การทำนายคุณภาพน้ำของคลองในกรุงเทพมหานครโดยใช้เทคนิคการทำเหมืองข้อมูล

นางศิริลักษณ์ อารีรัชชกุล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ
ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2555

PREDICTION OF WATER QUALITY FOR BANGKOK CANALS USING DATA MINING TECHNIQUES

Mrs. Sirilak Areerachakul

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Computer Science and Information Technology

Department of Mathematics and Computer Science
Faculty of Science
Chulalongkorn University

Academic year 2012

Thesis Title	PREDICTION OF WATER QUALITY FOR BANGKOK CANALS USING DATA MINING TECHNIQUES
By	Mrs. Sirilak Areerachakul
Field of Study	Computer Science and Information Technology
Thesis Advisor	Associate Professor Peraphon Sophatsathit, Ph.D.
Thesis Co-advisor	Professor Chidchanok Lursinsap, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

……………………………………………….. Dean of the Faculty of Science
(Professor Supot Hannongbua, Dr.rer.nat.)

THESIS COMMITTEE

……………………………………………….. Chairman
(Assistant Professor Saranya Maneeroj, Ph.D.)

……………………………………………….. Thesis Advisor
(Associate Professor Peraphon Sophatsathit, Ph.D.)

……………………………………………….. Thesis Co-advisor
(Professor Chidchanok Lursinsap, Ph.D.)

……………………………………………….. Examiner
(Assistant Professor Suphakant Phimoltares, Ph.D.)

……………………………………………….. Examiner
(Assistant Professor Rajalida Lipikorn, Ph.D.)

……………………………………………….. External Examiner
(Assistant Professor Supichai Tangjaitrong, Ph.D.)

ศิริลักษณ์ อารีรัชชกุล : การทำนายคุณภาพน้ำของคลองในกรุงเทพมหานครโดยใช้เทคนิคการทำเหมืองข้อมูล (PREDICTION OF WATER QUALITY FOR BANGKOK CANALS USING DATA MINING TECHNIQUES) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: รศ. ดร.พีระพนธ์ โสพัศสถิตย์ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม : ศ.ดร. ชิดชนก เหลือสินทรัพย์, 78 หน้า

น้ำที่มีการปนเปื้อนโดยปกติมักเกิดจากกิจกรรมต่างๆของมนุษย์ ปัญหาหลักที่นำมาศึกษาครั้งนี้มี 2 ประเด็น คือ 1 การวิเคราะห์หาแบบจำลองที่เหมาะสมในการจำแนกคุณภาพน้ำ 2 สร้างแบบจำลองที่มีประสิทธิภาพในการทำนายปริมาณค่าออกซิเจนที่มีผลต่อคุณภาพน้ำของคลองในกรุงเทพมหานคร โดยมีพารามิเตอร์ที่นำมาใช้ 13 ตัว พารามิเตอร์ดังกล่าวเก็บบันทึกข้อมูลมาจากสำนักการระบายน้ำกรุงเทพมหานครระหว่าง พ.ศ. 2546-2554 งานวิจัยนี้เสนอแบบจำลอง 2 ชนิด คือแบบจำลองต้นไม้ตัดสินใจ (CART) ซึ่งเป็นแบบจำลองที่เหมาะสมและมีประสิทธิภาพในการจำแนกประเภทคุณภาพน้ำของคลองในกรุงเทพมหานคร และแบบจำลองที่สองเป็นการนำเทคนิคการจัดกลุ่มมาจัดกลุ่มข้อมูลคุณภาพน้ำและผสมผสานกับโครงข่ายประสาทเทียมในการทำนายปริมาณค่าออกซิเจน ผลโดยใช้ตัววัดเปรียบเทียบประสิทธิภาพแบบจำลองด้วยค่าสัมประสิทธิ์สหสัมพันธ์ (R) ค่าความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAE) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ระหว่างเทคนิคที่ทำการศึกษาจะแสดงค่าความถูกต้องมากกว่าเทคนิคอื่นๆ ประโยชน์ที่ได้รับจากงานวิจัยนี้สามารถนำไปประยุกต์ใช้ในการวางแผนและรักษาคุณภาพน้ำให้มีประสิทธิภาพดีขึ้น

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์ ลายมือชื่อนิสิต
สาขาวิชา วิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก
ปีการศึกษา 2555 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม

5173862123: MAJOR COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

KEYWORDS : Water quality / Water management / Classification and regression tree / Multilayer perceptron neural network / Sub-space clustering / Clustering techniques / supervised and unsupervised neural networks

SIRILAK AREERACHAKUL: PREDICTION OF WATER QUALITY FOR BANGKOK CANALS USING DATA MINING TECHNIQUES. ADVISOR: ASSOC PROF. PERAPHON SOPHATSATHIT, Ph.D., CO-ADVISOR: PROF. CHIDCHANOK LURSINSAP, Ph.D., 78 pp.

Water is usually subject to contamination caused by human activities. The main problems being studied are two folds, 1) derivation of an appropriate model that can achieve accurate classification of water, and 2) construction of an efficient model in predicting the amount of Dissolved Oxygen (DO) that affects water quality of Bangkok's canals based on 13 water quality parameters. These parameters are collected from the data obtained from the Department of Drainage and Sewerage, Bangkok Metropolitan Administration during 2003-2011. This research proposes two prediction models. The Classification and Regression Tree (CART) model is established to classify water quality in Bangkok's canals. The second model is set up to cluster water quality parameters and integrated the clusters to predict DO with the help of an ANN model. Result comparisons by correlation coefficient (R), mean absolute error (MAE), and mean square error (MSE) between the proposed technique and other techniques show higher accuracy than the comparative techniques. Contributions of this work can be applied to better water quality planning and effective treatment.

Department : Mathematics and Computer Science Student's Signature ....

Field of Study : Computer Science and Information Technology Advisor's Signature ....

Academic Year: 2012 Co-advisor's Signature ....

Acknowledgments

I would like to take this opportunity to express my deepest appreciation to Associate Professor Dr. Peraphon Sophatsathit, who is my graduate advisor, for his generous advice, emotional support, encouragement, and patience on the work of this dissertation. I am also particularly grateful to my co-advisor, Professor Dr. Chidchanok Lursinsap, whose kindness, understanding, and valuable advice have enabled me complete dissertation.

I sincerely thank to the chairman and member of the committee of this research work, Assistant Professor Dr. Saranya Maneeroj and Assistant Professor Dr.Rajalida Lipikorn, Assistant Professor Dr. Suphakant Phimoltares, and Assistant Professor Dr. Supichai Tangjaitrong, for their valuable comments and personal efforts in reviewing my work.

I am grateful to the Advanced Virtual and Intelligence Computing (AVIC) Research Center for facilitating me to accomplish this research. I would like to thank the Department of Drainage and Sewerage, Bangkok Metropolitan Administration, for furnishing all the experimental data.

I am indebted to Suan Sunandha Rajabhat University for the scholastic financial support. Many thanks to Dr. Nathaporn Areerachakul and Assistant Professor Dr. Somkiat Korbuakaew for their support and advice in research work. I also thank to Mr. Prem Junsawang, friends and everyone who gave such a nice cheer up and helpful suggestion.

Finally, I would not be able to accomplish this work without the loves and supports from my beloved parents that have become the invaluable source of strengths along the duration of this study.

CONTENTS

List of Tables

Table Page

List of Figures

CHAPTER I

INTRODUCTION

Water will become the major constraining resource for sustainable development of large areas in the world. Classification and prediction water quality is an important component of water resource management and becomes the main basis for forward looking policies to relevant organizations. As water environment is a complex nonlinear system, many factors affect the water quality which often complicates the structure parameters and boundary conditions of the system in defining and establishing a representative water quality model.

1.1 Statement of the Problem

The quality of water resource is a subject of ongoing concern. Accurate assessment of long term water quality variations is also a challenging problem. One of the remedy to solve this problem is an increasing demand for monitoring water quality of many canals and rivers by regular measurement of various water quality parameters. Various municipal activities, ranging from commercial, industrial, and service that are initiated as a result of rapid expansion of the city and its population, are the culprit of accumulated environmental pollutions. Bangkok is the capital city as well as the economic center of Thailand. Water quality, in particular along the interconnecting canals of Bangkok, deteriorates to a point that the ecological system can no longer absorb such overwhelming pollutants into water resources [1]. At present, pollution level of the canals in most areas of Bangkok is very severe because the canals are still served sewers for direct discharge of waste water. Although there are legal regulations for large buildings on waste-water treatment or septic tanks in residential area, the problem still prevails. Some waste water is still discharged without treatment to public sewers which in turn drained into the canals [2]. The city uses Dissolved Oxygen (DO) which is one of the most common indicators of the aquatic ecosystem [3] to measure such leakage. The major impact can be classified into 3 types: [2]

- Impact to tourism activities: Bangkok is the capital city and the center of tourism. Therefore, deterioration of canal water has a direct impact to tourism activity. As pollution in canals is in the inner area of Bangkok, it certainly gives negative impression to the tourists who travel and stay in Bangkok.

- Impact to aquatic life: Basically, water pollution causes deaths of aquatic creatures either by its toxicity or reducing DO concentration. The toxicity may come from high concentration of sulfide or ammonia. It is observed that when water is polluted, more species of organisms are affected since they are generally more sensitive than fishes. Changes in specie mutation and counts are common. In severe cases, where DO concentration becomes zero, all aquatic creatures die. This phenomenon occurs in the highly polluted canals of Bangkok.

- Impact to public health: Poor sanitary conditions prevail in many parts of Bangkok. Records of some water related diseases generally associated with sanitary conditions are discussed in the Annual Epidemiological Surveillance.

Effects to lessen the above impact are attempted. Several techniques and tools for water quality management are applied, such as traditional statistics and data mining techniques. The latter has become extremely popular for classification and prediction in a number of application areas, power generation, bioinformatics water resources, and environment science.

This research proposes data mining technique to classify water quality based on surface water quality standards and derives methods to enhance performance efficiency of DO prediction model. This is one of the important water quality parameter which define different water quality measures depending on the designated area.

Additionally, this parameter can provide useful information for better planning and watershed management of canals in Bangkok.

1.2 Related Works

Water resource is of tremendous significance in our daily life and protection of the water resource is equivalent to the protection of our own life. However, with the development of modern society and increase of the population, a series of severe problems related to the water pollution has emerged and attracted more attention than ever before. The investigation on the water quality can be carried out in two procedures: classification and prediction. Recently, many researchers have reported about water quality prediction [4-6]; and lots of forecasting methods are proposed such as multivariate linear regression, time series method, decision tree, and so on. Park [7] combined two different artificial neural network (ANN) models to pattern and predict ecological status and water quality of target ecosystems. ANN was also used to model and predict the water quality index for rivers in Malaysia [8].

Fernández, E [9], in the Puerto Mallarino clear water plant located in Cali, conducted a series of numerical experiments which used optimal doses of coagulants for water treatment. The experiments demonstrated the closeness of fit in real time that could be achieved with data sets collected using ANN.

Li-hua Chen [10] applied ANN to classify water quality of the Yellow River. Within the period from 2003 to 2005 (high water, normal water and low water) 63 samples were collected and the measurement of 10 chemical variables of the Yellow River of Gansu were carried out. These variables are DO, Chemical Oxygen Demand (COD), Non-ion Ammonia (NHx), Volatilization Hydroxybenzene (OH), Cyanide (CN), etc. The results of all measurements for handling different chemo-informatics methods were employed: these were basic statistical methods that utilized uniform design to determinate the data set according to the water quality standards. Multi-layer perceptron (MLP) neural network and probabilistic neural networks (PNN) were used to classify the water quality of different sampling sites and different sampling time. The correlation between class of water quality and chemical measurements was determined. The model between the water quality classes and chemical measurements was built, and these models could quickly, completely, and accurately classify the water quality of the Yellow River.

There are several methods previously proposed to compute DO concentration in streams based on deoxidation process [11], rivers [12], and lakes [13]. However, these water quality models are often complex and costly in nature and data demanding [14]. ANN has been successfully used as an analysis tool in the field of water quality prediction and forecasting [15]. Palani, et al. [16] applied a neural network model for the prediction and forecasting of selected seawater quality variables. Soyupak, et al. [17] used a neural network approach to compute the pseudo steady state time and space dependent DO concentration in three separate reservoirs having different characteristics based on limited number of input variables. Sengorur, et al. [18] used a feed-forward neural network (FNN) to estimate the DO from limited input data. Kuo, et al. [19] applied ANN model to predict the DO in the Te-Chi reservoir (Taiwan). The correlation coefficients between predicted and observed DO values for training and test data sets were 0.75 and 0.72, respectively. Singh, et al. [20] computed DO and BOD levels in the Gomti river (India) by using a 3-layer FNN with back-propagation learning. The coefficients of predicted and observed DO values for training, validating, and testing data sets were 0.70, 0.74, and 0.76, respectively. Moreover, sensitivity analysis was used to select the relevant input parameters. Then FNN was applied to predict the DO in the Gruza reservoir, Serbia [21]. It can be seen that the accuracy of these results are not high enough for practical use. This may be caused by the improper training process.

1.3 Objectives

1. To develop a methodology to classify surface water quality in canals
2. To classify surface water quality in canals based on surface water quality standards.
3. To develop a methodology that can significantly improve prediction of DO in canals.

1.4 Scope of work

In this dissertation, the scope of work is constrained as follows:

1. Obtaining the data of water quality from the Department of Drainage and Sewerage Bangkok Metropolitan.
2. Using Data mining techniques to classify water quality based on surface water quality standard.
3. Integrating unsupervised and supervised neural networks to predict DO parameter.

1.5 Dissertation Organization

This dissertation is organized as follows. Chapter 2 provides necessary background for classify and prediction surface water quality. Chapter 3 describes algorithms and data collection. The results and discussion are given in Chapter 4. Finally, Chapter 5 concludes the work and suggests future extension.

CHAPTER II

RELATED BACKGROUND

This chapter provides the basic knowledge of water quality and theoretical background on classification and prediction water quality.

2.1 Knowledge of Water Quality Model

Water quality models can be applied to many different types of water system, including streams, rivers, canals, lakes, reservoirs, estuaries, coastal waters and oceans. The models describe the main water quality processes, and typically require the hydrological and constituent inputs (the water flows or volumes and the pollutant loadings). These models include terms for dispersive and/or advective transport depending on the hydrological and hydrodynamic characteristics of the water body, and terms for the biological, chemical, and physical reactions among constituents. Advective transport dominates in flowing rivers. Dispersion is the predominant transport phenomenon in estuaries subject to tidal action. Lake-water quality prediction is complicated by the influence of random wind directions and velocities that often affect surface mixing, currents, and stratification. For this and other reasons, obtaining reliable quality predictions for lakes is often more difficult than for streams, canals, rivers, and estuaries. In coastal waters and oceans, large-scale flow patterns and tide are the most important transport mechanisms. The development and application of water quality models is both a science and an art. Each model reflects the creativity of its developer, the particular water quality management problems and issues being addressed, the available data for model parameter calibration and verification, the time available for modelling and associated uncertainty, and other considerations. The fact that most, if not all, water quality models cannot accurately predict what actually happens does not detract from their value. Even relatively simple models can help managers understand the real world prototype and estimate at least the relative, if not actual, change in water quality associated with given changes in the inputs resulting from management policies or practices [22].

The recent deployment of in situ instrumentation in rivers, streams, and creeks nationwide, as well as real-time data reporting via satellite communication technology furnishes a wealth of data that had never before in the past. Special techniques such as data mining can utilize this vast base of data for pattern recognition and machine learning, so as to make accurate predictions.

2.2 Artificial Neural Networks

Artificial neural networks (ANN) is a mathematical algorithm for information processing and knowledge acquisition through the learning processes. In the design of network calculations, there is a massive interconnection of simple computing cells called neurons which will be used to imitate the human brain's capability. General descriptions of artificial neural networks can be found in various existing literatures [23] [24]. This section briefly describes ANNs in comparison with the human brain as follows.

Historically, ANN is a mathematical model or computational model based on biological neural networks. ANN can be considered as parallel distributed processors made up of simple processing units. Its functionality is to store experimental knowledge and to make it available for later use. ANN resembles the human brain in two aspects:

1. Knowledge is required by the network from its environmental through learning process.

2. Inter-neuron connection strength defined by weight value is used to store the acquired knowledge.

Human brain is basically composed of a specific type of cells, known as biological neurons. These cells provide us the abilities to remember, think, and apply our past experience to make decisions or solve problems. The brain is activated as a result of these biological neurons and the connections between them. Figure 2.1 depicts the two interconnected biological neurons. Each biological neuron comprises of four basic components, namely, dendrite, soma (cell body), axon, and synapses. Input and output signals to the soma of a biological neuron are transmitted along the axon and dendrite, while the synaptic resistance controls the strength of the signal. These neurons naturally learn to produce a particular signal by adjusting the synaptic resistance. Neurons that are connected to the others constitute an enormous network called a neural network. The synaptic resistance is formally referred as the weight of neuron.

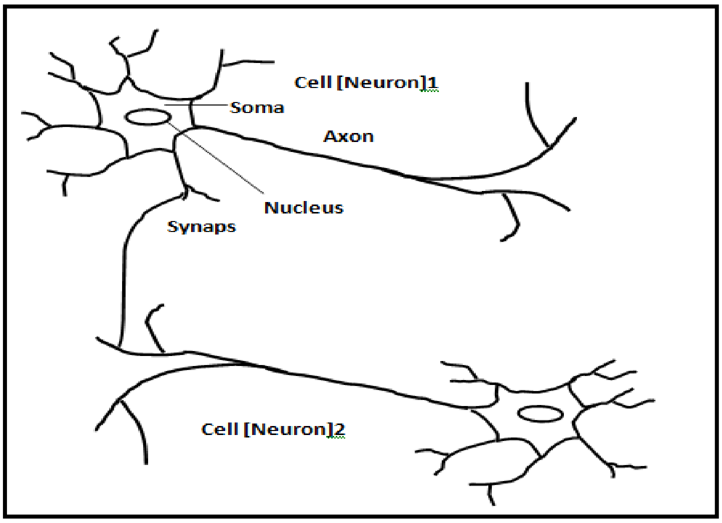


Figure 2.1: Two interconnected biological neurons.

ANN resembles the biological neuron in a way that each component forming an artificial neuron has a consistent functionality in comparison with biological neuron as follows: soma is consistent with a neuron node, axon is consistent with output, dendrite is consistent with an input, and synapse is consistent with a weight.

2.2.1 Classification of Artificial Neural Networks

For simplicity, this literature will hereafter refer to “artificial neural network” as “network” or “ANN” and “artificial neuron” as “neuron.” The general procedure of learning process or learning rule in a network is equivalent to adjusting weights of the network. Networks can be classified into three broad categories [23] based on the learning types as

1. Supervised learning: The learning rule or weights adjustment is referred to the set of input-target pairs called training set applied to the network. The learning objective is to produce the output as close to the target of the same input as possible by weight adjustment.

2. Unsupervised learning: The learning without a teacher does not require a target associated with each input pattern in the training data set. It explores the underlying structure, or correlates between patterns in the data, and organizes patterns into finite categories from these correlations.

3. Reinforcement learning: It is similar to supervised learning except that, instead of being provided with the target for each network input, the output of the network is assigned to grades or scores. Grade or score is the measure of the correctness of network outputs over a sequence of inputs.

The ANN with proposed learning rules have been widely used for solving seven classes of the challenging problems in sciences and engineering [25], namely, (i) pattern classification such as character and speech recognitions, (ii) clustering or categorization including data mining, data compression, and exploratory data analysis, (iii) prediction and forecasting, (iv) function approximations, (v) optimization, (vi) content-addressable memory or associative memory, which is extremely desirable in building multimedia information databases, and (vii) system control, such as engine idle-speed control. For applications on prediction problems, function approximations, pattern classification, data compression and control, the supervised learning neural network gains more popularity among researchers. In this work, the back-propagation algorithm, which is one of the important classes in supervised learning category, is used to model the rainfall-runoff relationship in the designated study areas.

2.2.2 A Neuron Model

A neuron model comprises of five basic elements, as shown in Figure 2.2. These elements are mathematically described as input vector, weight values, summer, activation function, and output of neuron.

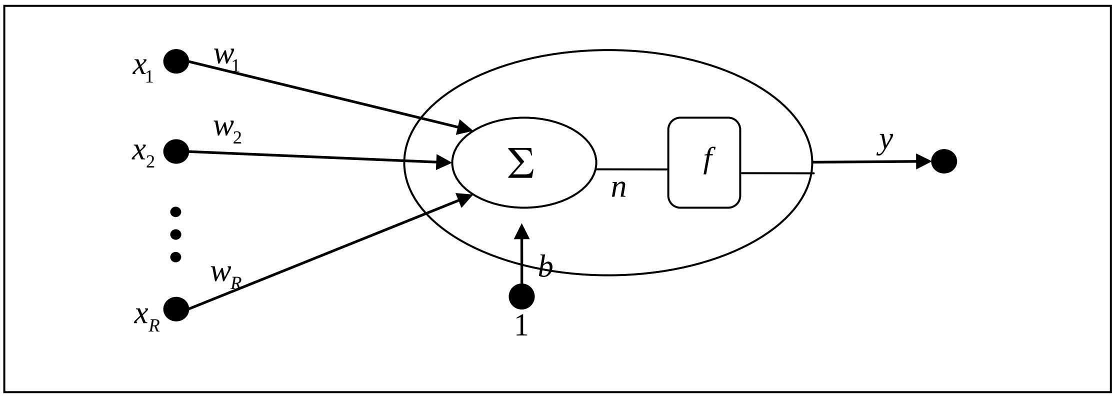


Figure 2.2: A neuron model.

An R-dimensional input vector $\vec{x}$ of a neuron is characterized by $x_1, x_2, ..., x_R$. Generally, this input can be derived from the input data or the output signals produced from other neurons. Weights $w_1, w_2, ..., w_R$ and a bias b are adjustable parameters. Bias b can be considered as weight of which the input is 1. A summer $\sum$ is an operator for generating net input *n* for an activation function f. The net input *n* is a weighted sum of all input fed into the neuron. The activation function f could be either linear or nonlinear function. A neuron produces only one output *y*, which can be written as

$$y = f(n) \text{ , where } n = \sum_{i=0}^{R} w_i x_i \qquad \text{for } i = 0 \text{ ; } w_0 = b \text{ and } x_0 = 1. \qquad (2.1)$$

Generally, the output of a neuron is defined in terms of the net input n. Three basic types of the activation functions are

1. threshold function or heaviside function defined by

$$f(n)=\begin{cases}1 & if \quad n\geq 0\\ 0 & if \quad n<0\end{cases} \quad , \tag{2.2}$$

as shown in Figure 2.3

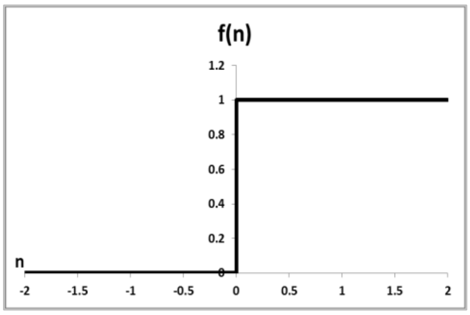


Figure 2.3: Threshold function.

2. piecewise-linear function defined by

$$f(n)=\begin{cases}1 \ , & n\geq 1\\ n \ , & 1>n>0\\ 0 \ , & n\leq 0\end{cases} \tag{2.3}$$

as shown in Figure 2.4

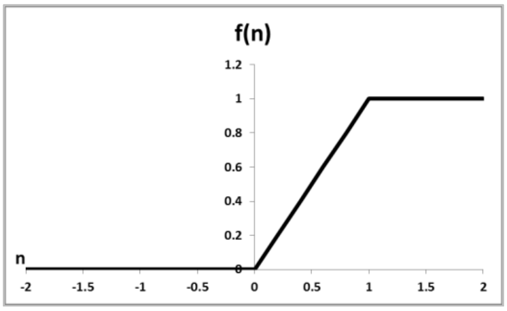


Figure 2.4: Piecewise-linear function.

3. logistic sigmoid function defined by

$$f(n) = \frac{1}{1+e^{-n}} \tag{2.4}$$

as shown in Figure 2.5

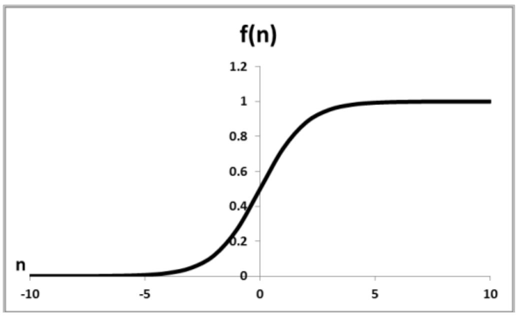


Figure 2.5: Logistic sigmoid function.

2.2.3 Multilayer Perceptrons

For supervised learning model, network convergence depends on learning parameters (weights and biases) that must be adjusted based on the result comparison of the network output and the target. One of the well-known neural networks for this learning objective is the multilayer feed-forward neural networks, which are commonly referred as the multilayer perceptrons. The feed-forward network structure typically consists of three layers: input layer, hidden layer, and output layer.

Input layer is the first layer consisting of neurons, called source nodes, which are derived from elements of the input pattern. The source nodes constitute the input signal for the neurons in the second layer.

Hidden layer is the second layer containing the computational nodes designed for signal from the input layer. In this layer, the weights and biases are adjusted by the calculation of which the output is passed on to the layer below it. The next layer can be another hidden layer or the output layer.

Output layer is the last layer of the network. It consists of computational nodes that are responsible for producing the network output.

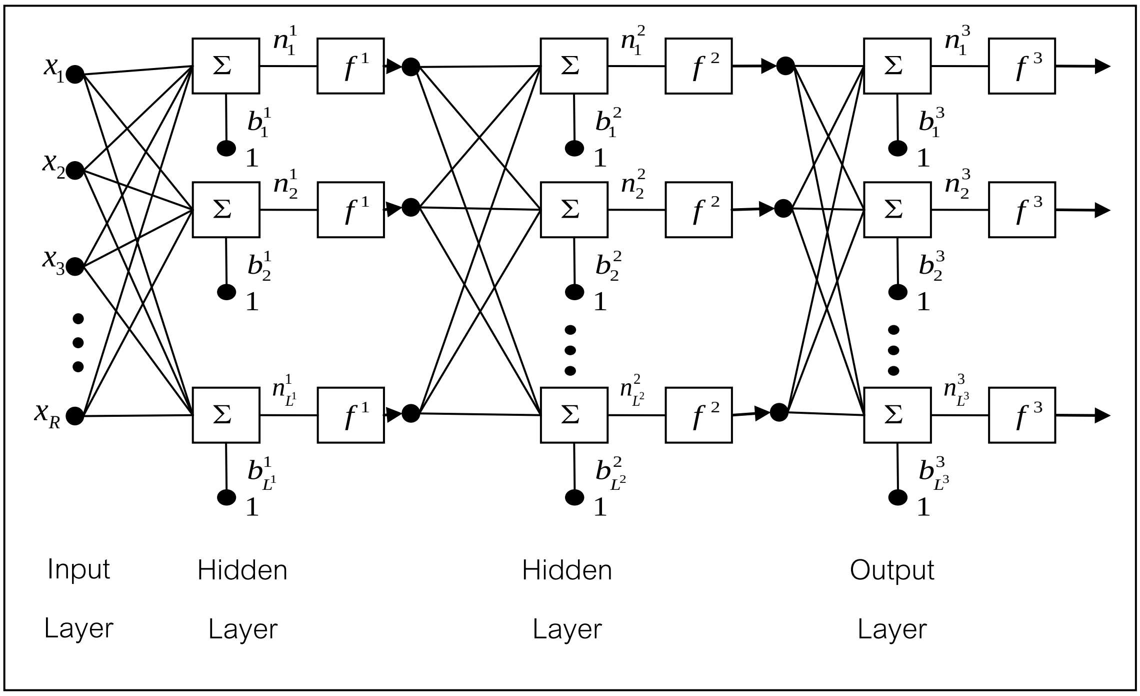


Figure 2.6: Structure of multilayer perceptrons.

When every neuron in each layer is connected to every other neuron in the next layer, the network is called a fully connected network as shown in Figure 2.6. If, however, some of these connections are missing, the network is said to be partially connected.

2.2.4 Learning Process

In the learning process of artificial neural network, there are learning rules based on weight and bias adjustment. The adjustment process is derived by the minimization of mean square error between the targets and corresponding outputs of the network.

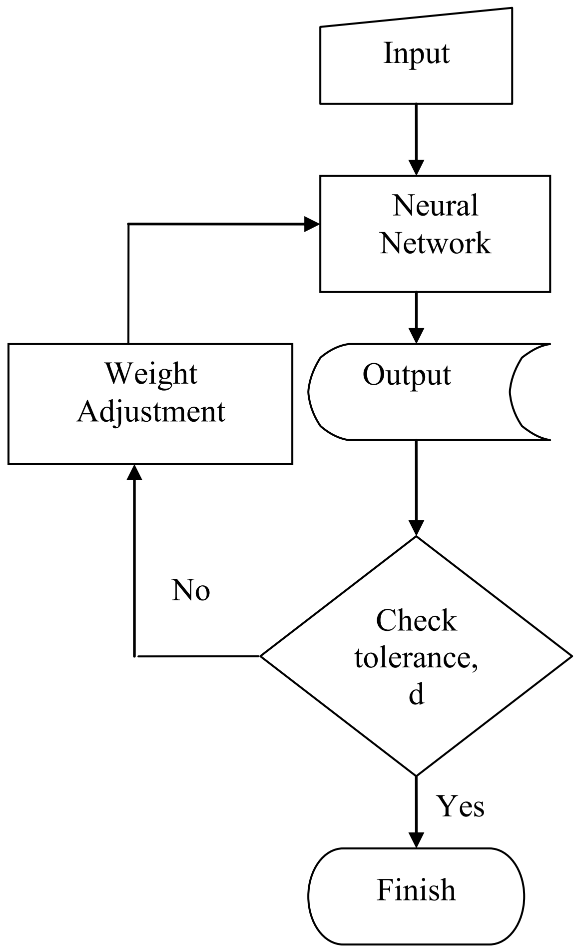


Figure 2.7: Diagram of weight adjustment process.

Figure 2.7 illustrates the diagram of weight adjustment process in the back-propagation algorithm. Firstly, the weights of the network are initialized. For each epoch, the input is fed into the network to generate the network output. Then, the difference, between the outputs and the corresponding targets by adopting statistical criterion (mean square error), is computed. If the difference is more than a tolerance d, then the weight adjustment of computational nodes is performed. The name of back-propagation is derived from the direction of weights adjustment, which proceeds backward from the output layer to the first hidden layer in the network. The process is performed iteratively until the difference between the targets and the corresponding network output is less than the tolerance d. Finally, the set of weights from the converged network is tested on another data set for validation. There are two kinds of signals, associated with the weights adjustment process, namely functions and error signals. The function signal is a signal generated from input nodes or output of neurons in hidden layers. This signal propagates forward through the network and transforms to the network outputs. The error signal originates at an output layer and propagates backward through the network.

To improve the learning process, the back-propagation algorithms are adopted in the weight adjustment process [26-30]. There is a variation of back propagation algorithms. The standard method is a gradient descent algorithm. The limit of this algorithm is that its learning time is practically slow due to the impact of learning rate [27]. There is no theoretically learning rate setting criterion to achieve the optimal learning rate that compromise between the high learning speeds and the minimization of the risk of divergence. Therefore, faster algorithms have been proposed and developed to avoid this difficulty for the past few years. Some of these faster algorithms can be classified by the two main approaches based on the order of Taylor's series expansion including first-order methods and second-order methods. First-order methods are derived from the analysis of the performance of the standard steepest descent algorithm such as resilient back-propagation algorithm. Second-order methods utilize standard numerical optimization techniques based on a quadratic model such as conjugate gradient, Quasi-Newton, and Levenberg-Marquardt algorithms. In both cases, iterative techniques are applied to minimize the cost function.

In this dissertation, the Levenberg-Marquardt algorithm is used in the weight adjustment process for the network training. This algorithm is one of the most efficient learning algorithms for neural network. The main advantage of the Levenberg-Marquardt is its speed of convergence [27] [31-32].

2.2.5 Levenberg-Marquardt Algorithm

The Levenberg-Marquardt (LM) algorithm [23] [31-32] is originally based on the minimization problem of a performance function $F(\vec{x})$, which is the sum of square functions. It is derived from Newton's method for the estimation of Hessian matrix, which is the second derivatives of the performance function, in terms of the Jacobian matrix.

The Newton's method for optimizing a performance function $F(\vec{x})$ is described as follows:

$$\Delta\vec{x}_k = -H_k^{-1}\vec{g}_k, \tag{2.5}$$

where H_k^{-1} is the inverse of Hessian matrix (H_k) at the k^{th} iteration. The Hessian matrix (H_k) characterized by $H_k \cong \nabla^2 F(\vec{x}_k)$, and $\vec{g}_k$ is gradient at the k^{th} iteration given by $\vec{g}_k \cong \nabla F(\vec{x}_k)$.

The sum of square functions $F(\vec{x})$ can be written as

$$F(\vec{x}) = \frac{1}{2}\sum_{i=1}^{N} v_i^2(\vec{x}) = \frac{1}{2}\vec{v}^T\vec{v}. \tag{2.6}$$

Here $\vec{x} = [x_1\ x_2 \ldots x_n]^T$ and $\vec{v} = [v_1\ v_2\ \ldots\ v_N]^T$, where $(\ )^T$ is the transpose of vector or matrix.

The j^{th} element of the gradient can be expressed as

$$\frac{\partial F(\vec{x})}{\partial x_j} = \sum_{i=1}^{N} v_i\,(w)\frac{\partial v_i}{\partial x_j}. \tag{2.7}$$

Therefore, the gradient can be written in matrix form as

$$\vec{g}(\vec{x}) = \nabla F(\vec{x}) = J^T(\vec{x})\vec{v}(\vec{x}), \tag{2.8}$$

where the Jacobian matrix is

$$\mathrm{J}(\vec{\mathrm{x}}) = \begin{bmatrix} \frac{\partial v_1}{\partial x_1} & \frac{\partial v_1}{\partial x_2} & \cdots & \frac{\partial v_1}{\partial x_n} \\ \frac{\partial v_2}{\partial x_1} & \frac{\partial v_2}{\partial x_2} & \cdots & \frac{\partial v_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial v_N}{\partial x_1} & \frac{\partial v_N}{\partial x_2} & \cdots & \frac{\partial v_N}{\partial x_n} \end{bmatrix}. \tag{2.9}$$

The k,j^{th} element of the Hessian matrix is given by

$$\mathrm{H}_{kj} = [\nabla^2 F(\vec{\mathrm{x}})]_{kj} = \sum_{i=1}^{N} [v_i(\vec{\mathrm{x}}) \frac{\partial^2 v_i(\vec{\mathrm{x}})}{\partial x_k \partial x_j} + \frac{\partial v_i(\vec{\mathrm{x}})}{\partial x_k} \frac{\partial v_i(\vec{\mathrm{x}})}{\partial x_j}]. \tag{2.10}$$

The final form of the Hessian matrix can then be rewritten in matrix notation as

$$\mathrm{H}(\vec{\mathrm{x}}) = \nabla^2 F(\vec{\mathrm{x}}) = \mathrm{J}(\vec{\mathrm{x}})^{\mathrm{T}} \mathrm{J}(\vec{\mathrm{x}}) + \mathrm{S}(\vec{\mathrm{x}}), \quad \text{where } \mathrm{S}(\vec{\mathrm{x}}) = \sum_{i=1}^{N} v_i(\vec{\mathrm{x}}) \frac{\partial^2 v_i(\vec{\mathrm{x}})}{\partial x_k \partial x_j}. \tag{2.11}$$

For the Gauss-Newton method, it is assumed that $\mathrm{S}(\vec{\mathrm{x}}) \approx 0$. The Hessian matrix can then be approximated by

$$\mathrm{H}(\vec{\mathrm{x}}) = \nabla^2 F(\vec{\mathrm{x}}) = \mathrm{J}(\vec{\mathrm{x}})^{\mathrm{T}} \mathrm{J}(\vec{\mathrm{x}}). \tag{2.12}$$

From (2.10) - (2.12), we obtain

$$\Delta \vec{\mathrm{x}}_k = -[\mathrm{J}^{\mathrm{T}}(\vec{\mathrm{x}}_k) \mathrm{J}(\vec{\mathrm{x}}_k)]^{-1} \mathrm{J}^{\mathrm{T}}(\vec{\mathrm{x}}_k) \vec{\mathrm{v}}(\vec{\mathrm{x}}_k). \tag{2.13}$$

One serious difficulty of the Gauss-Newton method occurs when the Hessian matrix is not invertible at some iteration. To resolve this, the approximate Hessian matrix has to be slightly modified by

$$G = H + \mu I \text{ , where } \mu \text{ is a positive constant.} \tag{2.14}$$

The question here is to show that the matrix G can be made invertible. Let the eigenvalues and eigenvectors of H be $\{\lambda_1, \lambda_2, ..., \lambda_s\}$ and $\{z_1, z_2, ..., z_s\}$, respectively. Then

$$Gz_i = [H + \mu I]z_i = (\lambda_i + \mu)z_i \quad , \tag{2.15}$$

where $\mu > 0$ and I is the identity matrix.

The eigenvalues of G are $\lambda_i + \mu$. It should be noted that both matrices G and H possess the same eigenvectors. The matrix G can be made positive definite by increasing μ until $\lambda_i + \mu > 0$ for all i and hence the matrix is invertible. Following (2.4), the Levenberg-Marquardt algorithm can be expressed as

$$\Delta\vec{x}_k = -[J^T(\vec{x}_k)J(\vec{x}_k) + \mu_k I]^{-1} J^T(\vec{x}_k)\vec{v}(\vec{x}_k). \tag{2.16}$$

As μ_k increases, the learning rate decreases. This leads to the steepest descent algorithm:

$$\vec{x}_{k+1} = \vec{x}_k - \frac{1}{\mu_k} J^T(\vec{x}_k)\vec{v}(\vec{x}_k) = \vec{x}_k - \frac{1}{2\mu_k} \nabla F(\vec{x}_k) \ , \tag{2.17}$$

for large μ_k.

On the other hand, the algorithm reduces to the Gauss-Newton method as μ_k decreases to zero.

The algorithm begins by setting small value of μ_k, i.e. $\mu_k = 0.01$. For the k^{th} iteration, if $F(\vec{x})$ obtained from computation does not decrease, then μ_k is multiplied by a constant $\delta > 1$. Since a small weight adjustment in (2.17) is taken in the direction of steepest descent, $F(\vec{x})$ should eventually decrease. Once $F(\vec{x})$

decreases, the μ_k is divided by δ in the next iteration. This algorithm is basically the Gauss-Newton, which can provide faster convergence [32].

2.3 Classification and Regression Tree

Classification and Regression Trees (CART) is a classification method which uses historical data to construct so-called decision trees. Decision trees are then used to classify new data. In order to use CART, one needs to know number of classes a priori [33]. For building decision trees, CART uses learning sample which is a set of historical data with pre-assigned classes for all observations. CART algorithm will search for all possible variables and all possible values in order to find the best split in question that splits the data into two parts with maximum homogeneity. The process is then repeated for each of the resulting data fragments. CART can easily handle both numerical and categorical variables. Among other advantages of CART method is its robustness to outliers. Usually the splitting algorithm will isolate outliers in individual node or nodes. An important practical property of CART is that the structure of its classification or regression trees is invariant with respect to monotone transformations of independent variables. One can replace any variable with its logarithm or square root value, the structure of the tree will not change. The term partitioning refers to the fact that the dataset is split into sections or partitioned.

2.3.1 Tree Growing Process

The basic idea of tree growing is to choose a split among all the possible splits at each node so that the resulting child nodes are the "purest." In this algorithm, only univariate splits are considered. That is, each split depends on the value of only one predictor variable. All possible splits consist of possible splits of each predictor. If X is a nominal categorical variable of I categories, there are $2^{I-1}-1$ possible splits for this predictor. If X is an ordinal categorical or continuous variable with K different values, there are $K-1$ different splits on X. A tree is grown starting from the root node by repeatedly using the following steps on each node.

1. Find each predictor's best split. For each continuous and ordinal predictor, sort its values from the smallest to the largest. For the sorted predictor, go through each value from top to examine each candidate split point or v, to determine the best candidate, that is, if x $\leq$ v, the case goes to the left child node, otherwise, goes to the right. The best split point is the one that maximize the splitting criterion the most when the node is split according to the criterion. The definition of splitting criterion is in later section. For each nominal predictor, examine each possible subset of categories or A, to find the best split, that is, if x$\in$ A, the case goes to the left child node, otherwise, goes to the right.

2. Find the node's best split. Among the best splits found in step 1, choose the one that maximizes the splitting criterion.

3. Split the node using its best split found in step 2 if the stopping rules are not satisfied.

2.3.2 Splitting Criteria and Impurity Measures

At node t, the best split s is chosen to maximize a splitting criterion Δi(s,t). When the impurity measure for a node can be defined, the splitting criterion corresponds to a decrease in impurity.

2.3.2.1 Categorical Dependent Variable

If Y is categorical, there is Gini splitting criteria. At node t, let probabilities $p(j,t)$, $p(t)$ and $p(j/t)$ be estimated by

$$p(j,t)=\frac{\pi(j)N_{w,j}(t)}{N_{w,j}}, \tag{2.18}$$

where $p(j,t), j=1,...,J$ is the probability of a case in class j and node t,

$N_{w,j}=\sum_{n\in h} w_n f_n I(y_n=j)$, $N_{w,j}(t)=\sum_{n\in h(t)} w_n f_n I(y_n=j)$ with $I(a=b)$ being indicator function taking value 1 when $a=b$ and 0 otherwise.

$h = \{x_n, y_n\}_{n=1}^{N}$ is the whole learning sample, $h(t)$ is the learning samples that fall in node t .

$$p(t) = \sum_j p(j,t), \tag{2.19}$$

where $p(t)$ is the probability of a case in node t.

$$p(j/t) = \frac{p(j,t)}{p(t)} = \frac{p(j,t)}{\sum_j p(j,t)}, \tag{2.20}$$

where $p(j/t), j = 1,...,J$ is the probability of a case in class j given that it falls into node t.

2.3.2.2 Gini Criterion

The Gini impurity measure at a node t is defined as

$$i(t) = \sum\nolimits_{i,j} C(i/j)p(i/t)p(j/t), \tag{2.21}$$

where $C(i/j)$ is the cost of miss-classifying a class j case as a class i case. Clearly $C(j/j) = 0$.

$$\Delta i(s,t) = i(t) - p_L i(t_L) - p_R i(t_R), \tag{2.22}$$

where p_L and p_R are probabilities of sending a case to the left child node t_L and to the right child node t_R respectively. They are estimated as $p_L = p(t_L)/p(t)$ and $p_R = p(t_R)/p(t)$.

Note that when user-specified costs are involved, the altered priors can be used to replace the priors (optional). When altered priors are used, the problem is considered as if no costs were involved. The altered prior is defined as $\pi'(j) = \frac{C(j)\pi(j)}{\sum_j C(j)\pi(j)}$, where $C(j) = \sum\nolimits_i C(i/j)$.

2.3.3 Stopping Rules

Stopping rules control if the tree growing process should be stopped or not. The following stopping rules are used:

• If a node becomes pure; that is, all cases in a node have identical values of the dependent variable, the node will not be split.

• If all cases in a node have identical values for each predictor, the node will not be split.

• If the current tree depth reaches the user-specified maximum tree depth limit value, the tree growing process will stop.

• If the size of a node is less than the user-specified minimum node size value, the node will not be split.

• If the split of a node results in a child node whose node size is less than the user specified minimum child node size value, the node will not be split.

• If for the best split s^* of node t, the improvement $\Delta I(s^*,t) = p(t)\Delta i(s^*,t)$ is smaller than the user-specified minimum improvement, the node will not be split.

2.3.4 Missing Value Handling

If the dependent variable of a case is missing, this case will be ignored in the analysis. If all predictor variables of a case are missing, this case will also be ignored. If the case weight is missing, zero, or negative, the case is ignored. If the frequency weight is missing, zero, or negative, the case is ignored. The surrogate split method is otherwise used to deal with missing data in predictor variables. Suppose that $X^* < s^*$ is the best split at a node. If value of X^* is missing for a case, the best surrogate split (among all non-missing predictors associated with surrogate splits) will be used to decide which child node it should go. If there are no surrogate splits or all the predictors associated with surrogate splits for a case are missing, the majority rule is used.

2.4 Clustering Techniques

Clustering is an important data mining technique that has a wide range of applications in many areas such as biology, medicine, market research, image processing, and geographical information systems, among others. In this research, the water qualities of canals in Bangkok were grouped according to their characteristic forming clusters. The clustering process was carried out using a K-means algorithm and Fuzzy c-means algorithm.

2.4.1 *K*-means algorithm

K-means algorithm [34] is a cluster analysis technique used as a partitioning method. It uses a similarity measure to cluster a group of data. Usually, the similarity measure is based on L_1-norm or L_2-norm. *K*-means algorithm is composed of the following steps [35]. Suppose a given data set must be clustered into *K* groups.

1. Generate the locations of *K* center points in the data space. These *K* points represent the initial centroids.
2. Assign each data point to the closest centroid.
3. Recalculate the locations of the *K* centroids with respect to the assigned data groups from step 2.
4. Repeat steps 2 and 3 until all centroids no longer change their locations.

The aim of this algorithm is to find *K* centroids for the *K* clusters. The algorithm minimizes the following objective function:

$$J = \sum_{j=1}^{K}\sum_{i=1}^{N}\left\| v_i - c_j \right\|^2 \qquad (2.23)$$

where v_i is the i^{th} data point, c_j is the j^{th} cluster centroid, $\|*\|$ is L_2-norm between a data point v_i and the cluster centroid c_j , and N is the number of data points.

2.4.2 Fuzzy c-mean algorithm

Fuzzy c-mean (FCM) is a popular clustering technique for which a data point is allowed to assign to two or more clusters based on its clusterwise membership degrees. This method is frequently applied in pattern recognition [36]. The algorithm minimizes the following objective function:

$$J = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m}\left\| v_i - c_j \right\|^2 \qquad (2.24)$$

where N is the number of data points, C is the number of clusters, m is any real number greater than 1, u_{ij}^{m} is the degree of membership of v_i in the cluster j , v_i is the i^{th} data, c_j is the center of the j^{th} cluster and $\|*\|$ is any norm expressing the similarity between any measured data and the center.

Fuzzy partitioning is carried out through an iterative optimization of the objective function as shown in Eq 2.19. The updates on membership degree u_{ij}^{m} and the c_j cluster centroid are given by

$$u_{ij}^{m} = \frac{1}{\sum_{t=1}^{C}\left(\frac{\left\| v_i - c_j \right\|}{\left\| v_i - c_t \right\|}\right)^{\frac{2}{m-1}}} \qquad (2.25)$$

and

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m v_i}{\sum_{i=1}^{N} u_{ij}^m} \tag{2.26}$$

This iteration stops when

$$\max\left\{\left\| u_{ij}^m(t+1) - u_{ij}^m(t) \right\|\right\} < \varepsilon \qquad (2\ 27)$$

where N is a terminating criterion set between 0 and 1 and $u_{ij}^m\ (t)$ is the membership degree at iteration step t. Eventually, this procedure converges to a local minimum or a saddle of point of J. The algorithm is composed of the following steps:

1. Initialize matrix $U(0)$ = $\left[u_{ij}^m\right]$ and set t = 0.

2. At t-step, calculate the centers vectors **C**(t) = [**c**$_1$(t); **c**$_2$(t),…, **c**$_C$(t)]^T based on $U(t)$.

3. Update $U(t)$ and set $t = t+1$ t = t + 1.

4. **If** max $\left\{\left\| u_{ij}^m(t+1) - u_{ij}^m(t) \right\|\right\} < \varepsilon$ **then** STOP **otherwise** return to step 2.

2.4.3 Compactness and Separation Quality Measures

The criteria widely accepted for partitioning a data set into a number of clusters are the degree of separation of clusters and their compactness. The optimum case implies that all parameters lead to the partitions as close as possible in terms of similarity to the real partitions of the data set [37]. A reliable quality assessment index should consider both the compactness and the separation; one of the quality measures that can be used in clustering is described as follows [37-38].

The compactness of the i^{th} spatial data set $V_i = \left\{v_{ij} \in R^n \mid 1 \le j \le N_i\right\}$, determined by σ (V_i), is computed as follows:

$$\sigma(V_i) = \frac{1}{N_i}\sum_{j=1}^{N_i}\left(v_{ij} - \overline{v_i}\right)^T\left(v_{ij} - \overline{v_i}\right), \tag{2.28}$$

where $\overline{v_i} = \dfrac{1}{N_i}\sum_{j=1}^{N_i} v_{ij}$.

Assume the set of clusters $\{V_i\}_{i=1}^{K}$ such that V = $V_1 \cup V_2 \cup ... \cup V_K$. After clustering, the following condition $V_i \cap V_j = \phi, \mathrm{i} \neq \mathrm{j}$ is expected to obtain. The total compactness of spatial data set with respect to K clusters, denoted by σ_t , is defined as

$$\sigma_t = \sum_{i=1}^{K} \left\| \sigma(V_i) \right\| . \qquad (2.29)$$

The average compactness of C clusters, Comp, is:

$$Comp = \frac{\sigma_t}{K} . \qquad (2.30)$$

The average scattering of data set compactness, $Scatt\_Comp$, is:

$$Scatt\_Comp = \frac{Comp}{\left\| \sigma(V) \right\|} . \qquad (2.31)$$

The more compact the clusters are, the smaller the $Scatt\_Comp$ becomes. Thus, for a given spatial data set, a smaller $Scatt\_Comp$ indicates a better clustering scheme. The distance between clusters is defined by the average distance between the centroid of specified clusters, that is,

$$d = \frac{\sum_{i=1}^{K} \sum_{j=1}^{K} \left\| \overline{v_i} - \overline{v_j} \right\|}{K(K-1)} . \qquad (2.32)$$

The larger d implies the higher degree of separation among clusters. A quality measure for clustering is defined as follows:

$$CD = \frac{Scatt\_Comp}{d} . \qquad (2.33)$$

The CD index measures the distance between the two clusters for the hierarchical clustering [39]. The definition of CD indicated that both criteria of "good" clustering (i.e., compactness and separation) are properly combined, enabling reliable evaluation of clustering results [38]. The small value of CD indicates that all the clusters in clustering scheme are overall compact and separated.

CHAPTER III

PROPOSED METHODOLOGY

3.1 Concepts of Proposed Solution

The characteristic of the canals in Bangkok is that they consist of many sites, are very large and complex system. So, this research demonstrates CART Model to classify water quality based on surface water quality standards. In addition, it provides the prediction DO parameter by applying clustering technique ANN with to arrive at better prediction results than without clustering technique. The proposed solution contributes to providing useful information for better planning and watershed management of canals in Bangkok.

3.2 Methodology of Classifying Water Quality

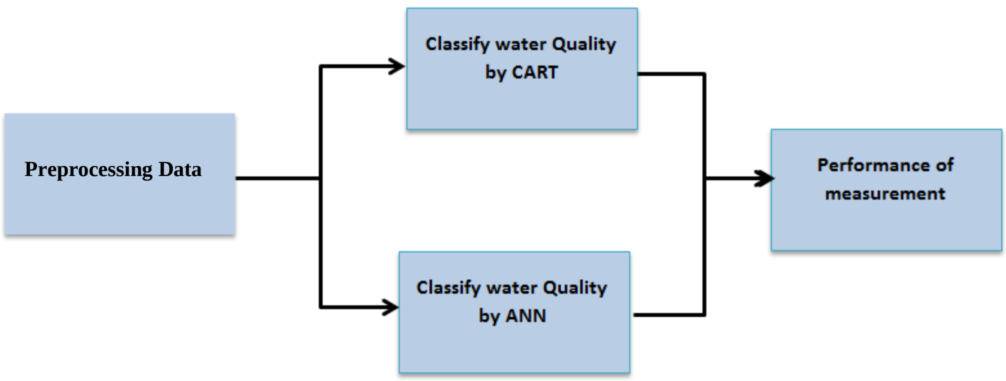


Figure 3.1: Research methodology of classifying water quality

Figure 3.1 shows a methodology for classifying water quality based on surface water quality standards. The objective is to find an optimal model to classify water quality of Bangkok's canals which constitute volume of data to be processed. The CART model classifies water quality while the ANN model adds the prediction of DO to attain finer and more accurate water quality.

The construct of ANN and CART algorithms are described below CART analysis encompasses four basic steps. The first step is tree building, during which a tree is built using recursive splitting of nodes. Each resulting node is assigned a predicted class based on the distribution of classes in the learning data set which would occur in that node and the decision cost matrix. The assignment of a predict class to each node occurs regardless of that node is subsequently split into child nodes. The second step determines stopping criteria for the tree building process. At this point, a maximal tree has been produced which probably overfits the information contained within the learning data set. The third step is tree pruning which results in the creation of a sequence of simpler trees by cutting off increasingly important nodes. The fourth step is optimal tree selection. The tree which fits the information in the learning data set but does not overfit the information is selected from among the sequence of pruned trees.

ANN analysis consists of two important functions: as pattern classifiers and as nonlinear adaptive filters. A general network consists of a layered architecture, an input layer, one or more hidden layers and an output layer. The Multilayer perceptron (MLP) is used extensively to solve a number of different problems, including pattern recognition and interpolation. Each layer is composed of neurons, which are interconnected with each other by weights. In each neuron, a specific mathematical function called the activation function accepts input from previous layers and generates output for the next layer. Both algorithms are summarizes below.

Algorithm1: Constructing Classification Tree

1. Partition the *D*-dimensional feature vector set $\mathbf{V}$ into training set $\mathbf{V}^{(tr)}$ and testing set $\mathbf{V}^{(te)}$.
2. Apply the CART Algorithm to construct a classification tree *T* based on the set $\mathbf{V}^{(tr)}$.
3. For each $\mathbf{v}_i \in \mathbf{V}^{(te)}$ do
4. Test and measure the prediction accuracy of $\mathbf{v}_i$ with the corresponding tree *T*.
5. Endfor

Algorithm2: Neural Network for Classification

1. Partition the *D*-dimensional feature vector set into training set $\mathbf{V}^{(tr)}$ and testing set $\mathbf{V}^{(te)}$.
2. Set the first and final structures of neural network and called *FirstNumHid* and *LastNumHid*, respectively.
3. For *NumHid* = *FirstNumHid* , *FirstNumHid*+1, … , *LastNumHid* do
4. Construct the MLP with *NumHid* hidden neurons based on the set $\mathbf{V}^{(tr)}$.
5. For each $\mathbf{v}_i \in \mathbf{V}^{(te)}$ do
6. Test and measure the prediction accuracy of $\mathbf{v}_i$ with the obtained MLP with *NumHid* hidden neurons.
7. Endfor
8. Endfor

3.3 Method of Predicting DO

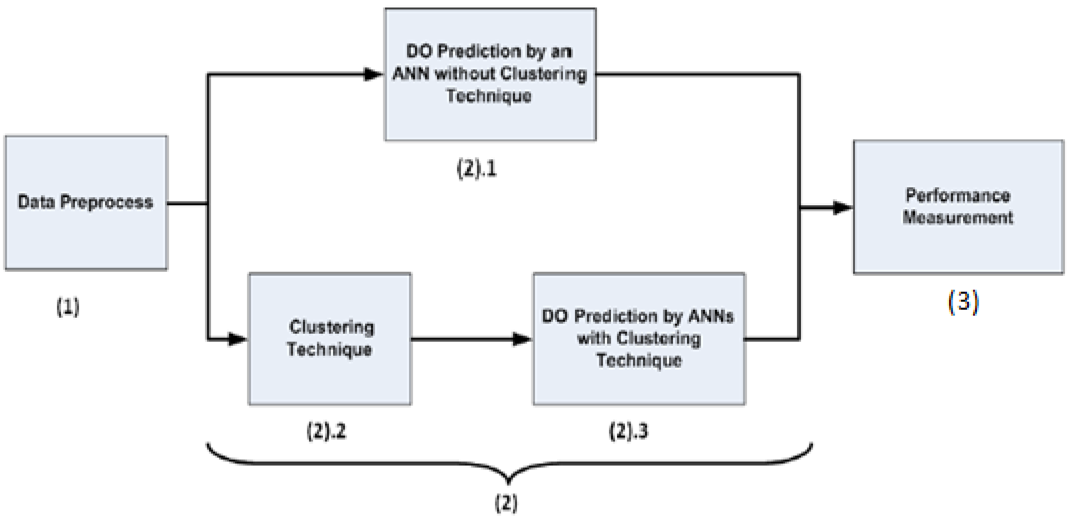


Figure 3.2: The research methodology for 1-month DO prediction

From Figure 3.2, a 1-month DO prediction of canals in Bangkok was considered. The canals of Bangkok compose of many sites that form a complex water network system. The main problem is to predict the amount of DO in the current month based on 13 water quality parameters, namely temperature (Temp), pH value (pH), Hydrogen Sulfide (H_2S), DO, BOD, Chemical Oxygen Demand (COD) Suspended Solids (SS), Total Kjeldahl Nitrogen (TKN), Ammonia Nitrogen (NH_3N), Nitrite Nitrogen (NO_2N), Nitrate Nitrogen (NO_3N), Total Phosphorous (T-P) and Total Coliform (TC) collected from the previous month. These parameters can be considered as a T-tuple vector and are treated as the input features in the predicting process, where T is the number of parameters used as prediction inputs.

The prediction problem is transformed into a problem of constructing a manifold of a predicting function or functional approximation in a high dimensional space. To attain the prediction accuracy as high as possible, the error associated with the constructed manifold must be minimized. Error minimization is based on the observation that the constructed manifold can be easily represented as a set of composite activation functions of all neurons in a supervised neural network. During manifold construction, the amount of weight adjusted is controlled by the total error computed from the entire training data set scattering in the data space. Therefore, it is rather difficult to minimize the local error of the constructed manifold. To resolve the effect, all data in the data space must be locally clustered and used to construct a local manifold.

Let $\mathrm{v}_i = \left[e_{i,1}, e_{i,2}, ..., e_{i,13}\right]^T$ be the feature vector formed at time i and e_{ij} be the j^{th} feature of vector i. An amount of DO at time i, denoted by o_i, can be written as follows:

$$o_i = f\left(\mathrm{v}_i\right). \tag{3.1}$$

The set of o_i represents the whole manifold of data space. To construct the function $f(.)$ with minimum error for every o_i, the whole manifold must be partitioned into several sub-manifolds to eliminate the error side-effect caused by irrelevant feature vectors and their distribution. Each sub-manifold must be constructed by a function

$f(\mathrm{v}_i)$ for all v_i distributed within the region of the sub-manifold. This observation leads to the following sub-problems.

1. How can the set of v_i for $1 \le k \le N$ are clustered into several groups so that the function approximated from each group has a minimum error with respect to each target o_i ?

2. For a given group of v_i obtained in problem 1, how can the sub-manifold by function $f(.)$ be constructed or approximated with a minimum error?

The first problem concerns the technique of clustering the feature vectors according to their similarity. This may imply that the training feature space is partitioned into several connected sub-spaces. It is rather difficult to estimate the best number of clusters in the algorithm. Assume the number of cluster to be a constant K. The actual value of K will be discussed in Chapter 4. For the second problem, the feature vectors in each sub-space are trained by using a feed-forward neural network to construct the manifold of the sub-space. The overview of the proposed algorithm to construct the DO predicting function is as follows.

Algorithm 3: Constructing DO Predicting Function

1. Normalize of feature vectors by using Algorithm 4.
2. Partition feature vector set $\mathbf{V}$ into training set $\mathrm{V}^{(tr)}$ and testing set. $\mathrm{V}^{(te)}$
3. Apply K-mean clustering algorithm to group feature vectors $\mathrm{V}^{(tr)}$ into K clusters, i.e. $\left\{\mathrm{V}_1^{(tr)}, \mathrm{V}_2^{(tr)}, \ldots, \mathrm{V}_K^{(tr)}\right\}$.
4. For each $\mathrm{V}_i^{(tr)}, 1 \le i \le K$ do
5. Get the distribution boundary in each dimension from each feature vector in $\mathrm{V}_i^{(tr)}$.
6. Train $\mathrm{V}_i^{(tr)}$ by using a feed-forward neural network.
7. Endfor
8. For each $\mathrm{v}_i \in \mathrm{V}^{(te)}$ do
9. Identify the corresponding neural network of step 6 by using the value of each feature in v_i and the distribution boundary obtained from step 5.

10. Test and measure the prediction accuracy of $\mathbf{v}_i$ with the corresponding network in step 6.
11. **Endfor**
12. Apply fuzzy c-mean clustering algorithm to group feature vectors $\mathbf{V}^{(tr)}$ into K clusters, i.e. $\left\{\mathrm{V}_1^{(tr)}, \mathrm{V}_2^{(tr)}, \ldots, \mathrm{V}_K^{(tr)}\right\}$.
13. **For** each $\mathrm{V}_i^{(tr)}, 1 \leq i \leq C$ **do**
14. Get the distribution boundary in each dimension from each feature vector in $\mathrm{V}_i^{(tr)}$.
15. Train $\mathrm{V}_i^{(tr)}$ by using a feed-forward neural network.
16. **Endfor**
17. **For** each $\mathrm{v}_i \in \mathrm{V}^{(te)}$ **do**
18. Identify the corresponding neural network of step 15 by using the value of each feature in v_i and the distribution boundary obtained from step 14.
19. Test and measure the prediction accuracy of v_i with the corresponding network in step 15.
20. **Endfor**

Algorithm 4: Normalizing Feature Values

1. **For** $1 \leq j \leq 13$ **do**
2. Let $a = \max_{1 \leq i \leq N} \left\{e_{i,j}\right\}$.
3. Let $b = \min_{1 \leq i \leq N} \left\{e_{i,j}\right\}$.
4. **For** each e_{ij} $, 1 \leq i \leq N$ **do**
5. Compute $e_{ij} = \dfrac{e_{i,j} - b}{a - b}$.
6. **Endfor**
7. **Endfor**

The following three performance measures, i.e., correlation coefficient (R), mean absolute error (MAE), and mean square error (MSE) are used to evaluate the performance of the propose technique. Definitions of these measures are as follows.

$$R = \frac{\sum(Q_o - M_o)(Q_p - M_p)}{\sqrt{\sum(Q_o - M_o)^2 \sum(Q_p - M_p)^2}}, \tag{3.2}$$

$$MAE = \frac{1}{N}\sum\left|Q_0 - Q_p\right|, \tag{3.3}$$

and

$$MSE = \frac{1}{N}\sum(Q_o - Q_p)^2, \tag{3.4}$$

where Q_0 and Q_p are the observed and predicted values, N is the total number of data, M_o and M_p are the mean of the observed and predicted values.

CHAPTER IV

EXPERIMENTS AND RESULTS

This chapter describes the construction and experimentation of the optimal model for classifies surface water quality and a methodology that significantly improves the prediction of DO in canals. The experimental results follow subsequently.

4.1 Water Quality data

The data set used in this research were monthly water quality data obtained from the Department of Drainage and Sewerage, Bangkok Metropolitan Administration, during the years 2003-2011. There were 276 sites covering 155 canals. A total data in the experiments consists of 13 parameters. The unit of each surface water quality parameter is shown in Table 4.1 [2] and Figure 4.1 [1] shows the network of the canals.

Table4.1: LIST OF SURFACE WATER QUALITY PARAMETERS.

Name of parameters	Unit of measurement
Temperature	Celsius
pH value	Standard Units
Hydrogen Sulfide	Milligrams per liter
Dissolved Oxygen	Milligrams per liter
Biochemical Oxygen Demand	Milligrams per liter
Chemical Oxygen Demand	Milligrams per liter
Suspended Solids	Milligrams per liter
Total Kjeldahl Nitrogen	Milligrams per liter
Ammonia Nitrogen	Milligrams per liter
Nitrite Nitrogen	Milligrams per liter
Nitrate Nitrogen	Milligrams per liter
Total Phosphorous	Milligrams per liter
Total Coliform	Most Probable Number per 100 Milliliter

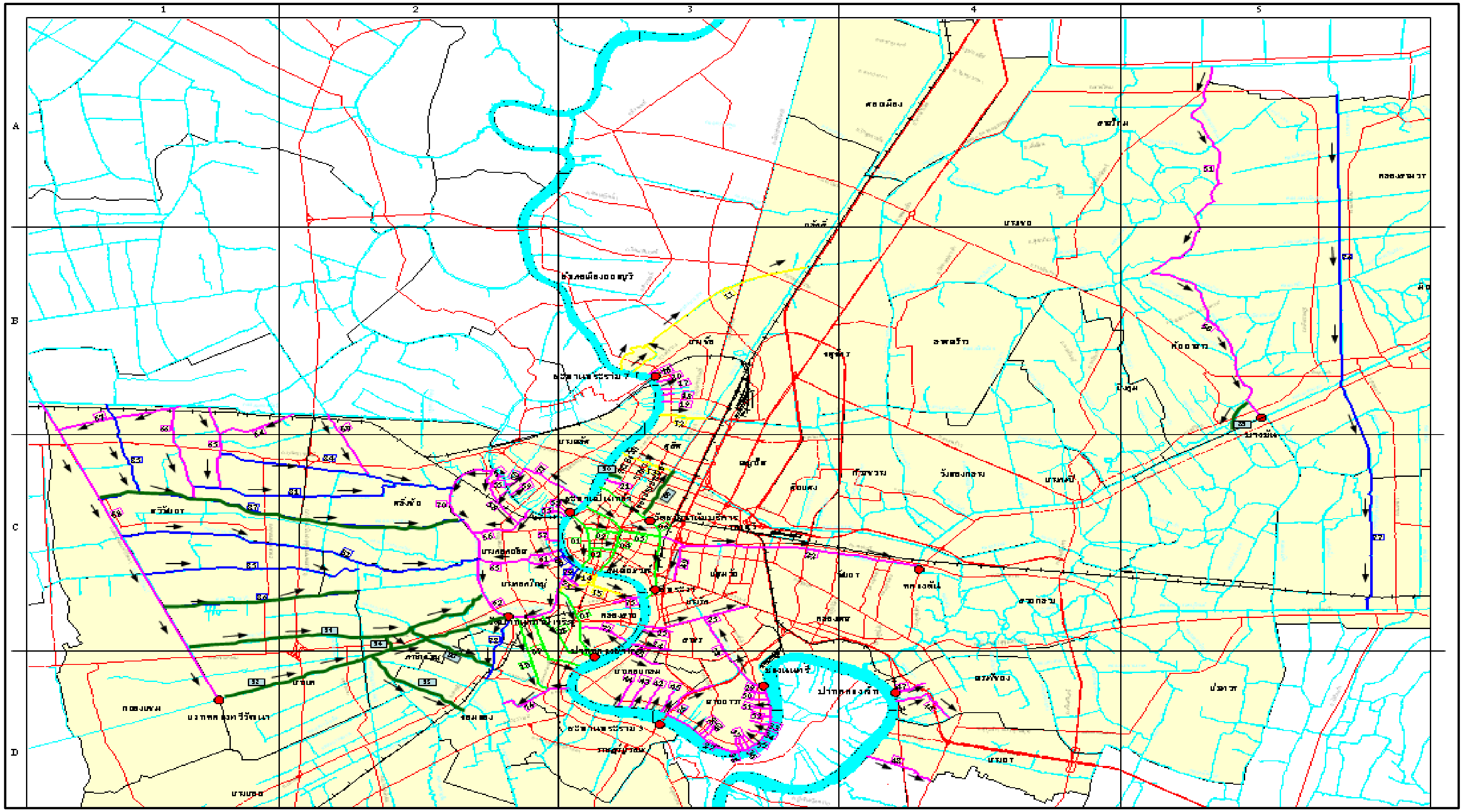


Figure 4.1: A map of canal in Bangkok

4.2 Classifying Surface Water Quality based on Surface Water Quality Standards

A period of five years water quality data of canal in Bangkok from 2003 to 2007 were used in the experiment. The main water quality indices include six parameters namely, pH, DO, BOD, NO3N, NH3N and T-Coliform. There are 11,820 samples for the analysis in water quality classification. In this work, the holdout and three-folds cross validation methods were used to evaluate the classification accuracy of the CART and MLP neural network. For holdout method, the data set was partitioned into two disjoint subsets called training and test subsets. In this work, the ratio of the train and test subsets was 60:40. This implied that, from the data records, there would be 7,092 and 4,728 records for the training and test subsets, respectively. For K-folds cross validation method, the data set was partitioned into K disjoint subsets. The K-1 subsets were chosen for training and the remaining subsets were chosen for testing. Thus, the learning procedure was executed three times on different test subsets. The advantage of K-folds cross validation method, over the simple training and testing data splitting, is repeated use of the entire available data for both building a learning machine and testing it [40]. In this work, the three-fold cross validation was used for model evaluation.

4.2.1 Surface Water Quality Standards

Many parameters can influence the surface water quality. Six parameters are selected for the data of Bangkok's canals. The surface water quality can be classified as in Table 4.2 [2]. Generally, surface water quality can be divided into five classes. Class I is extra clean fresh surface water resources used for conservation that is not necessary passed through water treatment processes. This class requires only ordinary processes for pathogenic destruction and ecosystem conservation where basic organisms can breed naturally. Class II is very clean fresh surface water resources used for consumption that requires ordinary water treatment processes before being used by aquatic organisms in conservation, fisheries, and recreation. Class III is medium clean fresh surface water resources used for consumption, but is passed through an ordinary treatment process before use. Class IV is fairly clean fresh surface water resources used for consumption, but required special water treatment processes before use. Class V is the source which is not within class I to class IV and is used for navigation.

Table 4.2: SURFACE WATER QUALITY STANDARDS.

Pollutants Index	Class				
	I	II	III	IV	V
pH (mg/l)	n	5-9	5-9	5-9	>9
DO (mg/l)	n	6	4	2	<2
BOD (mg/l)	n	1.5	2	4	>4
NO_3N (mg/l)	n	5	5	5	>5
NH_3N (mg/l)	n	0.5	0.5	0.5	>0.5
T-Coliform (MPN)/100	n	5000	20000	>20000	>20000

Remark: n is natural of water

4.2.2 Classification and Regression Tree Model

A sample decision trees that generated from the k-folds cross validation of classification and regression tree algorithm is shown in Figure 4.2. Classification and regression tree starts with parent node which is BOD. The independent parameters contain of 5 parameters, namely, pH, DO, NO_3N, NH_3N and T-Coliform. The CART procedure examines all possible independent variables and selects one that results in binary group. The obtained classification rules in, this study [41], are sequentially given as follows:

1. **If** "BOD $\geq$4.05 " **then** "Class = 5".

2. **If** "BOD < 4.05 " **and** "NH_3N $\geq$ 0.53 " **then** "Class = 5".

3. **If** “BOD < 4.05 ” **and** “NH_3N < 0.53” **and** “DO < 1.95” **then** “Class = 5”.

4. **If** “2.95 ≤ BOD < 4.05” **and** “NH_3N < 0.53” **and** “DO ≥ 1.95” **and** “T-Coliform < 9350 ” **then** “Class = 4”.

5. **If** “BOD <2.95” **and** “NH_3N < 0.53” **and** “1.95 ≤ DO < 3.05” **and** “T-Coliform < 9350 ” **then** “Class = 4”.

6. **If** “BOD <2.95” **and** “NH_3N < 0.53” **and** “DO ≥3.05” **and** “T-Coliform < 9350 ” **then** “Class = 3”.

7. **If** “BOD <2.95” **and** “NH_3N < 0.53” **and** “DO ≥ 1.95” **and** “T-Coliform ≥ 9350 ” **and** “NO_3N ≥5.375 ” **then** “Class = 5”.

8. **If** “BOD <2.95” **and** “NH_3N < 0.53” **and** “DO ≥ 1.95” **and** “ T-Coliform ≥20500 ” **and** “NO_3N <5.375 ” **then** “Class = 4”.

9. **If** “BOD <2.3” **and** “NH_3N < 0.53” **and** “DO ≥ 1.95” **and** “9350 ≤T-Coliform < 20500 ” **and** “NO_3N <5.375 ” **then** “Class = 3”.

10. **If** “2.3 ≤ BOD <4.05” **and** “NH_3N < 0.53” **and** “DO ≥ 1.95” **and** “9350 ≤T-Coliform < 20500 ” **and** “NO_3N <5.375 ” **and** “ pH<6.875 ” **then** “Class = 3”.

11. **If** “2.3 ≤ BOD <4.05” **and** “NH_3N < 0.53” **and** “DO ≥ 1.95” **and** “9350 ≤T-Coliform < 20500 ” **and** “NO_3N <5.375 ” **and** “ pH ≥ 6.875 ” **then** “Class = 3”..

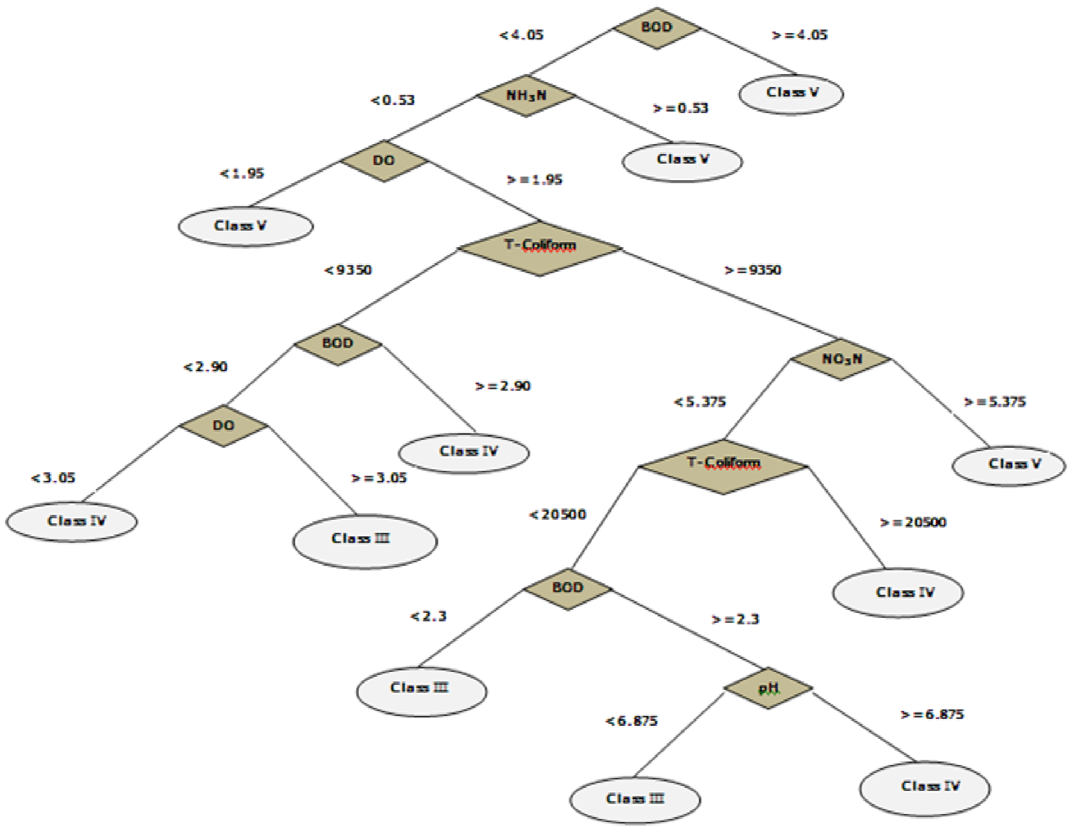


Figure 4.2: Example of Classification and Regression Tree Model

4.2.3 MLP Neural Network Model

The Levenberg-Marquardt algorithm uses input vectors and corresponding target vectors to train the neural network. All training records were fed into the network to learn the potential relationships between water quality indices and their corresponding categories. Accordingly, the 6 input layer nodes represented 6 water quality indices, while the 5 output layer nodes represented the 5 different class categories. The trained neural network determined an output representing the specific class for each water quality index. Test samples were used to verify the model classification ability through many experimental investigations. The number of hidden nodes that provided the optimal result was 4 nodes. The architecture of the network is 6-

4-5 as shown in Figure 4.3. The target mean square error (MSE) is 0.001 after 5000 iterations.

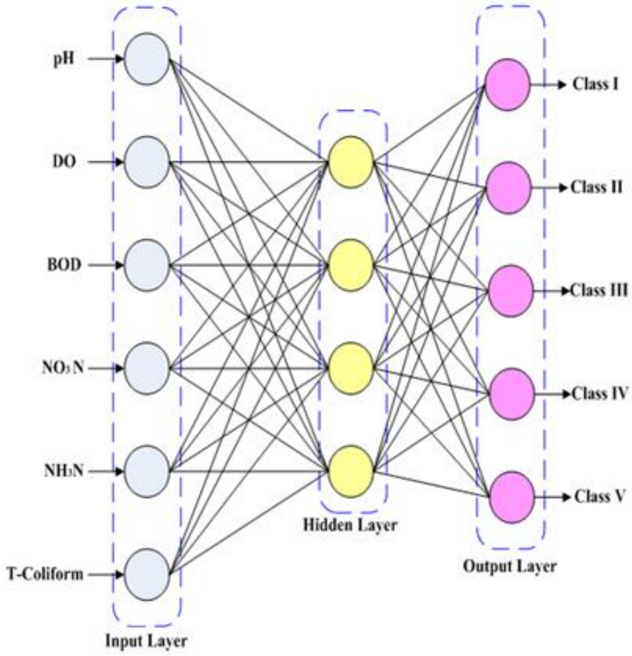


Figure 4.3: The architecture of the network 6-4-5.

The results of percentage accuracy comparison of canal water quality classification between CART and MLP neural network with holdout and three-folds cross validation are given in Table 4.3. It can be seen that classification and regression tree and multilayer perceptron neural network yielded higher performance after applied three-folds cross validation. The CART better correctly classified than multilayer perceptron neural network having the accuracy to be 99.96%. Application of the CART and MLP neural network indicates that it is robust and remarkably improves the efficiency of the classification of water pollutions which are useful for planner and watershed management nutrient loading, sedimentation and also water treatment process.

Table4.3: RESULTS OF PERCENTAGE ACCURACY IN THE TEST SET.

Method / Accuracy Percentage (%)	CART	MLP
Holdout	99.75	97.31
three-folds Cross Validation	99.96	98.82

4.3 Prediction of DO Parameter

To evaluate the merit of the proposed technique, two separated experiments of DO prediction were conducted. The first experiment was to predict without clustering. In this experiment, only one neural network was deployed to construct the predicting function. The second experiment was to the prediction with a priori clustering as discussed in Chapter 3. Details and results of each experiment are summarized in the following sections.

4.3.1 Prediction of DO without Clustering

In the first experiment, the 13 water quality parameters already collected were used to form the set of input-target pairs. The total 13,846 input-target pairs were separated into training and test subsets at the ratio 70:30. For the prediction model construction, the relevant input determination for predicting the DO in the current month t based on the collected 13 water quality parameters of the previous month $t-1$ was performed. The correlation values between each parameter in the month $t-1$, and the DO in month t were computed as shown in Table 4.4. The primary test for input determination was conducted. From Table 4.4, it is shown that the parameters SS and NH_3N provide two least values of the absolute values of the correlation. Three groupings of number of input parameters, called $NumInp$, were empirically evaluated as follows: i) $NumInp$ = 13 parameters selected as inputs, ii) $NumInp$ = 12 parameters except SS selected as inputs, and iii) $NumInp$ = 11 parameters except SS and NH_3N selected as input. For model evaluation of each group of inputs, the separation process was repeated ten times by a random selection with replacement. Then the MLP model was created for each of the ten training subsets and executed by a neural network with LM optimization. The sigmoid transfer function was used as an activation function of the hidden and output neurons. A trial and error strategy was applied, setting the number of epochs to 12,000 accordingly. The MSE measure was used for optimization, which was set to 0.001. The number of hidden neurons varied from 13 - 80 for each training subset. Since determining the optimal structure for each of 10 training subsets was time consuming only one of the 10 training subsets was used as the data set to obtain the optimal structure of MLP. The number of training subsets, test subsets of input-target pairs, and the optimal structure of MLP neural network for each type of input are shown in Table 4.5.

From Table 4.5, it is shown that $NumInp$ = 13 with 13-76-1 NN structure give the highest correlation coefficient. The architecture of the network is shown in Figure 4.4, where the 13 water quality parameters are selected as inputs of NN and considered to be relevant parameters for prediction of DO with clustering.

Table 4.4: CORRELATIONS BETWEEN THE 13 PARAMETERS IN THE PREVIOUS MONTH $t-1$ AND THE DO IN CURRENT MONTH t.

Parameters in the month $(t-1)$	Correlation with DO in the month (t)
Temp	0.191
pH	0.493
H_2S	0.425
DO	-0.169
BOD	-0.279
COD	0.073
SS	0.004
TKN	-0.179
NH_3N	0.037
NO_2N	0.259
NO_3N	-0.464
T-P	0.145
TC	-0.218

Table 4.5: NUMBER OF TRAINING AND TEST RECORDS, AND STRUCTURE OF NEURAL NETWORK WITHOUT CLUSTERING FOR EACH TYPE OF INPUT WITH CORRESPONDING CORRELATION COEFFICIENTS (R).

Number of training records	Number of test records	Structure of neural network	Correlation Coefficient (R)
9,695	4,151	13-76-1 ($NumInp = 13$)	0.62 ± 0.043
9,695	4,151	12-76-1 ($NumInp = 12$)	0.58 ± 0.041
9,695	4,151	11-74-1 ($NumInp = 11$)	0.53 ± 0.124

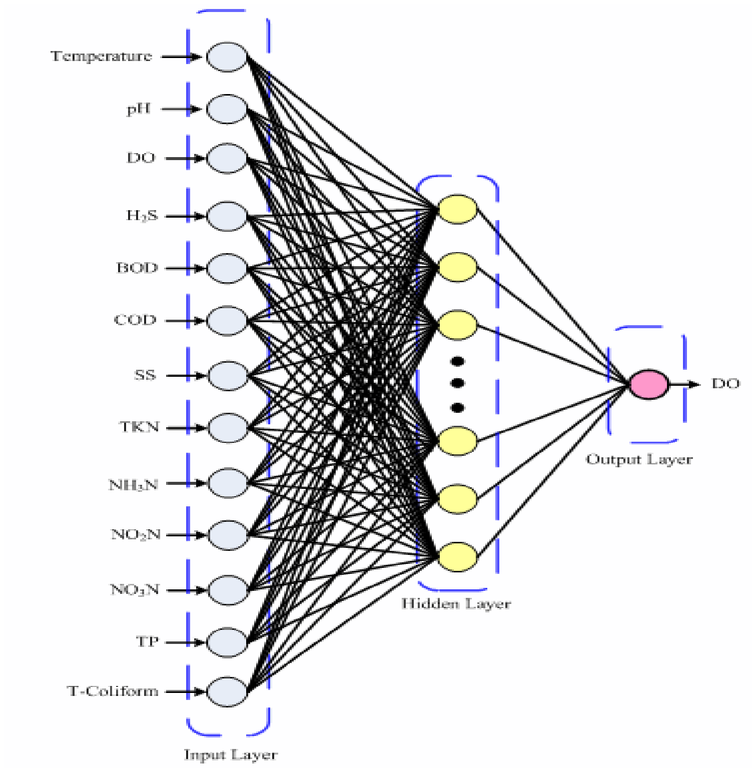


Figure 4.4: The architecture of the network 13-76-1.

4.3.2 Prediction of DO with clustering

For cluster analysis, the average of each water quality parameter on each site was computed. K-means and Fuzzy c-mean algorithms were employed to generate the K cluster sites. Then K ANN models were constructed to predict DO in K clusters based on 13 water quality parameters for each site.

4.3.2.1 Site Clustering

In this study, compactness and separation were used as measures to indicate the capability of generated centroids for representing the corresponding vectors in the clusters in both K-mean and Fuzzy c-mean algorithms. The numbers of clusters K were arbitrarily designated. As a result, different means were obtained among clusters [42]. The maximum number of initial clusters (K) was set to 6, and the CD values in equation (2.33) were used to evaluate the compactness and separation quality. The smaller the value of CD was obtained, the better compactness and separation of the clusters could be achieved. Therefore, the clustering configuration and the number of clusters could be established by choosing the result with the smallest CD value. The same CD values of K clusters by K-mean and Fuzzy c-mean algorithms are shown in Table 4.6.

Table 4.6 demonstrates the comparisons between CD values of *K*-mean and Fuzzy c-mean clustering results based on the number of clusters (K) from 2 to 6 clusters. It can be seen that, for each number of clusters (K), the *K*-mean result has a lower CD value than that of the Fuzzy c-mean result. Therefore, *K*-mean algorithm provides a better performance than Fuzzy c-mean algorithm in all the data set in this experiment. In addition, the CD value from the K-mean algorithm is the lowest when K is equal to 5, is 0.385. Thus, the number of clusters in K-mean algorithm was set to 5 for every site.

Additional information for mean comparisons of water quality parameters in five clusters is presented in Table 4.7. The mean value of water quality reflects some forms of relationship among parameters i each cluster. In Table 4.7, cluster 2 shows the lowest DO and the highest BOD. It means that the volume of oxygen contained in the water and the amount of oxygen held by the water depend on the water temperature, salinity, and pressure. The amount of DO often determines the number and type of organisms living in that body of water. This cluster denotes the worst quality of surface water. High COD also indicates bad water quality. As H_2S naturally occurs by decayed organic matters, this implies that a high value of H_2S also indicates poor water quality, having rotten egg smell. Similarly, TKN, the combination of NH_3N and organic nitrogen in the sample, are measured in milligrams per liter (mg/l). NO_3N originated from fertilizers, industrial sources (non-point sources), and municipal discharges (point sources) can contaminate the surface water run-off. TC indicates the amount of microorganisms in the water. The high TC means high microorganisms or bacteria and bad water quality. Table 4.7 shows the best water quality in cluster 4 and the worst water quality in cluster 2.

Table 4.6: THE CD VALUE WITH VARIOUS NUMBER OF CLUSTER BY *K*-MEANS AND FUZZY C-MEANS ALGORITHMS.

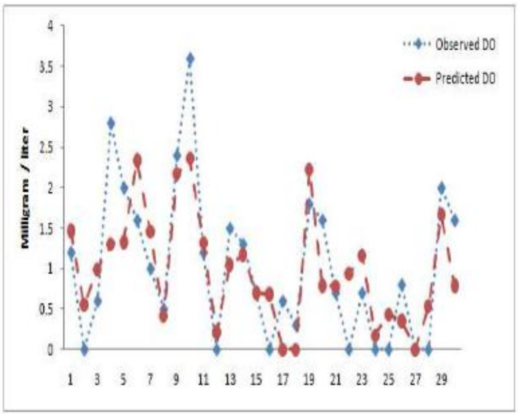
Number of Cluster (K)	*K*-means Clustering	Fuzzy c-means Clustering
K=2	0.859	1.107
K=3	0.462	0.486
K=4	0.391	0.422
K=5	**0.385**	**0.429**
K=6	0.397	0.452

Table 4.7: DESCRIPTIVE MEAN OF WATER QUALITY PARAMETERS IN FIVE CLUSTERS.

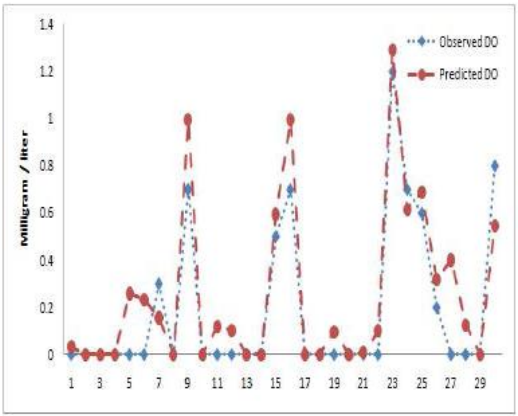
Water Quality Parameters	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Temp	29.35	29.42	29.41	29.49	30.29
pH	7.05	7.04	7.03	7.06	7.37
DO	1.50	0.28	0.73	2.42	2.41
H_2S	0.06	0.68	0.38	0.01	0.03
BOD	9.80	26.27	14.39	4.84	8.46
COD	43.37	72.61	51.52	30.70	67.57
SS	24.97	19.61	19.26	24.36	22.03
TKN	6.49	10.45	7.96	4.19	5.55
NH_3N	1.85	3.89	2.54	0.75	2.26
NO_2N	0.13	0.06	0.09	0.12	0.25
NO_3N	1.95	2.15	2.10	1.89	0.85
T-P	0.73	1.35	0.97	0.30	1.31
T.Coliform	1.06E+08	6.97E+08	2.56E+08	1.76E+07	4.85E+06
Number of sites	77	39	48	49	63

4.3.2.2 DO prediction by K ANN models

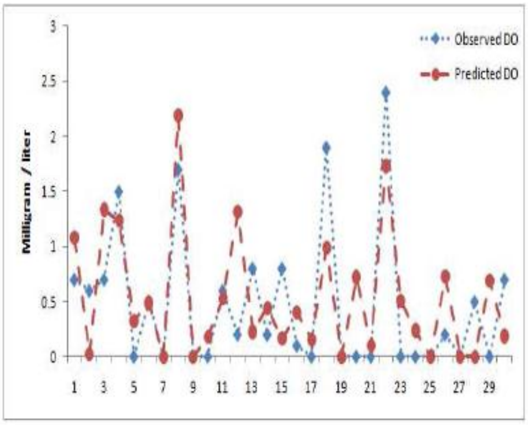
From the previous section, five-site clusters were established and data in each site cluster was independently learned by each MLP network. For model evaluation, the same training and testing subsets discussed in Section 4.3.1 were used to create the model and to evaluate the performance. The same type of activation function in the hidden and output layers as discussed in Section 4.3.1 was adapted in this experiment. The training process was terminated by using the mean square error measure and the number of epochs. The values of these stopping criteria are shown in Table 4.8.



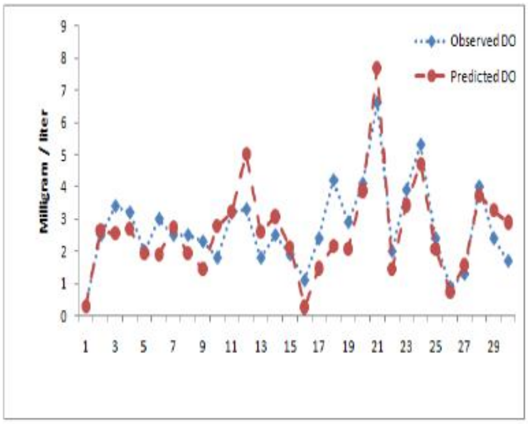
(a) Cluster 1



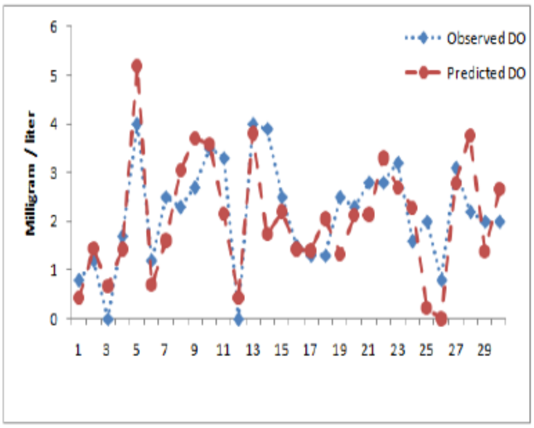
(b) Cluster 2



(c) Cluster 3



(d) Cluster 4



(e) Cluster 5

Figure 4.5: Comparison between predicted DO and observed DO in five clusters.

The same selected training set as in Section 4.3.1 was used to determine the appropriate number of hidden neurons in five MLP models. These numbers are shown in Table 4.9. The number of hidden neurons varied from 13 to 80. After the optimal number of hidden neurons in each MLP network was obtained, each MLP was trained by using the remaining training data with 10 times of weight initialization. Figures 4.5a-4.5e illustrated the comparison of predicted value and observed value of DO with 30 records in each cluster.

Table 4.8: PARAMETER SETTING OF GOAL OF MINIMIZING MEAN SQUARE ERROR AND THE NUMBER OF EPOCHS OF FIVE CLUSTERS.

Cluster	Goal of minimizing mean square error	Number of epochs
1	0.001	10,000
2	0.005	8,000
3	0.003	8,000
4	0.003	8,000
5	0.001	10,000

Table 4.9: NUMBER OF TRAINING AND TEST RECORDS AND STRUCTURE OF NEURAL NETWORK WITH FIVE CLUSTERS.

Cluster	Number of Training records	Number of test records	Structure of neural network
1	2,744	1,175	13-72-1
2	1,347	576	13-25-1
3	1,706	731	13-38-1
4	1,702	729	13-46-1
5	2,196	940	13-69-1

4.3.3 Comparison between Two Approaches

The results of DO prediction with clustering technique and DO prediction without clustering technique on ten data sets were compared by using the average standard deviation of three measures previously mentioned. DO prediction with clustering technique provides better average prediction values than those of DO prediction without clustering technique in all three measures as shown in Table 4.10. Furthermore, the standard deviation values of three measures from the prediction with clustering technique are very small when being compared with the results from the prediction without clustering technique. This implies that the DO prediction by ANN with clustering technique is more stable than DO prediction by ANN without clustering technique.

Table 4.10: THE AVERAGE PERFORMANCE VALUE WITH STANDARD DEVIATION OF THREE MEASUREMENTS BETWEEN DO PREDICTION WITHOUT CLUSTERING TECHNIQUE AND DO PREDICTION WITH CLUSTERING ON TEN DATA SETS.

Type of Prediction	R	MAE	MSE
DO prediction without clustering	0.62 ± 0.043	0.88 ± 0.078	1.58 ± 0.363
DO prediction with clustering	**0.83 ± 0.002**	**0.63 ± 0.010**	**0.75 ± 0.022**

It can be seen that the clustering technique integrated with ANN improved the DO prediction significantly. Since a large number of canals in Bangkok obtained from various distinctive areas constituted high complexity of the data physically, only ANN technique could not deal with this magnitude of complexity. Therefore, the clustering technique was applied to manage the complexity due to the very large number of canals by clustering the data with the same characteristic based on compactness and degree of separation to the same group. Then, ANNs were applied to construct the prediction model for each cluster. Moreover, the cluster of sites obtained empirically from clustering technique conformed to the geography of the studied area as shows in Appendix B.

CHAPTER V

CONCLUSION

5.1 Conclusion

This dissertation proposes an automated methodology that can quickly and efficiently classify the water quality of canals in Bangkok based on surface water quality standard. The proposed method also, achieves highly accurate prediction of the amount of Dissolved Oxygen (DO) in Bangkok's canals. In addition, a new technique is proposed to enhance the prediction accuracy by transforming the prediction problem into the problem of constructing a set of sub-manifolds of predicting function by deploying unsupervised and supervised neural networks. The data set used in this study were monthly water quality data obtained from the Department of Drainage and Sewerage, Bangkok Metropolitan Administration, during the years 2003-2011. There were 276 sites covering 155 canals. The parameters were collected and used in the experiments consist of 13 parameters, namely temperature, pH value (pH), Hydrogen Sulfide (H2S), DO, BOD, Chemical Oxygen Demand (COD) Suspended Solids (SS), Total Kjeldahl Nitrogen (TKN), Ammonia Nitrogen (NH_3N), Nitrite Nitrogen (NO_2N), Nitrate Nitrogen (NO_3N), Total Phosphorous (T-P) and Total Coliform (TC).

Two experiments were conducted. The first experiment applied 6 parameters, namely, pH, DO, BOD, NH_3N, NO_3N and TC to classify based on parameters in surface water quality standard. A K-folds cross validation was applied to classification and regression tree (CART) and multilayer perceptron (MLP) neural network using the Levenberg-Marquardt algorithm. The results indicate that the CART performs with high accuracy classification percentage of 99.96%, while the MLP neural network shows percent accuracy of 98.82. These encouraging results may be applied to automate water quality classifications. The second experiment used data collected from

Bangkok canals during the years 2007-2011, The proposed procedure for predicts the amount of DO by employing a neural functional approximation based on 13 water quality parameters. The entire data set was partitioned into several connected sub-spaces using clustering techniques, and the data in each sub-space were used to construct a local manifold in the form of a neural network. Prediction error of the clustering technique is less than half of the error obtained from the prediction without clustering. Because the applicability of these results are not limited to the Bangkok area, such a generation of cluster analysis and neural functional approximation becomes a useful and efficient tool for managing natural resources and maintaining compliance with water management regulations and policies.

5.2 Future Work

- The water quality classification model can be further enhanced with the help of Geographic Information System (GIS) to illustrate different areas of surface water quality, efficiently used in wider water quality management project.

- Integration of unsupervised and supervised neural network model of prediction DO parameters can predict other water quality parameters.

- The results obtained from neural network model should be compared with other models.

REFERENCES

[1] Department of Drainage and Sewerage Bangkok Metropolitan Administration, water quality management [online]. 2008. Available from: http://dds.bangkok.go.th/wqm/thai/home.html [2008, July 18]

[2] Department of Pollution Control Bangkok, water pollution [online]. 2009. Available from: http://www.pcd.go.th [2009, April 12]

[3] Streeter, H.W and Phelps, E.B. A study of the pollution and natural purification of the Ohio river III. United states public health service, No.146, 1925.

[4] H.-Y. Wang. Assessment and prediction of overall environmental quality of Zhuzhou City, Hunan Province, China. Journal of Environmental, (2002) : 329-340.

[5] D.-P. Hamilton, and S.-G. Schladow, "Prediction of water quality in lakes and reservoirs. Part 1 - Model description," Ecological Modeling, vol.96 (1997) : 91-110.

[6] J. Fang, L.-K Dai, Rapid Detection of Chemical Oxygen Demand Using Least Square Support Vector Machines, in Proceedings of the 5th World Congress on Intelligent Control and Automation, Hangzhou. P.R. China, pp.3810-3813, June, 2004.

[7] Y.-S. Park, et al, Patterning and predicting aquatic macro invertebrate diversities using artificial neural network, Water Research, vol.37(2003) : 1749-1758.

[8] L.-Y. Khuan, N. Hamzah, and R. Jailani, "Prediction of water quality index (WQI) based on artificial neural networks (ANN)," in Shah Alam, 2002 Student Conference on Research and Development Proceeding, Malaysia, pp.157-161, 2002.

[9] Fernández, E. and Gálvis, A. "Artificial Neural Networks Model Used for Clear Water Treatment Plans" Seminario Internacional La Hidroinformática en la Gestión Integrada de los Recursos Hídricos, Columbia, 2000.

[10] Li-hua Chen and Xiao-yun Zhang, "Application of Artificial Neural Networks to Classify Water Quality of the Yellow River" Fuzzy Info and Engineering, ASC 54, Springer-Verlag Berlin Heidelberg, China, pp.15–23, 2009.

[11] Butcher, J., and Covington, S. Dissolved-Oxygen analysis with temperature dependence. Journal of Environment. 121,10 (1995) : 756-759.

[12] Cox, A. A review of dissolved oxygen modeling techniques for lowland rivers. The Science of the Total Environment, pp.314-316, 2003.

[13] Garcia, A., Revilla, A., Medina, R., and Alvarez, C., Juanes, A. A model of predicting the temporal evolution of dissolved oxygen concentration in shallow estuaries. Hydrobiologia., pp.205-211, 2002.

[14] Palani, S., Shie-Yui, S., Tkalich, P., and Palanichamy, J. Development of a neural network model for dissolved oxygen in sea water. Indian Journal of Marine Sciences. 38(2009) : 151-159.

[15] Rankovic, V., Radulovic, J., Radojevic, I., and Ostojic, A., Comic, L. Neural network modeling of dissolved oxygen in the Gruza reservoir. Ecol. Model. 221(2010) : 1239-1244.

[16] Palani, S., Liong, S.Y., and Tkalich, P. An ANN application for water quality forecasting. Marine Pollut. Bull. 56, pp.1586-1597, 2008.

[17] Soyupak, S., Karaer, F., Gurbuz, H., Kivrak, E., Senturk, E., and Yazici, A. A neural network-based approach for calculating dissolved oxygen profiles in reservoirs.Neural Comput. Appl. 12, pp.166-172, 2003.

[18] Sengorur, B., Dogan, E., Koklu, R., and Samandar, A. Dissolved oxygen estimation using artificial neural network for water quality control. Fresen. Environ. Bull.15 (9a), pp.1064-1067, 2006.

[19] Kuo, J.T., Hsieh, M.H., Lung, W.S., and She, N. Using artificial neural network for reservoir eutrophication prediction. Ecol. Model. 200, pp.171-177, 2007.

[20] Singh, K.P., Basant, A., Malik, A., and Jain, G. Artificial neural network modeling of the river water quality a case study. Ecol. Model, 220 (2009) : 888-895.

[21] Rankovic, V., Radulovic, J., Radojevic, I., Ostojic, A., and Comic, L. Neural network modeling of dissolved oxygen in the Gruza reservoir. Ecol. Model. 221(2010) : 1239-1244.

[22] Thomann, R.V. and Mueller, J.A. Principles of surface water quality modelling and control. New York, Harper and Row, 1987.

[23] Hagan, M., Demuth, H., and Beale, M. Neural Network Design, Boston, MA: PWS Publishing, 1996.

[24] Haykin, S. Neural network, a comprehensive foundation. 2nd ed. New York: Macmillan College Publishing Company, 1994.

[25] Jain, A.K., Mao, J. and Mohiuddin, K.M. Artificial neural networks: A tutorial. IEEE, 1996.

[26] Furundzic, D. Application example of neural network for time series analysis rainfall-runoff modeling. Signal Processing 64: pp.383-396, 1998.

[27] Maier, H.R. and Dandy, G.C. Determining inputs for neural network models of multivariate time series. Microcomputers in Civil Engineering 12 (5): pp.353-368, 1997.

[28] Rajurkar, M.P., Kothyari, U.C. and Chaube, U.C. Modeling of the daily rainfall-ruuoff relationship with artificial neural network. Journal of Hydrology 285 (2004) : 96-113.

[29] Sajikumar, N., and Thandaveswara, B.S. A non-linear rainfall-runoff model using an artificial neural network. Journal of Hydrology 216(1999) : 32-35.

[30] Warnerm, B. and Misra, M. Understanding neural networks as statistical tools. American Statistician 50 (4): pp.284-293, 1996.

[31] Hagan, M.T. and Menhaj, M.B. Training feedforward networks with the Marquardt algorithm. IEEE Transactions on Neural Network 5(6): pp.989-993, 1994.

[32] Singh, V., Gupta, I. and Gupta, H.O. ANN-based estimator for distillation using Levenberg-Marquardt approach. Engineering Applications of Artificial Intelligence, 2006.

[33] Breiman L., Friedman J., Olshen R., and Stone C.. Classification and Regression Trees. Wadsworth Int. Group, 1984.

[34] MacQueen, J., Foster, I., and Kesselman, C. Some Methods for classification and Analysis of Multivariate Observations: Procee C., The Grid:Blueprint for a New Computing Infrastructure. In: Proc. Fifth Berkeley Symp on Math. Statist and Prob.1, pp.281-297, 1967.

[35] Musavi, S., and Golabi, M. Application of Artificial Neural Networks in the River Water Quality Modeling:Karoon River,Iran, Journal of Applied Sciences, Asian Network for Scientific Information, pp.2324-2328, 2008.

[36] Bezdek, J. Fuzzy partition and relations: an axiomatic basic for clustering. Fuzzy Sets and Systems 1, pp.111-127, 1978.

[37] Halkidi, M., and Vazirgiannis, M. Cluster validity assessment: finding the optimal partitioning of a dataset. In: Proc. ICDM Conference. San Jose,Californaia, USA, pp.187-194, 1996.

[38] Binbib, H., Fang, T., and Guo, D. Quality assessment and uncertainty handling in spatial data mining. In: Proc. 12th Conference on Geoinformations, Sweden, 2004.

[39] Halkidi, M., Bastistakis, Y., and Vazirgiannis, M. On Clustering Validation Techniques, Journal of Intelligent Information Systems, 17, pp.107-145, 2001.

[40] Anguita, D., Ridella, S., and Rivieccio, F. "K-folds Generalization Capability Assessment for Support Vector Classifiers", Proceeding of International Joint Conference on Neural Network, Canada, pp.855-858, 2005.

[41] S.Areerachakul and S.Sanguansintukul "Classification and Regression tree and MLP Neural Network to Classify Water Quality of Canals in Bangkok, Thailand", International Journal of Intelligent Computing Research (IJICR), Volume 1, Issue 1/2 (2010) : 43-50.

[42] Kanungo, T., Mount, D., Netanyahu, N., Piatko, D., and Silverman, R. An Efficient Kmeans Clustering Algorithm Analysis and Implementation. IEEE Transaction on Pattern Analysis and Machine Intelligence. 24(7), 2002.

APPENDICES

Appendix A

Sample of Water Quality Data

Source: Department of Drainage and Sewerage Bangkok Metropolitan Administration, dated January 1, B.E.2552 (2009).

Date	No.	Color	temp	pH	DO	H2S	BOD	COD	SS	TKN	NH3N	NO2	NO3	TP	T.Coliform
12 ม.ค. 52	12	yellow green	26	6.8	0.4	0.0	8	37	16	7.3	0.0	0.04	1.3	0.8	2.4E+06
12 ม.ค. 52	13	yellow green	26	6.9	0.4	0.0	8	32	14	6.0	0.3	0.06	2.1	0.6	3.9E+05
12 ม.ค. 52	14	yellow green	26	6.9	0.6	0.0	7	33	17	7.8	0.0	0.07	1.2	0.7	2.4E+06
12 ม.ค. 52	15	brown	26	7.0	3.0	0.0	4	22	22	3.4	0.0	0.07	1.6	0.1	3.0E+04
12 ม.ค. 52	21	gray	26	6.8	0.0	0.1	27	71	18	10.6	0.3	0.05	1.3	1.3	1.1E+08
12 ม.ค. 52	31	gray	26	6.9	0.0	0.1	19	45	13	9.5	0.0	0.11	1.2	1.0	2.4E+07
12 ม.ค. 52	41	yellow green	26	7.0	1.5	0.0	4	26	14	7.3	0.0	0.09	3.3	0.4	2.8E+05
12 ม.ค. 52	42	yellow green	26	7.1	3.0	0.0	4	16	24	4.5	0.0	0.12	2.1	0.2	2.4E+07
12 ม.ค. 52	43	brown	26	7.1	2.8	0.0	4	11	14	1.7	0.0	0.01	1.9	0.1	4.0E+04
12 ม.ค. 52	51	brown	26	7.1	4.0	0.0	4	16	14	7.3	0.0	0.09	1.8	0.1	3.0E+04
12 ม.ค. 52	52	gray	26	6.8	0.0	0.2	17	25	10	9.0	0.6	0.08	1.3	0.8	2.4E+07
12 ม.ค. 52	53	gray	26	6.9	0.0	0.2	17	40	13	9.5	0.0	0.07	1.3	1.0	2.4E+07
8 ม.ค. 52	61	black	30	6.7	0.0	0.1	13	28	6	14.6	0.6	0.10	1.8	1.0	2.4E+07
8 ม.ค. 52	62	black	30	6.9	1.2	0.0	11	19	5	6.7	0.3	0.19	2.1	0.8	1.1E+07
8 ม.ค. 52	63	black	30	7.0	3.1	0.0	3	36	6	1.1	0.0	0.09	2.1	0.1	2.8E+04
8 ม.ค. 52	72	natural	30	7.0	2.0	0.0	5	15	9	0.6	0.0	0.18	2.8	0.1	1.1E+07
8 ม.ค. 52	73	brown	30	6.9	2.0	0.0	6	23	9	2.2	0.0	0.04	1.8	0.3	2.4E+06
8 ม.ค. 52	74	brown	30	7.0	2.6	0.0	4	10	20	2.2	0.0	0.03	1.8	0.1	7.0E+04
8 ม.ค. 52	75	natural	30	7.0	2.2	0.0	4	19	14	3.9	0.0	0.04	1.9	0.1	1.1E+05
8 ม.ค. 52	76	natural	30	7.0	3.3	0.0	4	18	13	2.8	0.0	0.05	1.8	0.3	3.0E+04
5 ม.ค. 52	81	yellow green	28	6.8	1.8	0.0	2	8	11	9.5	0.6	0.03	2.2	0.1	7.5E+05
5 ม.ค. 52	82	yellow green	28	6.9	1.5	0.0	2	16	5	6.7	0.8	0.06	2.4	0.2	1.1E+07
5 ม.ค. 52	83	yellow green	28	7.0	3.3	0.0	5	20	17	4.5	0.0	0.11	2.2	0.2	2.1E+06
5 ม.ค. 52	85	green	28	7.0	0.4	0.0	5	21	2	9.5	0.8	0.16	3.8	0.3	1.1E+07
9 ม.ค. 52	90	brown	29	7.2	4.6	0.0	3	15	20	7.8	0.0	0.07	1.3	0.0	3.0E+04
6 ม.ค. 52	91	yellow green	28	7.1	2.7	0.0	6	48	3	2.8	0.0	0.60	6.7	1.0	1.5E+05
6 ม.ค. 52	92	yellow green	28	7.1	1.6	0.0	7	37	4	6.7	0.0	0.53	5.8	1.0	2.4E+06
6 ม.ค. 52	93	gray	28	7.2	0.0	0.1	16	62	61	6.7	0.0	0.63	4.6	1.1	1.1E+07
6 ม.ค. 52	94	yellow green	28	7.1	1.9	0.0	7	46	2	2.8	0.0	0.54	3.1	0.3	4.6E+06
9 ม.ค. 52	95	yellow green	29	7.0	1.2	0.0	5	44	9	6.2	0.0	0.19	1.7	0.3	1.1E+07
9 ม.ค. 52	96	yellow green	29	7.1	2.4	0.0	3	37	11	6.7	0.0	0.21	1.7	0.2	3.0E+04
9 ม.ค. 52	97	yellow green	29	7.1	1.8	0.0	3	34	6	3.9	0.0	0.30	1.8	0.2	3.0E+04
6 ม.ค. 52	98	yellow green	28	7.1	2.1	0.0	7	51	2	3.9	0.0	0.68	3.6	0.3	1.5E+06
9 ม.ค. 52	99	brown	29	7.1	3.4	0.0	4	19	21	6.2	0.6	0.11	1.2	0.0	3.0E+04
9 ม.ค. 52	99.1	brown	29	7.3	4.5	0.0	3	16	41	5.0	0.0	0.10	1.2	0.0	7.0E+04
5 ม.ค. 52	101	gray	28	7.2	1.0	0.0	10	50	13	7.8	0.6	0.42	3.6	1.0	1.2E+06
12 ม.ค. 52	112	yellow green	26	7.0	0.7	0.0	5	23	9	14.0	0.3	0.07	1.1	1.0	7.5E+05
12 ม.ค. 52	113	yellow green	26	7.1	0.9	0.0	3	24	6	12.9	0.0	0.10	1.3	0.6	3.0E+04
12 ม.ค. 52	114	yellow green	26	7.1	0.0	0.3	11	27	14	10.1	0.8	0.00	1.6	0.8	2.4E+08

12 ม.ค. 52	115	yellow green	26	7.1	0.0	0.3	11	33	9	11.2	0.6	0.11	1.3	0.9	1.1E+08
12 ม.ค. 52	116	yellow green	26	7.0	1.3	0.0	3	21	6	4.5	0.3	0.06	1.3	0.3	9.0E+04
12 ม.ค. 52	117	yellow green	26	7.1	2.6	0.0	3	19	18	2.2	0.6	0.09	1.5	0.1	4.0E+04
12 ม.ค. 52	118	black	26	7.0	0.0	0.3	12	34	12	11.8	0.3	0.12	1.5	0.9	2.4E+07
12 ม.ค. 52	119	black	26	7.0	0.0	0.3	11	36	11	9.0	0.6	0.14	1.6	0.9	2.4E+08
5 ม.ค. 52	121	yellow green	28	7.1	2.7	0.0	3	15	14	6.2	0.0	0.23	1.9	0.1	7.0E+04
5 ม.ค. 52	122	yellow green	28	7.0	1.0	0.0	3	21	5	5.6	0.0	0.12	4.0	0.2	1.2E+06
5 ม.ค. 52	123	yellow green	28	7.0	1.1	0.0	3	16	2	6.7	0.0	0.12	2.9	0.2	2.3E+04
14 ม.ค. 52	131	gray	24	7.1	2.8	0.0	13	22	2	6.7	0.0	0.01	2.9	0.4	1.1E+07
14 ม.ค. 52	132	gray	24	7.1	0.0	0.2	24	31	5	7.8	0.6	0.12	2.1	0.8	2.4E+08
14 ม.ค. 52	133	gray	24	7.1	0.0	0.6	21	29	6	5.0	0.3	0.14	2.8	0.8	2.4E+08
14 ม.ค. 52	141	green	24	6.9	0.0	0.5	29	44	6	4.5	0.0	0.16	1.7	0.8	4.3E+07
14 ม.ค. 52	142	green	24	7.1	1.2	0.0	12	22	12	3.9	0.0	0.17	2.0	0.5	1.2E+06
14 ม.ค. 52	143	green	24	7.1	0.0	0.3	17	21	4	5.0	0.3	0.19	2.0	0.7	2.1E+08
14 ม.ค. 52	144	gray	24	6.9	0.0	0.2	31	53	21	5.6	0.0	0.02	1.7	0.9	1.5E+07
14 ม.ค. 52	145	gray	24	7.0	0.0	0.2	42	58	9	5.0	0.3	0.02	1.6	1.1	9.3E+06
14 ม.ค. 52	146	green	24	7.1	1.0	0.0	8	18	5	4.5	0.0	0.14	2.4	0.4	1.1E+07
22 ม.ค. 52	151	gray	28	6.8	0.0	0.3	39	30	16	12.3	0.0	0.01	1.5	1.4	1.1E+09
22 ม.ค. 52	152	gray	28	6.9	0.0	0.5	29	34	11	2.8	0.0	0.01	1.2	1.3	4.6E+08
22 ม.ค. 52	161	gray	28	6.9	1.0	0.0	11	19	12	4.5	0.0	0.01	1.0	0.6	2.4E+06
21 ม.ค. 52	162	natural	26	7.1	2.1	0.0	3	12	5	10.6	0.0	0.02	1.8	0.2	2.4E+06
13 ม.ค. 52	171	green	26	7.1	2.8	0.0	11	37	13	3.9	0.6	0.03	1.8	0.9	2.1E+06
13 ม.ค. 52	173	green	26	7.1	1.1	0.0	11	58	7	12.3	0.6	0.04	1.9	0.9	2.4E+06
13 ม.ค. 52	174	gray	26	6.8	0.0	0.4	28	76	8	12.3	0.6	0.10	3.1	1.5	4.6E+08
13 ม.ค. 52	175	gray	26	7.0	0.0	7.0	12	36	6	10.6	0.0	0.11	3.8	1.4	2.0E+07
13 ม.ค. 52	181	green	26	7.1	1.6	0.0	11	54	9	9.5	0.0	0.08	3.0	0.9	1.1E+07
13 ม.ค. 52	182	green	26	7.2	1.1	0.0	10	51	8	9.5	0.0	0.12	3.9	1.0	2.4E+07
13 ม.ค. 52	183	gray	26	6.9	0.0	3.0	21	53	7	17.4	0.8	0.14	2.9	1.5	2.1E+08
13 ม.ค. 52	184	gray	26	7.0	0.0	3.0	19	59	10	9.0	0.0	0.13	2.4	1.5	2.4E+08
22 ม.ค. 52	191	yellow green	28	7.1	2.2	0.0	4	10	8	3.9	0.0	0.03	1.9	0.2	1.4E+05
22 ม.ค. 52	192	yellow green	28	7.0	1.9	0.0	4	12	23	10.6	0.0	0.01	1.3	0.2	2.0E+05
22 ม.ค. 52	201	yellow green	28	7.2	1.0	0.0	3	28	20	6.7	0.0	0.01	1.2	0.2	3.0E+04
22 ม.ค. 52	202	yellow green	28	7.2	2.2	0.0	3	27	11	4.5	0.0	0.01	1.0	0.2	3.0E+04
22 ม.ค. 52	211	green	28	7.1	0.8	0.0	3	13	5	11.2	0.8	0.05	1.8	0.2	3.0E+04
22 ม.ค. 52	212	green	28	7.2	1.4	0.0	3	18	8	6.2	0.0	0.02	1.5	0.1	7.0E+04
19 ม.ค. 52	221	natural	26	6.9	1.4	0.0	3	28	8	5.0	0.0	0.08	2.0	0.4	4.0E+04
19 ม.ค. 52	222	natural	26	6.9	1.4	0.0	3	19	6	10.6	0.0	0.09	2.1	0.4	7.0E+04
15 ม.ค. 52	231	natural	26	7.2	3.7	0.0	3	15	22	3.4	0.0	0.01	1.6	0.1	3.0E+04
15 ม.ค. 52	232	natural	26	7.2	3.0	0.0	3	16	17	5.0	0.3	0.01	1.7	0.1	3.0E+04
21 ม.ค. 52	241	green	26	6.9	2.0	0.0	3	11	12	11.8	0.8	0.05	1.9	0.1	6.4E+05

19 ม.ค. 52	251	natural	26	7.1	2.3	0.0	3	14	21	3.4	0.0	0.02	2.3	0.1	3.9E+05
19 ม.ค. 52	252	natural	26	7.0	2.5	0.0	3	17	33	8.4	0.0	0.04	2.1	0.2	1.5E+06
19 ม.ค. 52	261	natural	26	7.1	2.5	0.0	3	13	17	3.9	0.6	0.01	2.0	0.1	3.0E+04
19 ม.ค. 52	262	natural	26	7.1	2.6	0.0	3	13	17	3.9	0.0	0.01	2.0	0.1	3.0E+04
15 ม.ค. 52	263	natural	26	7.2	2.2	0.0	4	25	17	3.4	0.0	0.01	2.2	0.1	7.0E+04
15 ม.ค. 52	264	natural	26	7.2	2.6	0.0	3	22	7	3.4	0.0	0.01	2.1	0.1	4.0E+04
15 ม.ค. 52	265	natural	26	7.3	2.4	0.0	3	22	13	4.5	0.0	0.01	2.2	0.1	3.0E+04
5 ม.ค. 52	271	green	28	7.2	1.1	0.0	9	44	2	6.2	0.0	0.19	1.9	1.0	2.1E+06
5 ม.ค. 52	272	green	28	7.2	1.0	0.0	9	41	9	4.5	0.0	0.40	4.9	1.0	2.1E+06
5 ม.ค. 52	281	gray	28	7.3	0.0	3.0	20	65	9	14.6	0.8	0.05	1.8	1.3	2.1E+08
5 ม.ค. 52	282	gray	28	7.4	0.0	0.2	13	51	14	10.1	1.4	0.63	3.8	1.0	2.0E+07
22 ม.ค. 52	291	gray	28	7.0	0.0	0.3	29	32	14	3.9	0.0	0.02	1.3	1.1	1.1E+08
22 ม.ค. 52	292	black	28	7.0	0.0	1.3	28	31	13	12.3	0.0	0.02	1.4	1.3	2.4E+08
14 ม.ค. 52	301	green	26	6.9	0.0	0.2	58	75	13	10.6	0.6	0.01	2.0	1.2	2.4E+08
14 ม.ค. 52	303	green	26	6.9	0.0	0.1	19	28	22	5.0	0.0	0.09	2.5	0.8	2.4E+08
20 ม.ค. 52	311	green	26	6.8	1.7	0.0	2	5	4	2.8	0.0	0.04	1.9	0.2	3.0E+04
20 ม.ค. 52	312	green	26	6.9	1.5	0.0	2	35	4	5.6	0.6	0.06	2.0	0.2	4.0E+04
20 ม.ค. 52	321	natural	26	7.0	2.2	0.0	3	9	5	2.8	0.0	0.01	1.8	0.1	3.0E+04
20 ม.ค. 52	322	natural	26	7.1	2.7	0.0	3	8	5	7.3	0.3	0.04	1.9	0.1	3.0E+04
20 ม.ค. 52	331	gray	26	6.9	0.0	0.3	15	21	4	6.7	0.0	0.02	2.0	0.8	1.1E+08
20 ม.ค. 52	332	gray	26	6.9	0.0	0.1	9	15	3	10.6	0.6	0.02	2.1	0.6	4.6E+07
20 ม.ค. 52	341	green	26	7.0	1.0	0.0	2	13	4	4.5	0.6	0.01	2.0	0.4	7.0E+04
20 ม.ค. 52	342	green	26	7.0	0.8	0.0	3	8	2	3.4	0.0	0.01	1.9	0.4	4.0E+04
14 ม.ค. 52	351	green	26	7.0	0.0	3.0	45	57	7	10.1	0.0	0.01	1.9	1.1	2.4E+08
5 ม.ค. 52	371	green	28	7.4	4.2	0.0	10	107	12	8.4	0.6	0.02	1.3	1.0	7.5E+05
5 ม.ค. 52	381	black	28	7.2	3.7	0.0	3	14	3	5.0	0.0	0.01	3.8	0.0	2.3E+04
13 ม.ค. 52	382	natural	26	7.2	4.3	0.0	6	14	16	2.8	0.0	0.01	1.7	0.1	3.0E+04
19 ม.ค. 52	391	natural	26	7.1	2.7	0.0	4	16	23	7.3	0.6	0.01	1.5	0.1	3.0E+04
19 ม.ค. 52	392	natural	26	7.0	1.0	0.0	11	28	9	1.7	0.6	0.01	1.6	0.9	2.4E+06
19 ม.ค. 52	393	natural	26	7.1	1.8	0.0	3	17	13	4.5	0.6	0.02	1.6	0.2	3.0E+06
19 ม.ค. 52	394	natural	26	7.1	2.9	0.0	3	17	21	6.7	0.0	0.01	1.6	0.1	3.0E+04
19 ม.ค. 52	395	natural	26	7.1	2.5	0.0	3	14	18	3.9	0.0	0.11	2.1	0.1	4.0E+04
15 ม.ค. 52	401	natural	26	7.2	2.0	0.0	3	19	14	2.2	0.0	0.01	2.3	0.1	3.0E+04
15 ม.ค. 52	402	natural	26	7.3	3.0	0.0	3	17	24	3.4	0.0	0.01	1.7	0.1	3.0E+04
20 ม.ค. 52	411	natural	26	7.1	2.7	0.0	4	16	18	2.2	0.6	0.01	2.0	0.2	2.3E+05
20 ม.ค. 52	421	natural	26	7.1	2.1	0.0	4	7	16	11.2	0.6	0.01	2.1	0.1	4.0E+04
20 ม.ค. 52	422	natural	26	7.1	2.3	0.0	3	18	21	6.2	0.0	0.01	1.9	0.1	7.0E+04
20 ม.ค. 52	423	natural	26	7.1	1.9	0.0	3	12	32	3.9	0.0	0.06	2.1	0.1	3.0E+04
21 ม.ค. 52	431	yellow green	28	7.0	1.7	0.0	2	42	4	10.6	0.6	0.01	2.5	0.2	4.0E+04
21 ม.ค. 52	432	yellow green	28	7.1	1.5	0.0	2	1638	6	1.7	0.0	0.06	2.5	0.2	3.0E+04

21 มี.ค. 52	433	yellow green	28	7.0	0.7	0.0	8	34	8	9.0	0.0	0.10	2.6	0.7	1.1E+07
21 มี.ค. 52	434	yellow green	28	7.0	0.7	0.0	8	28	7	6.2	0.0	0.11	2.9	0.5	2.1E+06
21 มี.ค. 52	441	yellow green	28	7.1	0.7	0.0	7	26	5	5.0	0.3	0.08	2.9	0.5	4.6E+06
21 มี.ค. 52	451	natural	28	7.8	3.7	0.0	2	1678	162	3.9	0.0	0.01	2.0	0.1	3.0E+05
5 มี.ค. 52	461	natural	28	7.1	1.5	0.0	5	33	9	4.5	0.0	0.07	1.9	0.2	3.9E+05
5 มี.ค. 52	462	natural	28	7.3	4.7	0.0	5	19	2	3.4	0.0	0.15	1.5	0.1	1.1E+04
5 มี.ค. 52	463	natural	28	7.2	1.2	0.0	7	35	9	7.3	0.0	0.09	1.7	0.5	1.1E+07
21 มี.ค. 52	471	natural	28	7.8	3.5	0.0	2	1658	211	11.2	0.0	0.01	1.8	0.1	3.0E+04
21 มี.ค. 52	472	natural	28	7.8	2.9	0.0	2	1690	192	6.7	0.0	0.02	1.8	0.1	3.0E+04
21 มี.ค. 52	473	natural	28	7.9	3.6	0.0	2	1687	137	13.4	0.0	0.01	1.8	0.2	4.0E+04
5 มี.ค. 52	481	black	28	7.2	0.0	1.0	19	68	5	12.3	0.0	0.10	1.7	1.2	1.1E+08
21 มี.ค. 52	491	green	26	7.3	2.6	0.0	2	40	14	4.5	0.0	0.02	1.5	0.1	7.0E+04
6 มี.ค. 52	501	yellow green	28	7.1	1.3	0.0	9	42	6	10.6	0.8	0.72	3.1	1.2	6.4E+05
13 มี.ค. 52	511	natural	26	7.3	4.7	0.0	2	17	22	7.8	0.3	0.01	1.7	0.1	1.1E+05
13 มี.ค. 52	512	natural	26	7.1	1.6	0.0	4	27	11	6.2	0.8	0.01	1.4	0.2	7.5E+05
13 มี.ค. 52	513	natural	26	7.1	1.3	0.0	5	37	7	5.0	0.6	0.01	1.5	0.5	1.2E+06
13 มี.ค. 52	514	gray	26	7.1	1.0	0.0	10	43	10	6.7	0.3	0.02	1.5	0.9	1.1E+07
13 มี.ค. 52	515	gray	26	7.1	3.0	0.0	10	37	9	6.7	0.3	0.02	1.5	0.9	2.4E+06
13 มี.ค. 52	516	natural	26	7.2	3.4	0.0	4	27	26	9.5	0.0	0.02	1.8	0.1	2.3E+05
21 มี.ค. 52	521	green	26	7.0	0.0	0.6	30	67	10	9.0	0.0	0.01	1.8	1.5	2.4E+08
21 มี.ค. 52	522	gray	26	7.1	1.3	0.0	13	53	12	3.9	0.6	0.02	1.7	1.2	1.1E+07
21 มี.ค. 52	523	natural	26	7.0	0.0	0.6	16	56	11	6.7	0.0	0.02	1.6	1.0	2.1E+08
21 มี.ค. 52	531	gray	26	7.1	0.0	0.1	25	59	10	2.2	0.0	0.01	1.6	1.4	2.4E+08
21 มี.ค. 52	541	black	26	7.1	0.0	0.1	12	42	12	5.0	0.0	0.01	1.6	0.9	3.0E+05
21 มี.ค. 52	551	green	26	7.1	1.2	0.0	4	15	6	6.7	0.3	0.01	1.7	0.3	2.1E+06
21 มี.ค. 52	561	brown	26	7.2	2.7	0.0	4	7	17	3.4	0.0	0.02	1.5	0.1	9.0E+04
13 มี.ค. 52	571	natural	26	7.2	2.8	0.0	4	22	16	6.2	0.0	0.01	1.5	0.2	4.6E+06
13 มี.ค. 52	572	natural	26	7.2	2.8	0.0	3	25	9	3.9	0.0	0.01	1.5	0.2	1.2E+06
8 มี.ค. 52	581	gray	28	7.2	0.0	0.5	28	71	11	10.6	0.0	0.01	1.5	1.7	1.1E+09
8 มี.ค. 52	582	gray	28	7.1	0.8	0.0	14	37	5	9.5	0.8	0.01	1.7	1.0	2.4E+06
8 มี.ค. 52	591	yellow green	28	7.4	4.5	0.0	14	97	28	7.8	0.0	0.01	1.8	1.5	4.6E+06
8 มี.ค. 52	592	yellow green	28	7.2	0.0	0.3	36	96	17	6.2	0.0	0.01	1.8	1.8	2.4E+07
8 มี.ค. 52	601	gray	28	7.2	0.0	0.3	43	95	18	4.5	0.0	0.01	1.9	1.6	3.9E+06
8 มี.ค. 52	602	green	28	7.3	0.0	0.2	42	162	22	3.9	0.0	0.01	1.8	1.5	1.1E+08
14 มี.ค. 52	611	green	26	7.0	0.0	0.3	46	61	8	7.8	0.3	0.01	1.9	1.0	1.1E+09
14 มี.ค. 52	612	green	26	7.0	0.0	0.2	33	73	8	4.5	0.0	0.01	2.0	1.0	2.4E+07
14 มี.ค. 52	621	green	26	7.2	2.1	0.0	4	10	6	2.8	0.0	0.01	2.9	0.1	4.6E+06
22 มี.ค. 52	631	black	28	7.0	0.0	0.4	35	39	20	5.0	0.0	0.02	1.2	1.1	1.1E+09
22 มี.ค. 52	641	black	28	7.0	0.0	0.2	26	35	16	8.4	0.0	0.01	1.1	1.3	4.6E+07
22 มี.ค. 52	642	black	28	7.0	0.0	0.3	27	36	9	12.3	0.0	0.01	1.0	1.0	2.4E+07

14 ม.ค. 52	651	green	26	7.2	1.3	0.0	9	33	40	3.9	0.0	0.01	3.1	0.2	1.5E+05
14 ม.ค. 52	661	green	26	7.2	2.3	0.0	4	11	5	4.5	0.0	0.01	2.1	0.1	1.2E+06
22 ม.ค. 52	671	natural	28	7.2	1.7	0.0	3	13	17	6.2	0.0	0.02	1.6	0.1	4.0E+04
22 ม.ค. 52	672	natural	28	7.2	1.4	0.0	3	16	8	3.4	0.0	0.01	1.9	0.2	4.0E+04
7 ม.ค. 52	691	green	30	6.9	0.0	0.3	8	28	9	9.0	0.0	0.13	1.5	0.8	1.1E+09
7 ม.ค. 52	692	green	30	7.2	0.8	0.0	19	68	17	9.5	0.8	0.03	1.9	2.0	2.1E+06
7 ม.ค. 52	701	green	30	7.4	1.3	0.0	11	60	17	10.1	0.6	0.21	2.2	1.1	2.1E+05
7 ม.ค. 52	711	green	30	7.4	2.6	0.0	3	23	3	5.0	0.0	0.14	2.5	1.0	1.2E+06
7 ม.ค. 52	712	green	30	7.1	1.3	0.0	4	20	8	11.2	0.6	0.20	3.1	1.1	9.3E+05
8 ม.ค. 52	721	gray	30	7.0	1.7	0.0	5	24	3	6.9	0.3	0.04	1.8	0.2	2.4E+06
8 ม.ค. 52	722	gray	30	7.1	0.0	0.1	48	70	19	10.6	0.0	0.05	1.8	1.7	2.3E+06
5 ม.ค. 52	741	yellow green	28	7.1	4.0	0.0	3	17	2	3.9	0.0	0.10	3.3	0.2	2.4E+05
5 ม.ค. 52	751	green	28	7.3	2.9	0.0	6	23	12	5.0	0.0	1.20	1.3	1.0	1.5E+06
5 ม.ค. 52	752	green	28	7.1	2.7	0.0	6	26	1	6.2	0.0	0.58	6.7	1.0	1.1E+07
22 ม.ค. 52	761	green	28	7.1	1.2	0.0	5	12	5	4.5	0.0	0.04	1.2	0.2	2.4E+06
22 ม.ค. 52	762	green	28	7.1	0.3	0.0	15	25	5	5.0	0.0	0.02	1.3	0.7	2.1E+06
22 ม.ค. 52	771	yellow green	28	7.1	1.8	0.0	4	12	6	6.2	0.0	0.01	1.5	0.2	2.1E+06
9 ม.ค. 52	781	yellow green	28	6.9	0.6	0.0	17	52	17	9.0	1.1	0.02	1.9	0.9	2.8E+04
9 ม.ค. 52	782	yellow green	28	7.0	2.2	0.0	5	42	3	7.3	0.0	0.14	1.7	0.3	3.0E+04
21 ม.ค. 52	791	natural	26	7.1	0.7	0.0	8	23	5	6.2	0.6	0.02	1.4	0.6	1.1E+04
22 ม.ค. 52	801	natural	28	7.0	2.8	0.0	3	11	20	5.0	0.0	0.01	1.5	0.1	1.4E+05
21 ม.ค. 52	811	natural	26	7.1	1.5	0.0	2	15	9	8.4	0.0	0.04	1.5	0.1	1.2E+06
21 ม.ค. 52	821	natural	26	7.1	1.5	0.0	3	15	8	6.7	0.6	0.02	1.2	0.2	7.0E+04
21 ม.ค. 52	831	natural	26	7.2	2.3	0.0	3	18	6	9.5	0.0	0.06	1.4	0.1	3.0E+04
20 ม.ค. 52	841	natural	26	7.1	2.5	0.0	2	19	31	4.5	0.0	0.01	2.0	0.1	3.0E+04
20 ม.ค. 52	851	natural	26	7.1	1.3	0.0	2	13	19	3.4	0.6	0.01	2.0	0.1	4.0E+04
9 ม.ค. 52	861	green	28	7.1	0.8	0.0	10	41	6	6.2	0.0	0.11	1.9	1.1	7.0E+04
9 ม.ค. 52	871	yellow green	28	7.1	2.2	0.0	8	42	8	13.4	0.8	0.06	1.7	0.6	3.0E+04
9 ม.ค. 52	872	yellow green	28	7.1	2.0	0.0	6	65	4	9.5	0.8	0.04	1.7	0.6	7.0E+04
9 ม.ค. 52	873	yellow green	28	7.1	2.3	0.0	6	32	4	8.4	0.0	0.28	1.8	0.5	4.0E+04
9 ม.ค. 52	881	yellow green	28	7.1	2.2	0.0	5	33	3	5.0	0.0	0.30	1.7	0.3	1.1E+05
7 ม.ค. 52	891	yellow green	28	7.2	1.8	0.0	4	33	5	7.3	0.8	0.29	1.6	0.5	3.0E+04
7 ม.ค. 52	892	yellow green	28	7.2	1.7	0.0	5	44	13	10.6	0.6	0.04	1.7	0.5	3.0E+04
7 ม.ค. 52	901	yellow green	28	7.2	1.6	0.0	4	31	24	3.4	0.0	0.03	1.6	0.4	7.0E+04
7 ม.ค. 52	902	yellow green	28	7.1	2.0	0.0	4	31	8	12.3	0.0	0.04	1.7	0.3	3.0E+04
7 ม.ค. 52	911	yellow green	28	7.2	2.5	0.0	5	28	11	5.6	0.8	0.06	1.9	0.3	1.1E+07
9 ม.ค. 52	921	yellow green	29	7.1	2.5	0.0	3	22	8	4.5	0.0	0.18	1.7	0.0	3.0E+04
15 ม.ค. 52	931	yellow green	26	7.1	2.9	0.0	3	25	31	6.7	0.0	0.14	1.8	0.1	3.0E+10
6 ม.ค. 52	941	green	30	7.0	2.7	0.0	5	25	15	4.5	0.0	0.10	1.6	0.1	3.0E+04
6 ม.ค. 52	951	green	30	7.0	2.9	0.0	5	46	9	6.7	0.8	0.10	1.6	0.1	3.0E+04

6 ม.ค. 52	961	brown	30	7.1	3.1	0.0	4	23	6	3.4	0.0	0.03	1.4	0.1	3.0E+04
6 ม.ค. 52	971	green	30	6.9	1.2	0.0	5	30	20	3.9	0.0	0.06	1.6	0.1	3.0E+04
6 ม.ค. 52	981	natural	30	6.9	1.2	0.0	3	17	2	3.4	0.0	0.02	1.9	0.1	3.0E+04
6 ม.ค. 52	991	brown	30	7.0	2.1	0.0	3	7	10	4.5	0.0	0.05	1.6	0.1	3.0E+04
15 ม.ค. 52	1001	brown	26	7.2	4.1	0.0	5	26	7	10.6	0.6	0.01	1.4	0.1	9.0E+04
15 ม.ค. 52	1002	brown	26	7.4	4.1	0.0	3	9	34	11.8	0.6	0.01	1.4	0.1	3.0E+04
15 ม.ค. 52	1011	natural	26	7.0	3.4	0.0	2	16	8	9.5	0.0	0.01	1.0	0.0	3.0E+04
15 ม.ค. 52	1021	natural	26	7.1	2.9	0.0	3	17	10	2.8	0.0	0.01	0.9	0.0	3.0E+04
22 ม.ค. 52	1031	natural	28	7.1	3.0	0.0	3	5	7	3.9	0.0	0.01	1.2	0.1	3.0E+04
22 ม.ค. 52	1041	natural	28	7.1	1.4	0.0	6	29	10	7.8	0.0	0.01	1.5	0.5	7.0E+04
22 ม.ค. 52	1051	natural	28	7.1	2.0	0.0	4	10	8	5.6	0.0	0.01	1.2	0.2	1.5E+06
22 ม.ค. 52	1061	natural	28	7.1	1.0	0.0	5	16	9	3.4	0.0	0.01	1.3	0.3	6.4E+04
7 ม.ค. 52	1071	yellow green	29	7.0	0.0	0.4	13	35	9	2.2	0.8	0.07	1.5	1.0	3.9E+07
7 ม.ค. 52	1081	yellow green	29	7.2	1.5	0.0	6	37	9	9.0	0.0	0.03	2.0	0.7	2.3E+05
7 ม.ค. 52	1082	yellow green	29	7.2	1.0	0.0	7	44	5	10.6	0.0	0.35	2.2	0.6	2.1E+06
7 ม.ค. 52	1083	yellow green	29	7.2	1.2	0.0	5	39	11	11.8	1.4	0.08	1.7	0.4	9.0E+04
7 ม.ค. 52	1091	black	29	7.2	0.0	4.0	28	66	6	9.0	0.8	0.06	2.4	1.1	2.4E+08
7 ม.ค. 52	1101	yellow green	29	7.1	1.7	0.0	5	36	16	6.7	0.0	0.01	1.8	0.4	3.0E+04
7 ม.ค. 52	1102	yellow green	29	7.1	2.2	0.0	5	36	6	6.2	0.0	0.03	1.8	0.5	2.8E+05
12 ม.ค. 52	1111	green	25	7.3	1.9	0.0	5	50	10	14.6	0.6	0.04	1.3	0.4	4.0E+04
12 ม.ค. 52	1112	green	25	7.4	2.5	0.0	7	39	12	3.9	0.0	0.03	1.6	0.4	3.0E+04
12 ม.ค. 52	1121	green	25	7.2	1.1	0.0	6	41	9	12.9	0.8	0.02	1.7	0.7	1.4E+05
12 ม.ค. 52	1131	green	25	7.4	1.7	0.0	13	74	19	19.0	2.0	0.03	1.6	1.3	1.2E+06
12 ม.ค. 52	1141	yellow green	25	7.0	0.0	0.3	34	81	20	16.2	0.8	0.04	2.4	1.3	1.1E+09
12 ม.ค. 52	1151	green	25	7.2	0.5	0.0	10	45	9	17.9	0.6	0.05	1.8	1.2	1.1E+07
12 ม.ค. 52	1161	green	25	7.4	2.3	0.0	10	47	20	19.6	1.1	0.02	1.9	0.9	7.0E+04
7 ม.ค. 52	1171	yellow green	29	7.4	2.0	0.0	8	41	13	6.2	0.0	0.06	1.8	1.1	7.0E+04
13 ม.ค. 52	1181	green	26	7.2	1.7	0.0	7	35	17	5.6	0.3	0.06	1.7	0.5	1.1E+07
13 ม.ค. 52	1191	green	26	7.2	0.0	0.3	14	49	9	6.2	0.0	0.01	1.7	1.1	2.4E+08
13 ม.ค. 52	1201	green	26	7.2	4.4	0.0	4	38	16	8.4	0.3	0.15	2.3	0.1	3.0E+04
13 ม.ค. 52	1211	green	26	7.1	2.5	0.0	1	33	11	9.0	0.0	0.22	2.8	0.1	4.0E+04
13 ม.ค. 52	1221	green	26	7.1	2.4	0.0	3	27	3	3.9	0.0	0.28	2.9	0.1	3.0E+04
13 ม.ค. 52	1231	brown	26	6.9	1.0	0.0	12	46	9	5.0	0.3	0.06	1.6	0.6	1.1E+07
13 ม.ค. 52	1241	brown	26	7.0	0.9	0.0	12	28	12	6.2	0.0	0.01	1.5	0.7	4.6E+06
15 ม.ค. 52	1251	yellow green	26	7.2	1.9	0.0	3	21	6	8.4	0.6	0.02	7.8	0.6	4.0E+04
15 ม.ค. 52	1261	gray	26	7.0	0.0	0.2	25	31	4	5.0	0.0	0.05	2.6	1.1	1.1E+08
15 ม.ค. 52	1271	yellow green	26	7.2	3.1	0.0	5	13	10	3.4	0.0	0.02	2.3	0.2	3.0E+04
16 ม.ค. 52	1281	green	26	6.9	0.8	0.0	16	88	20	3.4	0.0	0.02	1.9	1.3	4.6E+06
16 ม.ค. 52	1291	gray	26	6.9	0.0	0.2	25	45	4	4.5	0.0	0.04	2.1	1.3	1.1E+08
16 ม.ค. 52	1301	green	26	6.9	2.8	0.0	5	26	10	6.7	0.0	0.02	1.9	0.3	3.0E+04

16 ม.ค. 52	1302	green	26	6.9	1.5	0.0	3	21	6	5.6	0.0	0.03	2.0	0.4	3.0E+04
16 ม.ค. 52	1303	green	26	6.9	1.1	0.0	3	18	35	4.5	0.0	0.03	1.9	0.4	7.0E+04
16 ม.ค. 52	1311	black	26	7.1	0.0	0.4	53	138	99	5.0	0.0	0.06	1.9	1.9	1.1E+09
16 ม.ค. 52	1321	green	26	7.4	4.7	0.0	14	43	28	2.8	0.3	0.04	2.0	0.6	3.0E+04
9 ม.ค. 52	1331	green	30	7.2	0.0	0.2	12	109	3	11.8	0.6	0.03	1.8	1.3	2.4E+07
9 ม.ค. 52	1341	green	30	7.2	0.0	0.2	25	127	13	13.4	0.6	0.01	2.1	1.3	1.1E+08
9 ม.ค. 52	1351	green	30	7.2	0.5	0.0	15	61	7	9.0	0.0	0.04	2.0	1.0	2.4E+06
9 ม.ค. 52	1352	green	30	7.2	0.9	0.0	12	63	7	10.6	0.6	0.06	1.9	1.0	4.6E+06
9 ม.ค. 52	1361	black	30	7.1	0.0	0.1	27	85	11	5.6	0.6	0.02	1.9	1.5	2.4E+07
14 ม.ค. 52	1371	green	26	7.2	1.9	0.0	4	23	12	7.3	0.0	0.05	2.1	0.1	2.1E+06
15 ม.ค. 52	1381	green	26	7.2	2.1	0.0	5	20	5	3.4	0.0	0.03	2.1	0.2	7.0E+04
15 ม.ค. 52	1391	green	26	7.3	1.5	0.0	5	16	11	5.0	0.0	0.03	2.9	0.3	2.1E+06
15 ม.ค. 52	1401	green	26	7.3	1.8	0.0	6	15	5	6.0	0.0	0.01	3.0	0.3	1.2E+06
15 ม.ค. 52	1411	green	26	7.2	2.1	0.0	6	12	10	8.4	0.0	0.03	2.2	0.1	7.0E+04
15 ม.ค. 52	1412	green	26	7.3	1.8	0.0	4	15	5	6.7	0.0	0.04	2.2	0.2	1.4E+05
15 ม.ค. 52	1421	green	26	7.3	2.0	0.0	5	18	31	7.3	0.0	0.02	2.7	0.2	2.0E+05
15 ม.ค. 52	1431	green	26	7.2	1.1	0.0	6	23	18	3.4	0.0	0.01	2.7	0.4	1.1E+07
14 ม.ค. 52	1441	yellow green	26	7.2	1.1	0.0	6	13	9	7.3	0.6	0.09	4.2	0.4	1.2E+06
16 ม.ค. 52	1451	yellow green	26	7.3	1.3	0.0	12	37	10	8.4	0.0	0.02	1.8	1.3	2.1E+06
16 ม.ค. 52	1452	yellow green	26	7.3	2.0	0.0	10	39	21	4.5	0.0	0.03	1.6	1.3	1.4E+05
14 ม.ค. 52	1461	yellow green	26	7.3	2.5	0.0	5	15	12	4.5	0.0	0.02	2.6	0.1	1.5E+05
14 ม.ค. 52	1471	yellow green	26	7.3	2.7	0.0	4	9	8	2.8	0.0	0.01	2.6	0.1	2.8E+05
14 ม.ค. 52	1491	yellow green	26	7.1	1.8	0.0	10	18	4	5.0	0.0	0.14	2.2	0.3	1.1E+07
14 ม.ค. 52	1501	yellow green	26	7.2	1.6	0.0	6	14	13	2.8	0.0	0.01	4.1	0.3	2.1E+06
6 ม.ค. 52	1513	yellow green	28	7.6	2.0	0.0	10	50	9	10.6	0.0	0.27	1.6	1.2	2.0E+05
6 ม.ค. 52	1531	black	28	7.1	0.0	0.2	12	57	8	12.3	1.4	0.03	1.8	1.4	4.6E+07
6 ม.ค. 52	1541	yellow green	28	8.8	6.2	0.0	15	136	42	11.2	0.0	0.01	2.1	0.6	1.4E+05
6 ม.ค. 52	1551	yellow green	28	8.2	0.8	0.0	13	111	22	15.1	1.1	0.10	2.4	1.5	4.6E+06
6 ม.ค. 52	1561	black	28	7.2	0.0	0.2	13	63	16	8.4	0.0	0.12	1.9	1.4	1.1E+08
16 ม.ค. 52	1571	yellow green	26	7.3	1.0	0.0	12	42	18	9.0	0.0	0.01	1.6	1.2	1.2E+06
16 ม.ค. 52	1572	yellow green	26	7.3	1.3	0.0	19	45	11	8.4	0.8	0.01	1.7	1.2	1.1E+07
16 ม.ค. 52	1591	gray	26	7.4	0.0	0.5	26	44	9	4.5	0.0	0.01	1.9	1.3	1.2E+08
16 ม.ค. 52	1611	yellow green	26	7.2	2.5	0.0	16	39	12	3.4	0.0	0.06	2.0	1.0	1.1E+07
16 ม.ค. 52	1621	yellow green	26	7.2	3.3	0.0	9	38	25	7.2	0.3	0.02	1.8	0.7	3.0E+04
16 ม.ค. 52	1631	yellow green	26	7.4	1.8	0.0	12	62	13	3.9	0.0	0.02	1.6	0.8	1.4E+05
15 ม.ค. 52	1641	yellow green	26	7.2	2.1	0.0	7	15	9	9.5	0.0	0.01	2.6	0.4	9.0E+04
14 ม.ค. 52	1651	gray	26	7.2	0.0	0.2	67	71	12	3.9	0.0	0.01	2.2	1.5	2.4E+08

Appendix B

Bangkok's Canal Sites

Source: Department of Drainage and Sewerage Bangkok Metropolitan Administration, dated May 1, B.E.2551 (2008).

Cluster	ID	Name of Canals	Site of Canals
1	012	คลองคูเมืองเดิม	ปตร.ราชินี
1	013	คลองคูเมืองเดิม	หน้ากรมที่ดิน
1	021	คลองหลอดวัดราชนัดดา	หลัง กทม. 1
1	031	คลองหลอดวัดราชบพิธ	ถนนตีทอง
1	041	คลองรอบกรุง	สะพานผ่านฟ้า
1	042	คลองรอบกรุง	หลังตลาดนานา
1	052	คลองรอบกรุง	สะพานดำรงสถิต (ถนนเจริญกรุง)
1	053	คลองรอบกรุง	สะพานสมมตอมรมารค (ถนนบำรุงเมือง)
1	061	คลองมหานาค	สะพานดำรัส
1	062	คลองมหานาค	สะพานเจริญราษฎร์
1	063	คลองมหานาค	เจริญผล
1	072	คลองผดุงกรุงเกษม	สถานีสูบน้ำกรุงเกษม
1	081	คลองสามเสน	ปตร. สามเสน
1	082	คลองสามเสน	วัดโบสถ์
1	083	คลองสามเสน	อนุสาวรีย์ชัยสมรภูมิ
1	091	คลองแสนแสบ	ถนนอโศกดินแดง
1	092	คลองแสนแสบ	ปตร.แสนแสบ (ถ.เพชรบุรี ซ.ประสานมิตร)
1	094	คลองแสนแสบ	สะพานบางกะปิ
1	095	คลองแสนแสบ	วัดบำเพ็ญเหนือ
1	098	คลองแสนแสบ	สะพานประตูน้ำเวิลเทรดเซ็นเตอร์
1	101	คลองตัน	ปตร. คลองตัน
1	112	คลองเปรมประชากร	สี่แยกสะพานแดง
1	113	คลองเปรมประชากร	ตลาดบางซื่อ
1	114	คลองเปรมประชากร	ทัณฑสถานวัยหนุ่ม (ถ.งามวงศ์วาน)
1	115	คลองเปรมประชากร	เทศบาลสงเคราะห์หน้าวัดเสมียนนารี
1	116	คลองเปรมประชากร	ถนนเศรษฐศิริ (สะพานเกษะโกมล
1	121	คลองบางซื่อ	สะพานพิบูลย์สงคราม (ถ.ประชาราษฎร์ สาย 1)
1	271	คลองพระโขนง	ปตร. พระโขนง
1	272	คลองพระโขนง	สะพานพระโขนง (ถนนสุขุมวิท)

Cluster	ID	Name of Canals	Site of Canals
1	461	คลองประเวศบุรีรมย์	หน้าวัดลานบุญ
1	463	คลองประเวศบุรีรมย์	ถนนพัฒนาการ
1	516	คลองบางเขนใหม่	ถนนพิบูลย์สงคราม
1	571	คลองบางโพ	สน. บางโพ
1	572	คลองบางโพ	ถนนประชาราษฎร์สาย 1
1	691	คลองลาดยาว	ถนนวิภาวดี-รังสิต
1	781	คลองสองต้นนุ่น	หลังโรงงานปรับปรุงคุณภาพน้ำร่มเกล้า
1	782	คลองสองต้นนุ่น	ซ.รามคำแหง 209
1	861	คลองเจ็ก	ถนนสุขาภิบาล 3 ซ.รามคำแหง 149
1	871	คลองหลอแหล	ถนนสุขาภิบาล 3 ซ.รามคำแหง 157/3
1	872	คลองหลอแหล	ถนนสุขาภิบาล 3 ซ.รามคำแหง 162
1	873	คลองหลอแหล	ถนนสุขาภิบาล ตลาดสด ซีรอ เข้า ซ.รามคำแหง 162
1	881	คลองบางชัน	ถนนสุขาภิบาล 3 ตรงข้ามธาราการ์เด้น
1	891	คลองลาดบัวขาว	ถนนราษฎร์พัฒนา โรงเรียนสุเหร่าลาดบัวขาว
1	892	คลองลาดบัวขาว	วัดลาดบัวขาว
1	901	คลองลำนายโส	ถนนราษฎร์พัฒนา โรงเรียนสุเหร่าลำนายโส
1	902	คลองลำนายโส	วัดปากบึง
1	931	คลองไผ่เหลือง	ถนนสุวินทวงศ์ ม.รุ่งสินก่อสร้าง
1	1031	คลองส้มป่อย	สะพานพระราม 6 สถานีเฝ้าระวัง
1	1041	คลองซุง	ใต้สะพานพระราม 6 ขนานทางรถไฟบางซ่อน
1	1082	คลองสองห้อง	ถนนเฉลิมพระเกียรติ ร.9
1	1101	คลองตะเข้ขบ	ถนนเฉลิมพระเกียรติ ร.9 ซอย 79
1	1102	คลองตะเข้ขบ	ถนนลาดกระบัง ใต้ทางด่วน
1	1111	คลองปลัดเปรียง	ถนนเฉลิมพระเกียติ ร.9 มัสยิดซะห์หร่อตุลอิสลาม
1	1121	คลองมะขามเทศ	ถนนเฉลิมพระเกียรติ ร.9 ซอย 51
1	1151	คลองสาด	ประตูระบายน้ำคลองสาด ถ.ศรีนครินทร์
1	1161	คลองบ้านม้า	ถนนศรีนครินทร์
1	1171	คลองหัวหมาก	ถนนอ่อนนุช
1	1181	คลองตาพุก	ถนนลาดกระบัง ใกล้ซอย 4
1	1221	คลองหนองคา	หมู่บ้านมณสินี
1	1241	คลองหัวตะเข้	ถนนลาดกระบัง ใกล้ซอย 17 (ตลาดอุดมผล)
1	1251	คลองบึงบัว	ติดถนนเจ้าคุณทหาร
1	1321	คลองสอง	ซ.วัดเกาะ สามแยกโรงเรียนระเบียบวิทยา
1	1401	คลองวัดจันทร์	ถ.พระราม 3 หน้าสำนักงานเขตบางคอแหลม
1	1411	คลองบางโคล่น้อย	ถ.พระราม 3 โฮมโปร

Cluster	ID	Name of Canals	Site of Canals
1	1412	คลองบางโคล่น้อย	ถ.เจริญราษฎร์
1	1421	คลองบางโคล่	ถ.พระราม 3 โรงแรมมณเฑียรริเวอร์ไซด์
1	1431	คลองวัดไทร	ถ.พระราม 3 วัดไทร
1	1441	คลองวัดช่องนนทรี	ถ.พระราม 3 เข้าซอย 53
1	1451	คลองลำสาลี	ถ.กรุงเทพกรีฑา โรงเรียนลำสาลี
1	1452	คลองลำสาลี	ถ.กรุงเทพกรีฑา ซ.ประชาร่วมใจ
1	1471	คลองวัดปริวาส	ถ.พระราม 3
1	1491	คลองวัดดอกไม้	ถ.พระราม 3 วัดดอกไม้
1	1501	คลองโรงน้ำมัน	ถ.พระราม 3 ข้างปั๊มน้ำมันบางจาก
1	1541	คลองคอตัน	ถ.ปัญญา-เนเจอรัลปาร์ค
1	1572	คลองบ้านม้า	ถ.รามคำแหง (สุขาภิบาล 3) เลยซอย 127/1
1	1611	คลองทับช้างล่าง	ถ.กรุงเทพกรีฑา ซ.ภราดร (ชุมชนคลองทับช้างล่าง)
1	1621	คลองบัวคลี่	ถ.เลียบมอเตอร์เวย์ หน้าวัดลาดบัวขาว
1	1631	คลองหลอแหล	ถ.ราษฎร์พัฒนา ซ.มิสทีน
2	132	คลองสาธร	สถานทูตซาอุดิอาระเบีย
2	133	คลองสาธร	ติดถนนพระราม 4 ใกล้สะพานไทย-เบลเยี่ยม
2	142	คลองช่องนนทรี	ถนนนราธิวาสราชนครินทร์ ตัดแยกถนนจันทร์
2	143	คลองช่องนนทรี	ก่อนแยกสาธร
2	144	คลองช่องนนทรี	ตัดถนนสีลม
2	145	คลองช่องนนทรี	ตัดถนนสุรวงศ์
2	171	คลองห้วยขวาง	ชุมชนห้วยขวาง
2	173	คลองห้วยขวาง	ถนนสุทธิสารวินิจฉัย
2	174	คลองห้วยขวาง	ถนนรัชดาภิเษก
2	175	คลองห้วยขวาง	ถนนวัฒนธรรม ใกล้สถานทูตเกาหลี
2	281	คลองบางนา	หน้ากรมอุตุนิยมวิทยา (ถนนสุขุมวิท)
2	301	คลองไผ่สิงโต	ตลาดคลองเตย
2	303	คลองไผ่สิงโต	ข้างโรงงานยาสูบ
2	351		โรงสูบน้ำพระราม 4
2	481	คลองบางจาก	โรงกลั่นน้ำมันบางจาก (เขตพระโขนง)
2	581	คลองหลุมไผ่	ถนนลาดปลาเค้า
2	582	คลองหลุมไผ่	ซ. ลาดปลาเค้า 47
2	591	คลองกุ่ม	หมู่บ้านสหกรณ์ (ซ.28)
2	592	คลองกุ่ม	ถ. สุขาภิบาล 2
2	601	คลองพังพวย	ถ. สุขาภิบาล 1
2	611	คลองเตย	ปตร. คลองเตย

Cluster	ID	Name of Canals	Site of Canals
2	612	คลองเตย	อาคารทวิช
2	621	คลองวัดดอน	ซ. เจริญกรุง 57
2	651	คลองหัวลำโพง	หน้าเขตคลองเตย
2	701	คองพญาเวิก	ถนนลาดพร้าว (ซอยลาดพร้าว 5/1)
2	711	คลองน้ำแก้ว	ถนนลาดพร้าว (ซอยลาดพร้าว 35)
2	712	คลองน้ำแก้ว	ถนนรัชดาภิเษก
2	721	คลองสวนหลวง 1	ถนนพระราม 1 (แยกเจริญผล)
2	722	คลองสวนหลวง 1	ถนนเจริญเมือง
2	751	คลองนาซอง	ถนนประชาสงเคราะห์ ซอยพาณิชย์จำนงค์
2	1091	คลองวัดกระทุ่มเสือปลา	ถนนอ่อนนุช ซอย 67 วัดกระทุ่มเสือปลา
2	1291	คลองจรเข้	ถนนสุขาภิบาล 5
2	1311	คลองสามง่าม	หมู่บ้านเกาะแก้ววิลล่า ซอย 2 ถ.พหลโยธิน 54/1 แยก 4-7
2	1351	คลองบ้านหลาย	ถ.สุขุมวิท 101/1 ชุมชนหมู่บ้านศรัณจิต
2	1352	คลองบ้านหลาย	ซ.พึ่งมี 50 พึ่งมีแมนชั่น
2	1361	คลองเคล็ด	ถ.อุดมสุข 51 ซ.วชิรธรรมสาธิต 57
2	1371	คลองขวาง	ถ.เจริญกรุง 72
2	1561	คลองลำต้นนุ่น	ถ.เสรีไทย 73
2	1651	คลองมะนาว	ถนนรัชดาภิเษก พระราม 3 หน้า ปตร.คลองมะนาว
3	073	คลองผดุงกรุงเกษม	สถานีรถไฟกรุงเทพ
3	085	คลองสามเสน	หลังแฟลตดินแดง
3	093	คลองแสนแสบ	ซอยเทพลีลา
3	118	คลองเปรมประชากร	ถนนแจ้งวัฒนะ
3	119	คลองเปรมประชากร	ถนนเตชะตุงคะ ใกล้ สน.ดอนเมือง
3	122	คลองบางซื่อ	ถนนพหลโยธิน
3	123	คลองบางซื่อ	ถนนรัชดาภิเษก (รร.เจ้าพระยาปาร์ค)
3	131	คลองสาธร	ปตร. สาธร
3	141	คลองช่องนนทรี	ถนนนราธิวาสราชนครินทร์
3	146	คลองช่องนนทรี	หน้าโรงงานควบคุมคุณภาพน้ำช่องนนทรี
3	181	คลองลาดพร้าว	ถนนประชาอุทิศ
3	182	คลองลาดพร้าว	ร.ร.พิบูลย์อุปถัมภ์ (ถนนลาดพร้าว)
3	183	คลองลาดพร้าว	วัดบางบัว
3	184	คลองลาดพร้าว	ถนนเกษตร-นวมินทร์
3	282	คลองบางนา	สะพานบางนา (ถนนศรีนครินทร์)
3	371	คลองสำโรง	สะพานสำโรง (ถนนศรีนครินทร์)
3	501	คลองบางกะปิ	ปตร. บางกะปิ

Cluster	ID	Name of Canals	Site of Canals
3	512	คลองบางเขน	วัดทางหลวง
3	513	คลองบางเขน	ข้างทัณฑสถานบางเขน
3	514	คลองบางเขน	ชุมชนบางบัว
3	515	คลองบางเขน	ใกล้โรงงานปรับปรุงคุณภาพน้ำบางบัว
3	661	คลองกรวย	ซ. เจริญกรุง 71
3	692	คลองลาดยาว	ถนนรัชดาภิเษก
3	741	คลองบางกระบือ	ถนนสามเสน (บริษัท บุญรอดบริวเวอรี่
3	752	คลองนาซอง	ถนนพระราม 9 (สถานีสูบน้ำนาซอง)
3	1051	คลองบางซ่อน	ถนนประชาราษฎร์สาย 1 ระหว่างซอย 34 - 36
3	1061	คลองสาธารณะประโยชน์	ถนนประชาราษฎร์สาย 1 ใกล้ซอย 21
3	1071	คลองบึง	หน้าสำนักงานเขตสวนหลวง
3	1081	คลองสองห้อง	ถนนอ่อนนุชใกล้ซอย 61
3	1083	คลองสองห้อง	หน้าโรงขยะอ่อนนุช (สถานีย่อยขยะมูลฝอย)
3	1112	คลองปลัดเปรียง	ทางลัดเข้าซอยร่มเย็น (ถ.เฉลิมพระเกียรติ ร.9
3	1131	คลองหนองบอน	ถนนเฉลิมพระเกียติ ร.9 ซอย 26
3	1141	คลองตาช้าง	สถานีสูบน้ำคลองตาช้าง ถนนศรีนครินทร์
3	1191	คลองบัวลอย	ถนนลาดกระบัง ใกล้ซอย 22
3	1261	คลองลำชวดเตย	ตัดถนนฉลองกรุง ใกล้หมู่บ้านวรีย์เฮ้าส์
3	1281	คลองบัว	โรงเรียนบ้านคลองบัว ถนนสุขาภิบาล 5
3	1301	คลองออเงิน	ตัดถนนสุขาภิบาล 5
3	1302	คลองออเงิน	หมู่บ้านมายด์เพลส
3	1303	คลองออเงิน	โรงเรียนศิริวัฒน์วิทยา
3	1331	คลองบางอ้อ	ถ.รถไฟสายเก่า คลังน้ำมันบางจาก
3	1341	คลองเจ็ก	ถ.รถไฟสายเก่า แฟลตเอื้ออมรสุข
3	1381	คลองสวนหลวง	ปตร.สวนหลวง ถ.เจริญกรุง 76
3	1391	คลองบางคอแหลม	ถ.พระราม 3 ตรงข้ามธนาคารกสิกรไทย
3	1513	คลองครุ	ถ.เสรีไทย 67
3	1531	คลองเกร็ด	ถ.ปัญญา-เนเจอรัลปาร์ค
3	1551	คลองรหัส	ถ.เสรีไทย 57
3	1571	คลองบ้านม้า	ถ.รามคำแหง สามแยกบ้านม้า
3	1591	คลองสะพานสูง	ถ.รามคำแหง หมู่บ้านกรีนเบอร์วิลล์
4	014	คลองหลอด	อนุสาวรีย์แม่ธรณีมวยผม สนามหลวง
4	015	คลองคูเมืองเดิม	ปตร. พระปิ่นเกล้า
4	043	คลองรอบกรุง	ปตร. บางลำภู
4	051	คลองรอบกรุง	ปตร. โอ่งอ่าง

Cluster	ID	Name of Canals	Site of Canals
4	074	คลองผดุงกรุงเกษม	หน้ากรมวิเทศสหการ
4	075	คลองผดุงกรุงเกษม	ตลาดเทวราช
4	076	คลองผดุงกรุงเกษม	ปตร. เทเวศน์
4	096	คลองแสนแสบ	มีนบุรี (รร.สตรีวิทยามีนบุรี)
4	097	คลองแสนแสบ	ปตร.แสนแสบ (ของกรมชลประทาน)
4	099	คลองแสนแสบ	ถนนเลียบวารี
4	090	คลองแสนแสบ	ตลาดหนองจอก
4	099.1	คลองแสนแสบ	ถ.สังฆสันติสุข ซ.โรงเรียนสุเหร่าใหม่ (หนองจอก)
4	117	คลองเปรมประชากร	ถนนศรีอยุธยา (วัดเบญจมบพิตร)
4	231	คลองอ้อมนนท์	ท่าน้ำวัดโตนด
4	232	คลองอ้อมนนท์	ท่าน้ำวัดประชารังสรรค์
4	263	คลองบางกอกน้อย	ทำน้ำวัดชลอ (ถ.บางกรวย-ไทรน้อย)
4	264	คลองบางกอกน้อย	ซอยร่วมวงศ์พัฒนา (ถ.บางกรวย-ไทรน้อย)
4	265	คลองบางกอกน้อย	ท่าน้ำวัดอุทยาน (ถ.บางกรวย-จงถนอม)
4	381	คลองประปา	ในเขื่อนตลาดบางซื่อ
4	382	คลองประปา	นอกเขื่อน
4	391	คลองบางกอกใหญ่	ถนนเพชรเกษม
4	392	คลองบางกอกใหญ่	ร.ร.พิาณิชยการธนบุรี
4	393	คลองบางกอกใหญ่	วัดช่างเหล็ก
4	394	คลองบางกอกใหญ่	บางขุนนนท์
4	395	คลองบางกอกใหญ่	สะพานเจริญพาสน์
4	401	คลองบางกรวย	วัดกล้วย (ถ.บางกรวย-ไทรน้อย22)
4	402	คลองบางกรวย	วัดสำโรง (สะพานเฉลิมศักดิ์)
4	462	คลองประเวศบุรีรมย์	สำนักงานเขตลาดกระบัง
4	511	คลองบางเขน	ปตร. บางเขนเก่า
4	801	คลองวัดราชา	วัดราชา
4	911	คลองบึงขวาง	ถนนร่มเกล้า ชุมชนพัฒนาบึงขวาง
4	921	คลองทรายกองดิน	ถนนราษฎร์อุทิศ สุเหร่าทรายกองดิน
4	941	คลองลำหิน	ถ.คู้คลองสิบ สะพานข้ามคลอง
4	951	คลองลำหินฝั่งใต้	ถนนเลียบวารี สะพานคลองลำหินฝั่งใต้
4	961	คลองสิบเอ็ด	ถนนคลองสิบ คลองสิบเอ็ด สะพานคลองสิบเอ็ด
4	971	คลองหวังโต	ถนนมิตรไมตรี ชุมชนหนองจอก
4	981	คลองบึงแตงโม	ถนนมิตรไมตรี หมู่บ้าน เค.ซี.การ์เด้น
4	991	คลองลำต้นกล้วย	ถนนมิตรไมตรี ก่อนถึงสามแยก รพ.หนองจอก
4	1001	คลองลำผักชี	ถนนสุวินทวงศ์ สถานีย่อยสุวินทวงศ์ กฟน.

Cluster	ID	Name of Canals	Site of Canals
4	1002	คลองลำผักชี	ถนนฉลองกรุง หมู่บ้านฉลองกรุงริเวอร์วิว
4	1011	คลองอู่ตะเภา	ถนนสุวินทวงศ์ บ.มีนบุรีวิศวกรรมโยธา
4	1021	คลองลำปลาเน่า	ถนนฉลองกรุง จุดกลับรถใต้สะพานข้ามคลอง
4	1201	คลองลาดกระบัง	ถนนลาดกระบัง วัดลาดกระบัง
4	1211	คลองหนองตะกร้า	ถนนลาดกระบัง ใกล้ซอย 40
4	1231	คลองหนองปรือ	สวนพระนคร ใกล้สำนักงานเขตลาดกระบัง
4	1271	คลองลำกอไผ่	ตัดถนนฉลองกรุง
4	1461	คลองวัดใหม่	ถ.พระราม 3 ธ.กรุงศรีอยุธยาสำนักงานใหญ่
4	1641	คลองลำมะขาม	ถ.ฉลองกรุง
5	151	คลองบางไส้ไก่	ถนนเจริญนคร
5	152	คลองบางไส้ไก่	ถนนตากสิน
5	161	คลองสมเด็จเจ้าพระยา	สะพานท่าดินแดง
5	162	คลองสมเด็จเจ้าพระยา	หน้าโรงพยาบาลสมเด็จเจ้าพระยา
5	191	คลองบางน้ำชน	ถนนเจริญนคร
5	192	คลองบางน้ำชน	ถนนพระเจ้าตากสิน
5	201	คลองดาวคะนอง	ถนนเจริญนคร
5	202	คลองดาวคะนอง	ถ.พระเจ้าตากสิน หลังโรงเก็บขนมูลฝอยทางน้ำ
5	211	คลองบางขุนเทียน	ท่าน้ำวัดบางขุนเทียนนอก (ถ. จอมทอง 19)
5	212	คลองบางขุนเทียน	ถนนพระรามที่ 2
5	221	คลองภาษีเจริญ	ร.ร.วัดรางบัว (ถ.เพชรเกษม 33)
5	222	คลองภาษีเจริญ	ถนนมาเจริญ (ถนนเพชรเกษม 81)
5	241	คลองบางละมุด	ตรงข้าม รร.เทคโนโลยีพระรามหก)
5	251	คลองมอญ	ถนนอรุณอัมรินทร์
5	252	คลองมอญ	ถนนจรัญสนิทวงศ์
5	261	คลองบางกอกน้อย	ถนนอรุณอัมรินทร์
5	262	คลองบางกอกน้อย	ท่าน้ำวัดสุวรรณาราม
5	291	คลองสำเหร่	ถนนเจริญนคร
5	292	คลองสำเหร่	ถนนเทอดไทย
5	311	คลองแจงร้อน	ถนนราษฎร์บูรณะ
5	312	คลองแจงร้อน	ถนนสุขสวัสดิ์
5	321	คลองราษฎร์บูรณะ	ถนนราษฎร์บูรณะ
5	322	คลองราษฎร์บูรณะ	ถนนสุขสวัสดิ์
5	331	คลองบางปะกอก	ถนนราษฎร์บูรณะ
5	332	คลองบางปะกอก	ถนนสุขสวัสดิ์
5	341	คลองบางปะแก้ว	ถนนราษฎร์บูรณะ

Cluster	ID	Name of Canals	Site of Canals
5	342	คลองบางปะแก้ว	ถนนสุขสวัสดิ์
5	411	คลองมหาสวัสดิ์	หน้าวัดชัยพฤกษ์
5	421	คลองทวีวัฒนา	ปตร.ทวีวัฒนา ถนนศาลาธรรมสพน์
5	422	คลองทวีวัฒนา	แยกถนนเพชรเกษม 69
5	423	คลองทวีวัฒนา	สน. ศาลาแดง
5	431	คลองสนามชัย	วัดสิงห์
5	432	คลองสนามชัย	วัดเลา
5	433	คลองสนามชัย	สน. ท่าข้าม
5	434	คลองสนามชัย	วัดบางกระดี่
5	441	คลองหัวกระบือ	วัดหัวกระบือ
5	451	คลองขุนราชวินิตใจ	หน้า ปตร.คลองขุนฯ
5	471	คลองเชิงตาแพ	ปตร. เชิงตาแพ
5	472	คลองรางหอกหัก	ปตร. วัดลูกโค
5	473	คลองรางหอกหัก	ปตร. เรือสัญจร
5	491	คลองบางรัก	ถนนจรัญสนิทวงศ์
5	521	คลองบางยี่ขัน	วัดพระศิริไอยสวรรค์
5	522	คลองบางยี่ขัน	ถนนจรัญสนิทวงศ์
5	523	คลองบางยี่ขัน	ซอยบรมราชชนนี 2
5	531	คลองบางจาก	วัดเปาว์โลหิต (เขตบางพลัด)
5	541	คลองบางพลู	วัดภาณุรังสี
5	551	คลองบางพลัด	ถนนจรัญสนิทวงศ์
5	561	คลองพระครู	ถนนจรัญสนิทวงศ์
5	602	คลองพังพวย	หลังแฟลตคลองจั่น ซ.57
5	631	คลองต้นไทร	ถ. เจริญนคร 17
5	641	คลองวัดทองเพลง	วัดทองเพลง
5	642	คลองวัดทองเพลง	ถนนเจริญนคร
5	671	คลองด่าน	ถนนเทอดไทย
5	672	คลองด่าน	ท่าน้ำวัดนางนองวรวิหาร (ถ.วุฒากาศ)
5	761	คลองบางสะแก	ถนนเทอดไทย (สะพานเทศบาล 11)
5	762	คลองบางสะแก	ถนนเทอดไทย ซอย 33 (วัดบางสะแกใน)
5	771	คลองวัดอนงค์	หน้าวัดอนงค์
5	791	คลองบางลำภูล่าง	วัดเศวตฉัตร (คลองสาน)
5	811	คลองบุปผา	สน.บุปผาราม
5	821	คลองบางผักหนาม	ซอยจรัญฯ 41
5	831	คลองบางบัวหรุ	ซอยจรัญฯ 45
5	841	คลองควาย	ถนนพุทธมณฑลสาย 2
5	851	คลองบัว	ถนนสวนผัก 46

Appendix C

Surface Water Quality Standard in Thailand

Source: Department of Pollution Control Bangkok, dated April 12, B.E.2552 (2009), [online].

Available from: http://www.pcd.go.th/info_serv/reg_std_water05.html

Surface Water Quality Standards								
Parameter[1/]	Units	Statistics	Standard Value for Class[2/]					Methods for Examination
			Class1	Class2	Class3	Class4	Class5	
1. Colour,Odour and Taste	-	-	n	n'	n'	n'	-	-
2. Temperature	C°	-	n	n'	n'	n'	-	Thermometer
3. pH	-	-	n	5-9	5-9	5-9	-	Electrometric pH Meter
4. Dissolved Oxygen (DO)[2/]	mg/l	P20	n	6.0	4.0	2.0	-	Azide Modification
5. BOD (5 days, 20°C)	mg/l	P80	n	1.5	2.0	4.0	-	Azide Modification at 20°C, 5 days
6. Total Coliform Bacteria	MPN/100 ml	P80	n	5,000	20,000	-	-	Multiple Tube Fermentation Technique
7. Fecal Coliform Bateria	MPN/100 ml	P80	n	1,000	4,000	-	-	Multiple Tube Fermentation Technique
8. NO_3 -N	mg/l	-	n	5.0			-	Cadmium Reduction
9. NH_3 -N	mg/l	-	n	0.5			-	Distillation Nesslerization
10.Phenols	mg/l	-	n	0.005			-	Distillation,4-Amino antipyrene
11.Copper (Cu)	mg/l	-	n	0.1			-	Atomic Absorption -Direct Aspiration
12.Nickle (Ni)	mg/l	-	n	0.1			-	Atomic Absorption -Direct Aspiration
13.Manganese (Mn)	mg/l	-	n	1.0			-	Atomic Absorption -Direct Aspiration
14.Zinc (Zn)	mg/l	-	n	1.0			-	Atomic Absorption -Direct Aspiration
15.Cadmium (Cd)	mg/l	-	n	0.005* 0.05**			-	Atomic Absorption -Direct Aspiration
16.Chromium Hexavalent	mg/l	-	n	0.05			-	Atomic Absorption -Direct Aspiration
17.Lead (Pb)	mg/l	-	n	0.05			-	Atomic Absorption -Direct Aspiration
18.Total Mercury (Total Hg)	mg/l	-	n	0.002			-	Atomic Absorption-Cold Vapour Technique
19.Arsenic (As)	mg/l	-	n	0.01			-	Atomic Absorption -Direct Aspiration
20.Cyanide (Cyanide)	mg/l	-	n	0.005			-	Pyridine-Barbituric Acid
21.Radioactivity - Alpha - Beta	Becqurel/l	-	n	0.1 1.0			-	Gas-Chromatography
22.Total Organochlorine Pesticides	mg/l	-	n	0.05			-	Gas-Chromatography
23.DDT	µg/l	-	n	1.0			-	Gas-Chromatography
24.Alpha-BHC	µg/l	-	n	0.02			-	Gas-Chromatography
25.Dieldrin	µg/l	-	n	0.1			-	Gas-Chromatography
26.Aldrin	µg/l	-	n	0.1			-	Gas-Chromatography
27.Heptachlor & Heptachlorepoxide	µg/l	-	n	0.2			-	Gas-Chromatography
28.Endrin	µg/l	-	n	None			-	Gas-Chromatography

Remark : [1/] กำหนดค่ามาตรฐานเฉพาะในแหล่งน้ำประเภทที่ 2-4 สำหรับแหล่งน้ำประเภทที่ 1 ให้เป็นไปตามธรรมชาติ และแหล่งน้ำประเภทที่ 5 ไม่กำหนดค่า
[2/] ค่า DO เป็นเกณฑ์มาตรฐานต่ำสุด

P Percentile value

n naturally

n' naturally but changing not more than 3°C

* when water hardness not more than 100 mg/l as $CaCO_3$

** when water hardness more than 100 mg/l as $CaCO_3$

Based on Standard Methods for the Examination of Water and Wastewater recommended by APHA : American Public Health Association, AWWA : American Water Works Association and WPCF : Water Pollution Control Federation

Source : Notification of the National Environmental Board, No. 8, B.E. 2537 (1994), issued under the Enhancement and Conservation of National Environmental Quality Act B.E.2535 (1992) , published in the Royal Government Gazette, Vol. 111, Part 16, dated February 24, B.E.2537 (1994).

Classification	Objectives/Condition and Beneficial Usage
Class 1	Extra clean fresh surface water resources used for : (1) conservation not necessary pass through water treatment process require only ordinary process for pathogenic destruction (2) ecosystem conservation where basic organisms can breed naturally
Class 2	Very clean fresh surface water resources used for : (1) consumption which requires ordinary water treatment process before use (2) aquatic organism of conservation (3) fisheries (4) recreation
Class 3	Medium clean fresh surface water resources used for : (1) consumption, but passing through an ordinary treatment process before using (2) agriculture
Class 4	Fairly clean fresh surface water resources used for : (1) consumption, but requires special water treatment process before using (2) industry
Class 5	The sources which are not classification in class 1-4 and used for navigation.

BIOGRAPHY

Name: Mrs. Sirilak Areerachakul

Date of Birth: 28th July, 1976

Educations:

• M.Sc. Program in Computer Science, Faculty of Science and Technology, Rangsit University, Pathumthani, Thailand.

• B.Sc. Program in Applied Statistics, Faculty of Science and Technology, Suan Dusit Rajabhat University, Bangkok, Thailand.

Publications:

• Areerachakul, S. and Sanguansintukul, S. (2010). 'Classification and Regression and MLP Neural Network to Classify Water Quality of Canals in Bangkok, Thailand', International Journal of Intelligent Computing Research (IJICR), Volume 1, March 2010, pp 43-52.

•Areerachakul, S. and Sanguansintukul, S. (2010). 'Clustering Analysis of Water Quality for Canals in Bangkok, Thailand', The 2010 International Conference on Computational Science and Applications (ICCSA 2010), March 23-26, 2010, Fukuoka, Japan.

• Areerachakul, S. and Sanguansintukul, S. (2009) 'Water Quality Classification using Neural Networks: Case Study of Canals in Bangkok, Thailand' The 4t International Conference for Internet Technology and Secured Transactions (ICITST-2009) November 9-12, 2009, London, UK.

• Areerachakul, S and Sanguansintukul, S. (2009) 'A Comparison between the Multiple Linear Regression Model and Neural Networks for Biochemical Oxygen Demand Estimations' The Eight International Symposium on Natural Language Processing (SNLP2009). October 20-21, 2009. Bangkok.