Faculty of Engineering

MODELLING HOURLY RUNOFF USING ANN FOR SG. SARAWAK KANAN BASIN

CHONG KAH WENG

Bachelor of Engineering with Honours
(Civil Engineering)
2005

UNIVERSITI MALAYSIA SARAWAK

BORANG PENYERAHAN STATUS TESIS

Judul: **Modelling Hourly Runoff Using ANN for Sg. Sarawak Kanan Basin**

Sesi Pengajian: 2004 - 2005

Saya **CHONG KAH WENG**

(HURUF BESAR)

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

1. Tesis adalah hamilik Universiti Malaysia Sarawak.
2. Pusat Khidmat Maklumat Akademik, Universiti Malaysia Sarawak dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Membuat pengdigitan untuk membangunkan Pengkalan 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:
No 24, Jalan 4/37B,
Taman Bukit Maluri,
52100 Kuala Lumpur.

Ass. Prof. Dr. Nabil Bessaih
Nama Penyelia

Tarikh: 4/4/2005

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 ini perlu dikelaskan sebagai SULIT dan TERHAD.

The following Final Year Project Report:

Title : MODELLING HOURLY RUNOFF USING ANN FOR Sg. SARAWAK KANAN BASIN

Name : CHONG KAH WENG

Matrix Number : 6392

has been read and approved by :

ASSOCIATE PROFESSOR
DR. NABIL BESSAIH

Project Supervisor

04-04-05

Date

MODELLING HOURLY RUNOFF USING ANN FOR SUNGAI SARAWAK KANAN BASIN

CHONG KAH WENG

This project is submitted in partial of fulfilment of
the requirements for the degree of Bachelor of Engineering with Honours
(Civil Engineering)

Faculty of Engineering
UNIVERSITI MALAYSIA SARAWAK
2005

Dedicated to my family and beloved one

ACKNOWLEDGEMENTS

My sincerest appreciation must be extended to my supportive supervisor; Associate Professor Dr. Nabil Bessaih who has kindly given his support, advices, comments and suggestions all throughout the process of completing my project.

A special thanks to Mah Yau Seng and Kelvin Kuok who had been so helpful and supportive along the implementation of the project.

I also want to thank my beloved father and mother for all their moral and financial supports within this year, and also to my dearest brothers and sisters for all of their help and supports.

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 in the modelling hourly runoff for Sungai Sarawak. An Artificial Neural Network is undoubtedly a robust tool for forecasting various non-linear hydrologic processes, including develop a rainfall-runoff model. It is a flexible mathematical structure, which is capable to generalize patterns in imprecise or noisy and ambiguous input and output data sets. In this study, the ANNs were developed specifically to forecast the hourly rainfall-runoff for Buan Bidi Station. Distinctive networks were trained and tested using hourly data obtained from the DID Department in Kuching. Various training parameters were considered in order to gain the best model possible. The performances of the ANNs were evaluated based on the coefficient of correlation, R. The back propagation algorithm was adopted for this study. With the three months of training length data, the optimal model found in this study is the network using five hours of antecedent data, with the combination of learning rate and the number of neurons in the hidden layer of 0.8 and 150. This model generated the highest R Testing of 0.896 when trained with the scaled conjugate gradient algorithm (TRAINSCG). It has been found that the ANN has the potential to develop a rainfall-runoff model. After appropriate trainings, they are able to generate satisfactory results during both of the training and testing phases.

ABSTRAK

Kajian ini menganjurkan aplikasi Rangkaian Neural Buatan untuk modulasi kadaralir sejam di Sungai Sarawak. Rangkaian Neural Buatan merupakan satu alternatif yang efektif dalam meramalkan pelbagai proses hidrologi tidak linear, termasuk ramalan paras air di sungai-sungai. Ia merupakan struktur matematik yang fleksibel dimana ia berupaya membuat kesimpulan secara menyeluruh terhadap suatu bentuk keadaan yang kurang jelas, dengan set data input dan output yang kurang tepat. Dalam kajian ini, Rangkain Neural Buatan dibangunkan secara spesifik untuk modulasi kadaralir setiap jam bagi Stesen Buan Bidi. Rangkaian yang berbeza dilatih dan diuji menggunakan data setiap jam yang diperolehi daripada Jabatan Pengairan dan Saliran, Kuching. Pelbagai parameter latihan diambil kira untuk mencapai keputusan ramalan terbaik. Prestasi Rangkain Neural Buatan dinilai berdasarkan Pekali Perkaitan, R. Algoritma '*back propagation*' telah diaplikasikan dalam kajian ini. Data telah dilatih dengan tiga bulan dan nilai terbaik bagi R untuk fasa ujian telah dicapai oleh rangkaian yang menggunakan lima jam terdahulu, menggunakan *learning rate* 0.8 dan bilangan neuron 150. Rangkaian terbaik yang telah dilatih ialah *trainscg* dengan memberi Pekali Perkaitan R 0.896. Setelah melaksanakan latihan yang sesuai, keputusan yang memuaskan telah dicapai untuk kedua-dua fasa latihan dan ujian. Selain itu, kekuatan dan kelemahan rangkaian ini turut dibincangkan, berdasarkan keputusan yang telah diperolehi dalam kajian ini.

LIST OF CONTENT

LIST OF TABLE

LIST OF FIGURE

CHAPTER 1

INTRODUCTION

1.1 GENERAL

Human civilization has always been developed along rivers at the early time because of the need of irrigation for crops, water supply for communities and latter power generation. These advantages have been counterbalanced by the danger of floods, which will destroy properties, crops and sometimes even human's life. For civil engineers who are responsible for designing flood protection measures, they are required to plan engineering structures such as storage reservoirs, barrage and tidal control gates. Furthermore, as the flood wave passes through a river it is necessary to know how the storage varies with respect to time and distance for the design of river engineering works as well as for establishment and operation of flood warning systems by the civil authorities. For this purpose, predicting flood discharge magnitude accurately is very important. The technique of artificial neural networks (ANNs) has been found to be a powerful tool for solving different problems in a variety of applications including simulation for flood discharge magnitude.

1.2 APPLICATIONS OF ANN IN RAINFALL-RUNOFF MODELING

Information about rainfall and runoff is needed for hydrologic engineer to design and apply in management purpose. But to determine the relationship between the rainfall and runoff for a watershed is one of the most important problems faced by hydrologist and engineers. This relationship is known to be highly complex. In additional to rainfall, runoff is dependent on numerous factors such as initial soil moisture, land use, watershed topography, evaporation, infiltration, distribution, duration of the rainfall and etc.

A number of researchers have investigated the potential of neural networks in modelling watershed runoff based on rainfall inputs. In preliminary study, Halff et al. (1993) designed a three-layer feedforward ANN in observation rainfall hyetographs were applied as inputs and hydrograph recorded by the U.S. Geological Survey (USGS) at Bellvue, Washington as outputs. A total of five storms and five nodes in the hidden layer were considered, of which data from four storms used for the training while the fifth storm were used for testing the performance. The study opened up several possibilities for rainfall-runoff application using neural networks.

Hjelmfelt and Wang (1993a-c) developed a neural network to compute runoff hydrograph for a watershed using linear superposition and appropriate summation of unit hydrograph ordinates and runoff excesses. Rainfall and runoff data from 24 large storm events were chosen from Godwater Creek watershed (12.2 km^2) in central Missouri to train and test the ANN. The inputs to the ANN were sequences of rainfall data. Its outputs were rainfall excesses. The resulting network was shown to reproduce the unit hydrograph better than the one obtained through the standard gamma function representation. In a later study, Hjelmfelt and Wang (1996) compared this method with a regular three-layered artificial network with backpropagation. The conclusion is that the neural network was more suited for unit hydrograph computations.

Dawson and Wilby (1998) used a 3-layered backpropagation network to determine runoff over the catchments of the River Amber and River Mole that are prone to floods. ANN inputs included past flows and averages of past rainfall and flow values. The ANN output consisting of predicting future flows at 15 minutes intervals up to a lead-time of 6 hours. The result shows that ANN performed about as well as existing forecasting system that required more information. When compared with actual flows, the ANN appeared to overestimate low flows for River Mole.

Tokar and Johnson (1999) reported that the ANN models when provided higher training and testing accuracy when compared with regression and simple conceptual models. Their goal was to forecast daily runoff for the Little Pantuxent River, Maryland, with daily precipitation, temperature and snowmelt equivalent serving in inputs. It was found that the selection of training data has a large impact in accuracy of prediction. ANN had the highest prediction accuracy when trained on wet and dry data.

The first categories of studies are those where ANNs were trained and tested using existing models. These studies may be viewed as providing a 'proof of concept' analysis for ANNs. The results laid the foundation for future ANN use by demonstrating they are indeed capable of replicating model behaviour and providing sufficient data for the training.

Most ANN based studies fall into the second category, those that have been used in observe rainfall-runoff data. These studies provide a more comprehensive evaluation of ANN performance and are capable of establishing ANNs as possible tools for modelling rainfall-runoff. While most studies report that ANNs have resulted in superior performance, but they still do not providing any useful insight for watershed processes. More creative use of ANNs in modelling the rainfall-runoff process will be needed in the future.

1.3 OBJECTIVE

The objective of this project is to develop a rainfall-runoff model for the upper of Sungai Sarawak Kanan basin, which is located in Sungai Sarawak basin by using relatively new technique-Artificial Neural Network method.

CHAPTER 2

LITERATURE REVIEW

2.1 ARTIFICAIL NEURAL NETWORK (ANN)

The advantage of ANN is that given sets of input-output pairs, the network is capable of recognizing the patterns in data without any understanding of the actual phenomena. Even if the data is noisy and contaminated with errors, ANNs have been known to identify the underlying rule. That's the reason hydrologist suggest ANNs may be well suited to the problems of estimation and prediction in rainfall-runoff.(ASCE 2000)

ANN models have been used successfully to model complex non-linear input-output relationships in an extremely interdisciplinary field. Hydrologists had undertaken several studies proven the potential of this model in rainfall-runoff process. The result have shown a good performance for time-series modelling of nonlinear rainfall-runoff relationship and neural network could predict runoff accurately, with good agreement between the experimental and predicted values (Sobri Harun).

2.2 APPLICATION IN RAINFALL-RUNOFF MODELING

Sobri Harun et al. (1996) applied ANNs in daily rainfall-runoff modelling for the estimation of inflows into the Pedu and Muda reservoirs in Kedah, Malaysia. Rainfall and net inflow records of 14 years (1971-1984) were used for model calibration and 3 years (from January 1985 to December 1987) for testing. The results from ANN simulation were compared with calculated runoff using multiple regression equation.

Three types of ANNs models were developed for same target monthly runoff value (January 1980), but with different input nodes namely:

a. Model NN1 (8-10-1) : from June 1979 to Dec 1980.
b. Model NN2 (6-10-1) : March, April, May, Nov 1979 and Jan 1980.
c. Model NN3 (5-12-1) : all 5 nodes from the five rain gauge stations namely Kg. Pinang, Naka, Kuala Nerang, Pedu and Muda, monthly rainfall input is January 1980.

Three layer feedforward neural networks with backpropagation learning algorithm are used. The activation function used is sigmoid function, the learning rate is 0.05 and the momentum constant is 0.9. The original rainfall and runoff data are normalized into the range of 0.1 to 0.9.

The performance criteria used is coefficient of efficiency (R^2). Results obtained from multiple regression show that the observed runoff is almost similar to the calculated runoff with R^2=0.8. Meanwhile Model NN1 and NN2 give R^2=0.50 and 0.62 respectively with small ratio of input nodes to hidden nodes. However model NN3 needs a larger number of hidden nodes, but it only achieve R^2=0.30. Model NN3 can be improved by the introduction of an intervention node.

The good performances of models NN1 and NN2 in inflow estimation show that ANNs have capability to compute with statistical modelling. Results also show that model NN3 manages to produce a reliable estimation with 5 input nodes from 5 rain gauge stations. Therefore it can be concluded that ANNs have tie potential to learn spatially rainfall data from different locations.

Dibike and Solomatine (1999) investigated the use of ANN for daily river flow prediction in Apure river basin (southwest part of Venezuela) and the navigation channel between Puente Remolini and Bruzual (Solomatine and Torres, 1996). This river basin consists of 40,000km^2 rural catchments divided into channels and drainage area. The available data are daily weighted average rainfall and average monthly evapotranspiration over Bruzual sub-basin, daily and weekly runoff at the station at Bruzual from 1981 to 1985.

The five years input-output data was divided into training and verification periods. The weekly data of the first three years (1981-1983) was used for model calibration and the remaining two years data (1984-1985) for model verification. The input variables for this two networks are rainfall including concurrent and antecedent rainfalls { P(t), P(t-1), P(t-2) ... ... P(t-n) }, previous runoff { Q(t-1), Q(t-2), ... Q(t-n) } and evapotranspiration.

Two types of ANN architectures, namely multi-layer perceptron (MLP) and radial basis function network (RBF) were implemented. Different combination of input patterns was tried. These are:

a) concurrent rainfall and evapotranspiration.
b) concurrent rainfall, antecedent rainfalls and evapotranspiration.
c) concurrent rainfall, antecedent rainfalls, antecedent runoffs and evapotranspiration.

Model efficiency, R^2 defined by Nash and Sutcliffe (1970) and root mean square error (RMSE) were used to evaluate the performance of ANNs. The performances of these networks were compared with a conceptual rainfall-runoff model (conceptual tank model) developed in Japan.

The maximum possible model efficiency for MLP network was 86.5% and 41.1% for training and verification periods respectively when the network input pattern consist of concurrent rainfall and evapotranspiration. The performance of neural network is improved with the increase in number of antecedent rainfall in the input patterns. The optimal performance was found when a concurrent and four antecedent rainfalls were used as input with efficiencies of 98.4% and 91.2% for the training and verification periods respectively. However, the best configuration of input patterns for MLP network is two antecedent runoffs, four antecedent rainfalls, one concurrent rainfall and one evapotranspiration data where the network achieves efficiencies of 98.8% and 94.3% for training and verification periods respectively.

Table 2.1: The Effect of Input Rainfall Pattern On The Efficiency of The RBF Network. (Note: Mx=Input pattern consisting of x number of antecedent and concurrent rainfall data)

Input Pattern	Training		Verification	
	RMSE	Efficiency, R^2 (%)	RMSE	Efficiency, R^2 (%)
M1	7.95	82.0	12.82	43.1
M2	7.84	83.6	12.00	50.1
M3	7.80	83.8	10.80	59.4
M4	7.29	85.8	9.58	68.2
M5	5.99	90.7	9.20	70.7
M6	5.69	91.4	9.21	70.6
M7	5.36	92.8	9.40	69.2

Table 2.1 shows the effect of input rainfall pattern on the efficiency of RBF network. The best RBF network performance was obtained with

four antecedent rainfalls, one concurrent rainfall, two runoffs and one evapotranspiration data where the efficiencies being achieved are 90.8% and 80.7% for the training and verification periods respectively.

MLP network shows slightly better performance both in training and verification periods than the RBF network. However, backpropagation network needs relatively longer time to tune the training parameters and train the network.

It was also found that ANNs performed better than conceptual tank model. This was proved by model efficiency of 98.4% on training data and 91.1% on the verification data obtained by ANN with the appropriate input pattern, corresponding values of 95.9% and 80.2% were found with the properly calibrated conceptual tank model.

Elshorbagy, Simonovic and Panu (2000) have used ANNs to predict the daily runoff of Red River in southern Manitoba, Canada. These catchments experienced a major flood in April 1997. These Floods occurred because of

a) High water content in the soil at freeze-up time
b) Heavy snowpack accumulation during winter
c) Rapid melting of winter snowpack, possibly in combination with spring rainfall.