POTENTIAL OF MULTIVARIATE ANALYSIS IN SUBSTANCE ABUSE RESEARCH

A. Fazillah[1,*], H. Juahir[1], E. Toriman[2], N. Mohamad[3], M. K. A. Kamarudin[1] and M. Mohamad[4]

[1]East Coast Environmental Research Institute, Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Terengganu, Malaysia

[2]School of Social Development and Environmental Studies, Faculty of Social Science and Humanities, National University of Malaysia, 43600 Bangi, Selangor, Malaysia

[3]Medical Centre, Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Terengganu, Malaysia

[4]Faculty of Economics and Management Sciences, Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Terengganu, Malaysia

ABSTRACT

Substance abuse has a close association with health and social problems. This study demonstrates the usefulness of multivariate statistical analysis in substance abuse studies. Applying multivariate analyses such as the factor analysis, principal component analysis, cluster analysis and discriminant analysis in other studies gives a view on how this method can be applied in the substance abuse field. To date, there are limited studies on applying this powerful method in this study. However, looking forward on the capability of the multivariate analysis on handling and interpreting the complexity of the data, some potential goals that can be achieved in substance abuse study can be highlighted.

Author Correspondence, e-mail: fazillah1988@gmail.com

These are identifying the possible factors associated with substance abuse, characterising the spatial and temporal variation of substance abuse and developing a substance abuse index.

Keywords: multivariate analysis; substance abuse; factor analysis; principal component analysis; cluster analysis; discriminant analysis.

1. INTRODUCTION

Interpreting the data through the analysis of raw data is important for each research. A tool and method for analysis should be done correctly without losing any data, even the simplest data, because it will disturb a result and may cause an error. Numerous researchers have used mathematical methods to analyse their data. There is no denying that in each study, the data collected has to be analysed in order to get information and interpret the complexity of data. Through a comprehensive analysis, we believe that this is helpful for a better understanding of the study and will lead us to greater comprehension on the matter. One powerful statistical analysis that is widely applied in many areas of study is the multivariate analysis. A wide range of studies have been done using the multivariate analysis for better interpretation of the complexity of data as well as a better understanding of their studies [1-2]. Most of these have been in the environmental field. Literature contains successes in using multivariate analysis in such fields as hydrochemical studies [3-4] water quality [2, 5] and groundwater assessment [6-7]. Recently, some other fields of study have seen growing interest in applying this method to their studies, like public health. We can see that this method is very convenient for application in substance abuse study.

Substance abuse is one of the most complicated health and social problems faced by many countries [8]. It is globally known that substance abuse is destructive to humanity and should be controlled internationally. This problem has become considerable in social and health spheres worldwide [9]. Although various attempts have been made from the macro level to micro level, there are still millions of people reported as being involved in substance abuse problems, and what is really worrying is the abuse of this substance involves adolescents and it is expected that this will significantly increase in the future [10]. A lot of studies have investigated substance abuse problems among adolescents [10-12]. Other studies showed that

substance abuse is associated with age [13] and the majority of substance abusers start using illicit substances when they were adolescents [14-15].

Furthermore, Malaysia is not excused from facing this problem which has becomes one of the country's main problems. The country's tagline "Dadah Musuh Negara" meaning that substance as a scourge of the nation [16] clearly shows that the government has emphasized eradicating this problem. Substance abuse in Malaysia is closely related to the economic factor. According to history, the emergence of substance (drug) in Malaysia has begun since before Independence Day and it involved labourers. At the beginning, the substance was used among labourers and fisherman for treating asthma. However, overuse of this substance caused a strong dependence to the substance which is acknowledged as addiction [17]. A lot of effort has been done to tackle this problem including stressing the dangers of drugs via media, providing rehabilitation centres, providing counselling and so on. Additionally, many non-governmental organizations have tried taking care of this problem by setting up a treatment centres. However, until now the problem still remains.

In recent years, the multivariate analysis has been successfully used in many fields. As a response to that, this paper brings forward the capability of using the multivariate analysis in interpreting data through the review of some studies which use this method. Furthermore, it is believed that it has the ability to help researchers in effort of solving substance abuse problems. The performance of the multivariate analysis in various fields is reviewed in this paper to give a picture on how it can be applied in substance abuse management.

2. POTENTIAL MULTIVARIATE ANALYSIS OF SUBSTANCES ABUSE PROBLEM

The multivariate analysis has been widely applied to a variety of fields of study to analyse and interpret data. It is a statistical method used to analyse multiple dependent variables and multiple independent variables, hence offering a valuable tool for evaluation and interpreting data [18]. Multivariate analysis has served as a robust tool to resolve complex data processing and interpretation [19]. A lot of studies in the environmental field, particularly in water quality studies were carried out using the multivariate analysis. The multivariate analysis is capable to organize and simplify the data to output useful information [20]. Among famous

multivariate analysis techniques used in many studies are factor analysis (FA), principal component analysis (PCA), cluster analysis (CA) and discriminant analysis (DA).

PCA and FA are an unbiased method used to reduce the dimensionality of the data into smaller numbers and identify the variables which explain the most variation in the sample [17, 21-22], hence summarizing data with minimal losses of information [15]. PCA and FA try to explain the correlation between variables in terms of underlying factors which are not directly observable [23]. Usually, before a further analysis will be carried out, the PCA is applied to reduce the dimensionality of the database and concurrently point out the most relevant information for the data set [24]. In the PCA process, the data will transform the original variable to a new uncorrelated variable called the principal component (PC) [25]. Then, the eigenvalue that is greater than one will be used to determine which PC is meaningful [26]. After that, further maximization of variability using the varimax rotation will be conducted to ensure that each variable is only correlated to a PC [27-28] by reducing the variables that are less significant and increasing the variables that are more significant [29]. Then, the participation of the original variables in PCs generated is given by loadings (factor loading) [30]. The loading value is clearly classified into strong, moderate and weak as > 0.75, 0.75-0.50 and 0.50-0.30 respectively [31]. Nevertheless, in certain studies, a value greater than 0.5 could be considered significant [28]. However, the function of the PCA and FA is similar, therefore some studies use both PCA/FA in their writing.

CA is unbiased, the unsupervised pattern recognition technique of multivariate statistical analysis is used to group the data into a group based on similarities of their characteristics and differentiate them from the other groups based on their dissimilarities [32-34] without any a priori knowledge [35]. A common approach of cluster analysis which is widely used in multivariate analysis is hierarchical agglomerative clustering (HCA) that provides an intuitive similarity relationship between the entire set of data and presents the results in a dendogram (tree diagram) [36]. The dendogram is an output, summarizing the cluster results in a visual form which presents a picture of the groups and their proximity with a dramatic reduction in the dimensionality of the original data [32]. The Euclidean distance usually used to measure similarities among clusters and distance can be presented by the difference in analytical values

from the sample [37]. Some HCA techniques are single linkage (nearest neighbour), complete linkage (furthest neighbour), average linkage (between and within groups) and Ward's methods [6, 33].

DA is an unbiased multivariate statistical analysis used to classify objects into mutually exclusive and exhaustive groups on the basis of a set of independent variables. The linear combination of the independent variables found by means of these techniques will discriminate the groups in such a way that the misclassification error rates are minimized [38-39]. The variance-covariance between the classes is maximized while the variance-covariance within the classes is minimized under simultaneous consideration of all types analysed to achieve the objective [40]. The discriminant function is a mathematical function used to maximise discrimination among the groups via combining all the information from independent variables [41]. Additionally, DA will be carried out via standard mode, stepwise forward mode and backward stepwise mode to obtain a classification matrix value that will determine the best significance of DA.

Usually, a combination of some multivariate analysis techniques will be used to get the best interpretation result in a study. The application of multivariate analysis in many fields has assured us that this method may be helpful in substance abuse research, especially in analysing and interpreting complex data. Some potential for applying this method in substance abuse study through reviewing other studies is shown below.

3. DETERMINATION OF THE POSSIBLE FACTORS ASSOCIATED WITH SUBSTANCE ABUSE PROBLEM

Due to the ability of the multivariate analysis in analysing and interpreting data, many fields of study have applied this method on their studies. One view is that this could be applied in substance abuse research and identify the possible factors associated with substance abuse problems. Some studies have proven that multivariate analysis help to identify the factors contributing to the problem as such in environmental studies. The most common type of multivariate analysis used for identifying the sources of the problems are correlation analysis, principal component analysis (PCA) and cluster analysis (CA) [42] without the loss of any

interpretation [43].

In [30] states that many studies use PCA to identify major factors influencing groundwater quality and classify them into a group. Based on the study by [44], the major gradients and principal patterns between wet and dry seasons in their study area were determined using the PCA. Besides that, a study by [18] on groundwater quality has employed the PCA to identify the factors influencing geochemical processes in the aquifer system. In contrast with the study done by [45], they preferred FA to identify the sources of potential pollution. This is similar with to the study by [46] in which they used FA to determine contributing sources of trace metals in the Dil Deresi Stream. Some other studies use this method to check the reliability of their newly developed instruments [47-49].

In [50] stated that in understanding the water quality data, the most superior multivariate techniques that can be used are CA, PCA, FA and DA. However, in [51] used CA, DA and PCA to analyse the dataset of water quality and identify the potential of pollution sources. Meanwhile, in [52] has used FA/PCA and CA to identify the anthropogenic sources which affect the river water in the study area. In [7] applied the PCA and CA coupled with a Pearson correlation to identify the possible factors that influence the water system. In [53] identified the sources of heavy metal in the study area using PCA and CA then grouped the parameter to the groups based on their similarities. In a study done by [54], they suggest that PCA and FA are very useful in identifying important surfaces of water quality monitoring stations and parameters.

Using PCA and HCA combining geochemical modelling techniques, the anthropogenic sources affecting groundwater quality in a Mediterranean coastal aquifer, Mersin-Erdemli basin in Turkey was identified by [55]. For [6], in a hydrological study used PCA followed by CA to determine the surface water origin and effect of wastewater discharge into the stream. In the study, PCA was used to determine the influence of groundwater and wastewater on streams then the CA indicated the association between different streams according to their characteristics. A study by [56] integrated PCA and HCA to classify groundwater samples and identify geochemical processes controlling groundwater geochemistry. Other than that, some studies on air pollution used PCA coupled with HCA for managing air pollution in the study area [57, 27].

All the environmental studies above have pointed out that PCA and FA can help to classify water quality and identify major factors influencing water quality. Supported by another multivariate analysis such as CA and DA, it is suitable for classifying the association between a variable and other factors. Therefore, we believe that multivariate analyses such as PCA, FA, DA and DA could be applied in substance abuse management in order to recognise the sources of the problem including the environmental surroundings, socioeconomic background and sociocultural background.

4. CHARACTERISING SPATIAL AND TEMPORAL VARIATION OF SUBSTANCE ABUSE

On the other hand, multivariate analysis has been widely used to characterize and evaluate spatial and temporal variation [58]. According to [59], the most effective multivariate statistical analysis used to determine the spatial and temporal distribution in the water quality field, besides identifying the latent pollution factors are Pearson correlation, FA, PCA, CA, DA and positive matrix factorization (PMF). In [1, 60] used a factor analysis and PCA to explore the variation of the data. Meanwhile, in [61] used PCA and CA to investigate and assess the water quality, thus discriminating between the relative magnitude of anthropogenic and natural influences on river water quality. Another study on the identification of coastal water quality done by [34] used PCA to determine the spatial and temporal patterns of trophic status. In [62] only used a PCA to analyse the temporal evolution of groundwater composition in an alluvial aquifer. Similarly, [63] analysed the spatial and temporal variation of water quality in catchment areas using PCA for a better understanding of the processes governing water quality. Some studies prefer FA coupling CA to conduct spatial and temporal variation [20, 64-66].

In [67] integrated the CA, DA, PCA and FA to interpret the spatiotemporal variation of water quality in Xingxi River with a dataset spanning 5 years (2002-2006) of continual monitoring. They believed that all these multivariate analyses were very helpful in identifying the important the water quality monitoring site parameters. In [68] integrated some multivariate analysis to characterize the spatial and temporal trends in water quality latent pollution sources in Qiandao Lake where they used CA to divide the sampling sites and month. Then, they performed a

further analysis using DA to identify the most significant variables to assess the variation trends in water quality. Finally, PCA was performed to detect the unidentified, latent pollution sources in Qiandao Lake.

A study by [69] used PCA to get information about the sample distribution and optimal variables for cluster analysis in the study of bedrock groundwater water quality in rural areas in Korea. However, a previous study done by [32] found out that DA was the best method for data reduction and pattern recognition for both spatial and temporal analysis of water quality compared to FA, PCA and CA.

The previous studies above have shown that some multivariate analysis can be used to analyse the spatial and temporal distribution of the study. Some of them use a single method and also integrated a few methods to present the data. Thus, through identifying the spatial and temporal distribution of substance abuse using multivariate analysis, the high risk areas could be determined and further planning to monitor the problems could be implemented.

5. DEVELOPING SUBSTANCE ABUSE INDEX

A previous study by [70] used a factor analysis to develop a quality of life index through the determination of weighing the dimensions with the variance of the respective factors which were generated by a factor analysis. Using [70] as a reference, [71] applied the same method to develop an index of objective QOL in Addis Ababa, Kirkos sub city. Another study by [72] used factor analysis to construct a quality of life in their study area. Based on their study, QOL index was developed using following equation:

$$QOL = \sum_{1}^{n} F_i W_i$$

where n is the number of factors selected, F_i is factor I score and W_i is the percentage of variance factor i explains. By that, each factor score viewed as one aspect of QOL. Then, factors score can be used as a single index indicating the aspect with the factors associated. Referring to the similarity function of the FA and PCA and using the data collected for substance abuse in the selected study area, it is impossible to produce an index of substance abuse whether by using FA or PCA.

6. CONCLUSION

In general, this paper reflects a diverse use of multivariate analysis in various studies and it cannot be denied that most of them are in environmental studies. However, the application of multivariate analysis in other fields is growing due to more powerful methods on evaluating, analysing and interpreting the complexity of data that rival traditional methods. Some benefits to applying this method in substance abuse research have been identified. For example, determining the possible factors associated with substance abuse problem, characterising spatial and temporal variation of substance abuse and developing a substance abuse index.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] Singh K P, Malik A, Sinha S. Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques-A case study. Analytica Chimica Acta, 2005, 538(1):355-374

[2] Magyar N, Hatvani I G, Székely I K, Herzig A, Dinka M, Kovács J. Application of multivariate statistical methods in determining spatial changes in water quality in the Austrian part of Neusiedler See. Ecological Engineering, 2013, 55:82-92

[3] Jiang Y, Wu Y, Groves C, Yuan D, Kambesis P. Natural and anthropogenic factors affecting the groundwater quality in the Nandong karst underground river system in Yunan, China. Journal of Contaminant Hydrology, 2009, 109(1):49-61

[4] Kim K H, Yun S T, Choi B Y, Chae G T, Joo Y, Kim K, Kim H S. Hydrochemical and multivariate statistical interpretations of the spatial controls of nitrate concentrations in a

shallow alluvial aquifer around oxbow lakes (Osong area, central Korea). Journal of Contaminant Hydrology, 2009, 107(3):114-127

[5] Li S, Li J, Zhang Q. Water quality assessment in the rivers along the water conveyance system of the Middle Route of the South to North Water Transfer Project (China) using multivariate statistical techniques and receptor modeling. Journal of Hazardous Materials, 2011, 195:306-317

[6] Mencíó A, Mas P J. Assessment by multivariate analysis of groundwater-surface water interactions in urbanized Mediterranean streams. Journal of Hydrology, 2008, 352(3):355-366

[7] Molla M A, Saha, N., Salam S M A, Rakib-uz-Z M. Surface and groundwater quality assessment based on multivariate statistical techniques in the vicinity of Mohanpur, Bangladesh. International Journal of Environmental Health Engineerin, 2015, 4(1):1-9

[8] Rusdi A R, Noor Zurani M H R, Muhammad M A Z, Mohamad H H. A fifty-year challenge in managing drug addiction in Malaysia. Journal of Health and Translational Medicine, 2008, 11(1):3-6

[9] Al-Haqwi A I. Perception among medical students in Riyadh, Saudi Arabia, regarding alcohol and substance abuse in the community: A cross-sectional survey. Substance Abuse Treatment, Prevention, and Policy, 2010, 5(1):1-6

[10] Fauziah I, Mohamad M S, Chong S T, Manaf A A. Substance abuse and aggressive behavior among adolescents. Asian Social Science, 2012, 8(9):92-97

[11] Hemovich V, Crano W D. Family structure and adolescent drug use: An exploration of single-parent families. Substance Use and Misuse, 2009, 44(14):2099-2113

[12] Choate P W, Doan A. Adolescent alcoholism and drug addiction: The experience of parents. Behavioral Sciences, 2015, 5(5):461-476

[13] Jaladin R A M. Demographic determinants of the drug abuse problem among secondary. Jurnal Antidadah Malaysia, 2007, 1(2):155-172

[14] De Silva P V, Fonseka P. Drug addicts and their behaviour related to drug addiction among the institutionalized addicts of the Galle District. Galle Medical Journal, 2008, 13(1):9-13

[15] Mahmud R, Inoue N, Sen R. Assessment of irrigation water quality by using principal component analysis in an arsenic affected area of Bangladesh. Journal of Soil and Nature, 2007,

1(2):8-17

[16] Ali J, Hassan S, Karim N A A. Kos ekonomi penyalahgunaan dadah. Jurnal Antidadah Malaysia, 2008, 2008:1-14

[17] Jalilian F, Motlagh F Z, Amoei M R, Hatamzadeh N, Gharibnavaz H, Alavijeh M M. Which one support (family, friend or other significant) is much more important to drug cessation? A study among men Kermanshah addicts, the West of Iran. Journal of Addiction Research and Therapy, 2014, 5(174):1-4

[18] Machiwal D, Jha M K. Identifying sources of groundwater contamination in a hard-rock aquifer system using multivariate statistical analyses and GIS-based geostatistical modeling techniques. Journal of Hydrology: Regional Studies, 2015, 4:80-110

[19] 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

[20] Reghunath R, Murthy T R 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

[21] Petrişor A I, Ianoş I, Iurea D, Văidianu M N. Applications of principal component analysis integrated with GIS. Procedia Environmental Sciences, 2012, 14:247-256

[22] Bell N, Schuurman N, Hayes M V, Kawachi I, Berkman L, Dunn J, Hayes M. Using GIS-based methods of multicriteria analysis to construct socio-economic deprivation indices. International Journal of Health Geographics, 2007, 6(1):1-19

[23] Yu S, Shang J, Zhao J, Guo H. Factor analysis and dynamics of water quality of the Songhua River, northeast China. Water, Air, and Soil Pollution, 2003, 144(1):159-169

[24] Gosav S, Praisler M, Birsa M L. Principal component analysis coupled with artificial neural networks-a combined technique classifying small molecular structures using a concatenated spectral database. International Journal of Molecular Sciences, 2011, 12(10):6668–6684

[25] Sârbu C, Pop H F. Principal component analysis versus fuzzy principal component analysis: A case study: The quality of Danube water (1985-1996). Talanta, 2005,

65(5):1215-1220

[26] Kim J O, Mueller C W. Introduction to factor analysis: What it is and how to do it. California: SAGE, 1987

[27] Lu W Z, He H D, Dong L. Performance assessment of air quality monitoring networks using principal component analysis and cluster analysis. Building and Environment, 2011, 46(3):577-583

[28] Yidana S M. Hydrochemical characterization of aquifers using sequential multivariate analyses and geographic information systems in a tropical setting. Journal of Environmental Engineering, 2011, 137(4):258-272

[29] Kebede Y K, Kebedee T. Chapter 5: Application of principal component analysis in surface water-quality monitoring. In P. Sanguansat (Ed.), Principal component analysis-engineering applications. Rijeka: InTech, 2012, pp. 83-100

[30] Halim M A, Majumder R K, Nessa S A, Oda K, Hiroshiro Y, Jinno K. Arsenic in shallow aquifer in the eastern region of Bangladesh: Insights from principal component analysis of groundwater compositions. Environmental Monitoring and Assessment, 2010, 161(1-4):453-472

[31] Liu C, Lin K, Kuo Y. Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan, Science of the Total Environment, 2003, 313(2):77-89

[32] Alberto W D, María del P D, María V A, Fabiana P S, Cecilia, H A, María de los Á B. Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquía River Basin (Córdoba-Argentina). Water Research, 2001, 35(12):2881-2894

[33] Saraçli S, Dogan N, Dogan I. Comparison of hierarchical cluster analysis methods by cophenetic correlation. Journal of Inequalities and Applications, 2013, 1(203):1-8

[34 Wu M L, Wang Y S, Sun C C, Wang H, Dong J D, Yin J P and Han S H. Identification of coastal water quality by statistical analysis methods in Daya Bay, South China Sea. Marine Pollution Bulletin, 2010, 60(6):852-860

[35] Bottegoni G, Rocchia W, Recanatini M, Cavalli A. ACIAP, Autonomous hierarchical

agglomerative Cluster Analysis based protocol to partition conformational datasets. Bioinformatics, 2006, 22(14):58-65

[36] Iscen C F, Emiroglu Ö, Ilhan S, Arslan N, Yilmaz V and Ahiska S. Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat Lake, Turkey. Environmental Monitoring and Assessment, 2008, 144(1-3):269-276

[37] Avdullahi S, Fejza I, Tmava A. Evaluation of groundwater pollution using multivariate statistical analysis: A case study from Burimi area, Kosovo. Journal of Biodiversity and Environmental Sciences, 2013, 3(1):95-102

[38] Lambrakis N, Antonakos A, Panagopoulos G. The use of multicomponent statistical analysis in hydrogeological environmental research. Water Research, 2004, 38(7):1862-1872

[39] Vaselli O, Buccianti A, De Siena C, Bini C, Coradossi N, Angelone M. Geochemical characterization of ophiolitic soils in a temperate climate: A multivariate statistical approach. Geoderma, 1997, 75(1):117-133

[40] Kowalkowski T, Zbytniewski R, Szpejna J, Buszewski B. Application of chemometrics in river water classification. Water Research, 2006, 40(4):744-752

[41] Voss A, Witt K, Kaschowitz T, Poitz W, Ebert A, Roser P, Bar K J. Detecting cannabis use on the human skin surface via an electronic nose system. Sensors, 2014, 14(7):13256-13272

[42] Bodrud-Doza M, Islam A R M T, Ahmed F, Das S, Saha N, Rahman M S. Characterization of groundwater quality using water evaluation indices, multivariate statistics and geostatistics in central Bangladesh. Water Science, 2016, 30(1):19-40

[43] Wang Y, Wang P, Bai Y, Tian Z, Li, Shao X, Mustavich L F, Li B L. Assessment of surface water quality via multivariate statistical techniques: A case study of the Songhua River Harbin region, China. Journal of Hydro-Environment Research, 2013, 7(1):30-40

[44] Chen J, Lu J. Effects of land use, topography and socio-economic factors on river water quality in a mountainous watershed with intensive agricultural production in East China. Plos One, 2014, 9(8):1-12

[45] Huang F, Wang X, Lou L, Zhou Z, Wu J. Spatial variation and source apportionment of water pollution in Qiantang River (China) using statistical techniques. Water Research, 2010,

44(5):1562-1572

[46] Pekey H, Karakaş D, Bakoğlu M. Source apportionment of trace metals in surface waters of a polluted stream using multivariate statistical analyses. Marine Pollution Bulletin, 2004, 49(9):809-818

[47] Zubaran C, Emerson J, Sud R, Zolfaghari E, Foresti K. The application of the drug user quality of life scale (DUQOL) in Australia. Health and Quality of Life Outcomes, 2012, 10(1):1-9

[48] Fonseca M J, Costa P, Lencastre L, Tavares F. A statistical approach to quantitative data validation focused on the assessment of students' perceptions about biotechnology. SpringerPlus, 2013, 2(496):1-13

[49] González-saiz F, García-valderrama T. The opiate treatment index (OTI) clinical interview: New evidence of its reliability and validity. Heroin Addiction and Related Clinical Problems, 2012, 14(2):19-34

[50] Zhao Y, Xia X H, Yang Z F, Wang F. Assessment of water quality in Baiyangdian Lake using multivariate statistical techniques. Procedia Environmental Sciences, 2012, 13:1213-1226

[51] Zhang Y, Guo F, Meng W, Wang X Q. Water quality assessment and source identification of Daliao River Basin using multivariate statistical methods. Environmental Monitoring and Assessment, 2009, 152(1-4):105-121

[52] Mendiguchía C, Moreno C, Galindo-Riaño M D, García-Vargas M. Using chemometric tools to assess anthropogenic effects in river water: A case study: Guadalquivir River (Spain). Analytica Chimica Acta, 2004, 515(1):143-149

[53] Bhuiyan M A H, Suruvi N I, Dampare S B, Islam M A, Quraishi S B, Ganyaglo S, Suzuki S. Investigation of the possible sources of heavy metal contamination in lagoon and canal water in the tannery industrial area in Dhaka, Bangladesh. Environmental Monitoring and Assessment, 2011, 175(1):633-649

[54] Ouyang Y. Evaluation of river water quality monitoring stations by principal component analysis. Water Research, 2005, 39(12):2621-2635

[55] Demirel Z, Guler C. Hydrogeochemical evolution of groundwater in a Mediterranean

coastal aquifer, Mersin-Erdemli basin (Turkey). Environmental Geology, 2006, 49(3):477-487

[56] Cloutier V, Lefebvre R, Therrien R, Savard M M. Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. Journal of Hydrology, 2008, 353(3):294-313

[57] Pires J C M, Sousa S I V, Pereira M C, Alvim-Ferraz M C M, Martins F G. Management of air quality monitoring using principal component and cluster analysis-Part II: CO, NO_2 and O_3. Atmospheric Environment, 2008, 42(6):1261-1274

[58] Garizi A Z, Sheikh V, Sadoddin A. Assessment of seasonal variations of chemical characteristics in surface water using multivariate statistical methods. International Journal of Environmental Science and Technology, 2011, 8(3):581-592

[59] Chen J, Li F, Fan Z, Wang Y. Integrated application of multivariate statistical methods to source apportionment of watercourses in the Liao River Basin, Northeast China. International Journal of Environmental Research and Public Health, 2016, 13(10):1-27

[60] Zhao J, Fu G, Lei K, Li Y. Multivariate analysis of surface water quality in the Three Gorges area of China and implications for water management. Journal of Environmental Sciences, 2011, 23(9):1460-1471

[61] Razmkhah H, Abrishamchi A, Torkian A. Evaluation of spatial and temporal variation in water quality by pattern recognition techniques: A case study on Jajrood River (Tehran, Iran). Journal of Environmental Management, 2010, 91(4):852-860

[62] Helena B, Pardo R, Vega M, Barrado E, Fernandez J M, Fernandez L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water Research, 2000, 34(3):807-816

[63] Selle B, Schwientek M, Lischeid G. Understanding processes governing water quality in catchments using principal component scores. Journal of Hydrology, 2013, 486:31-38

[64] Varol M. Dissolved heavy metal concentrations of the kralkizi, dicle and batman dam reservoirs in the Tigris river basin, turkey. Chemosphere, 2013, 93(6):954-962

[65] Lei D, Yongzhang Z, Jin M, Yong L, Qiuming C, Shuyun X, Hongfu W. Using multivariate statistical and geostatistical methods to identify spatial variability of trace elements in agricultural soils in Dongguan City, Guangdong, China. Journal of China University of

Geosciences, 2008, 19(4):343-353

[66] Shyu G S, Cheng B Y, Chiang C T, Yao P H, Chang T K. Applying factor analysis combined with kriging and information entropy theory for mapping and evaluating the stability of groundwater quality variation in Taiwan. International Journal of Environmental Research and Public Health, 2011, 8(4):1084-1109

[67] Wang X, Cai Q, Ye L, Qu X. Evaluation of spatial and temporal variation in stream water quality by multivariate statistical techniques: A case study of the Xiangxi River basin, China. Quaternary International, 2012, 282:137-144

[68] Gu Q, Zhang Y, Ma L, Li J, Wang K, Zheng K, Zhang X, Sheng L. Assessment of reservoir water quality using multivariate statistical techniques: A case study of Qiandao Lake,China. Sustainability, 2016, 8(3):1-17

[69] Kim K H, Yun S T, Park S S, Joo Y, Kim T S. Model-based clustering of hydrochemical data to demarcate natural versus human impacts on bedrock groundwater quality in rural areas, South Korea. Journal of Hydrology, 2014, 519:626-636

[70] 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

[71] 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

[72] Schyns P, Boelhouwer J. Measuring quality of life in Amsterdam. 2002, https://www.researchgate.net/profile/Peggy_Schyns/publication/255568319_Measuring_Quality_of_Life_in_Amsterdam_from_the_viewpoint_of_participation/links/54aa75ab0cf2bce6aa1d2028.pdf