

Spatial Assessment of Water Quality Patterns using Environmetric Techniques: A Case Study in Muda River Basin (Malaysia)

¹Shah Christirani Azhar*

¹Pusat Asasi Sains Pertanian, Universiti Putra Malaysia,

43400 Serdang, Selangor, Malaysia

*Corresponding author: christirani@upm.edu.my

Abstract

River pollution impact human health, environment and the sustainable development. This study was conducted to identify spatial patterns and the main parameters affecting the water pollution within nine monitoring stations in the Muda River basin (Malaysia) over a 16-year database (1998–2013). Environmetric techniques were applied to the dataset. These combined Cluster Analysis, Discriminant Analysis, and Multiple Linear Regression. The Cluster Analysis showed that the monitoring stations divided into two separate groups based on similarities features of water quality while Discriminant Analysis validated these groups. Furthermore, the Multiple Linear Regression analysis showed that the significant parameters contributing to variability the Water Quality Index was biological oxygen demand, chemical oxygen demand and ammonia nitrogen. This was due to the point-source pollution, particularly from rubber factory. Therefore, the results provided information to support future water pollution control strategies.

Keywords: spatial assessment; surface water quality; environmetric techniques

INTRODUCTION

Water quality in many rivers worldwide has degraded due to natural and anthropogenic activities. This can be attributed to water pollution and improper management of water resources [1]. Water resources of poor water quality adversely impact sustainable development. Located in the north western part of Peninsular Malaysia, the Muda River basin is one of the most valuable water resources for agriculture [2]. Besides that, the Muda River supplies water for domestic, industrial, and irrigation purposes in the northern state of Peninsular Malaysia [3]. Agriculture related operations like animal and crop farming cover approximately 55% of the total land area of this basin. Meanwhile, about 35% of the catchment area is still covered by forests. Therefore, agro-based industries such as rubber processing and palm oil factories, as well as logging, are the primary economic activities in the area. The Muda River has been radically affected by unsustainable human activities like agricultural and logging activities [3].

Agriculture has both indirect and direct effects on the quality of surface water and is one of the primary activities that cause deterioration of water quality. Agricultural development has led to the increase of nutrients in the river system associated with consumption of inorganic fertilizers [4]. Although agriculture represents the primary source of diffused water pollution, forests are also a source of major concern. The main water quality problems in forest areas are increased turbidity and sedimentation [5] due to logging activities. Besides that, industrial waste from plants is a typical source of pollution. In most developing countries, urban and industrial waste is disposed of without adequate treatment [6]. Rubber effluent negatively impacts the water quality through biological oxygen demand, ammonia, chemical oxygen demand, conductivity, and phosphate [7, 8]. Several studies have reported Malaysia's rivers are polluted by untreated rubber effluent [9, 10, 7].

In Malaysia, the Water Quality Index (WQI) is used to assess the water quality status of rivers. WQI can give effective feedback on the status of water quality to policy makers and environmental managers'. The method of calculating the WQI is explained in the Department of Environment report [11]. Briefly, it is calculated by the concentration of dissolved oxygen (DO), chemical oxygen demand (COD), biochemical oxygen demand (BOD), pH, suspended solids (SS), and ammonia nitrogen (NH₃-N). Furthermore, the WQI classifies water quality based on suitability for some target uses including domestic consumption, aquatic propagation and fisheries, livestock drinking, recreation and agriculture activity [11]. The method used for computing the WQI in Malaysia is similar to those employed in other countries [12].

In the environmental study, the environmetric techniques approach has often been used to deal with complex and vast amounts of data. The datasets from the environmental monitoring program are classified, characterized, modelled and interpreted using environmetric techniques. The technique is also able to identify parameters that affect the water system. In this study, the environmetric techniques including Cluster Analysis (CA), Discriminant Analysis (DA) and Multiple Linear Regressions (MLR) are applied. The CA classifies the large data into classes which have identical features within the group, but are contrary between the groups [13]. In previous studies, MLR is applied to determine the main parameters influencing the deterioration of water quality [13, 14, 15]. The aim of this study is to determine the spatial pattern of water quality within the basin and identify the significant parameters which contribute to the WQI variability in the Muda River basin.

MATERIALS AND METHODS

Study areas

The Muda River basin is located in northwest Peninsular Malaysia (Fig. 1) and covers an area of about 4210 km². The average annual temperature is 27°C while the mean annual precipitation in the whole basin is 2300 mm. The study area experiences two rainy seasons in a year: one from April to May and another from September to November [16]. The major land uses in the basin are agriculture and forestry (Fig. 1). The water quality in the Muda River basin is classified as Class II [11]. The Muda River, which has a length of 180 km, is the main river in the basin. The Muda River belongs to the state of Kedah and Penang [16] and is the primary source of water in both states.

Water quality data

The dataset covered the period 1998–2013 and contained the values of selected monitored water quality parameters (WQP) for nine monitoring stations of the main river and the tributaries. The secondary data of water quality were obtained from the Department of Environment (DOE), Malaysia. The locations of the river monitoring stations are presented in Fig. 1. The datasets taken for this study comprise of six WQP: DO, COD, BOD, SS, pH and NH₃-N. Descriptive statistics of the 16-year dataset are summarized in Table 1. The lowest mean WQI value (58.77) was reported for MD02 while the other monitoring stations (MD01, and MD03-MD09) had an average WQI value within the range of 77.41-90.68. The highest concentrations of BOD, COD, and NH₃-N and the lowest levels of DO were reported at station MD02.

Cluster analysis (CA)

CA is the analysis used to classify objects that have the same features, while the objects that have distinct features are placed in another group [13]. CA begins by assuming all objects form into different groups and the pair-wise distances between groups are calculated [15,13]. After that, the new groups are formed by linking the pair having the smallest distance [13]. The similarity levels are then measured by Ward's method and a Euclidean distance during the classification process [14]. A dendrogram illustrates the classification of the monitoring stations (objects).

Discriminant analysis (DA)

Discriminant analysis (DA) is a classification method that determines the variables that discriminate between two or more naturally occurring classes/groups. It is a multivariate statistical modelling technique that can be used to classify objects into mutually exclusive groups [12]. The DA was executed in the stepwise mode through the raw water quality. In the stepwise mode, variables are included step by step, starting from the most relevant parameters until no substantial change is obtained.

Multiple linear regressions (MLR)

Multiple regressions can be applied to forecast the variability between the two variables. General multiple linear regression models have k independent variables, and there are n observations. Thus, the regression model can be written, as given in Eq.1 [17].

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ki} + \varepsilon_i \quad (1)$$

With $i = 1, \dots, n$ and β_1 are the regression coefficients, x_1 as independent variables, and ε is the error associated with the regression. The best regression model can be determined by the coefficient of determination (R^2) and Root Mean Square Error (RMSE). R^2 measured the goodness of the fit of MLR [18]. The best linear model is the model with the largest R^2 and smallest RMSE.

RESULTS AND DISCUSSION

Spatial variations of the water quality

Cluster analysis was performed using six water quality parameters (DO, BOD, COD, pH, NH₃-N, and SS) to study the spatial variations in water quality. The CA could distinguish the monitoring stations according to water quality characteristics. The dendrogram (Fig. 2) shows that the monitoring stations in the Muda River basin can be classified into two main groups. Group 1 comprise of stations MD01 and MD03–MD09. While Group II, comprises of MD02. DA confirmed the two spatial groups produced by CA (Wilk's lambda = 0.31, df = 2, $p < 0.05$). From previous section, the descriptive statistics in Table 1 shows the lowest mean WQI value (58.77) was reported for station MD02. These preliminary finding indicate that station MD02 is the most contaminated station in Muda River basin from the year 1998-2013.

Contribution of water quality parameters to the deterioration of water quality

The Multiple Linear Regression (MLR) was conducted to determine the main parameters which contribute to the WQI variability in the Muda River basin. The dependent variable was the WQI (representing water quality status), while the independent variables were the six water quality parameters employed in sub-index calculation of WQI (viz. DO, BOD, COD, pH, SS, and NH₃-N). The coefficient of determination (R^2) and root mean square error (RMSE) are shown in the proposed equation, as given in Eq. 2 and Eq. 3.

a) Group I:

$$\text{WQI} = 68.57 + 3.10(\text{DO}) - 0.66(\text{BOD}) - 0.17(\text{COD}) - 0.06(\text{SS}) - 0.70(\text{pH}) - 3.13(\text{NH}_3\text{-N}) \quad (2)$$

($R^2 = 0.80$ RMSE = 2.31)

b) Group II:

$$\text{WQI} = 71.24 + 3.56(\text{DO}) - 0.27(\text{BOD}) - 0.05(\text{COD}) - 0.06(\text{SS}) - 2.63(\text{pH}) - 0.08(\text{NH}_3\text{-N}) \quad (3)$$

($R^2 = 0.91$, RMSE = 5.88)

The regression model for Group II had the highest coefficient of determination ($R^2 = 0.90$, $p < 0.05$), implying that 90% of the variations in the WQI for Group II are explained by the six water quality parameters employed. Apart from DO, the WQI is negatively influenced by the water quality parameters

(viz. BOD, COD, SS, pH, and NH₃-N). In the case of Group I, the model comprised of the six water quality parameters ($R^2 = 0.89$, $p < 0.05$), which means that these six parameters explain about 89% of the variations in the WQI for Group I. The concentrations of BOD, COD, SS, and NH₃-N have negative effects on the WQI value, in contrast to DO and pH.

Furthermore, the standardised coefficient values for both groups are illustrated in the bar charts (Fig.3). The result show that, in both groups, the increasing concentration of BOD, COD, SS and NH₃-N lead to adverse impacts on water quality status. The scatterplot graph between WQP and WQI with R^2 values are illustrated in Fig. 4. The R^2 (0.35-0.67, $p < 0.05$) shows that BOD, COD and NH₃N are the highest contributors to the deterioration of water quality in Group II. However, the impact of BOD, COD, SS and NH₃N on WQI in Group I is small ($R^2 = 0.07-0.15$, $p < 0.05$), similar to the effect of concentration of SS in Group II ($R^2 = 0.10$, $p < 0.05$). The findings demonstrate that COD, BOD and NH₃-N are the significant parameters which contribute to the WQI variability in the Muda River basin, particularly in Group II (MD02).

The presence of a rubber factory adjacent to station MD02 has a significant adverse impact on the water quality of the river. The production of rubber products from natural rubber consumes large amount of water and generates large quantities of effluent [7]. A large volume of untreated wastewater was discharged into the river system and deteriorated its water quality. The rubber effluent increases the concentrations of BOD, NH₃-N, and COD [7, 8]. The effluent discharged from latex rubber factories consists mainly of carbonaceous organic materials, nitrogen, and sulphate due to the use of acids in latex coagulation. In the rubber effluent, the bulk of nitrogen is in the form of NH₃-N as a result of the use of ammonia in preservation of the latex. Restoration efforts should focus on station MD02 because contaminants from this area are primarily responsible for the poor water quality at Muda River basin. Highest priority should be given to the reduction of point-source pollution from the rubber industry and surface water restoration within the basin.

CONCLUSION

The result of Cluster Analysis (CA) divided the nine water quality monitoring stations into two different groups based on similarity of water quality characteristics. The groups were confirmed by DA. Furthermore, Multiple Linear Regression (MLR) was conducted to determine the main parameters which contribute to the WQI variability in the Muda River basin. The results showed that COD, BOD and NH₃-N are the main parameters influencing the deterioration of water quality, particularly in Group II (MD02). These pollutants are effluent of a nearby rubber factory. It is proposed that all stakeholders should cooperate to produce an action plan that can reduce latex plant effluent discharge directly into the river. If carried out effectively, this plan would enhance the water quality for the long term.

ACKNOWLEDGEMENTS

The authors would like to thank the Department of the Environment (DOE) for their permission to use their water quality data for this study.

REFERENCES

- [1] Prasanna, M.V., Praveena, S.M., Chidambaram, S., Nagarajan, R., Elayaraja, A. (2012). *Environmental Earth Sciences*, 67, 1987-2001.
- [2] Lee, K.F. A background study: Economic benefits of the Muda Water Catchment. Presented at the National Conference on Economic Evaluation of Forest Resources, Kuala Lumpur, Malaysia, November 11-13, 2009.

- [3] Ab. Ghani, A., Ali, R.A., Zakaria, N.A., Abu Hasan, Z., Chang, K.C., Ahmed, M.S.S. (2010). *International Journal River Basin Management*, 8, 25-37.
- [4] Wu, Y., Chen, J. (2013). *Ecological Indicators*, 32, 294-304.
- [5] Neary, D.G., Smethurst, P.J., Baillie, B.R., Petrone, K.C., Cotching, W.E., Baillie, C.C. (2010). *Journal of Soils and Sediments*, 10(4), 652-670.
- [6] Ma, J., Ding, Z., Wei, G., Zhao, H., Huang, T. (2009). *Journal of Environmental Management*, 90(2), 1168-1177.
- [7] Runguang, N., Babel, S. (Treatment of natural rubber processing wastewater by combination of ozonation and activated sludge process, Presented at the International Conference on Environmental Research and Technology (ICERT 2008), Penang, Malaysia, May 28-30, 2008.
- [8] Arimoro, F.O. (2009). *Chemosphere*, 77(3), 440-449.
- [9] Hutagalung, R.L. (2003). Use of carbohydrate residues in Malaysia. Department of Animal Sciences, Universiti Pertanian Malaysia, Selangor.
- [10] Tekasakul, P., Tekasakul, S. (2006). *Journal of Aerosol Research*, 21(2), 122-129.
- [11] Department of Environment (DOE) Malaysia Environmental Quality Report 2013, Department of Environment, Ministry of Natural Resources and Environment, Petaling Jaya, Malaysia, 2014.
- [12] Gazzaz, N.M., Yusoff, M.K., Ramli, M.F., Aris, A.Z., Juahir, H. (2012). *Marine Pollution Bulletin*, 64(4), 688-698.
- [13] Pires, J.C.M., Sousa, S.I.V., Pereira, M.C., Alvim-Ferraz, M.C.M., Martins, F.G. (2008). *Atmospheric Environment*, 42, 1249-1260,
- [14] Lau, J., Hung, W.T., Cheung, C.S. (2009). *Atmospheric Environment*, 43, 769-777.
- [15] Dominick, D., Juahir, H., Latif, M.T.C., Zain, S.M., Aris, A.Z. (2012). *Atmospheric Environments*, 60, 172-181.
- [16] Julien, P.Y., Ab. Ghani, A., Zakaria, N.A., Abdullah, R., Chang, C.K. (2010). *Journal of Hydraulic Engineering*, 136, 251-261.
- [17] Kovac-Andric, E., Brana, J., Gvozdic, V. (2009). *Ecological Informatics*, 4(2), 117-122.
- [18] Ilten, N., Selici, T. (2008). *Environmental Monitoring and Assessment*, 140, 259-266.