      |      |      |      |     |      |      |        |      |        |
| **NIGHT** |    |      |      |      |      |      |     |      |      |        |      |        |
| SP1     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP2     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP3     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP4     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP5     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP6     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP7     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP8     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP9     |    |      |      |      |      |      |     |      |      |        |      |        |
| SP10    |    |      |      |      |      |      |     |      |      |        |      |        |

| Limit   | 5-7 | 6.0-9.0 | 3  | 25  | 50  | 0.04 | 0.3 | 5   | 5   | 10    | 5000 | 5000  |
|---------|-----|---------|----|-----|-----|------|-----|-----|-----|------|------|------|

Note: below limit; above limit
Figure 2. Pareto charts of major contributors toward Danau Kota Lake water pollution during both sampling periods.

Pareto charts in Figure 2 highlight the major pollutants which contribute to the serious water problems in Danau Kota Lake. Those pollutants are BOD, OG, AN and TC with high cumulative percentage recorded of 66% and 74% both sampling periods. Therefore, the present study has drawn the need of local authority to pay particular attention in controlling these pollutants, particularly on the untreated direct discharge into the Danau Kota Lake water system from the surrounding area.

3.2. Spatial classification of Danau Kota Lake water quality

The 15 water quality parameters observed at 10 sampling stations in Danau Kota Lake were classified spatially using HACA. Dendograms in Figure 3 give two clusters during day time period and three clusters during night time period. Based on day time dendogram, Cluster 1 is consisting of eight sampling points namely SP1, SP2, SP3, SP4, SP5, SP7, SP9, and SP10. In the meantime, Cluster 2 is consisting of two sampling points which are SP6 and SP8. For the time being, based on the night sampling period, Cluster 1 consists of eight sampling points which are SP1, SP3, SP4, SP6, SP7, SP8, SP9 and SP10. Besides that, Cluster 2 and Cluster 3 consist of one sampling point which is SP2 and SP5 respectively. Therefore, in consideration for future sampling plan, it is sufficient to collect samples from only two (2) sampling points which represent each cluster during day time period as well as three (3) sampling points which represent three (3) clusters during the night time period.
3.3. Spatial classification of Danau Kota Lake sediment based on metal variations.

The spatial classification of Danau Kota Lake sediment was performed based on the concentration of 22 metals. However, only 17 metals were taken into consideration for further classification analysis, as other five (5) were removed due to no significant variation throughout the sampling period. Dendrogram in Figure 4 shows that there are three (3) clusters were formed spatially based on the similarity of metals concentration in Danau Kota Lake sediment. The clusters can be interpreted as High Contaminated Sampling Point (HCSP) which consists of station SP2 and SP4; Moderate Contaminated Sampling Point (MCSP) which comprises of station SP5; as well as Low Contaminated Sampling Point (LCSP) which involves of station SP3 and SP10 (Figure 5). Based on the spatial cluster analysis on metals concentration in the sediment sample; in order to monitor the future metals contamination in Danau Kota Lake sediment, it is strongly suggested to collect the sediment samples at three (3) stations which represents three (3) formed clusters. As a consequence, by practicing this method, the future sampling cost and time can be reduced. Other than that, the spatial distribution of metals in Danau Kota Lake sediment has emphasized the polluted area was within the middle part of the lake area. Thus it is strongly presuming that metals input from the surrounding area was accumulated during the lake sedimentation process towards the middle part of the lake area.
3.4. Spatial variation of water quality based on different water depth

Table 2 and Figure 6 show the spatial distribution of lake water quality according to the water depth. Based on the observed data, pH and DO were enriched at the surface water but depleted at the bottom water. On the other hand, the concentration of TDS was low at the surface water but high at the bottom water of Danau Kota Lake. In the meantime, temperature and turbidity did not show any particular trend. Additionally, data analysis has shown the lowest concentration of pH and DO as well as an increased of TDS was observed at the deepest depth of Danau Kota Lake water column. Obviously, the difference of pH levels recorded was corresponded to different water levels or depth of Danau Kota Lake as it could be due to the increased of carbon dioxide (CO₂) level from the respiratory activities and the decomposition of organic materials below the thermocline. Hence the saturated concentration of CO₂ accumulates in the lower strata of the lake. However, the concentration of DO increases as the pressure increases. DO saturation decreases by 10% per meter as depth of lake increases, resulting from the hydrostatic pressure. To these, the fact is that, the concentration of dissolved oxygen decreases downward to the lake bottom [35,38].
In the common notion of water quality, the high levels of TDS lead to the increasing of water temperature and decreasing of DO concentration. It is commonly attributed to the absorbance of more heat from solar radiation by the suspended particles [35]. The absorbed heat is then directly transferred to the surrounding water through the conduction process. The warmer water unlikely has potential to hold as much as DO compare to cold water, resulting in a significant decline in DO levels. In addition, as the surface temperature increases, it can cause the stratification, or layering of lake-water. Furthermore, whenever the lake water stratifies, the upper and lower layers will form different layers. No mixture between the layers. The frequent occurrence of decomposition and respiration in the lower layers of lake-water creates the condition of hypoxic (low level of DO). Therefore, it will aggravate the living communities or organisms in the lake to survive.

Based on data analysis, temperature and turbidity are not statistically significant (p>0.05). Temperature tends to deplete in the deeper water depth, in contrary with turbidity where it was spatially increased from the water surface downward to the lake bottom.

![Figure 6. Spatial water quality of Danau Kota Lake at different water depth.](image)
Table 2. Mann-Kendall non parametric test of trend on Danau Kota Lake water quality based on different water depth.

| Variable | Kendall's tau | S    | Var(S) | Trend | p-value |
|----------|---------------|------|--------|-------|---------|
| pH       | -0.905        | -19.000 | 0.000 | ↓     | 0.003   |
| Temp.    | -0.524        | -11.000 | 0.000 | NT    | 0.136   |
| DO       | -0.905        | -19.000 | 0.000 | ↓     | 0.003   |
| Turbidity| 0.238         | 5.000  | 0.000  | NT    | 0.562   |
| TDS      | 0.714         | 15.000 | 0.000  | ↑     | 0.030   |

Note: ↑-upward trend; ↓-downward trend; NT-no trend

3.5. Temporal Variation of Danau Kota Lake Water Quality

3.5.1. The variation of water quality parameters temporally using DA.

In order to study the temporal variation of water quality data, DA was applied to the raw data post grouping of Danau Kota Lake water quality. The collected water quality data was grouped into two main clusters/groups, defined by day and night time sampling period. Both groups were treated as the dependent variables, while the water quality parameters were treated as the independent variables. DA was performed through modes of standard, forward stepwise, and backward stepwise modes with calculated DFA of 94.4% (11 discriminant variables), 72.2% (one discriminant variable), and 94.4% (three discriminant variables), respectively (Figure 8 and Table 3). The forward stepwise mode has shown that only pH found to be the significant variable. This indicates that pH have high variation in terms of their temporal distribution. pH shows a lower value during the day time (µ Day = 7.12) but slightly higher during the night time (µ Night = 8.13). On the other hand, the backward stepwise mode of which included TSS (µ Day = 20 mg/L; µ Night = 7 mg/L) and TN (µ Day = 5.05 mg/L; µ Night = 4.48 mg/L) as the second and third parameter show a high temporal variation. Both parameters show slightly higher concentrations during day time period compared to night time period. This pattern was anticipated due to the watering activities of landscape plants and ornaments home during daytime.

Figure 7. Plot of discriminant functions for temporal variation of Danau Kota Lake water quality.
Apart from that, the residual of chemical fertilizer with high contents of nitrogen and plant soil transported by water residual during the watering activities has been flushed into the drain and finally directly goes into the lake systems. In the meantime, the remaining eight water quality parameters depicted no significant difference ($p>0.05$) between both sampling periods. Box and whisker plots of pH, TSS and TN recorded during the sampling period are shown in Figure 8. These parameters which yielded the highest variability (the most significant ($p<0.05$)) was analysed using backward stepwise of the DA and the output will be used for further discussion.

### Table 3. Classification matrix of temporal DA for Danau Kota Lake water quality.

| Time  | DA     | Total | % correct |
|-------|--------|-------|-----------|
| Day   | 9      | 9     | 100.00%   |
| Night | 1      | 8     | 88.89%    |
| Total | 10     | 18    | 94.44%    |

**Standard mode DA**

| Time | DA     | Total | % correct |
|------|--------|-------|-----------|
| Day  | 7      | 9     | 77.78%    |
| Night| 3      | 9     | 66.67%    |
| Total| 10     | 18    | 72.22%    |

**Stepwise forward DA**

| Time | DA     | Total | % correct |
|------|--------|-------|-----------|
| Day  | 9      | 9     | 100.00%   |
| Night| 1      | 8     | 88.89%    |
| Total| 10     | 18    | 94.44%    |

**Stepwise backward DA**

| Time | DA     | Total | % correct |
|------|--------|-------|-----------|
| Day  | 7      | 9     | 77.78%    |
| Night| 3      | 9     | 66.67%    |
| Total| 10     | 18    | 72.22%    |
3.5.2. Identification of the possible pollution sources using PCA

PCA was employed to the data set to compare the compositional patterns between the examined water parameters and to identify the factors that influence the variation of lake water quality. Three PCs were obtained with eigenvalues larger than one summing almost 87.6% of the total variance in the data set. Corresponding varimax factors (VFs), variable loadings, and variance explained are presented in Table 4.

Table 4. Loadings of varimax-rotated PCs for water quality data collected in Danau Kota Lake.

| Variables | VF1  | VF2  | VF3  |
|-----------|------|------|------|
| pH        | -0.379 | 0.306 | -0.127 |
| BOD       | 0.978 | -0.011 | -0.176 |
| COD       | 0.973 | 0.059 | 0.024 |
| TSS       | 0.707 | -0.535 | -0.418 |
| O&G       | 0.98 | -0.008 | -0.138 |
| AN        | -0.627 | 0.197 | 0.558 |
| TN        | 0.205 | 0.955 | -0.056 |
| TP        | 0.976 | -0.123 | -0.058 |
| Ch-a      | -0.116 | -0.076 | 0.917 |
| Coliform  | 0.964 | 0.171 | -0.195 |
| E.coli    | 0.91 | 0.299 | -0.252 |

| Eigenvalue | 7.056 | 1.481 | 1.102 |
| Variability (%) | 64.147 | 13.464 | 10.014 |
For daytime, VF1 is accounted for 64.2% of the total variance, reveals the strong positive loadings on the variables of BOD (0.98), COD (0.97), TSS (0.71), O&G (0.98), TP (0.98), coliform (0.95) and E.coli (0.91). Strong positive loadings on BOD and COD prove the influence of organic pollutants from the point sources. These pollutants are commonly related to the anthropogenic pollution sources such as wastewater or sewage treatment plants, domestic wastewater and industrial effluents [39-41].

TSS can be originated from both wastewater treatment plants (point source, PS) and polluted non-point source (NPS) [42-43]. Solids are also known to be one of the most common contaminants found in urban storm water. Commonly they are originated from many sources which include the erosion of pervious surfaces and dust; litter and other particles deposited on the impervious surfaces from anthropogenic activities and atmospheric deposition. Additionally, the soil erosion at construction sites and robust development is counted as a he major source of TSS into the lake water column. Therefore, the urban waterways create more TSS which lead to the increase of turbidity and reduce the penetration of light in the lake water column. Consecutively, low penetration of light in the water cause limitation in the growth of desirable aquatic plants.

Other than that, TSS which settling inside the lake ecosystem as the bottom deposits will lead to the sedimentation, alter and eventually destroys the habitat of many fish species and bottom-dwelling organisms [39-40,44]. Besides that, the settling down of TSS on the lake bed surface provides a medium for the accumulation, transport and storage of other pollutants including nutrients, pesticides, metals as well as OG pollutants. According to Bastian [45], the typical concentration of TSS in urban runoff is substantially higher than that in treated wastewater. Therefore, the construction produces the highest loading of TSS over other urban land use categories. In the meantime, OG pollutant sources are most possibly derive from roads, driveways, parking lots, vehicle maintenance areas, gas stations, illicit dumping into storm drains and domestics cooking wastes [46].

TP is the main nutrients found in urban storm water. The major sources of nutrients in urban storm water are lawn fertilizers, atmospheric deposition, automobile exhaust, soil erosion, animal waste and detergents [47]. TP measures the total amount of phosphorus in both the organic and inorganic forms. The degree to which nitrogen and phosphorus are present in a lake can determine the trophic status and the amount of algal biomass produced. Excess nutrients tend to increase the primary biological productivity. The major impact associated with nutrient over enrichment is excessive growth of algae that leads to nuisance algal blooms and eutrophic conditions [48-49]. A secondary impact is the residual negative effect of decomposing algae in the form of sediment oxygen demand that depletes dissolved oxygen concentrations, particularly in bottom waters [50].

TC and E. coli which enters the lake water system are most possibly related to municipal sewage and wastewater treatment plants as well as domestic wastewater [51]. The presence of coliform and E. coli in lake water consumes a large amount of oxygen. As the amount of available DO decreases, they will undergo anaerobic fermentation processes which leading to the production of ammonia and organic acids [52]. The hydrolysis of these acidic materials causes a decrease of water pH values [53].

VF2 explains 13.5% of the total variance with a strong loading on TN. TN pollution is a crucial to the entire Klang Valley region. The main caused are identified from the untreated excreta in sewage directly discharged into the lake. In addition, the TN presence in the environment from the decomposition of the organic matter, such as food residue discarded directly into the waterways [54].

VF3 explains 10% of the total variance with 0.917 strong loading on Chl-a. The high loading of the Chl-a could be derived from the abundance of suspended algae within the lake. The accumulation of the Chl-a in the water can reduce the dissolve oxygen that harmful to the various species of living organism [50].
4. Conclusion and recommendation

4.1. Conclusion

The spatial and temporal study of Danau Kota Lake is important in order to enhance the water quality of the lake as it is located in the middle of Kuala Lumpur city center. The present study is important as it could aid the Kuala Lumpur City Council to come up with a pragmatic and precise decision making, conserve and preserve Danau Kota Lake water column. CA has classified the polluted area in Danau Kota Lake into three (3) groups namely HCSP, MCSP and LCSP. In the meantime, the concentration of DO decreased downward to the lake bottom, which is aggravating the living communities in the lake. It could be due to the decomposition and respiration in the lower layers of lake-water. Therefore, it creates the condition of hypoxic (low dissolved oxygen levels). The output of temporal DA backward stepwise mode reveals that three variables namely pH, TSS and TN has influenced the lake water quality during the night time period. Through PCA, strong positive loadings were shown by BOD (0.98), COD (0.97), TSS (0.71), O&G (0.98), TP (0.98), coliform (0.95) and E.coli (0.91). The strong positive loadings on these variables could possibly derived from the anthropogenic pollution sources such as wastewater or sewage treatment plants, domestic wastewater and industrial effluents. Since Danau Kota Lake is a public attraction lake, it is important to maintain the physical view of the lake water. High level of turbidity lead to a murkier lake water which is unpleasant view for the visitors. Additionally, high turbidity also cause limitation in sunlight penetration into the water column which impact the growth of desirable aquatic plants. Therefore, the present study could help Kuala Lumpur City Council to overcome the lake water quality problem with expeditiously, cost saving and efficient manners.

4.2. Recommendation

The Kuala Lumpur City Council should implement the water quality monitoring program to Danau Kota Lake and other lakes within the Klang Valley in the agenda of future plan towards achieving the environment sustainability. By maintaining the water quality of the lake, the lake itself safe for any form of outdoor water activities and aesthetically preserved. There are few recommendations should be taken into Kuala Lumpur City Council consideration based on the present study findings, as follow: (i) On-ground verification as to reconfirm the point source of contamination flowing into the lake for the next course of treatment. The flow has to be treated before flowing into the lake; (ii) Control the point-sources prior the flow being treated.

To find the best water treatment technology to solve the water quality problem efficiently in order to meet the Class IIB NWQS.

5. Acknowledgement

Kuala Lumpur City Hall (DBKL) funded this research through a research grant (grant numbers: I1017-0008). The authors would like to thank DBKL for the research grant and the East Coast Environmental Research Institute (ESERI) for the assistance given during sampling and analysis.

References

[1] Chapman D V 1996 Water quality assessments: a guide to use of biota, sediments and water in environmental monitoring.
[2] Brönmark C and Hansson L A 2002 Environmental issues in lakes and ponds: current state and perspectives. *Environmental conservation*, 29(3), 290-307.
[3] Zalewski M 2007 Ecohydrology as a concept and management tool. *Managing Water in Diverse Ecosystems to Ensure Human Well-being*. United Nations University-International Network on Water, Environment and Health (UNU-INWEH), Hamilton, pp39-50.
[4] Matthews G B, Matthews R A and Hachmöller B 1991 Mathematical analysis of temporal and spatial trends in the benthic macroinvertebrate communities of a small stream. *Canadian Journal of Fisheries and Aquatic Sciences*, 48(11), 2184-2190.

[5] Yang G, Ma R, Zhang L, Jiang J, Yao S, Zhang M and Zeng H D 2010 Lake status, major problems and protection strategy in China. *Journal of Lake Sciences*, 22(6), 799-810.

[6] Simeonov V, Stratis J A, Samara C, Zachariadis G, Voutsa D, Anthemidis A and Kouimtzis T 2003 Assessment of the surface water quality in Northern Greece. *Water research*, 37(17), 4119-4124.

[7] Massart D L, Vandeginste B G M, Buydens L M C, De Jong S, Lewi P J and Smeyers-Verbeke J 1998 *Handbook of chemometrics and qualimetrics; data handling in science and technology, parts A and B.*, Elsevier: Amsterdam.

[8] Zhou F, Liu Y and Guo H 2007 Application of multivariate statistical methods to water quality assessment of the waterscources in Northwestern New Territories, Hong Kong. *Environmental Monitoring and Assessment*, 132(1-3), 1-13.

[9] Wang Y, Wang P, Bai Y, Tian Z, Li J, Shao X and Li BL 2013 Assessment of surface water quality via multivariate statistical techniques: a case study of the Songhua River Harbin region, China. *Journal of Hydro-Environment Research*, 7(1), 30-40.

[10] Jiang L, Li Y, Zhao X, Tillotson M R, Wang W, Zhang S, Sarpong L, Asmaa Q and Pan B 2018 Parameter uncertainty and sensitivity analysis of water quality model in Lake Taihu, China. *Ecological Modelling*, 375, 1-12.

[11] Hughes D A and Slaughter A R 2016 Disaggregating the components of a monthly water resources system model to daily values for use with a water quality model. *Environmental Modelling & Software*, 80, 122-131.

[12] Tundisi J G, Santos M A and Menezes C F S 2003 Tucuruí reservoir and hydroelectric power plant. Sharing Experiences and Lessons Learned in Lake Basin Management, Burlington, Vermont. *Management Experiences and Lessons Learned Brief*, 1, 1-20.

[13] Luo Y, Yang K, Yu Z, Chen J, Xu Y, Zhou X and Yang Y 2017 Dynamic monitoring and prediction of Dianchi Lake cyanobacteria outbreaks in the context of rapid urbanization. *Environmental Science and Pollution Research*, 24(6), 5335-5348.

[14] Wei Z, Han Y, Li M, Yang K, Yang Y, Luo Y and Ong SH 2017 A small UAV based multi-temporal image registration for dynamic agricultural terrace monitoring. *Remote Sensing*, 9(9), 904.

[15] Missaghi S, Hondzo M and Melching C 2013 Three-dimensional lake water quality modeling: sensitivity and uncertainty analyses. *Journal of environmental quality*, 42(6), 1684-1698.

[16] Ismail A, Toriman ME, Juahir H, Zain SM, Habir NLA, Retnam A and Azid A 2016 Spatial assessment and source identification of heavy metals pollution in surface water using several chemometric techniques. *Marine pollution bulletin*, 106(1), 292-300.

[17] Mohamad M, Ali A M, Ab Ghani N I, Abdullah A R and Mokhlis S 2012 Positioning Malaysia as a tourist destination based on destination loyalty. *Asian Social Science* 9(1), 286-292.

[18] Awang Z, Afthanorhan A, Mohamad M and Asri M A M 2015 An evaluation of measurement model for medical tourism research: The confirmatory factor analysis approach. *International Journal of Tourism Policy* 6(1), 29-45.
[19] Mohamad M, Ghani N I A, Mamat M and Mamat I 2014 Satisfaction as a mediator to the relationships between destination image and loyalty. *World Applied Sciences Journal* 30(9), 1113-1123.

[20] Kashif M, Sansi S Z M, Awang Z and Mohamad M 2016 EXQ: measurement of health care experience quality of Malaysian settings: A contextualist perspective. *International Journal of Pharmaceutical and Healthcare Marketing* 10(1), 27-47.

[21] Ong M C, Joseph B, Shazili N A M, Ghazali A and Mohamad M N 2015 Heavy Metals Concentration in Surficial Sediments of Bidong Island, South China Sea off the East Coast of Peninsular Malaysia. *Asian Journal of Earth Sciences*, 8(3), 74

[22] Ghazali A, Shazili N A M, Bidai J and Shaari H 2016 The spatial distribution of Al, Fe, Cu, Cd and Pb in the surface sediment of Brunei Bay, Borneo during the southwest and northeast monsoons. *Journal of Sustainability Science and Management*, 11, 93-106

[23] Bidai J, Adiana G and Shazili N A M 2016 Particle sizes effect of the elements in the South China Sea sediment off Pahang coastal during the Northeast monsoon (pre-) and the Southwest monsoon (post-) periods. *Environmental Earth Sciences*, 75(7), 614

[24] Adiana G, Shazili N A M and Marinah M M 2011 Cadmium, manganese and lead distribution in the South China Sea off the South Terengganu Coast, Malaysia during post-monsoon and pre-monsoon. *Journal of Sustainability Science and Management*, 6(2), 181-192

[25] Massart D L and Kaufman L 1983 The interpretation of analytical data by the use of cluster analysis. New York: Wiley.

[26] McKenna J E Jr 2003 An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. *Environmental Modelling & Software*, 18(2), 205–220.

[27] Adiana G, Juahir H, Joseph B and Shazili N A M 2017 Tracing the sources of lead (Pb) in Brunei Bay, Borneo by using integrated spectrometry ICP-MS and chemometric techniques. *Marine pollution bulletin*, 123(1-2), 232-240

[28] Willet P 1987 *Similarity and clustering in chemical information systems*. New York: Research Studies Press, Wiley.

[29] Adams M J 1998 The principles of multivariate data analysis. *In* Ashurst PR & Dennis MJ (Eds), *Analytical methods of food authentication*, Blackie Academic & Professional, London, pp350.

[30] Otto M 1998 Multivariate methods. *In* R. Kellner, J.M.Mermet, M. Otto, & H. M. Widmer (Eds.), *Analyticalchemistry*. Wenheim: Wiley-VCH.

[31] Forina M, Armanino C and Raggio V 2002 Clustering with dendograms on interpretation variables. *Analytica Chimica Acta*, 454, 13–19.

[32] Singh K P, Malik A and Sinha S 2005 Water quality assessment and apportionment of pollution sources of Gomti River (India) using multivariate statistical techniques: A case study. *Analytica Chimica Acta*, 35, 3581–3592.

[33] Johnson R A and Wichern D W 1992 *Applied multivariate statistical analysis* (3rd ed.). Prentice-Hall Int.:New Jersey.

[34] Juahir H, Sharifuddin M Z, MohdKamil Y, TengkuHanidza T I, MohaArmi A S, Toriman M E & Mokhtar M 2011 Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques. *Environmental Monitoring Assessment* 173, 625-641.

[35] Reghunath R, Murthy S T R and Raghavan B R 2002 The utility of multivariate statistical techniques in hydrogeochemical studies: An example from Karnataka, India. *Water Research*, 36, 2437–2442.
[36]   Adiana G, Joseph B, Shazili N A M, Mohamad M and Juahir H 2018 Spectrometry and chemometrics interpretation of metals in the Sulu and Celebes Seas of Malaysian waters during PMSE 09’ Expedition. Journal of Engineering and Applied Sciences 13(7), 1745-1756.

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

[38]   Liu C W, Lin K H and Kuo Y M 2003 Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan. The Science of the Total Environment, 313, 77–89.

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

[40]   Juahir H, Zain S M, Aris A Z, Yusof M K, Samah M A A and Mokhtar M 2010 Hydrological trend analysis due to land used changes at Langat River Basin. Environment Asia 3, 20-31.

[41]   Mohd K A K, Mohd E T and Nur H S 2015 Classification of tropical river using chemometrics technique: Case study in Pahang River, Malaysia. Malaysian Journal of Analytical Sciences 19(5), 1001-1018.

[42]   Dolan D M 1993 Point source loadings of phosphorus to Lake Erie: 1986-1990. Journal of Great Lakes Research 19(2), 212-223.

[43]   Ha S R and Bae M S 2001 Effects of land use and municipal wastewater treatment changes on stream water quality. Environmental Monitoring and Assessment 70(1-2), 135-151.

[44]   Kamarudin M K A, Toriman M E, Rosli M H, Juahir H, Aziz N A A, Azid A and Sulaiman W N A 2015 Analysis of meander evolution studies on effect from land use and climate change at the upstream reach of the Pahang River, Malaysia. Mitigation and Adaptation Strategies for Global Change 20(8), 1319-1334.

[45]   Bastian R K 1997 Biosolids management in the United States. Water Environment & Technology 9(5), 45-50.

[46]   Rozell D J and Reaven S J 2011 Water pollution risk associated with natural gas extraction from the Marcellus Shale. Risk Analysis 32(8), 1382-1393.

[47]   Dodds W K & Smith V H 2016 Nitrogen, phosphorous and eutrophication in streams. International Society of Limnology 6, 155-164.

[48]   Dodds W K, Bouska W W, Eitzmann J L, Pilger T J, Pitts K L, Riley A J, Schloesser J T and Thombrugh D J 2009 Eutrophication of U.S. freshwaters: Analysis of potential economic damages. Environmental Science and Technology 43(1):12-19.

[49]   Smith V H, Tilman G D and Nekola J C 1999 Eutrophication: impacts of excess nutrient inputs on freshwater, marine and terrestrial ecosystems. Environmental Pollution 100(1-3), 179-196.

[50]   Withers P J A and Jarvie H P 2008 Delivery and cycling of phosphorous in rivers: A review. Science of the Total Environment 400(1-3), 379-395.

[51]   Frenzel S A and Couvillion C S 2002 Fecal-indicator bacteria in streams along a gradient of residential development. Journal of the American Water Resources Association 38(1), 265-273.

[52]   Mara D 2005 Pond process design - a practical guide. In Andy Shilton (Ed). Pond Treatment Technology IWA Publishing, UK, pp100-131.

[53]   Alrumman S A, El-kott A F and Keshk S M A S 2016 Water pollution: Source and treatment. American Journal of Environmental Engineering 6(3), 88-98.
[54] Hamilton A J, Stagnitti F, Xiong X, Kreidl S L, Benke K K and Maher P 2007 Wastewater irrigation: The state of play. *Vadose Zone Journal* 6(4), 823-840.