Faculty of Engineering

# ESTIMATING MISSING PRECIPITATION TO OPTIMIZE PARAMETERS FOR PREDICTION OF DAILY WATER LEVEL USING ARTIFICIAL NEURAL NETWORK

Dayang Suhaila Binti Awang Suhaili



Bachelor of Engineering with Honours
(Civil Engineering)
2006

**Universiti Malaysia Sarawak**
Kota Samarahan

Fk

## BORANG PENYERAHAN TESIS

Judul: ESTIMATING MISSING PRECIPITATION TO OPTIMIZE PARAMETERS FOR PREDICTION OF DAILY WATER LEVEL USING ARTIFICIAL NEURAL NETWORK

**SESI PENGAJIAN: 2002 - 2006**

Saya DAYANG SUHAILA BINTI AWANG SUHAILI
**(HURUF BESAR)**

mengaku membenarkan laporan projek ini disimpan di Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dengan syarat-syarat kegunaan seperti berikut:

1. Hakmilik kertas projek adalah di bawah nama penulis melainkan penulisan sebagai projek bersama dan dibiayai oleh UNIMAS, hakmiliknya adalah kepunyaan UNIMAS.
2. Naskhah salinan di dalam bentuk kertas atau mikro hanya boleh dibuat dengan kebenaran bertulis daripada penulis.
3. Pusat Khidmat Maklumat Akademik, UNIMAS dibenarkan membuat salinan untuk pengajian mereka.
4. Kertas projek hanya boleh diterbitkan dengan kebenaran penulis. Bayaran royalti adalah mengikut kadar yang dipersetujui kelak.
5. * Saya membenarkan/tidak membenarkan Perpustakaan membuat salinan kertas projek ini sebagai bahan pertukaran di antara institusi pengajian tinggi.
6. ** Sila tandakan (✓)

| | | |
|---|---|---|
| ☐ | SULIT | (Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972). |
| ☐ | TERHAD | (Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/ badan di mana penyelidikan dijalankan). |
| ☑ | TIDAK TERHAD | |

Disahkan oleh

(TANDATANGAN PENULIS) | (TANDATANGAN PENYELIA)

Alamat tetap: 4/41 Kampung Kaba,
95400 Saratok, SARAWAK

PUAN ROSMINA AHMAD BUSTAMI
( Nama Penyelia )

Tarikh: 17 · 5 · 2006

Tarikh: 12 June 2006

CATATAN * **Potong yang tidak berkenaan.**
** **Jika Kertas Projek ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/ organisasi berkenaan dengan menyertakan sekali tempoh kertas projek. Ini perlu dikelaskan sebagai SULIT atau TERHAD.**

The following Final Year Project Report:

Title : ESTIMATING MISSING PRECIPITATION TO OPTIMIZE PARAMETERS FOR PREDICTION OF DAILY WATER LEVEL USING ARTIFICIAL NEURAL NETWORK

Author's Name : DAYANG SUHAILA BINTI AWANG SUHAILI

Matric No. : 8135

has been read and approved by:

Puan Rosmina binti Ahmad Bustami

Supervisor

12 June 2006

Date

# ESTIMATING MISSING PRECIPITATION TO OPTIMIZE PARAMETERS FOR PREDICTION OF DAILY WATER LEVEL USING ARTIFICIAL NEURAL NETWORK



DAYANG SUHAILA BINTI AWANG SUHAILI

This project is submitted in partial fulfillment of the requirements for
Bachelor of Engineering with Honors
(Civil Engineering)

Faculty of Engineering
UNIVERSITI MALAYSIA SARAWAK
2006

***Dedicate to my family and my beloved one...***

## ACKNOWLEDGEMENT

Fullest gartitude to God, the most Gracious and Merciful for giving me a chance and strength to complete this project on time.

I wish to express my greatest appreciation to my supportive supervisor, Puan Rosmina Ahmad Bustami for the invaluable guidance and kindly support, advices, comments and suggestions throughout the process of completing my project.

I also want to thank my beloved family for all moral and financial supports within this year.

Last but not least, thank you to all my friends who have shared their suggestions and evaluations of this script.

Finally, I am deeply grateful to those individuals who involve directly and indirectly all throughout the process of doing this project.

Thank you so much.

## ABSTRACT

This study proposes the application of Artificial Neural Network (ANN) in predicting missing precipitation to predicting daily water level for Sg. Bedup station located in Batang Sadong Basin, Sarawak. ANN is undoubtedly a strong tool for forecasting various non-linear hydrologic processes, including the missing precipitation and water level prediction. ANN was chosen based on its ability to extract the relation between the inputs and outputs of a process without the physics known explicitly. In this study, the ANN was developed specifically to predict the daily missing precipitation and data simulated are utilized to optimize prediction accuracy for daily water level. Typical networks were trained and tested using daily data obtained from the Drainage and Irrigation Department (DID) Kota Samarahan. Various training parameters were considered in order to gain the best prediction possible. The performances of the ANN were evaluated based on the coefficient of correlation, R. The back propagation algorithm was adopted for this study. The optimal model for predicting missing data found in this study is the network with the combination of learning rate and the number of neurons in the hidden layer of 0.2 and 60. This model generated the highest coefficient of correlation value of 0.964 when trained with the The Resilient Back propagation (trainrp). It has been found that the ANN has the potential to solve the problems of estimation missing precipitation in predicting daily water level. After appropriate trainings, they are able to generate satisfactory results during both of the training and testing phases.

**ABSTRAK**

Kajian ini mengaplikasikan penggunaan *Artificial Neural Network* (ANN) untuk meramal data curah hujan yang tidak lengkap dan meramal paras air untuk Sungai Bedup. ANN merupakan salah satu alternatif yang efektif dalam meramal pelbagai proses hidrologi yang tidak seragam. Ini termasuklah meramal data curah hujan yang tidak lengkap dan meramal paras air sungai-sungai. ANN dipilih berdasarkan kebolehan untuk mengekstrak hubungan antara proses input dan output tanpa menggunakan kaedah fizik. Dalam kajian ini, ANN dibangunkan secara terperinci untuk meramal data curah hujan yang tidak lengkap dan data yang diramal digunakan untuk meramal paras harian air untuk Stesen Sungai Bedup. Rangkaian yang berbeza dilatih dan diuji dengan menggunakan data setiap hari yang diperolehi daripada Jabatan Pengairan dan Saliran, Kota Samarahan. Pelbagai parameter latihan diambil kira untuk mencapai keputusan ramalan yang terbaik. Prestasi ANN dinilai berdasarkan Pekali Perkaitan, R. Algoritma '*back propagation*' telah diaplikasikan dalam kajian ini. Nilai terbaik bagi R untuk fasa ujian bagi meramal data curah hujan yang tidak lengkap telah dicapai oleh rangkaian yang menggunakan '*learning rate*' 0.2 dan bilangan neuron 60. Rangkaian ini telah dilatih dengan '*trainrp*'. Setelah melaksanakan latihan yang sesuai, keputusan yang memuaskan telah dicapai untuk kedua-dua fasa latihan dan ujian.

# LIST OF CONTENTS

## CHAPTER 2

## LITERATURE REVIEW



## CHAPTER 3

## METHODOLOGY

# LIST OF FIGURE

# LIST OF TABLE

# ABBREVIATIONS

| | | |
|---|---|---|
| ANN | - | Artificial Neural Network |
| DID | - | Department of Irrigation and Drainage |
| HEC-HMS | - | Hydrologic Engineering Centre-Hydrologic Modeling System |
| IVF | - | Index of Volumetric Fit |
| MLP | - | Multiple Layer Perceptron |
| NWS | - | National Weather Service |
| OLS | - | Orthogonal Least Square Algorithm |
| POP | - | The probability of precipitation |
| QPF | - | Quantitative Precipitation Forecast |
| R | - | Coefficient of Correlation |
| RBF | - | Radial Basis Function |
| SOFM | - | Organising Feature Maps |
| Trainrp | - | The resilient back propagation Algorithm |
| Trainscg | - | The scaled Conjugate Gradient Algorithm |
| Traincgf | - | The Fletcher-Reeves Update |

# CHAPTER 1

# INTRODUCTION

## 1.1 INTRODUCTION

A prediction of high water condition is one of the most essential hydrological tasks for a river basin management and is generally performed by means of traditional conceptual and deterministic models using predicted precipitation. It is important to predict water level because water level is significance to the ecosystem along the river basin especially for low-lying area that always flooding. Besides, river basin plays an important role in development and economical aspect such as agricultural and fisheries thus making the task of predicting water level become significant.

A precise estimation of water level needs an accurate estimation of the runoff from a given precipitation event and an accurate hydraulic model to predict the water level for a given discharge. The use of precipitation data is essential and fundamental to the rainfall-runoff process. The precipitation data are the driving force in the relationship. The accuracy of the precipitation data at a point (i.e., at the rain gauge) is extremely significant to all the remaining use of the data. After acquiring a set of point precipitation data, it is necessary to first verify the data before using it for analysis or design. The data set should be checked for consistency and for missing data. The missing data should be replaced if possible and for the inconsistent data, they should be adjusted. Thus, it is proposed that Artificial Neural Network (ANN) is used to estimate missing precipitation to optimize parameters for prediction daily water level.

## 1.2 SELECTION OF ARTIFICIAL NEURAL NETWORK

In the new era of technology, Artificial Neural Network (ANN) is one of those words that are getting fashionable. ANN has become an increasingly popular field of research in many branches of science. These include computer engineering and computer science, signal processing, information theory, and physics. Besides that ANN also apply extensively in the hydrological field.

The reason why ANN becomes fashionable is because they are able to approximate any function to any degree of accuracy given internal nodes (Sandhu and Finch, 1996). Furthermore, ANN was chosen based on its ability to generalized patterns in imprecise or noisy and ambiguous input and output data sets. Mathematically, an ANN may be treat as a universal approximator. ANN has an ability to learn from example and generalize and it makes ANN possible to solve a complex problem applied in hydrology today such as pattern recognition, nonlinear modeling, classification, association, control, and other.

Lately, it is found that ANN is a strong tool for modeling many of the nonlinear hydrologic processes. ANN is suitable to perform a kind of function fitting by using multiple parameters on the existing information and predict the possible relationships in the coming future, if significant variables are known, without knowing the exact relationships. This sort of problem includes rainfall-runoff prediction, water level and discharge relations, flow and sediment transport, water quality prediction etc. Besides that ANN also filling or restoring of missing data in a time series can be considered as a kind of prediction.

Generally, ANN is one of the most popular data-driven techniques attributed by various authors to machine learning, data mining, soft computing etc. An ANN is an information processing system that roughly replicates the behavior of a human brain by emulating the operations and connectivity of biological neurons (Tsoukalas and Uhrig, 1997). It performs a human-like reasoning, learns the attitude and stores the relationship of the processes on the basis of a representative data set that already exists. ANN has certain performance characteristics resembling biological neural network of the human brain because ANN is a massively parallel-distributed information processing. ANN has been developed as a generalization of mathematical models of human cognition or neural biology. A neural network is characterized by its architecture that represents the pattern of connection between nodes, its method of determining the connection weight and the activation function (Fausett 1994). Caudill presented a comprehensive description of neural networks in a series of papers (Caudill, 1987, 1988, 1989).

For the traditional models, a great deal of detailed data it is required, for example, topographical maps, river networks and characteristics, soil characteristics, rainfall and runoff data. Frequently, for model calibration, these data are not available and pose a great difficulty. In addition, a sufficiently long lead-time for forecasting is required to take the necessary flood evacuation measures. For dissemination of flood information and other flood evacuation measures, computational speed of the models used are of absolute significance.

In hydrology field, the problems are not clearly understood or are too complex for an analysis using traditional methods. Even when such models are available, they have to rely on assumptions that make ANN more attractive. The presents of noise in the inputs and outputs is handled by an ANN without severe loss of accuracy because of distributed processing within the network. This, along with the nonlinear nature of the activation function, truly enhances the generalizing capabilities of ANN and makes them desirable for a large class of problems in hydrology. Hence, the application of ANN in hydrology for predicting the water level is a great alternative in order to achieve the best result possible.

## 1.3 OBJECTIVE OF THE STUDY

The objective of this study is to estimate missing precipitation in predicting daily water level using Artificial Neural Network for Sungai Bedup station located in Batang Sadong Basin, Sarawak.