Spatial pattern analysis for water quality in free-surface constructed wetland

Reza Mohammadpour, Syafiq Shaharuddin, Chun Kiat Chang, Nor Azazi Zakaria and Aminuddin Ab Ghani

ABSTRACT

Free-surface constructed wetlands are known as a low-energy green technique to highly decrease a wide range of pollutants in wastewater and stormwater before discharge into natural water. In this study, two spatial analyses, principal factor analysis and hierarchical cluster analysis (HACA), were employed to interpret the effect of wetland on the water quality variables (WQVs) and to classify the wetland into groups with similar characteristics. Eleven WQVs were collected at the 17 sampling stations twice a month for 13 months. All sampling stations were classified by HACA into three clusters, with high, moderate, and low pollution areas. To improve the water quality, the performance of Cluster-III (micropool) is more significant than Cluster-I and Cluster-II. Implications of this study include potential savings of time and cost for long-term data monitoring purposes in the free-constructed wetland.

Key words | constructed wetland, pollution removal, principal factor analysis, spatial pattern, water quality

Reza Mohammadpour (corresponding author)
Syafiq Shaharuddin
Chun Kiat Chang
Nor Azazi Zakaria
Aminuddin Ab Ghani
River Engineering and Urban Drainage Research Centre (REDAC),
Universiti Sains Malaysia,
Engineering Campus,
Seri Ampangan,
14300 Nibong Tebal,
Penang,
Malaysia
E-mail: *reza564@gmail.com*

INTRODUCTION

Due to urban areas and effect of industrial activities, pollution remains as an important issue of human life and ecology. Lack of plants in municipal regions often results in industrial runoff, with the highly polluted water being discharged into natural water. Through surface runoff, domestic and agricultural waste discharges continue to directly pollute rivers without treatment. There is a wide range of treatment techniques that can be employed to remove pollution and improve the quality of water to meet effluent standards.

Constructed wetlands are designed to highly decrease a wide range of pollutants and wastewater before discharge into rivers and other water resources. The structure of constructed wetlands is similar to natural wetland with low hydraulic retention time and high organic sediments. The performance of free-surface wetlands (FSWs) to enhance water quality (WQ) and reduce runoff pollution was reported in several studies (Brix 1997; Shaharuddin *et al.* 2013). Storm water runoff is recognized as a source of pollution for the FSW, and the characterization of runoff pollution is necessary to WQ management and habitat degradation (Zakaria *et al.* 2003; Greenway 2004).

The FSW as a low-energy green technique can provide ecosystem facilities such as food, control of floods, reduction of water pollution, and recreational and educational benefits (Vymazal 2011). Generally, in these kinds of wetlands, three parts can be recognized: the inlet, macrophyte and open water area (Zakaria *et al.* 2003). The macrophyte area (wetland plants) has a high effect on the wetland ecosystem and WQ, which was previously investigated in different areas (Brix 1997; Kadlec & Wallace 2008).

Surface WQ analysis has become a main concern in water resource and environmental systems (Zhang *et al.* 2013). In environmental and ecological problems, since the number of water quality variables (WQVs) is extensive a robust spatial analysis, such as principal factor analysis (PFA) and hierarchical cluster analysis (HACA), is essential to interpret the relation between physical–chemical parameters (Singh *et al.* 2004; Gazzaz *et al.* 2012). The ability of different parts of a FSW to remove pollution is very

doi: 10.2166/wst.2014.343

important to understand the performance of wetlands and improve the WQ, which is less documented in previous research.

In this study, HACA was used to categorize the FSW into different clusters with similar characteristics. The correlation between WQVs in each cluster was determined using PFA. The relation between WQVs was used to understand the effect of each cluster on the improvement of WQ. Finally, the results indicate the ability of each cluster to remove pollution and improve WQ.

Table 1 | Plant species and the water depth in the USM wetland

Site	Dominant plant species	Water depth (m)
W1	*Hanguana malayana, Lepironia articulata*	0.25–0.3
W2	*Hanguana malayana, Typha angustifolia*	0.27–0.32
W3	*Lepironia articulata, Eleocharis variegata*	0.51–0.62
W4	*Hanguana malayana, Lepironia articulata, Eleocharis variegata*	0.47–0.54
W5	*Lepironia articulata*	0.51–0.64
W6	*Lepironia articulata*	0.31–0.54
Micropool (MA, MB & MC)	Without plant	2.48–2.54

MATERIALS AND METHODS

The FSW in the Universiti Sains Malaysia (USM) was selected as a case study in this research. In total, 11 WQVs were collected at 17 sampling stations twice a month for 13 months (from December 2010 to December 2011). Therefore, a total of 442 datasets were collected in this time period. The WQVs were temperature, pH, dissolved oxygen (DO), conductivity, total suspended solids (TSS), nitrite, nitrate, ammoniacal nitrogen (AN), biochemical oxygen demand (BOD), chemical oxygen demand (COD), and phosphate. The sampling stations were selected to consider a range of plants and water depths (Table 1) and included the inlet, six stations in the macrophyte area (W1 to W6), nine points in the micropool (MA1 to MC3), and the outlet (Figure 1).

All collected data were employed in the multivariable spatial analysis. The variables employed in PFA and HACA were standardized to the z-scale using the following equation (Kowalkowski *et al.* 2006):

$$Z_{ij} = \frac{O_{ij} - \mu}{\sigma} \tag{1}$$

where Z_{ij} is standardized value; O_{ij} is observed data; σ is the standard deviation; and μ is the mean value of observed data.

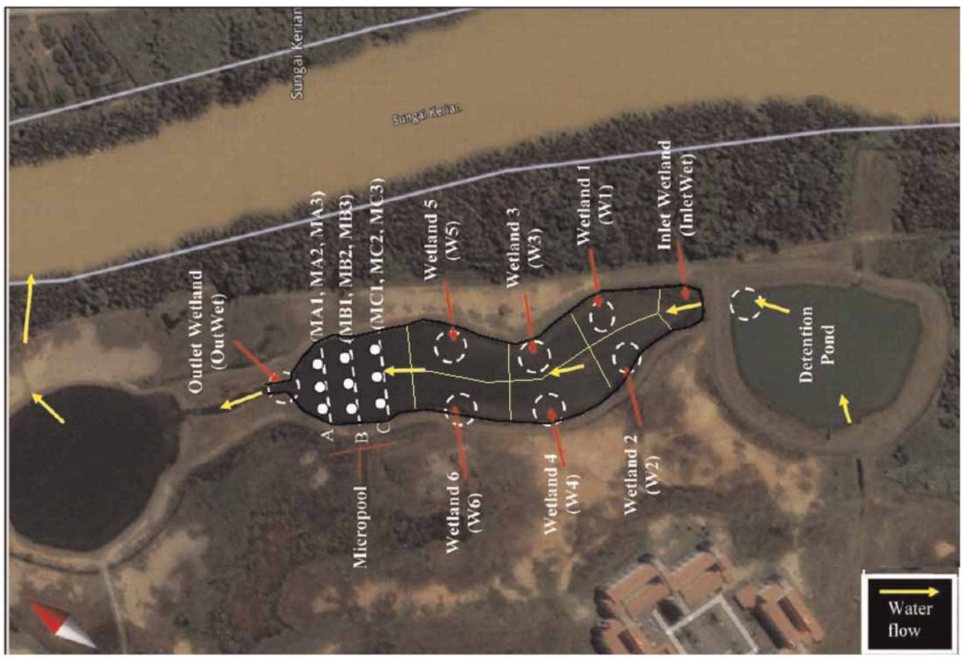

Figure 1 | Sampling point in the USM wetland.

Principal factor analysis

Factor analysis (FA) is a robust technique which can be used to find underlying constructs or factors which describe the correlations among a set of variables. Principal component analysis (PCA) was employed to factor extraction and matrix rotation (Boyacioglu & Boyacioglu 2008). Hereafter the FA is referred to as PFA.

Two criteria, the Kaiser-Meyer-Olkin (KMO) and Barlett's tests of sphericity, were chosen to find the PFA in this study. The KMO is a measure of sampling adequacy, which produces value between 0 and 1 (Table 2). Bartlett's test of sphericity examines whether the correlation matrix is an identity matrix (the null hypothesis). To confirm that the PFA variables are correlated, the null hypothesis should be rejected.

In the next step, to reduce the number of extracted factors, the factor loadings should be evaluated via eigenvalues. The factor loadings were developed to determine the relationships between the extracted factors and the WQVs. As a rule of thumb, a factor loading ≥ 0.6 is recognized as strong while a factor loading less than 0.4 is categorized as weak (Lambrakis *et al.* 2004; Gazzaz *et al.* 2012).

Hierarchical cluster analysis

Cluster analysis (CA) or clustering is an unsupervised data analysis in order to ultimately categorize the dataset into different classes (clusters) with similar characteristics (Templ *et al.* 2008). The agglomerative clustering analysis can be proposed as the most common linear clustering approach among all hierarchical clustering analysis (HACA) strategies (Shrestha & Kazama 2007). The tree diagram (dendrogram) is chosen to show the similarities between clusters as a visual summary (Templ *et al.* 2008). In this study, the standardized data, Ward's method and squared Euclidean distance were employed in the agglomerative clustering analysis (Singh *et al.* 2004).

Table 2 | Interpretation of KMO value (Kaiser 1974)

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

RESULT AND DISCUSSION

Principal factor analysis

The KMO provides a coefficient of 0.87, which illustrates the number of data is meritorious (Table 2). Also, Bartlett's test of sphericity with a chi-square of 2,836 ($\rho = 0.000 < 0.05$ and df $= 55$) reveals that the WQ data met the sphericity assumption and rejected the null hypothesis. These findings reflect that the PFA can be used to explain the WQVs and the collected data are valuable (Mcneil *et al.* 2005).

The PFA extracted three significant factors with eigenvalues roughly bigger than one for each factor, which described about 73.34% of the variance of the data (Table 3). The correlation between extracted factors and the WQVs can be estimated using the factor loadings. Strong factor loadings (bigger than 0.60) are in bold in Table 3.

The variation of the first factor is around 47.94% for the WQ dataset. The WQVs and their factor loadings on this factor are TSS (0.85), nitrate (0.85), phosphate (0.84), nitrite (0.83), AN (0.81), BOD (0.80), and COD (0.74). The high factor loading in the first factor can be interpreted as a high correlation between WQVs. The good relation between TSS and nitrogen components, especially nitrate (0.85) and phosphate (0.84), can be due to a combination of sedimentation

Table 3 | Rotated component matrix[a]

	Component (factor)		
Item	1	2	3
TSS	**0.85**	0.13	0.22
Nitrate	**0.85**	0.09	0.14
Phosphate	**0.84**	0.09	0.19
Nitrite	**0.83**	0.24	– 0.02
AN	**0.81**	0.09	0.24
BOD	**0.80**	0.03	– 0.07
COD	**0.74**	– 0.19	0.26
pH	0.00	**0.87**	0.19
Conductivity	0.22	**0.84**	– 0.07
Temperature	0.32	0.01	**0.83**
DO	0.02	0.54	**0.62**
Eigen value	5.273	1.844	0.95
Initial variance (%)	47.936	16.768	8.640
Cumulative variance (%)	47.936	64.704	73.344
Total variance (%)			73.344

Extraction method: PCA. Rotation method: varimax with Kaiser normalization.
[a]Rotation converged in five iterations.

and nitrification. Generally, the suspended solid consists of the inorganic fraction (silts, clays, etc.) and an organic fraction (algae, zooplankton, bacteria and detritus) which enter the wetland through runoff and settles there. The macrophytes uptakes nutrient through their root system (Lin *et al.* 2002). The concentration of nutrient decreases through the process of nitrification and denitrification as well as through the nutrient uptake through the plants (Brix 1997). The relation between TSS and phosphate in the first factor reflects that the phosphate and organic nitrogen are settled by suspended sediment on the wetland bottom. Loading BOD and COD with nitrogen components in the first factor can be related to the concentration of nitrite and nitrate in the wetland. In the process of nitrification, the ammonium undergoes biological oxidation and converts to nitrite and nitrate. Then, the nitrogenous BOD, which is a portion of BOD, increases with increasing nitrogen compounds. Furthermore, loading BOD in first factor can also be due to the effect of TSS. An excess of TSS could mean a production of higher levels of BOD, which would deplete the DO (T. Lundquist, Lawrence Berkeley National Laboratory, personal communication).

The second factor explains about 16.77% of the variance in the dataset (Table 3). The result indicates a positive factor loading for two variables, pH (0.87) and conductivity (0.84). Loading pH and conductivity in same factor can be expressed as the effect of algae in the wetland. The algae growth is closely related to light intensity, total nitrogen, total phosphorus and water temperature (Scholz 2010). The presence of algae and submerged macrophytes has a significant effect on the pH (Reddy & Delaune 2004). The photosynthesis of algae has a direct effect on pH. On the other hand, the conductivity in terms of high ion concentrations has a strong effect on distributions of individual algae (Sigee 2005; Hamed 2008). Therefore, the macrophyte area with shallow water and high concentration of nutrient (Table 1) provides a good environment to grow different kinds of algae, which influences both pH and conductivity. However, other parameters such as time of day and minerals affect wetland pH.

The third factor receives a high loading from temperature (0.83) and DO (0.62) and describes 8.64% of the variance in the WQ dataset. Loading these variables on the same factor may be due to the photosynthesis process. During the days, the sunshine increases the wetland temperature, and this condition contributes to increased submerged plant photosynthesis. Then, a large amount of DO releases in the macrophyte area, which leads to increasing DO (Luyiga & Kiwanuka 2003).

This finding reflects that just one variable represented by the first, second and third factor can be used as an indicator to estimate WQ in the wetland. Obviously, easily measured parameters such as TSS, conductivity and DO could be used as candidates.

Hierarchical cluster analysis

All 11 WQVs were used in HACA to classify the wetland into different zones with similar WQ characteristics. Figure 2(a) illustrates that the HACA classified all 17 sampling stations into three statistically significant clusters.

The first cluster (Cluster-I), in the upper part of the wetland, is formed by two parts of the wetland, inlet and W1. The Cluster-II is located in the middle part of the wetland and included five parts of the wetland: W2, W3, W4, W5

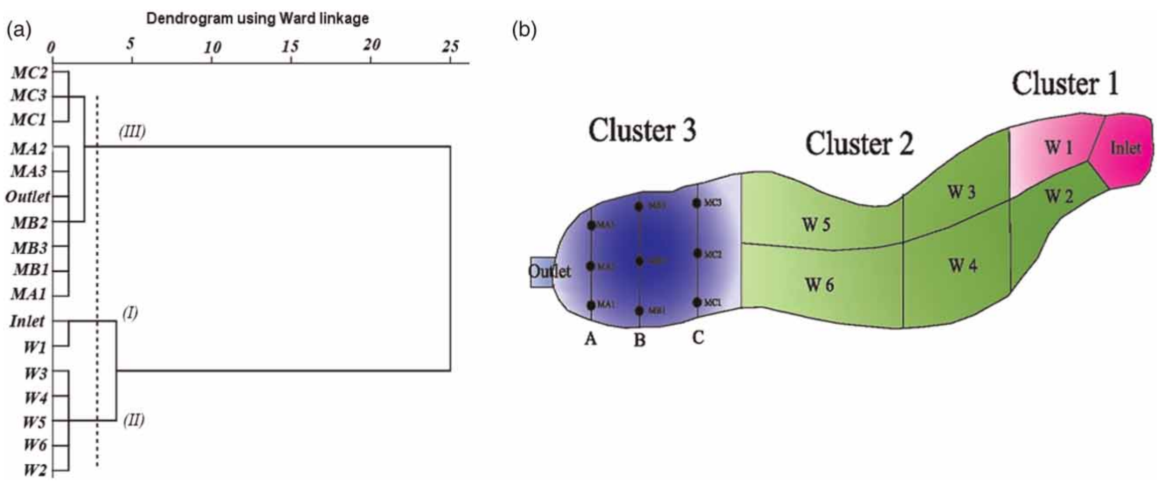

Figure 2 | Clustering of wetland. (a) Dendrogram; (b) location of three clusters.

and W6. The remaining part of the wetland, micropool and outlet, is classified as Cluster-III. The water depth and dominant plants are shown in Table 1. The location of three clusters in the wetland is shown in Figure 2(b). Table 4 shows the average value and reduction of WQVs in the clusters. The output from each cluster was chosen as input for the next cluster, and the inlet and outlet to the wetland were selected as input and output for Cluster-I and Cluster-III respectively. The removal percentage in each cluster is shown in Table 5 and Figure 3. The negative and positive value reflects increase and reduction for the variable respectively.

Increase of temperature in Cluster-I of 0.14 °C and reduction in Cluster-II (0.46 °C) and III (1.63 °C) is due to low water depth in Cluster-I in comparison with other clusters (Table 1). The pH is increased in both Cluster-I (0.6%) and Cluster-II (0.7%) and decreased in Cluster-III (0.8%). Overall, the wetland has no considerable effect on the pH (Figure 3).

The DO is increased in Cluster-I (1.3%) and Cluster-II (2.3%) while it is decreased dramatically in Cluster-III (7.9%). It can be due to photosynthesis by the high concentration of plants and algae in Clusters I and II. The DO decreases with decreasing percentage of plants in Cluster-III. Another reason for reduction of DO is the nitrification process

Table 4 | Average value and reduction of WQVs in each cluster

	Average value of WQVs				Reduction of WQVs		
WQVs	**Inlet**	**Cluster-I**	**Cluster-II**	**Cluster-III**	**Cluster-I**	**Cluster-II**	**Cluster-III**
Temperature (°C)	31.87 ± 1.47	32.01 ± 1.45	31.55 ± 1.32	29.92 ± 1.23	– 0.14 ± 0.40	0.46 ± 0.62	1.63 ± 1.34
pH (μS/cm)	7.63 ± 0.73	7.67 ± 0.72	7.71 ± 0.71	7.63 ± 0.67	– 0.04 ± 0.10	– 0.04 ± 0.27	0.08 ± 0.53
DO (mg/l)	8.25 ± 0.94	8.36 ± 0.97	8.5 ± 0.83	7.79 ± 0.39	– 0.11 ± 0.17	– 0.14 ± 0.72	0.71 ± 0.94
Conductivity	137 ± 27.89	139.04 ± 27.56	140.29 ± 25.53	132.58 ± 23.7	– 2.04 ± 1.12	– 1.25 ± 12.75	7.71 ± 8.32
Nitrite (mg/l)	0.035 ± 0.01	0.033 ± 0.01	0.016 ± 0.00	0.007 ± 0.00	0.001 ± 0.00	0.02 ± 0.01	0.01 ± 0.00
Nitrate (mg/l)	2.94 ± 0.65	3.18 ± 0.58	2.48 ± 0.66	1.13 ± 0.48	– 0.24 ± 0.28	0.7 ± 0.42	1.35 ± 0.40
Phosphate (mg/l)	0.41 ± 0.07	0.4 ± 0.06	0.28 ± 0.04	0.16 ± 0.03	0.00 ± 0.04	0.13 ± 0.05	0.11 ± 0.03
AN (mg/l)	0.3 ± 0.06	0.31 ± 0.06	0.26 ± 0.05	0.15 ± 0.02	– 0.01 ± 0.02	0.05 ± 0.03	0.12 ± 0.04
BOD (mg/l)	2.73 ± 0.49	3.01 ± 0.32	2.73 ± 0.28	2.00 ± 0.30	– 0.28 ± 0.24	0.28 ± 0.40	0.73 ± 0.32
COD (mg/l)	28.23 ± 4.3	29.08 ± 4.12	24.19 ± 2.61	15.42 ± 3.36	– 0.85 ± 2.00	4.88 ± 3.22	8.77 ± 2.84
TSS (mg/l)	27.65 ± 5.15	28.13 ± 4.47	21.94 ± 3.04	8.77 ± 3.93	– 0.48 ± 1.82	6.19 ± 3.53	13.17 ± 2.77

Table 5 | Removal percentage in each cluster

	Cluster-I (%)		Cluster-II (%)		Cluster-III (%)		Total removal (%)	
WQVs	**Average**	**SD***	**Average**	**SD**	**Average**	**SD**	**Average**	**SD**
Temperature	– 0.5	1.2	1.4	1.9	5.1	4.0	6.0	4.1
pH	– 0.6	1.5	– 0.7	3.8	0.8	7.3	– 0.5	8.7
DO	– 1.3	2.0	– 2.3	8.7	7.9	11.9	4.3	12.1
Conductivity	– 1.6	1.0	– 1.6	10.7	5.4	4.6	2.2	12.3
Nitrite	3.2	7.1	48.1	15.0	29.3	15.3	80.6	8.6
Nitrate	– 10.3	16.3	25.1	18.0	47.8	16.4	62.6	11.3
Phosphate	– 0.4	7.5	31.8	10.1	28.4	7.9	59.8	8.0
AN	– 4.3	9.5	17.5	14.4	37.3	10.9	50.5	8.4
BOD	– 11.7	10.9	8.8	12.8	28.0	13.7	25.1	14.1
COD	– 3.4	7.3	16.7	9.5	31.8	10.8	45.1	10.2
TSS	– 2.4	6.7	21.7	10.6	49.1	13.7	68.4	12.7

*SD = standard deviation.

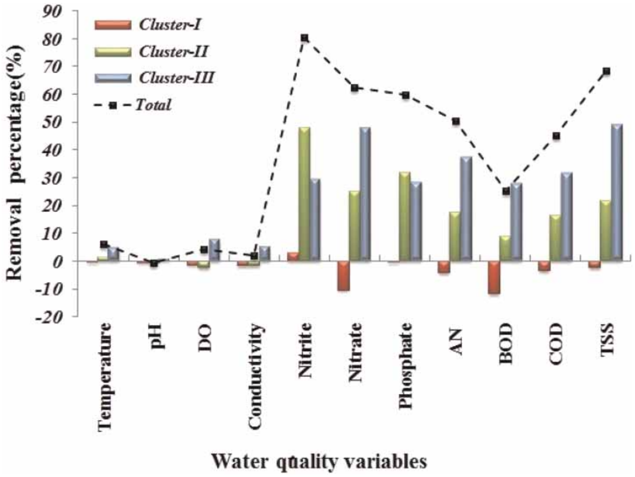

Figure 3 | Removal percentage of WQVs in the wetland.

in Cluster-III. It is a very high oxygen demanding process. The variation of conductivity in three clusters was observed to be similar to DO and pH, which can be due to algae effect.

The wetland reduces a high percentage of nitrite (80.6%) with 3.2, 48.1 and 29.3% for Cluster-I, Cluster-II and Cluster-III respectively. The Cluster-II with high macrophyte area removes more percentage of nitrite in the wetland.

Cluster-III has the highest effect on reduction of nitrate and AN in the wetland. A removal of 47.8 and 37.3% for nitrate and AN respectively can be due to a combination of nitrification and denitrification processes. In Cluster-II with high concentration of plants, removal of nitrate and AN with a percentage of 25.1 and 17.5 respectively is due to plant uptake. In total, nitrogen removal due to denitrification is higher than that by plant and algae assimilation (Metcalf *et al.* 1991). Due to decomposition processes, in Cluster-I the nitrate and AN are increased by a percentage of 10.3 and 4.3 respectively. Overall, the wetland removed around 63% of nitrate, which is a considerable value in comparison with phosphate (59.8%), AN (50.5%), BOD (25.1%), COD (45.1%) and TSS (68.4%).

Decomposition of organic matter can be the reason for increase of phosphate by 0.4% in Cluster-I. Although the plant uptakes phosphorus, ammonia and nitrate for growing, the decomposition of organic matter releases phosphorus and nutrient back to the water (Kadlec & Wallace 2008). The percentage of phosphate removal in Cluster-I and Cluster-II was 31.8 and 28.4 respectively.

The variation of BOD is similar to COD in all clusters, with a total percentage removal of 25.1 and 45.1 for BOD and COD respectively. Cluster-III shows a high percentage reduction of these parameters. In Cluster-I, increase of BOD and COD by 11.7 and 3.4% may be interpreted as a concentration of organic matter and macro-invertebrates (Shaharuddin *et al.* 2013).

TSS was observed to increase in Custer I and decrease in both clusters II and III. Sedimentation is a main cause of decreased suspended solid in the wetlands. Cluster-III with high water depth (Table 1) and low velocity could remove a high percentage of sediment (49.1%) compared to Cluster-II (21.7%). Increase of TSS in Cluster-I is due to decomposition processes and concentrations of zooplankton, phytoplankton and macroinvertebrate which directly increase the organic suspended material.

As shown in Figure 3, Cluster-I removed a low percentage of nitrite (3.2%) and increased DO (1.3%) but other pollutants are increased in this cluster. Although Cluster-II has a significant effect to improve the quality of water and remove high percentages of nitrite, the performance of Cluster-III is better than Cluster-II, with a high reduction in nitrate, AN, BOD, COD and TSS. Therefore, the HACA classified the wetland into three categories with different physicochemical characteristics and pollution levels. In total, the wetland removed a high percentage of nitrite (80.6%), TSS (68.4%), nitrate (62.6%) and phosphate (59.8%).

The results of HACA indicate that for quick estimation of the WQ in the wetland, the WQVs can be collected regularly from just one station in each cluster since the whole cluster can be represented just by one station in the cluster. Consequently, collecting data from only three parts of the wetland rather than 17 will accurately reflect the spatial dimension of the WQ in the entire wetland and reduce the time, costs and effort to collect samples, without losing significant information.

CONCLUSIONS

Two multivariable spatial analyses, PFA and HACA, were employed to recognize the latent structure of WQ in the FSW and to classify 17 sampling stations into groups with similar characteristics. The PFA provided three latent factors, including 11 WQVs to describe about 73.34% of total variation of the WQ dataset. All sampling stations were classified by HACA into three clusters, Cluster-I, Cluster-II and Cluster-III, with different physicochemical characteristics and pollution levels. In total, the wetland was able to remove a high percentage of nitrite (80.6%), TSS (68.4%), nitrate (62.6%) and phosphate (59.8%). The results indicate that Cluster-III with high removal percentage is more effective to improve WQ in comparison with other clusters. Since the sampling stations have similar characteristics, WQVs can be collected just from one station in each cluster, without losing significant information, in the wetland. This finding contributes to save time, costs and effort for collection of data.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial assistance from the Ministry of Education under Long Term Research Grant (LRGS) No. 203/PKT/672004 entitled 'Urban Water Cycle Processes, Management and Societal Interactions: Crossing from Crisis to Sustainability'. This is a subproject entitled 'Sustainable Wetland Design Protocol for WQ Improvement' (Grant number: 203/PKT/6724002).

REFERENCES

Boyacioglu, H. & Boyacioglu, H. 2008 Water pollution sources assessment by multivariate statistical methods in the Tahtali Basin, Turkey. *Environmental Geology* **54**, 275–282.

Brix, H. 1997 Do macrophytes play a role in constructed treatment wetlands? *Water Science and Technology* **35**, 11–17.

Gazzaz, N. M., Yusoff, M. K., Ramli, M. F., Aris, A. Z. & Juahir, H. 2012 Characterization of spatial patterns in river water quality using chemometric pattern recognition techniques. *Marine Pollution Bulletin* **64**, 688–698.

Greenway, M. 2004 Constructed wetlands for water pollution control-processes, parameters and performance. *Developments in Chemical Engineering and Mineral Processing* **12**, 491–504.

Hamed, A. 2008 Biodiversity and distribution of blue-green algae/cyanobacteria and diatoms in some of the Egyptian water habitats in relation to conductivity. *Australian Journal of Basic Applied Science* **2**, 1–21.

Kadlec, R. H. & Wallace, S. D. 2008 *Treatment Wetlands*. 2nd edn, CRC Press, Boca Raton, FL, USA.

Kaiser, H. F. 1974 An index of factorial simplicity. *Psychometrika* **39**, 31–36.

Kowalkowski, T., Zbytniewski, R., Szpejna, J. & Buszewski, B. 2006 Application of chemometrics in river water classification. *Water Research* **40**, 744–752.

Lambrakis, N., Antonakos, A. & Panagopoulos, G. 2004 The use of multicomponent statistical analysis in hydrogeological environmental research. *Water Research* **38**, 1862–1872.

Lin, Y.-F., Jing, S.-R., Wang, T.-W. & Lee, D.-Y. 2002 Effects of macrophytes and external carbon sources on nitrate removal from groundwater in constructed wetlands. *Environmental Pollution* **119**, 413–420.

Luyiga, S. & Kiwanuka, S. 2003 Plankton composition, distribution and significance in a tropical integrated pilot constructed treatment wetland in Uganda. *Water Science and Technology* **48**, 241–248.

Mcneil, V. H., Cox, M. E. & Preda, M. 2005 Assessment of chemical water types and their spatial variation using multi-stage cluster analysis, Queensland, Australia. *Journal of Hydrology* **310**, 181–200.

Metcalf and Eddy Inc., revised by Tchobanoglous, G. & Burton, F. L. 1991 *Wastewater Engineering – Treatment, Disposal, and Reuse*. 3rd edn, McGraw-Hill, New York.

Reddy, K. R. & Delaune, R. D. 2004 *Biogeochemistry of Wetlands: Science and Applications*. CRC Press, Boca Raton, FL, USA.

Scholz, M. 2010 *Wetland Systems: Storm Water Management Control*. Springer Verlag, Berlin.

Shaharuddin, S., Zakaria, N. A., Ab. Ghani, A. & Chang, C. K. 2013 Performance evaluation of constructed wetland in Malaysia for water security enhancement. *Proceedings of 2013 IAHR World Congress*, China.

Shrestha, S. & Kazama, F. 2007 Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. *Environmental Modelling & Software* **22**, 464–475.

Sigee, D. 2005 *Freshwater Microbiology: Biodiversity and Dynamic Interactions of Microorganisms in the Aquatic Environment*. Wiley, Chichester, UK.

Singh, K. P., Malik, A., Mohan, D. & Sinha, S. 2004 Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)–a case study. *Water Research* **38**, 3980–3992.

Templ, M., Filzmoser, P. & Reimann, C. 2008 Cluster analysis applied to regional geochemical data: Problems and possibilities. *Applied Geochemistry* **23**, 2198–2213.

Vymazal, J. 2011 Enhancing ecosystem services on the landscape with created, constructed and restored wetlands. *Ecological Engineering* **37**, 1–5.

Zakaria, N. A., Ab Ghani, A., Abdullah, R., Mohd. Sidek, L. & Ainan, A. 2003 Bio-ecological drainage system (BIOECODS) for water quantity and quality control. *International Journal of River Basin Management* **1**, 237–251.

Zhang, H., Sun, L., Sun, T., Li, H. & Luo, Q. 2013 Spatial distribution and seasonal variation of polycyclic aromatic hydrocarbons (PAHs) contaminations in surface water from the Hun River, Northeast China. *Environmental Monitoring and Assessment* **185**, 1451–1462.

First received 10 March 2014; accepted in revised form 22 July 2014. Available online 5 August 2014