

UNIVERSITI TEKNOLOGI MARA

**ARTIFICIAL NEURAL NETWORK
(ANN) APPROACH IN PREDICTING
ARSENIC AND MERCURY SPECIES IN
KINTA RIVER**

NORSHIDAH BINTI BAHARUDDIN

Thesis submitted in fulfillment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Applied Sciences

October 2014

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the result of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student : Norshidah Binti Baharuddin
Student's I.D. No. : 2009479682
Programme : Doctor of Philosophy
Faculty : Applied Sciences
Thesis Title : Artificial Neural Network (ANN) Approach In
Predicting Arsenic and Mercury Species In Kinta
River

Signature of Student : 

Date : October 2014

ABSTRACT

Surface water is most exposed to pollution from chemical, physical and biological contaminants by anthropogenic activities. Identifying the variables contributing to the deterioration of water quality is crucial and predicting the future status is vital in managing the ecosystem. As Kinta River is an ex-mining area, heavy metals contamination is expected to be the major pollution contributor particularly arsenic and mercury. Confirmation on the significant contribution of arsenic and mercury species was done by applying selected chemometrics techniques namely cluster analysis (CA) and principal component analysis (PCA) on the physicochemical variables monitored by the Department of Environment (DOE), Malaysia during the period from 1997 – 2006. Thirty physicochemical variables were selected as input variables for CA and PCA in an attempt to identify the significant variables by the factor loadings obtained from PCA. From CA, the physicochemical variables were classified into four main clusters based on the similarities and dissimilarities. PCA applied to the dataset resulted in ten varifactors with a total variance of 78.06%. Arsenic and mercury were classified as moderately significant variables that affect the Kinta River water with a factor loadings of 0.561 and 0.643, respectively. The pollution of the river due to these metals could be contributed by industrial discharge, agricultural activities and residential waste. Since the toxicity of arsenic and mercury depended on the specific species of the metals, speciation analysis is therefore important to study the toxicity of these metals species due to their significant risks to human health and to the environment. LC-ICP-MS offers several advantages for speciation analysis as it allows multi-element and multi-isotope detection, high sensitivity and a wide linear dynamic range of detection. Analytical validation carried out showed that the detection limits for As were 1.0 and 0.5 ppb for As (III) and As (V) respectively with recoveries of As (III) and As (V) in the range of 90 – 115%. The detection limits for Hg^{2+} and CH_3Hg^+ were found to be 1.0 ng L^{-1} and 1.5 $\mu\text{g L}^{-1}$, respectively with recoveries of Hg^{2+} and MeHg^+ in the range of 95 – 111%. Water samples taken from six sampling points selected from the DOE monitoring stations were analysed for arsenic and mercury species. This speciation analytical data was used in validating the artificial neural network (ANN) developed for predicting the concentration of these species in Kinta River. In developing the ANN model, two approaches were considered. The first approach, prediction was performed based on the original principal components (PCs) with thirty PCs as the input and the second approach, rotated principle components (RPCs) with eigenvalues greater than 1 with ten RPCs as inputs. It was observed that the predicted output of the first approach obtained a higher coefficient of determination (R^2) value of 1.0 for both As(V) and Hg^{2+} as compared to the second approach with R^2 of 0.915 and 0.984 for As (V) and Hg^{2+} , respectively. The sensitivity analysis with leave-one-out technique showed that both arsenic and mercury were found to have significant input-output relationship with contribution of 31.6% and 86%, respectively. The ANN model developed obtained a good agreement for As(V) with R^2 of 1, whilst relatively low correlation for Hg^{2+} with R^2 of 0.6528.

TABLE OF CONTENTS

	Page
AUTHOR'S DECLARATION	ii
ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	v
LIST OF TABLES	ix
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xiii
CHAPTER ONE: INTRODUCTION	1
1.1 BACKGROUND	1
1.2 PROBLEM STATEMENT	3
1.3 SIGNIFICANCE OF STUDY	4
1.4 SCOPE OF STUDY	5
1.5 OBJECTIVES OF STUDY	5
CHAPTER TWO: LITERATURE REVIEW	7
2.1 INTRODUCTION	7
2.2 WATER QUALITY OF KINTA RIVER	9
2.3 APPLICATION OF CHEMOMETRICS TECHNIQUES TO ENVIRONMENTAL DATA	9
2.3.1 Cluster Analysis (CA)	10
2.3.2 Principal Component Analysis (PCA)	11
2.4 SPECIATION ANALYSIS OF ARSENIC AND MERCURY	12
2.4.1 Arsenic Speciation in the Environment	12
2.4.2 Mercury Speciation in the Environment	14
2.4.3 High Performance Liquid Chromatography On-Line Coupled with Inductively Coupled-Plasma Mass Spectrometry (LC-ICP-MS)	16

2.5	APPLICATIONS OF ARTIFICIAL NEURAL NETWORKS (ANN) IN WATER QUALITY MODELING	17
2.5.1	The Multilayer Perceptron (MLP) Neural Network	19
2.6	ARTIFICIAL NEURAL NETWORK (ANN) IN ENVIRONMENTAL MODELLING	22
2.7	THE APPLICATIONS OF ANN IN PREDICTING MALAYSIA RIVER WATER QUALITY	24
2.8	ADVANTAGES AND DISADVANTAGES OF ANN	25
	CHAPTER THREE: METHODOLOGY	27
3.1	STUDY AREA	27
3.2	MONITORING STATIONS	27
3.3	DOE WATER QUALITY DATASETS	30
3.4	WATER SAMPLING	31
3.5	CHEMOMETRICS ANALYSIS	31
3.5.1	Cluster Analysis (CA)	31
3.5.2	Principal Component Analysis (PCA)	32
3.6	CHEMICALS AND REAGENTS	32
3.7	INSTRUMENTATION AND EQUIPMENT	33
3.8	PREPARATION OF SOLUTIONS	38
3.8.1	Preparation of Standard Solutions	38
3.8.2	Preparation of Spiked Samples	38
3.8.3	Preparation of Real Samples	38
3.8.4	Preparation of Mobile Phase	39
3.9	SPECIATION ANALYSIS PROCEDURE	39