MODELING WATER QUALITY PARAMETER BOD IN JOHOR RIVER AND MUAR RIVER

PUAH LIH CHII

A dissertation submitted in partial fulfilment of the

requirements for the award of the degree of

Master of Science (Mathematics)

Faculty of Science

Universiti Teknologi Malaysia

JULY 2015

To my beloved father: Puah Eng Chiew

To my beloved mother: Ng Bee Kee

To my dearest grandfather: Puah Kok Kuah

To my dearest grandmother: Tan Siew Tin

To my precious brothers: Puah Chee Hock & Puah Zi Jian

To my precious sisters: Puah Lee Sia & Puah Lee Lee

To my precious brother in law: Chua Tor Hee

To my precious sisters in law: Yong Siew Lee & Lee Seow Ching

And

My adorable nieces and nephews

Thank you for all your love and support

ACKNOWLEDGEMENT

First of all, I am very grateful to my supervisor, Dr. Haliza Binti Abdul Rahman for her continuous guidance and support throughout this research. Without her, this study would not have been possible. She has given me a lot of encouragement throughout the whole time.

I am very thankful to my parents, who have encouraging me throughout the whole process. Next, a big thank you to my brothers, sisters, brother in law and sisters in law for being there when I needed them.

I would like to extend my gratitude and appreciation to my fellow friends who have supported me in completing this research. I have really appreciated the friendship and support of so many people within the Faculty of Science.

Lastly, thanks to all people who have supported me, either directly or indirectly that encouraged me to come this far. Thank you to all and God bless.

ABSTRACT

Rivers are one of the important sources of water. Rivers also support a diverse flora and fauna. As Malaysia is a fast becoming an industrial country, many of the rivers have become polluted. Water pollution is very harmful to human, animals and water life. Therefore, water quality models are very important to describe and to predict the observed effects of a change in the river system. Biochemical oxygen demand (BOD) concentration data of Sungai Johor and Sungai Muar in Johor were fitted to two models, namely Streeter-Phelps and nonlinear regression (NLR) models, by using weighted least squares method and Gauss-Newton method respectively. BOD is one of the indicators for river. The data used in the study is BOD concentration data of Sungai Johor and Sungai Muar in Johor state from year 1981 to year 1990 and the data were obtained from Drainage and Irrigation Department Malaysia. The performance of the results was tested by mean squared error (MSE). The NLR model gives the smaller MSE values for Sungai Johor and Sungai Muar with values of 0.662319 and 2.367337 respectively. Conclusively, the results showed that NLR model is a better model than Streeter-Phelps model in estimating the parameters of BOD concentration data of Sungai Johor and Sungai Muar in Johor.

ABSTRAK

Sungai adalah salah satu sumber air yang penting. Sungai juga menjadi tempat tinggal pelbagai binatang dan tumbuhan. Oleh sebab Malaysia berkembang pesat menjadi sebuah negara perindustrian, terdapat banyak sungai yang telah tercemar. Pencemaran air sangat bahaya kepada manusia, binatang dan hidupan akuatik. Oleh itu, model kualiti air sangat penting untuk menerangkan dan menjangkakan perubahan daripada kesan yang diperhatikan dalam sistem sungai. Data kepekatan keperluan biokimia oksigen (BOD) daripada Sungai Johor dan Sungai Muar di Johor telah digunakan dalam dua model, iaitu model Streeter-Phelps dan model regresi tak linear, masing-masing yang menggunakan kaedah wajaran kuasa dua terkecil dan kaedah Gauss-Newton. BOD ialah salah satu penunjuk ukur untuk sungai. Data yang digunakan dalam kajian ini ialah data kepekatan keperluan biokimia oksigen daripada Sungai Johor dan Sungai Muar di negeri Johor dari tahun 1981 hingga tahun 1990. Data ini diperolehi daripada Jabatan Saliran dan Perparitan Malaysia. Model regresi tak linear telah memberikan keputusan nilai ralat purata kuasa dua (MSE) yang lebih kecil untuk Sungai Johor dan Sungai Muar, iaitu 0.662319 dan 2.367337. Kesimpulannya, keputusan tersebut telah menunjukkan bahawa model regresi tak linear ialah model yang lebih baik daripada model Streeter-Phelps dalam menjangkakan ukuran bagi data kepekatan keperluan biokimia oksigen bagi Sungai Johor dan Sungai Muar di Johor.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATIONS	MEANING
DOE	Department of Environment
BOD	Biochemical Oxygen Demand
INWQS	Interim National Water Quality Standards
WQI	Water Quality Index
NLR	Nonlinear Regression
PS	point sources
NPS	non-point sources
LSE	Least Squares Estimates
MSE	Mean Squared Error
WHO	World Health Organization
UNICEF	United Nations Children's Fund
NWRC	National Water Resources Council
APHA	American Public Health Association
AN	Ammoniacal Nitrogen
COD	Chemical Oxygen Demand
DO	Dissolved Oxygen
TSS	Total Suspended Solid
USEPA	U. S. Environmental Protection Agency

LIST OF SYMBOLS

SYMBOLS	MEANING
%	Percent
y	Response or dependent variable
x	Predictor or explanatory variable
km^3	Cubic kilometer
km^2	Square kilometer
km	Kilometer
°	Degree
m^3	Cubic meter
mg/L	Milligram per liter
L	Amount of first order BOD remaining in wastewater at time t, mg/L
k_1	BOD reaction rate constant, $time^{-1}$
t	Time
$f(t, \theta)$	Predicted BOD, mg/L
μ	Mean
n	Number of observations
σ	Standard deviation
B_t	BOD concentration level at time t, mg/L
B_0	Ultimate BOD, mg/L
E	Disturbance or error term

CHAPTER 1

INTRODUCTION

1.1 Introduction

A river is a large natural stream of water flowing into an ocean, lake, wetland, sea or other body of water and usually fed along its course by converging tributaries. Small rivers may sometimes be called by several other names, such as stream, creek, brook, rivulet and rill.

Rivers are one of the important sources of water for households, industry and agriculture, and are used for navigation and recreational activities, including boating, angling and walking. Rivers also support a diverse flora and fauna. Although rivers contain only about 0.0001% of the total amount of water in the world at any given time, they are vital carries of water and nutrients to areas all around the earth. Rivers also play a very important role in the water cycle, acting as drainage channels for surface water. The world's rivers drain nearly 75% of the land's surface.

Many rivers have been polluted because of the rapid growth in industrialization to support the country's growing population and economy. Domestic and industrial sewage, agricultural wastes have polluted the rivers. According to Chapra (1997) each of these uses will affect the water quality.

Industries discharge their liquid waste products into rivers. Agriculture with chemical fertilizers and pesticides usage contribute to river pollution as rainwater drains these chemicals into the rivers. The size of towns and cities will grow as the population grows. Hence, the amount of domestic wastes that are thrown into rivers

will increase and add the level of pollution. It will pose a serious health problem if people continue to depend on this polluted water from the river. Pollution of river will also affect the reproductive ability of animal and fish species in rivers thus causes extinction in future.

Recent studies show that most of the rivers from all over the world, including Malaysia, are polluted. Based on the Malaysia Environment Quality Report 2007, the Department of Environment (DOE) has described that one of the major pollutants is Biochemical Oxygen Demand (BOD). Data from DOE in 2004, based on BOD, 18 river basins were classified polluted, 37 river basins were slightly polluted and 65 river basins were in clean condition. According to the Malaysian River Classification and River Water Quality Monitoring Project Reports, the general trend of the overall river water quality in Malaysia is deteriorating. There are several factors, such as domestic and industrial sewage, agricultural wastes and discharge from livestock and heavy metal, will influence the river water quality which will lead to deterioration of river water quality (Abu Bakar and Dalilah, 2007). Figure 1.1 shows the water quality status for river basins of Peninsular Malaysia.

Figure 1.1 Water quality status for river basins of Peninsular Malaysia

According to the Interim National Water Quality Standards (INWQS), water quality of the main river basins in Malaysia was assessed based on the Water Quality Index (WQI) score using Roman numerals I, II, III, IV and V, where had been recommended as very clean, clean, moderate, slightly polluted and severely polluted respectively.

Nowadays several water quality models exist and have been applied in many regions from all over the world for various purposes. It has become a priority to model the water quality for simulating the level of river water pollution. Many researchers had created a large number of water quality models. The water quality models are not only for assessing water quality and detecting trends of water quality parameters, but for identifying the impact on water quality of the various potential alternative actions. Water quality models can be classified based on the type of approach, pollutant items, area of applications, nature etc. (Tsakiris and Alexakis, 2012). Many different types of models are used for different problems and purposes, appropriate model and data are chosen based on the purpose of the specific study.

In this research, two models are chosen to fit the data of the two rivers in Johor state. The two models are Streeter-Phelps model and nonlinear regression (NLR) model. Streeter-Phelps model and NLR model have been applied in many fields such as management science, physics, biology, electronics, engineering, economics and psychology, and in operations research. They have spread widely over queuing, finance and insurance.

A thorough grounding in linear regression (LR) is fundamental to understanding NLR. LR is the most widely used of all statistically techniques. LR is an approach for modeling the relationship between a dependent variable, y, and one or more explanatory variables denoted by x. The case of one explanatory variable is called simple linear regression. For more than one explanatory variable, the process is called multiple linear regressions (Freedman, 2009).

The basic idea of the NLR is almost the same as in the LR, namely to relate a response or dependent variable, *y*, to a vector of predictor variables or explanatory variables *x*. NLR is characterized by the fact that the prediction equation depends nonlinearly on one or more unknown parameters. On the other hand, LR is often used for building a purely empirical model. NLR usually arises when there are physical reasons for believing that the relationship between the response and the predictors follows a particular functional form (Smyth, 2002).

1.2 Background of the Study

River pollution is one the most common hazard in many countries in the world, which includes Malaysia. There are two main sources in contributing to the river water pollution, which are point sources (PS) and non-point sources (NPS). The point sources consist of detectable sources pollution component such as domestic waste water discharge and industrial waste water discharge. Non-point sources are undetectable pollution sources such as surface runoffs, agriculture activities and so on.

In recent times, the rapid growth in industrialization to support the country's growing population and economy had caused river pollutions due to domestic and industrial sewage, agricultural wastes. Activities such as deforestation due to illegal logging, mining, clearance of land for various usage like agriculture, housing and industrial purposes had ruined the ecosystem and increased the sediments in river. The sediments contain other materials such as organics matter, nutrients and toxics. It will increase the BOD level of river. BOD is one of the indicators for river pollution besides Ammoniacal Nitrogen (AN), Chemical Oxygen Demand (COD), Dissolved Oxygen (DO), Total Suspended Solid (TSS) and pH. The water quality model of the rivers is very important in order to solve the problem of deterioration of

river water. There are many types of water quality models due to wide variety of river system. Each type of river water body needs the appropriate type of model.

In this research, the study gives an understanding to the potential change in the data of BOD from year 1981 to year 1990. The study aims to benefit environment engineer, scientist, specialist and agency in designing and managing the river systems based on the expected BOD data over a certain period of time.

To estimate the parameters for the Streeter-Phelps model and NLR model, this study focuses on the weighted least squares and Least Squares Estimates (LSE) methods respectively. The advantage of using LSE is that the method is fairly simple and produces consistent estimators.

1.3 Problem Statement

A study has been done by DOE on 116 rivers nationwide. Some 10 percent of these rivers are heavily polluted or dead, 63 percent are polluted and only 27 percent are healthy. The study also showed that 70 percent of the pollution is caused by human activities such as deforestation, dumping and logging. Furthermore, arising of new development land use for agriculture and sand extractions may cause a lot of pollution to river water quality.

As Malaysia is a fast becoming an industrial country, many of the rivers have become polluted due to a lot of wastes that have been discharged into the rivers. Water pollution is very harmful to human, animals and water life.

In order to model the water quality of the river, two models were used in this study, namely Streeter-Phelps model and NLR. LR is a powerful method for analyzing data described by models which are linear in the parameters. It has a mathematical expression which relates the response to the predictor variables, and these models are usually nonlinear in the parameters. In such cases, linear regression techniques must be extended, which introduces considerable complexity.

In this study, information related to BOD concentration data is valuable towards designing water treatment structures.

1.4 Objectives of the Study

The main objectives of the study are:

1. to estimate the parameters of Streeter-Phelps model and NLR model using weighted least squares and LSE.
2. to compare which model is the best fit to the data of the two rivers in Johor state using Mean Square Error (MSE).

1.5 Scope of Study

In this study, Streeter-Phelps model and NLR model were used. The BOD concentration data for the two rivers in Johor state from year 1981 to year 1990 is analyzed. The values of the parameter estimations of the Streeter-Phelps and NLR

models had been estimated by using the weighted least squares and LSE methods. The Gauss-Newton method is used to estimate the parameters in the NLR.

Finally, to find the best model in this study, MSE will be utilized.

1.6 Significance of Study

Analysis of river water pollution is significant to many as it is beneficial for managing the consumption of water, improving the river basins and predicting the level of water pollution in Malaysia. Better estimation will assist environment engineer, scientist or specialist to design and manage the river water resources more efficiently. The analysis and information about river water quality can be used to evaluate and measure the level of water pollution so that the appropriate actions will be taken once the degree of the related river is polluted or severely polluted is found.

Moreover, environmental agencies will be able to predict the practical material that should be used for water treatment to improve the water quality. From the information received, the relevant authorities can make better decision to control water pollution and provide an alternative to better water quality management. Besides, it is hoped that the study may contribute towards preventing unnecessary costs and economic damages as well as avoiding danger and hazard as a result of river water pollution in the country.

1.7 Thesis Organization

This thesis is organized into five chapters. Chapter 1 is the introduction of this thesis. It highlights the background of the study and the problem statement. Objectives and scopes of this study are also mentioned in this chapter together with significance of study.

Chapter 2 outlines literature review of this study. It starts with a discussion on the water resources and river basins management in Malaysia, followed by an introduction of Sungai Johor in Kota Tinggi and Sungai Muar in Segamat which are the scope of this study. This chapter discusses the basic knowledge of BOD, the application of Streeter-Phelps model and NLR model in the water quality modeling done by earlier researchers. In this chapter, previous work and studies that are related to the research is highlighted.

Chapter 3 includes the methodology used to estimate the unknown parameters in Streeter-Phelps and NLR models. For the parameter estimation in the NLR model, the Gauss-Newton method is introduced and calculated. All the steps that are considered and discussed in Chapter 3 are summarized in the research framework in Section 3.3. A detailed explanation about the data sources of the study is also presented.

Chapter 4 presents the descriptive statistics of the BOD concentration data and the analysis of the results are reported. Next, the values of the parameter estimations of the Streeter-Phelps and NLR models are estimated using the weighted least squares and LSE methods. Finally, the performances of the two models are compared by using mean squared error (MSE).

Lastly, Chapter 5 discusses the conclusion of this dissertation. A few recommendations for the future studies are also presented at the end of the chapter.

REFERENCES

Abu Bakar Yang and Dalilah Haji Dali. (2007). *Alam Sekitar Anugerah Tuhan.* (3rd ed.) Malaysia: Jabatan Alam Sekitar.

Adamowski, J., Hiu, F. C., Prasher, S. O., Ozga-Zielinski, B. and Sliusarieva, A. (2012). Comparison of Multiple Linear and Nonlinear Regression, Autoregressive Integrated Moving Average, Artificial Neural Network, and Wavelet Artificial Neural Network Methods for Urban Water Demand Forecasting in Montreal, Canada, *Water Resources Research*, Vol. 48, W01528, doi: 10.1029/2010WR009945.

Baird, R. B. and Smith, R. K. (2002). *Third Century of Biochemical Oxygen Demand*. Water Environment Federation.

Bates, D. M. and Watts, D. G. (1988). *Nonlinear Regression Analysis and Its Applications*. John Wiley and Sons.

Bhargava, D. D. (1983). Most Rapid BOD Assimilation in Ganga and Yamuna Rivers. *Journal of Environmental Engineering, ASCE*. 109, 174-187.

Box, G. E. P. (1960). Fitting Empirical Data. *Annals of the New York Academy of Sciences* 86: 792-816.

Butts, T. A. and Kothandaraman, V. (1970). *Water and Sewage Works*. Vol. 117, No. 8, p. 276.

Camp. (1963). First Expanded BOD-DO Model. *Basic River Water Quality Models, IHP-V Project 8.1*, pp. 27-29.

Chapra, S. C. (1997). *Surface Water Quality Modeling*. McGraw–Hill Series in Water Resources and Environmental Engineering. New York: McGraw-Hill Companies, Inc.

Clark, J. W. and Viessman, W. (1965). *Water Supply and Pollution Control*. Intl. Textbook Company, pp. 387-390

Crabtree, R. W., Cluckie, I. D., Forster, C. F., and Crockett, C. P. (1986). A Comparison of Two River Quality Models. *Water Research*. 20:53-61.

Department of Environment Malaysia. (1985). Development of Water Quality Criteria and Standards for Malaysia.

Department of Environment Malaysia. (2007). Malaysia Environmental Quality Report 2006. Chapter 3: River Water Quality. Sasyaz Holdings Sdn Bhd, pp. 24.

Freedman, D. A. (2009). *Statistical Models: Theory and Practice*. Cambridge University Press. p. 26.

Gundelach, J. M. and Castillo, J. E. (1976). Natural Stream Purification Under Anaerobic Conditions. *J. Water Poll. Contr. Fed*. 48(7), 1753-1758.

Hach, C. C., Klein, R. L. and Gibbs, C. R. (1997). *Introduction to Biochemical Oxygen Demand.* USA: Hach Company.

Haider, H., Al, W. and Haydar, S. (2013). A Review of Dissolved Oxygen and Biochemical Oxygen Demand Models for Large Rivers. *Pak. J. Engg. & Appl. Sci*. Vol. 12 Jan., p. 127-142.

Hartley, H. O. (1961). The Modified Gauss-Newton Method for the Fitting of Non-linear Regression Functions by Least Squares. *Techmometrics* 3: 269-280.

Izamudin Mohamed. (2011). *Water Quality Assesment of Sungai Skudai.* Bachelor Thesis. Universiti Teknologi Malaysia, Skudai.

Jha, R., Ojha, C. S. P. and Bhatia, K. K. K. (2007). Development of Refined BOD and DO Models for Highly Polluted Kali River in India. *Journal of Environmental Engineering*. 133(8), 839-852.

Jolanki, G. (1997). Basic River Water Quality Models. *IHP-V, Technical Documents in Hidrology*. No. 13. UNESCO, Paris.

Kojiri, T. and Teramura, T. (2003). Integrated River Basin Environment Assessment and Planning Through Hybrid Processes. XI IWRA World Water Congress.

Le, T. V. (2005). *Water Quality Modeling for Unconventional BOD*. Master Thesis. Louisiana State University.

Li, J., Liu, H., Li, Y., Mei, K., Dahlgren, R. and Zhang, M. (2012). *Monitoring and Modeling Dissolved Oxygen Dynamic Through Continuous Longitudinal Sampling: A Case Study in Wen-Rui Tang River, Wenzhou, China.* Hydrological Processes.

Li, W. H. (1972). Effect of Dispersion on DO-sag in Uniform Flow. *Journal of Sanitary Engineering Division. American Society of Civil Engineers*. Vol. 98, No. SA1, pp. 169-182.

Maketab Mohamed. (2008). Water Quality Models in River Management. *Proceedings of the 1st Technical Meeting of Muslim Water Researchers Cooperation (MUWAREC).* December. Johor, Malaysia, 14-26.

Maldeniv, N., Strzepek, K. and Serumola, O. M. (2005). Water Quality Assessment and Modeling of an Effluent Dominating Stream. The Notwane River, Botswana. Environmental Monitoring and Assessment. 109, 97-121.

Marske, D. (1967). *Biochemical Oxygen Demand Data Interpretation Using Sum of Squares Surface*. PhD Thesis. University of Wisconsin-Madison.

Marske, D. M. and Polkowski, L. B. (1972). Evaluation of Methods for Estimating Biochemical Oxygen Demand Parameters. *Journal Water Pollution Control Federation*. Vol. 44, No. 10, pp. 1987-1999.

McCutcheon, S. C. (1989). *Water Quality Modeling. Vol. I Transport and Surface Exchange in Rivers*. CRC Press. Boca Raton. FL. P. 141.

Mcginley, M. (2009). *Water Profile for Malaysia*. Retrieved July 10, 2012, from Encyclopedia of Earth: http://www.eoearth,org/article/Malaysia.

Mondragon, P. F. (2003). *A Comparison of Nonlinear Regression Codes*. Master Thesis. New Mexico Institute of Mining and Technology Socorro, New Mexico.

Nas, S. S. and Nas, E. (2009). Water Quality Modeling and Dissolved Oxygen Balance in Streams: A Point Source Streeter-Phelps Application in the Case of the Harsit Stream. *CLEAN-Soil, Air, Water*. Vol. 37, Issue 1, pages 67-74.

Nemerow, N. L. (1974). *Scientific Stream Pollution Analysis*. McGraw-Hill.

NIST. (2013, October 10). e-Handbook of Statistical Methods. *Engineering Statistics Handbooks*. Retrieved May 5, 2015, from http://www.itl.nist.gov/div898/handbook/

Nur Atikah Mohd Rohani. (2012). *Study on Water Quality at Manjung River and Its Tributaries.* Bachelor Thesis. Universiti Teknologi Petronas, Perak.

Nurul Izzaty Binti Hj. Mohd Yunus. (2014). *Stochastic Modeling for River Pollution of Sungai Perlis.* Master Thesis. Universiti Teknologi Malaysia.

Omole, D. O. and Longe, E. O. (2012). Re-aeration Coefficient Modeling: A Case Study of River Atuwara in Nigeria. *Research Journal of Applied Sciences, Engineering and Technology. 4(10), 1237 – 1243.*

Penn, M. R., Pauer, J. J. and Mihelcic, J. R. (2009). Biochemical Oxygen Demand. *Environmental and Ecological Chemistry*. II.

Raja Zaharaton Raja Zainal Abidin. (2008). *Water Resources Management in Malaysia – The Way Forward*. Kuala Lumpur: The Board of Engineers Malaysia.

ReVelle, P. and ReVelle, C. (1988). *The Environment – Issues and Choices for Society:* Boston, Jones and Barlett, 749p.

Ritter, C. and Tanner, M. A. (1992). Facilitating the Gibbs Sampler: The Gibbs Stopper and the Griddy-Gibbs Sampler. *Journal of the American Statistical Association.* 87, 861-868.

Saffran, K., Cash, K., Hallard, K., Neary, B. and Wright, R. (2001). *CCME Water Quality Index1.0 User's Manual.* Canadian Water Quality Guidelines for the Protection of Aquatic Life. Canadian Environmental Quality Guidelines. Canadian Council of Ministers of the Environment.

Singh, B. (2004). *Determination of BOD Kinetic Parameters and Evaluation of Alternate Methods*. Master Thesis. Thapar Institute of Engineering & Technology, Patiala.

Smyth, G. K. (2002). Nonlinear Regression: *Encyclopedia of Environmetrics*. 3, 1405 – 1411.

Streeter, H. W. and Phelps, E. P. (1925). *A Study of the Pollution and Natural Purification of the Ohio River.* United States: Public Health Service.

Tebbutt, T. H. Y. and Berkun, M. (1976). Respirometric Determination of BOD. *Water Research*. Vol. 10, pp. 613-617.

Thomann, R. V. (1974). Systems Analysis and Water Quality Management. McGraw-Hill. New York, NY, p. 140.

Thomas, H. A. (1957). Hydrology and Oxygen Economy in Stream Purification. Seminar on Waste Water Treatment and Disposal. Boston Society of Civil Engineers. Boston. MA.

Tsakiris, G. and Alexakis, D. (2012). Water Quality Models: An Overview. *European Water.* 37, 33 – 46.

Young, J. C. and Clark, J. W. (1965). Second Order Equation for BOD. *Journal of the Sanitary Engineering Division Proceedings of the American Society of Civil Engineers*. Vol. 91, No. SA1, pp. 43-57.

Zaki Zainudin. (2010). Benchmarking River Water Quality in Malaysia. *Jurutera*. 12-15.