Gazzaz, N.M., Kamil, M., Juahir, H., Firuz, M., and Zaharin, A. 2015. Artificial neural network modeling of the water quality index using land use areas as predictors. Water Environment Research 82(2): 99-112.

Artificial Neural Network Modeling of the Water Quality Index Using Land Use Areas as Predictors

Nabeel M. Gazzaz[1,*], Mohd Kamil Yusoff[1], Mohammad Firuz Ramli[1], Hafizan Juahir[2], Ahmad Zaharin Aris[2]

ABSTRACT: This paper describes the design of an artificial neural network (ANN) model to predict the water quality index (WQI) using land use areas as predictors. Ten-year records of land use statistics and water quality data for Kinta River (Malaysia) were employed in the modeling process. The most accurate WQI predictions were obtained with the network architecture 7-23-1; the back propagation training algorithm; and a learning rate of 0.02. The WQI forecasts of this model had significant ($p < 0.01$), positive, very high correlation ($\rho_S = 0.882$) with the measured WQI values. Sensitivity analysis revealed that the relative importance of the land use classes to WQI predictions followed the order: mining > rubber > forest > logging > urban areas > agriculture > oil palm. These findings show that the ANNs are highly reliable means of relating water quality to land use, thus integrating land use development with river water quality management. *Water Environ. Res.*, **87,** 99 (2015).

KEYWORDS: artificial neural network, function approximation, three-layer perceptron, land use areas, water quality index, weighted arithmetic mean, unweighted harmonic square mean.

doi:10.2175/106143014X14062131179276

Introduction

The surface water quality in a region depends largely on the anthropogenic activities and on the types, areas, and distributions of land uses within the river catchments in the region. The river systems are particularly vulnerable to pollution as they are highly accessible to waste discharges, directly through surface runoff and indirectly through drains and tributaries (Singh et al., 2009; Tourbier, 1994). Nearly all land use activities directly and/or indirectly affect environmental parameters like topography, vegetation cover, soil stability, surface runoff, and river flow, which in turn modify the transport of water, organic matter, sediments, and numerous pollutants to surface water bodies (Johnson et al., 1997; Schindler, 1997). As such, health of a river system is essentially a function of the types of processes and interactions taking place on the landscape within watershed boundaries (Seitz et al., 2011). Hunsaker and Levine (1995) argued that land use change may be the single most influential factor impacting ecological resources since the different land uses differ in their individual contributions to point and non-point sources of pollution in terms of quantity and quality. They further illustrated that hydrologic changes to rivers and streams brought about by changes in land uses, habitat alteration, and non-point sources of pollution are potentially the most potent and widespread threats to the well-being of the riverine ecosystems.

Assessment of water quality entails evaluation of biological, chemical, and physical characteristics of water in reference to natural quality, human health effects, and intended uses (Fernández et al., 2004; Pesce and Wunderlin, 2000). Nevertheless, the water quality can be evaluated by a single parameter (e.g., dissolved oxygen (DO) or biochemical oxygen demand (BOD)) for specific interest or by a number of parameters selected carefully to reveal the pollution status of the river and reflect the quality of its water. However, since no individual parameter can express the water quality sufficiently, it is normally assessed by measuring a broad range of variables (e.g., electric conductivity (EC), pH, turbidity, DO, and the concentrations of a variety of pollutants, including pathogens, nutrients, organics, and metals). Consequently, the monitoring programs produce large amounts of data and these data require integration if the results are to be presented in a meaningful way to local decision makers, watershed managers, and the public. For this reason, water quality indices have been developed. The water quality index (WQI) is a single numeric score that describes the water quality condition at a particular location in a specific time (Kaurish and Younos, 2007; Pesce and Wunderlin, 2000).

Nowadays, many countries in the world use water quality indexing systems to evaluate the water quality status of their rivers (e.g., Argentina (Pesce and Wunderlin, 2000), China (Song and Kim, 2009), India (Sargaonkar and Deshpande, 2003), Spain (Sánchez et al., 2007), Turkey (Boyacioglu, 2007), and the USA (Cude, 2001)). However, many of the WQI computing methods entail somewhat meticulous transformations of the raw water quality parameters into sub-indices (e.g., Abrahão et al. (2007), Boyacioglu (2007), Cude (2001), Hanh et al. (2011), Liou et al. (2004), Pesce and Wunderlin (2000), Sánchez et al. (2007), Sargaonkar and Deshpande (2003), and Song and Kim (2009)). Many of the transformations employ different mathematical equations for different values, or ranges of values, of the majority of the parameters constituting the WQI. In addition, the classic approach to river water quality assessment uses manual or simple computerized calculations to compute the WQI. Manual calculation of such indices takes time and effort and may be occasionally associated with errors during sub-index calcula-

[1] Department of Environmental Sciences, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangur Darul Ehsan, Malaysia.

[2] Center of Excellence for Environmental Forensics, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 UPM Serdang, Selangur Darul Ehsan, Malaysia.

[*] Department of Environmental Sciences, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangur Darul Ehsan, Malaysia. E-mail address: NabeelMGazzaz@Yahoo.com.