UNIVERSITI MALAYSIA SARAWAK
UNIMAS

Faculty of Engineering

DAILY RAINFALL RUNOFF MODELING USING ARTIFICIAL NEURAL NETWORK FOR SUNGAI SARAWAK KANAN UPPER CATCHMENT

ONG KHIN KIAT

Bachelor of Engineering with Honours
(Civil Engineering)
2005

UNIVERSITI MALAYSIA SARAWAK

BORANG PENGESAHAN STATUS TESIS

Judul: **Daily Rainfall Runoff Modeling Using Artificial Neural Network For Sungai Sarawak Kanan Upper Catchment**

Sesi Pengajian: 2004 - 2005

Saya **ONG KHIN KIAT**

(HURUF BESAR)

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

1. Tesis adalah hakmilik Universiti Malaysia Sarawak.
2. Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Membuat pendigitan untuk membangunkan Pangkalan Data Kandungan Tempatan.
4. Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
5. ** Sila tandakan (√) di kotak berkenaan

[ ]	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: 12, Taman Mewah,
Jalan Arang,
93250 Kuching.

Assc. Prof. Dr. Nabil Bessaih

Nama Penyelia

Tarikh: 4/4/05

Tarikh: 04-04-05

CATATAN
* Tesis dimaksudkan sebagai tesis bagi Ijazah Doktor Falsafah, Sarjana dan Sarjana Muda
** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh tesis perlu dikelaskan sebagai SULIT atau TERHAD

APPROVAL SHEET

Report of Final Year Project as followed:

Title: DAILY RAINFALL RUNOFF MODELING USING ARTIFICIAL NEURAL NETWORK FOR SUNGAI SARAWAK KANAN UPPER CATCHMENT

Author: ONG KHIN KIAT

Matric Number: 7013

is hereby read and approved by:

Assc. Prof. Dr. Nabil Bessaih

Project Supervisor

04-04-05

Date

DAILY RAINFALL RUNOFF MODELING USING ARTIFICIAL NEURAL NETWORK FOR SUNGAI SARAWAK KANAN UPPER CATCHMENT

ONG KHIN KIAT

Tesis Dikemukakan Kepada

Fakulti Kejuruteraan, Universiti Malaysia Sarawak

Sebagai Memenuhi Sebahagian daripada Syarat

Penganugerahan Sarjana Muda Kejuruteraan

Dengan Kepujian (Kejuruteraan Sivil)

2005

To my beloved family and friends

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my utmost gratitude to my supervisor, Assoc. Prof. Dr. Nabil Bessaih for giving me invaluable advices, excellent guidance, encouragement and technical support in the completion of this thesis.

I would also like to express my appreciation to the Department of Irrigation and Drainage (DID) Sarawak for providing relevant data for model calibration.

Last but not least, I would like to thank my family and friends, who are always supportive and caring throughout this project.

ABSTRACT

This thesis reports the results predicted from artificial neural network (ANN) models for daily rainfall-runoff simulation, at Sungai Sarawak Kanan upper catchment. The system is monitored by four rainfall gauging stations namely Kampung Monggak, Krokong, Bau and Kampung Opar located upstream of the river system and one river stage gauging station namely Buan Bidi. Backpropagation network (BP) of multilayer perceptron (MLP) is used for daily runoff simulation. Input variables used are current rainfall, antecedent rainfall and antecedent runoff while the output is current runoff. Several networks were trained and tested using data obtained from Department of Irrigation and Drainage (DID) Sarawak. The effects of different types of training algorithms, different numbers of hidden neurons, different numbers of antecedent data and different numbers of hidden layers were investigated to find the optimal neural network. Judging on coefficient of correlation R, one layered training algorithm *trainoss* (R = 0.839) with 150 hidden neurons and 5 days backdated performed the best for the simulations. Therefore this make this study useful for heavy rainfall predictions and thus a good tool for flood warning.

ABSTRAK

Thesis ini melaporkan keputusan yang diperolehi daripada simulasi kadaralir harian dengan menggunakan model *artificial neural network (ANN)* di kawasan atasan Sungai Sarawak Kanan. Sistem ini terdiri daripada empat stesen kutipan air hujan yang terletak di Kampung Monggak, Krokong, Bau dan Kampung Opar pada bahagian muara sungai dan satu stesen ukur paras air yang bernama Buan Bidi. *Backpropagation network* (BP) dengan *multilayer perceptron* (MLP) digunakan untuk simulasi kadaralir harian. Input yang digunakan termasuk hujan semasa, hujan sebelumnya serta kadaralir sebelumnya manakala outputnya adalah kadaralir semasa. Beberapa *network* telah dilatih dan diuji dengan menggunakan data yang diperolehi daripada Jabatan Pengairan dan Saliran (JPS) Sarawak. Kesan-kesan yang disebabkan oleh *training algorithms*, bilangan *neurons*, bilangan *antecedent* data dan bilangan *hidden layers* telah disiasat untuk menentukan *neural network* yang paling optima. Berdasarkan pemalar *correlation* (R), satu lapisan *training algorithm trainoss* (R = 0.839) dengan 150 *neuron* dan 5 hari *backdated* telah mendapat keputusan yang terbaik. Maka ini menunjukkan kajian inin amat berguna untuk meramal hujan lebat pada masa depan dan seterusnya dapat bertindak sebagai satu amaran untuk banjir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDIXES

CHAPTER 1

INTRODUCTION

1.1 Background

In Sarawak, majority of its important industrial cities are located along rivers due to the need of water for human activities. However, the annual rainfall of Sarawak is very high, which is about 3500 mm per year, this makes Sarawak prone to flooding. In the last 40 years, there have been four significant hydrological events, which caused severe flooding throughout Sarawak River. During January and February of 1963, 2500 mm of heavy rainfall which lasted for 2 months occurred. In about the same period of year 1976, severe rainfall occurred once again and caused emergency evacuation of 2500 residents to safe areas. In that year, south of Kuching was flooded to a depth of 5 meters. In February 2003 and January 2004, excessive rainfalls occurred and caused Bau town severely flooded with water.

From the above events, we can see that droughts cause populations to be displaced and human suffering. Or at the other extreme, streams and rivers overflow into populated areas. Wildlife is displaced or destroyed. Damage to transportation systems occurs. And, unfortunately there is a long list of human sufferings resulting in excessive government expenditures when we are unable to predict excessive rainfalls. Thus, rainfall-runoff modeling systems are carried out to predict future runoffs in order to prevent floods.

Artificial Neural Networks (ANN) is introduced about 10 years ago and proved itself to be a useful technique in the hydrology field. The special characteristic of ANN is that when given sets of input-output pairs, the network is able to recognize the patterns in data and thus predicts the future outputs when given inputs. ANN *'tends to be data intensive'* (ASCE Task Committee, 2000). By giving sufficient data, ANN is renowned to be able to extract and adapt to the relation of a process through learning even if the data is 'noisy' and 'contaminated'. The above-mentioned characteristics of ANN make it suitable for most basins no matter how complex their characteristics are.

Several studies indicate that ANNs have proven to be a potentially useful tool in hydrologic modeling such as modeling of rainfall-runoff relationship (Hsu, Gupta and Sorooshian 1995, Mins and Hall 1996, Dawson and Wilby 1998); water demand forecasting (Zhang et. 1994); rainfall

forecasting (Luk, Ball and Sharma 1998); snowmelt-runoff modeling (Trent et al. 1993); assessment of stream's hydrologic and ecologic response to climate changes (Roger and Dowla 1994); sediment transport prediction (Poff et al. 1996); pier scour estimation (Tokar 1996); groundwater remediation (Markus 1997) and stage-discharge relationship (Bhattacharya and Solomatine 2000).

1.2 Artificial Neural Networks (ANNs)

The most basic elements of ANNs are called neurons, units, cells or nodes. Each neuron is linked to its neighbours with varying coefficients of connectivity, each with an associated weight that represents information used by the net to solve a problem. Neurons in an ANN are arranged in groups called layers. Each neuron in a layer operates in logical parallelism. Information is transmitted from one layer to others in serial operations (Hecht-Nielsen 1990). ANNs can be composed of one to many layers. Three basic layers of ANNs are:

a) Input layer

The layer that is responsible in introducing data to the network and then pass on to the hidden layer after multiplying by a weight.

b) Hidden layer

The layer where data are processed by adding up the weighted input received from each input nodes, associates it with a bias and then passes the result in through a nonlinear transfer function.

c) Output layer

The layer where the results of given inputs are produced.

The architecture of ANN is designed by weights between neurons that represent information used by the net to solve the problem, a transfer function that controls the generation of output in a neuron and learning laws that define the relative importance of weights for input to a neuron (Caudill 1987).

ANNs operate on the principle of learning from a training set. Before an ANN is trained, it doesn't have any prior knowledge about the problem. In order to solve the problem, they must be trained with a set of typical input/output pairs of data called the training set. At the beginning of training, the network weights are usually initialized with a set of random values.

Training consists of two major phases namely forward pass and reverse pass. The target output at each output node is compared with the network output. The difference is minimized by adjusting the weights and biases through some training algorithms. The objective of ANNs is to process the

information that is previously trained to generate satisfactory results. ANNs can learn from experience, generalize from previous examples to new ones and abstract essential characteristics from inputs containing irrelevant data. The main control parameters of ANNs model are interneuron connection (weights and biases). In most cases, the output layer had only one neuron, which is the runoff (Q).

After ANNs is trained, the relationship between inputs and outputs is encoded in the network. The final weight vector of a successfully trained ANNs represents its knowledge about the problem.

1.3 Project Objective

The objective of this project is to develop, train and test the rainfall-runoff model for the upper catchment of Sungai Sarawak Kanan basin, which is located in Sungai Sarawak basin by using Artificial Neural Networks, and thus in the process, the usefulness of Artificial Neural Networks in rainfall-runoff modeling is studied.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Relationship between rainfall and runoff is too complex and nonlinear. For rainfall, it depends on numerous factors such as initial moisture content, land use, watershed geomorphology, evaporation and infiltration. To gather all these data for computing rainfall is a difficult task. Thus, rainfall- runoff models are used to generate runoff. Over these years, researchers and hydrologists from all around the world look into the potential of artificial neural networks (ANNs) for rainfall- runoff modeling. Review of these studies is discussed in the following section.

2.2 Applications in Rainfall-Runoff Modeling

Shamseldin (1996) has used ANNs for daily rainfall-runoff modeling and then compared the results with simple linear model (SLM), seasonally based linear perturbation model (LPM) and nearest neighbour linear perturbation model (NNLPM). ANNs is tested on the synchronous rainfall and discharge data of six catchments from different geographical locations of the world. These six catchments are Sunkosi from Nepal, Yanbian from China, Shiquan from China, Brosna from Ireland, Bird Creek from USA and Wolombi Brook from Australia. Available data are divided for calibration (two third of total data) and verification (one third of total data).

The chosen form of ANNs is tested for four different input scenarios. These are:

a) The first inputs scenario {N1} are seasonal rainfall index ${}_sRI_i$ and current rainfall index RI_i.

b) The second inputs scenario {N2} are the seasonal expectation of the discharge ${}_sq_i$ and the corresponding seasonal rainfall index ${}_sRI_i$ as well as current rainfall index RI_i.

c) The third inputs scenario {N3} are current rainfall index RI_i, the mean nearest neighbours discharge ${}_nq_i$ and the corresponding rainfall index ${}_nRI_i$ which can be obtained from the mean rainfall

index and discharge of the RI band to which the current value of RI belongs. In this study, the number of the RI bands used is 250.

d) The forth input scenario {N4} are rainfall index RI_i, its seasonal expectation rainfall index ${}_sRI_i$, the mean nearest neighbours rainfall index ${}_nRI_i$, the discharge ${}_nq_i$ of the nearest neighbours and seasonal discharge expectation ${}_sq_i$.

Three-layer feedforward neural network with one hidden layer was chosen for training purpose. The transfer function used in hidden and output layers is the logistic function. Complex conjugate gradient algorithm is chosen for parameter estimation as it is generally faster and more efficient than the backpropagation algorithm.

The comparison of model performances are assessed in terms of mean square error (MSE), the R^2 modal efficiency criterion (Nash and Sutcliffe, 1970) and the results are shown in Table 2.1.

Catchments	Models	Calibration		Verification		Catchments	Models	Calibration		Verification	
		MSE	R^2 (%)	MSE	R^2 (%)			MSE	R^2 (%)	MSE	R^2 (%)
Sunkosi-1	SLM	3.157	82.37	4.248	81.81	Brosna	SLM	0.375	35.68	0.636	44.59
	LPM	1.439	91.97	2.135	90.86		LPM	0.181	68.93	0.277	75.88
	NNLPM	2.027	88.68	3.881	83.38		NNLPM	0.313	46.31	0.669	41.72
	N1	2.540	85.82	3.849	83.52		N1	0.361	38.15	0.645	43.81
	N2	1.340	92.52	2.155	90.77		N2	0.121	79.19	0.167	85.46
	N3	2.019	88.73	4.308	81.55		N3	0.325	44.28	0.692	39.64
	N4	1.073	94.01	2.273	90.27		N4	0.096	83.50	0.172	85.03
Yanbian	SLM	3.571	70.64	3.445	71.83	Bird Creek	SLM	3.226	58.90	1.472	-45.75
	LPM	2.277	81.82	2.870	76.53		LPM	3.012	61.63	1.369	-35.50
	NNLPM	2.589	78.70	3.778	69.10		NNLPM	1.290	83.58	0.771	23.60
	N1	3.140	74.18	3.327	72.79		N1	1.621	79.36	0.747	25.98
	N2	2.056	83.10	2.415	80.25		N2	1.544	80.33	1.050	-4.06
	N3	2.546	79.07	3.613	70.45		N3	1.385	82.36	0.886	12.20
	N4	1.623	86.65	2.656	78.28		N4	1.479	81.17	0.749	25.72
Shiquan-3	SLM	2.521	71.01	4.036	51.05	Wolombi Brook	SLM	1.830	44.85	1.081	-8.62
	LPM	2.206	74.64	4.195	49.13		LPM	1.709	48.51	1.146	-15.05
	NNLPM	1.289	85.18	2.390	71.01		NNLPM	0.697	78.99	0.282	71.65
	N1	1.619	81.38	2.717	67.04		N1	0.845	74.55	0.231	76.80
	N2	1.385	84.07	2.872	65.17		N2	0.505	84.78	1.127	-13.20
	N3	1.286	85.21	3.520	57.32		N3	0.593	82.13	0.479	51.84
	N4	1.113	87.20	4.080	50.51		N4	0.173	94.80	0.678	31.84

Table 2.1 Summary of results of the SLM, LPM, NNLPM and four neural network forms N1, N2, N3, and N4.

SLM performed the worst in all the six catchments in calibration period. However, in verification period, it is the worst only in the case of Bird Creek catchment where it has a negative R^2 value of -45.75%. These results confirmed that the SLM is unsatisfactory for these six test catchments.

In calibration period, one or other form of neural network N1, N2, N3 or N4 has the highest R^2 value except for Bird Creek catchment where its R^2 value of 82.36% is not notably different from 83.58% for the NNLPM. However, in verification period, one or other form of neural network has the highest R^2 in four out of the six test catchments especially Yanbian, Brosna,

Bird Creek and Wolombi Brook catchments. These results revealed that the neural network is an efficient tool for river flow forecasting.

The results also indicate that N1 has performed better than SLM for both calibration and verification periods except in the case of verification period for Brosna catchment. The R^2 value of 43.81% for N1 is just marginally lower than 44.59% for SLM. N2 has also shown better performance than LPM in all the six catchments in calibration period. Similarly in verification period, N2 is performing better than the LPM except for two catchments namely Sunkosi and Wolombi Brook catchments. These results indicated that N2 is much capable than LPM for extracting more information from the same input data.

In calibration period, N3 also shows better results than NNLPM in four out of six catchments namely Sunkosi, Yanbian, Shiquan and Wolombi Brook catchments. However in verification period, NNLPM has higher R^2 values than N3 in five out of six catchments except for Yanbian catachment.

Incorporation of the seasonal information of the rainfall and the discharge time series, improved the performance of neural network. N2 has shown better R^2 results than N1 for all the test catchments in calibration period. In verification period, N2 has performed better than N1 in three catchments. However in the case of the nearest neighbours information, N3

has also performed better than N1 in calibration period for all the six catchments. In verification period, this situation is reversed.

N4 has the highest R^2 efficiency values compared to N1, N2 and N3 in all the six catchments except for Bird Creek catchment in calibration period. In verification period, N4 has R^2 values which are in the vicinity of the maximum R^2 among other three forms of ANNs in four catchments namely Yanbian, Sunkosi-1, Brosna and Bird Creek.

On the other hand, N4 has the highest R^2 values in all the six test catchments except for Bird Creek catchment in calibration period when compared with SLM, LPM, and NNLPM models. Nonetheless in verification period, N4 has the highest R^2 values in three catchments namely Brosna, Yanbian and Bird Creek catchments.

Results also shown that N4 has not always yield the best performance when compared with SLM, LPM and NNLPM. This may be because the data used in this study does not support a more sophisticated model. It was concluded that ANNs has considerable potential in rainfall-runoff modeling but like all such models, has variable results.