

RAINFALL-RUNOFF MODELLING USING
ARTIFICIAL NEURAL NETWORK METHOD

NOR IRWAN BIN AHMAT NOR

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

Faculty of Civil Engineering
Universiti Teknologi Malaysia

AUGUST 2005

DECLARATION

“I declare that this thesis entitled “*Rainfall-runoff modelling using artificial neural network method*” is the result of my research except as cited in references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any degree”

Signature : [Signed]
Name of Candidate : NOR IRWAN BIN AHMAT NOR
Date : AUGUST, 2005

DEDICATION

“Dan sesungguhnya tiadalah seseorang itu memperoleh melainkan apa yang telah diusahakannya”

(Al-Najm: 39)

I pay my most humble gratitude to Allah Subhanahuwataala for blessing me with good health and spirit to undertake and complete this study.

To my beloved mother and father

ACKNOWLEDGEMENT

I would like to express my sincerest and deepest appreciation and thanks to my supervisors, Assoc. Prof. Dr. Sobri Bin Harun (UTM) and Prof. Ir. Dr. Amir Hashim Bin Mohd. Kassim (KUiTTHO) for their guidance and kind encouragement throughout the length of this research.

High gratitude I intend to the authorities of the Universiti Teknologi Malaysia, Skudai, Johor Darul Takzim. I would like also to express my gratitude and sincere thanks to the Ministry of Science, Technology, and Environmental that provided financial support during my study in the Universiti Teknologi Malaysia. I would like to thanks to the office staff of Sekolah Pengajian Siswazah (SPS) and Graduate Studies Committee, Faculty of Civil Engineering for their support and their good management for the students. My thanks also to the office staff of Hydrology Division, Department of Irrigation and Drainage (DID) Malaysia for providing me the data for my study and a good advice, and also to colleagues and friends who have given me invaluable assistance throughout my research work.

Most important of all, I am deeply indebted to my parent for providing me the peace of mind to pursue knowledge and at the same time being close at hand to render love, comfort, and support. My family has been the source of my perseverance with the research at times all seemed lost.

ABSTRACT

Rainfall and surface runoff are the driving forces behind all stormwater studies and designs. The relationship is known to be highly non-linear and complex that is dependent on numerous factors. In order to overcome the problems on the non-linearity and lack of information in rainfall-runoff modelling, this study introduced the Artificial Neural Network (ANN) approach to model the dynamic of rainfall-runoff processes. The ANN method behaved as the black-box model and proven could handle the non-linearity processes in complex system. Numerous structures of ANN models were designed to determine the relationship between the daily and hourly rainfall against corresponding runoff. Therefore, the desired runoff could be predicted using the rainfall data, based on the relationship established by the ANN training computation. The ANN architecture is simple and it considers only the rainfall and runoff data as variables. The internal processes that control the rainfall to runoff transformation will be translated into ANN weights. Once the architecture of the network is defined, weights are calculated so as to represent the desired output through a learning process where the ANN is trained to obtain the expected results. Two types of ANN architectures are recommended and they are namely the multilayer perceptron (MLP) and radial basis function (RBF) networks. Several catchments such as Sungai Bekok, Sungai Ketil, Sungai Klang and Sungai Slim were selected to test the methodology. The model performance was evaluated by comparing to the actual observed flow series. Further, the ANN results were compared against the results produced from the application of HEC-HMS, XP-SWMM and multiple linear regression (MLR). It had been found that the ANN could predict runoff accurately, with good correlation between the observed and predicted values compared to the MLR, XP-SWMM and HEC-HMS models. Obviously, the ANN application to model the daily and hourly streamflow hydrograph was successful.

ABSTRAK

Hujan dan airlarian permukaan merupakan daya penggerak kepada semua kajian dan rekabentuk berkaitan ributhujan. Diketahui umum bahawa perhubungan antara keduanya adalah taklinear dan kompleks yang mana bergantung kepada banyak faktor. Bagi menyelesaikan masalah akibat kekurangan maklumat dan ketaklinearan hubungan antara hujan dan airlarian, maka kajian ini memperkenalkan kaedah atau pendekatan rangkaian neural buatan (ANN) untuk memodelkan proses dinamik hubungan tersebut. Kaedah ANN bercirikan model 'kotak hitam' dan telah dibuktikan bahawa ianya boleh menghadapi proses taklinear dalam sistem yang kompleks ini. Pelbagai struktur bagi model ANN telah direkabentuk untuk mendapatkan perhubungan harian dan jam yang selaras dengan hubungan hujan dengan airlarian. Dengan itu, data airlarian sebenar boleh diramal menggunakan data hujan berdasarkan kepada hubungan yang telah dikenalpasti perkiraannya melalui proses latihan dalam ANN. Senibina ANN adalah mudah kerana ia mengambilkira data hujan dan airlarian sebagai pembolehubah. Proses dalaman yang mengawal transformasi hujan kepada airlarian dapat diterjemahkan melalui pemberat-pemberat pada ANN. Setelah senibina rangkaian ANN dikenalpasti dan pemberat-pemberat ditentukan, ia akan dapat menterjemahkan keluaran sebenar melalui proses pembelajaran yang mana ANN telah dilatih untuk mendapatkan keputusan seperti yang dijangkakan. Dua jenis senibina ANN telah dicadangkan iaitu kaedah rangkaian perseptron pelbagai lapisan (MLP) dan fungsi asas jejarian (RBF). Beberapa kawasan tadahan iaitu kawasan tadahan Sungai Bekok, Sungai Ketil, Sungai Klang dan Sungai Slim telah dipilih untuk menguji metodologi ini. Keupayaan model dinilai dengan membandingkannya dengan siri-siri aliran cerapan sebenar. Seterusnya, keputusan ANN ini dibandingkan dengan keputusan yang diperolehi dari aplikasi HEC-HMS, SWMM dan regresi linear berbilang (MLR). Didapati bahawa, ANN boleh meramalkan airlarian setepatnya dengan korelasi yang baik antara nilai cerapan sebenar dengan nilai ramalan berbanding model-model MLR, XP-SWMM dan HEC-HMS. Jelasnya, aplikasi ANN untuk permodelan hidrograf aliran sungai bagi sela masa harian dan jam dapat dilaksanakan dengan jayanya.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	DECLARATION	
	ACKNOWLEDGEMENTS	i
	ABSTRACT	ii
	ABSTRAK	iii
	TABLE OF CONTENTS	iv
	LIST OF TABLES	ix
	LIST OF FIGURES	xiv
	LIST OF SYMBOLS	xvii
	LIST OF APPENDICES	xxii
1	INTRODUCTION	
	1.1 Background of Study	1
	1.2 Statement of the Problem	5
	1.3 Study Objectives	8
	1.4 Research Approach and Scope of Work	9
	1.5 Significance of the Study	10
	1.6 Structure of the Thesis	11

2	LITERATURE REVIEW	
2.1	General	13
2.2	Rainfall-Runoff Process and Relationship	14
2.3	Review of Hydrologic Modelling	18
2.4	Rainfall-Runoff Models	22
2.5	Artificial Neural Network	27
	2.5.1 Basic Structure	30
	2.5.2 Transfer Function	32
	2.5.3 Back-propagation Algorithm	34
	2.5.4 Learning or Training	35
2.6	Neural Network Application	37
2.7	Neural Network Modelling in Hydrology and Water Resources	38
	2.7.1 Versatility of Neural Network Method	44
2.8	Bivariate Linear Regression and Correlation in Hydrology	45
	2.8.1 Fitting Regression Equations	48
2.9	Review on HEC-HMS Model	50
2.10	Review on XP-SWMM Model	55
2.11	Summary of Literature Review	57
3	RESEARCH METHODOLOGY	
3.1	Introduction	59
3.2	Multilayer Perceptron (MLP) Model	60
	3.2.1 Training of ANN	67
	3.2.2 Selection of Network Structures	69
3.3	Radial Basis Function (RBF) Model	70
	3.3.1 Training RBF Networks	71
3.4	Multiple Linear Regression (MLR) Model	74

3.5	HEC-HMS Model	76
3.5.1	Evaporation and Transpiration	77
3.5.2	Computing of Runoff Volumes	77
3.5.3	Modelling of Direct Runoff	80
3.6	XP-SWMM Model	85
3.7	Calibration of Distributed Models	89
3.8	Evaluation of the Model	90
3.8.1	Goodness of Fit Tests	90
3.8.2	Missing Data and the Outliers	93
3.9	The Study Area	94
3.9.1	Selection of Training and Testing Data	95
3.9.2	The Sungai Bekok Catchment	97
3.9.3	The Sungai Ketil Catchment	99
3.9.4	The Sungai Klang Catchment	101
3.9.5	The Sungai Slim Catchment	103
3.10	Computer Packages	106

4 RESULTS AND DISCUSSIONS

4.1	General	107
4.2	Results of the Multilayer Perceptron (MLP) Model	108
4.2.1	Results of Daily MLP Model	108
4.2.2	Results of Hourly MLP Model	117
4.2.3	Training and Validation	125
4.2.4	Testing	126
4.2.5	Robustness Test	128
4.3	Results of the Radial Basis Function (RBF) Model	128
4.3.1	Results of Daily RBF Model	129
4.3.2	Results of Hourly RBF Model	132

4.3.3	Training and Validation	135
4.3.4	Testing	136
4.3.5	Robustness Test	137
4.4	Results of the Multiple Linear Regression (MLR) Model	138
4.4.1	Calibration	138
4.4.2	Results of Daily MLR Model	143
4.4.3	Verification	146
4.4.4	Robustness Test	146
4.5	Results of the HEC-HMS Model	147
4.5.1	Calibration	148
4.5.2	Results of Daily HEC-HMS Model	152
4.5.3	Results of Hourly HEC-HMS Model	156
4.5.4	Verification	158
4.5.5	Robustness Test	159
4.6	Results of the SWMM Model	160
4.6.1	Calibration	161
4.6.2	Results of Daily SWMM Model	165
4.6.3	Results of Hourly SWMM Model	169
4.6.4	Verification	171
4.6.5	Robustness Test	172
4.7	Discussions on the Rainfall-Runoff Modelling	173
4.7.1	Basic Model Structure	176
4.7.2	Model Performance	184
4.7.3	Transfer Function and Algorithm	188
4.7.4	Robustness and Model Limitation	190
4.7.5	River Basin Characteristics	193
4.7.6	Time Interval	195

5	CONCLUSIONS AND RECOMMENDATIONS	
5.1	General	216
5.2	Conclusions	217
5.3	Recommendations for future work	220
	REFERENCES	223
	Appendices A-J	241-357

LIST OF TABLES

TABLE NO.	TITLE	PAGE
3.1	Infiltration rates by the soil groups	79
3.2	Rain Gauges used in calibration and verification of the models for Sg. Bekok catchment	98
3.3	Rain Gauges used in calibration and verification of the models for Sg. Ketil catchment	101
3.4	Rain Gauges used in calibration and verification of the models for Sg. Klang catchment	103
3.5	Rain Gauges used in calibration and verification of the models for Sg. Slim catchment	104
4.1(a)	Results of 3 Layer neural networks for Sg. Bekok catchment-using 100% of data sets in training phase	109
4.1(b)	Results of 3 Layer neural networks for Sg. Bekok catchment-using 50% of data sets in training phase	110

4.1(c)	Results of 3 Layer neural networks for Sg. Bekok catchment-using 25% of data sets in training phase	110
4.2(a)	Results of 4 Layer neural networks for Sg. Bekok catchment-using 100% of data sets in training phase	111
4.2(b)	Results of 4 Layer neural networks for Sg. Bekok catchment-using 50% of data sets in training phase	112
4.2(c)	Results of 4 Layer neural networks for Sg. Bekok catchment-using 25% of data sets in training phase	112
4.9(a)	Results of 3 Layer neural networks for Sg. Bekok catchment -using 100% of available data sets in training phase	118
4.9(b)	Results of 3 Layer neural networks for Sg. Bekok catchment -using 65% of available data sets in training phase	119
4.9(c)	Results of 3 Layer neural networks for Sg. Bekok catchment -using 25% of available data sets in training phase	119
4.10(a)	Results of 4 Layer neural networks for Sg. Bekok catchment -using 100% of available data sets in training phase	120
4.10(b)	Results of 4 Layer neural networks for Sg. Bekok catchment -using 65% of available data sets in training phase	121
4.10(c)	Results of 4 Layer neural networks for Sg. Bekok catchment -using 25% of available data sets in training phase	121

4.17(a)	Results of RBF networks for Sg. Bekok catchment -using 100% of data sets in training phase	129
4.17(b)	Results of RBF networks for Sg. Bekok catchment -using 50% of data sets in training phase	130
4.17(c)	Results of RBF networks for Sg. Bekok catchment -using 25% of data sets in training phase	130
4.21(a)	Results of RBF networks for Sg. Bekok catchment -using 25% of available data sets in training phase	133
4.21(b)	Results of RBF networks for Sg. Bekok catchment -using minimum data sets in training phase	133
4.25(a)	Results of MLR Model for Sg. Bekok catchment -using 100% of data sets in training phase	143
4.25(b)	Results of MLR Model for Sg. Bekok catchment -using 50% of data sets in training phase	144
4.25(c)	Results of MLR Model for Sg. Bekok catchment -using 25% of data sets in training phase	144
4.29(a)	Calibration Coefficients of Sg. Bekok catchment -using 100% of data	150
4.29(b)	Calibration Coefficients of Sg. Bekok catchment -using 50% of data	151

4.29(c)	Calibration Coefficients of Sg. Bekok catchment -using 25% of data	151
4.33(a)	Calibration Coefficients of Sg. Bekok catchment -using 25% of data	152
4.33(b)	Calibration Coefficients of Sg. Bekok catchment -using minimum data	152
4.37(a)	Results of HEC-HMS Model for Sg. Bekok catchment -using 100% of data sets in training phase	153
4.37(b)	Results of HEC-HMS Model for Sg. Bekok catchment -using 50% of data sets in training phase	154
4.37(c)	Results of HEC-HMS Model for Sg. Bekok catchment -using 25% of data sets in training phase	154
4.41(a)	Results of HEC-HMS Model for Sg. Bekok catchment -using 25% of data sets in training phase	157
4.41(b)	Results of HEC-HMS Model for Sg. Bekok catchment -using minimum data sets in training phase	157
4.45(a)	Calibration Coefficients of Sg. Bekok catchment -using 100% of data	163
4.45(b)	Calibration Coefficients of Sg. Bekok catchment -using 50% of data	163

4.45(c)	Calibration Coefficients of Sg. Bekok catchment -using 25% of data	164
4.49(a)	Calibration Coefficients of Sg. Bekok catchment -using 25% of data	165
4.49(b)	Calibration Coefficients of Sg. Bekok catchment -using minimum data	165
4.53(a)	Results of SWMM Model for Sg. Bekok catchment -using 100% of data sets in training phase	166
4.53(b)	Results of SWMM Model for Sg. Bekok catchment -using 50% of data sets in training phase	166
4.53(c)	Results of SWMM Model for Sg. Bekok catchment -using 25% of data sets in training phase	167
4.57(a)	Results of SWMM Model for Sg. Bekok catchment -using 25% of data sets in training phase	169
4.57(b)	Results of SWMM Model for Sg. Bekok catchment -using minimum data sets in training phase	170

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
2.1	A schematic outline of the different steps in the modelling process	25
2.2	Simple mathematical model of a neuron	29
2.3	A three-layer neural network with i inputs and k outputs	31
2.4	A threshold-logic transfer function	33
2.5	A hard-limit transfer function	33
2.6	Continuous transfer function: (a) the sigmoid, (b) the hyperbolic tangent	33
2.7	The gaussian function	33
2.8	Steps in training and testing	37
2.9	Typical HEC-HMS representation of watershed runoff	53
3.1	Structure of a MLP rainfall-runoff model with one hidden layer	61

3.2	Hyperbolic-tangent (tansig) activation function	64
3.3	The structure of RBF Model	71
3.4	The Sungai Bekok catchment area	99
3.5	The Sungai Ketil catchment area	100
3.6	The Sungai Klang catchment area	102
3.7	The Sungai Slim catchment area	105
4.1(a)	Daily results of 3-layer neural networks for Sg. Bekok catchment using 100% of data sets in training phase	199
4.1(b)	Daily results of 3-layer neural networks for Sg. Bekok catchment using 50% of data sets in training phase	200
4.1(c)	Daily results of 3-layer neural networks for Sg. Bekok catchment using 25% of data sets in training phase	201
4.2(a)	Daily results of 4-layer neural networks for Sg. Bekok catchment using 100% of data sets in training phase	202
4.2(b)	Daily results of 4-layer neural networks for Sg. Bekok catchment using 50% of data sets in training phase	203
4.2(c)	Daily results of 4-layer neural networks for Sg. Bekok catchment using 25% of data sets in training phase	204
4.9(a)	Hourly results of 3-layer neural networks for Sg. Bekok catchment using 100% of data sets in training phase	205

4.9(b)	Hourly results of 3-layer neural networks for Sg. Bekok catchment using 65% of data sets in training phase	206
4.9(c)	Hourly results of 3-layer neural networks for Sg. Bekok catchment using 25% of data sets in training phase	207
4.10(a)	Hourly results of 4-layer neural networks for Sg. Bekok catchment using 100% of data sets in training phase	208
4.10(b)	Hourly results of 4-layer neural networks for Sg. Bekok catchment using 65% of data sets in training phase	209
4.10(c)	Hourly results of 4-layer neural networks for Sg. Bekok catchment using 25% of data sets in training phase	210
4.17(a)	Daily results of RBF networks for Sg. Bekok catchment using 100% of data sets in training phase	211
4.17(b)	Daily results of RBF networks for Sg. Bekok catchment using 50% of data sets in training phase	212
4.17(c)	Daily results of RBF networks for Sg. Bekok catchment using 25% of data sets in training phase	213
4.21(a)	Hourly results of RBF networks for Sg. Bekok catchment using 25% of data sets in training phase	214
4.21(b)	Hourly results of RBF networks for Sg. Bekok catchment using min of available data sets in training phase	215

LIST OF SYMBOLS

net_j	-	a summation of weighted input for the j th neurons
W_{ij}	-	a weight from the i th neuron in the previous layer to the j th neuron in the current layer
X_i	-	the input form the i th to the j th neuron
x, y	-	the variables for their population linear regressions
b_1, b_2	-	the tangents of slope angles of the two regression lines
a_1, a_2	-	the intercepts
α	-	learning rate parameter
μ	-	momentum parameter
x_i	-	input rainfall variables
y_j	-	output signal from rainfall
y_in_j	-	sum of weighted input signals
w_{0j}	-	weight for the bias
w_{ij}	-	weight between input layer and hidden layer
$f(t)$	-	hyperbolic-tangent function
x_in_k	-	weighted input signals
$c_0^{(k)}$	-	weight for the bias
$c_j^{(k)}$	-	weight between second layer and third layer
$f(z_in_j)$	-	output signal from rainfall

z_j	-	input signal or rainfall
δ_k	-	error information term
$\Delta c_j^{(k)}$	-	weight correction term
$\Delta c_0^{(k)}$	-	bias correction term
t_k	-	target neural network output
$\bar{y}^{(k)}$	-	neural network output
δ_in_j	-	delta inputs
Δw_{0j}	-	bias correction term
$w_{ij}(new)$	-	updates bias and weights
$\Delta c_j^{(k)}(t+1)$	-	update weight for bias with momentum
$\Delta w_{ij}(t+1)$	-	update weight for backpropagation with momentum
η	-	learning rate
E_{\min}	-	minimum error
H	-	Hessian matrix
J	-	Jacobian matrix
E	-	sum of squares function
g	-	gradient
J^T	-	transposition of J
e	-	vector of network errors
w_k	-	vector of current weights and biases
g_k	-	current gradient
$y(t)$	-	runoff at the present time
$x(t)$	-	rainfall at present time
$x(t-i)$	-	rainfall at previous time
$y(x)$	-	output with input vector x
c	-	centre
\mathfrak{R}	-	metric

r_j	-	Euclidean length
ϕ	-	transfer function
T	-	transposition
I	-	interposes
\underline{y}	-	datum vector
$\underline{Y}^{(j)}$	-	radial centre
$\bar{y}^{(k)}$	-	output layer with linear combination of $\phi(r_j)$
y'	-	prediction of the actual output
x	-	input vector
y_i	-	actual output
n	-	length of input vector
p	-	set of input pattern stored
y_{ij}	-	desired output
y'_j	-	predicted output component
x^i	-	stored pattern
$W(x, x^i)$	-	the weight
D	-	distance function
σ_k	-	sigma value
N_j	-	the summation units computes
y	-	dependent variable
x_i	-	independent variables
a, b	-	constants
e	-	random variable
x_{ki}	-	value of independent variable x_k
n	-	number of observations
α, β	-	coefficients
S	-	summation of square function

P_{MAP}	-	total storm mean areal precipitation
$p_i(t)$	-	precipitation depth at time t at gage i
f_c	-	rate of precipitation loss
pe_t	-	the excess precipitation at time t
I_a	-	initial loss
P_e	-	accumulated precipitation
P	-	accumulated rainfall depth
S	-	potential maximum retention
A_i	-	the drainage area of subdivision i
Q_n	-	storm hydrograph ordinate
P_m	-	depth
U_{n-m+1}	-	dimensions of flow rate per unit depth
U_p	-	UH peak discharge
T_p	-	the time to UH peak
C	-	the conversion constant
t_c	-	time of concentration
I_t	-	average inflow to storage
O_t	-	outflow from storage
S_t	-	storage at time t
R	-	constant linear
C_A, C_B	-	routing coefficients
\overline{O}_t	-	average outflow
A	-	the drainage area
L	-	the distance from the upper end of the plane to the point of interest
n	-	the Manning resistance coefficient
S	-	dimensionless slope of the surface
N	-	basin roughness

Q_p	-	the peak discharge
t_p	-	the time to peak
C	-	constant
R^2	-	correlation of coefficient
Q_0	-	actual observed streamflow
Q_s	-	model simulated streamflow
n	-	is the number observed streamflow

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	Daily and hourly results of MLP model	241
B	Daily and hourly results of RBF model	259
C	Results of application of MLR model	267
D	Daily and hourly results of the HEC-HMS model calibration	272
E	Daily and hourly results of application of HEC-HMS model	277
F	Daily and hourly results of the SWMM model calibration	285
G	Daily and hourly results of application of SWMM model	290
H	Daily and hourly results of PBIAS	298
I	Figures illustrate the daily and hourly result of ANN models	301
J	The architecture of MLP network structures	352

CHAPTER 1

INTRODUCTION

1.1 Background of Study

Hydrologists are often confronted with problems of prediction and estimation of runoff, precipitation, contaminant concentrations, water stages, and so on (ASCE, 2000). Moreover, engineers are often faced with real situations where little or no information is available. The processes and relationship between rainfall and surface runoff for a catchment area require good understanding, as a necessary pre-requisite for preparing satisfactory drainage and stormwater management projects. In the hydrological cycle, the rainfall occurs and reaching the ground may collect to form surface runoff or it may infiltrate into the ground. The surface runoff and groundwater flow join together in surface streams and rivers which finally flow into the ocean. Most of hydrologic processes has a high degree of temporal and spatial variability, and are further plagued by issues of non-linearity of physical processes, conflicting spatial and temporal scales, and uncertainty in parameter estimates. That the reason why our understanding in many areas especially in hydrologic processes is far from perfect. So that empiricism plays an important role in modelling studies. Hydrologists strive to provide rational answers to problems that arise in design and management of water resources projects. As modern computers become ever more powerful, researchers continue testing and evaluating a new approach of solving problems efficiently.

A problem commonly encountered in the stormwater design project is the determination of the design flood. Design flood estimation using established methodology is relatively simple when records of streamflow or runoff and rainfall are available for the catchment concerned. The quantity of runoff resulting from a given rainfall event depends on a number of factors such as initial moisture, land use, and slope of the catchments, as well as intensity, distribution, and duration of the rainfall. Knowledge on the characteristics of rainfall-runoff relationship is essential for risk and reliability analysis of water resources projects. Since the 1930s, numerous rainfall-runoff models have been developed to forecast streamflow. For example, conceptual models provide daily, monthly, or seasonal estimates of streamflow for long term forecasting on a continuous basis. Sherman (1932) defined the unit graph, linear systems analysis has played an important role in relating input-output components in rainfall-runoff modelling and in the development of stochastic models of single hydrological sequences (Singh, 1982). The performance of a rainfall-runoff model heavily depends on choosing suitable model parameters, which are normally calibrated by using an objective function (Yu and Yang, 2000). The entire physical process in the hydrologic cycle is mathematically formulated in conceptual models that are composed of a large number of parameters (Tokar and Johnson , 1999).

The modelling technique approach used in the present study is based on artificial neural network methods in modelling of hydrologic input-output relationships. The rainfall-runoff models are developed to provide predicts or forecast rainfalls as input to the rainfall-runoff models. The observed streamflow was treated as equivalent to runoff. The previous data were used in the test set to illustrate the capability of model in predicting future occurrences of runoff, without directly including the catchment characteristics. Tokar and Markus (2000) believed that the accuracy of the model predictions is very subjective and highly dependent on the user's ability, knowledge, and understanding of the model and the watershed characteristic. Artificial intelligence (AI) techniques have given rise to a set of 'knowledge engineering' methods constituting a new approach to the design of high-performance software systems. This new approach represents an evolutionary change with revolutionary consequences (Forsyth, 1984). The

systems are based on an extensive body of knowledge about a specific problem area. Characteristically this knowledge is organized as a collection of rules, which allow the system to draw conclusions from given data or premises.

Application of neural networks is an extremely interdisciplinary field such as science, engineering, automotive, aerospace, banking, medical, business, transportation, defense, industrial, telecommunications, insurance, and economic. In the last few years, the subject of artificial neural networks or neural computing has generated a lot of interest and receives a lot of coverage in articles and magazine. Nowadays, artificial neural networks (ANN) methods are gaining popularity, as is evidenced by the increasing number of papers on this topic appearing in engineering and hydrology journals, conferences, seminars, and so on. This modelling tool is still in its nascent stage in terms of hydrologic applications (ASCE, 2000). Recently there are increasing number of works attempt to apply the neural network method for solving various problems in different branches of science and engineering. This highly interconnected multiprocessor architecture in ANN is described as parallel distributed processing and has solved many difficult computer science problems (Blum, 1992). Electrical Engineers find numerous applications in signal processing and control theory. Computer engineers and computer scientists find that the potential to implement neural networks efficiently and by applications of neural networks to robotics and it show promise for difficult problems in areas such as pattern recognition, feature detector, handwritten digit recognition, image recognition, etc. Manufacturers use neural networks to provide a sophisticated machine or instrument enabling the consumers to gain some benefit in a modern society and our life become comfortable and productive. In medical, the neural networks used to diagnose and prescribe the treatment corresponding to the symptoms it has been before. It is a tool to provide hydraulic and environmental engineers with sufficient details for design purposes and management practices (Nagy et. al., 2002). In other word, apparently neural network models are able to treat problems of different disciplines.

The main function of all artificial neural network paradigms is to map a set of inputs to a set of output. However, there are a wide variety of ANN algorithms. An

attractive feature of ANN is their ability to extract the relation between the inputs and outputs of a process, without the physics being explicitly provided to them. They are able to provide a mapping from one multivariate space to another, given a set of data representing that mapping. Even if the data is noisy and contaminated with errors, ANN has been known to identify the underlying rule (ASCE, 2000). Neural network can learn from experience, generalize from previous examples to new ones, and abstract essential characteristics from inputs containing irrelevant data (Fausett, 1994; Wasserman, 2000). Therefore, the natural behaviour of hydrological processes is appropriate for the application of ANN methods.

In this study, artificial neural network (ANN) methods were applied to model the hourly and daily rainfall-runoff relationship. The available rainfalls and runoffs data are from four catchments known as Sungai Bekok, Sungai Ketil, Sungai Klang, and Sungai Slim. An attractive feature of ANN methods is their ability to extract the relation between the inputs and outputs of process, without the physics being explicitly provided to them. The networks were trained and tested using data that represent different characteristics of the catchments area and rainfall patterns. The sensitivity of the network performance to the content and length of the calibration data were examined using various training data sets. Existing commercially available models used in modelling study were HEC-HMS and XP-SWMM. The performances of the ANN model for the selected catchments were investigated and comparison was made against the XP-SWMM, HEC-HMS and linear regression models. The performance of the proposed models and the existing models are evaluated by using correlation of coefficient, root mean square error, relative root mean square error, mean absolute percentage error and percentage bias.

1.2 Statement of the Problem

In many parts of the world, rapid population growth, urbanization, and industrialization have increased the demand for water. These same pressures have resulted in altered watersheds and river systems, which have contributed to a greater loss of life and property damages due to flooding. It is becoming increasingly critical to plan, design, and manage water resources systems carefully and intelligently. Understanding the dynamics of rainfall-runoff process constitutes one of the most important problems in hydrology, in order to predict or forecast streamflow for purposes such as water supply, power generation, flood control, water quality, irrigation, drainage, recreation, and fish and wildlife propagation. During the past decades, a wide variety of approaches, such as conceptual, has been developed to model rainfall-runoff process. However, an important limitation of such approaches is that treatment of the rainfall-runoff process as a realization of stochastic and statistical process means that only some statistical features of the parameters are involved. Therefore, what is required is an approach that seeks to understand the complete dynamics of the hydrologic process, capturing not only the overall appearance but also the intricate details.

The rainfall-runoff relationships are among the most complex hydrologic phenomena to comprehend due to the tremendous spatial and temporal variability of watershed characteristics, snow pack, and precipitation patterns, as well as a number of variables involved in modelling the physical processes (Tokar and Johnson, 1999). The modelling of rainfall-runoff relationship is very important in the hydraulics and hydrology study for new development area. The transformation of rainfall to runoff involves many highly complex components, such as interception, infiltration, overland flow, interflow, evaporation, and transpiration, and also non-linear and cannot easily calculate by using simple equation. The runoff is critical to many activities such as designing flood protection works for urban areas and agricultural land and assessing how much water may be extracted from a river for water supply or irrigation. Despite the complex nature of the rainfall-runoff process, the practice of estimating runoff as fixed percentage of rainfall is the most commonly used method in design of urban storm

drainage facilities, highway culverts, and many small hydraulic structures. The quantity of runoff resulting from a given rainfall event depends on a number of factors such as initial moisture, land use, and slope of the catchments, as well as intensity, distribution, and duration of the rainfall. Various well known currently available rainfall-runoff models have been successfully applied in many problems and catchments. Numerous papers on the subject have been published and many computer simulation models have been developed. All these models, however, require detailed knowledge of a number of factors and initial boundary conditions in a catchments area which in most cases are not readily available. However, the existing popular rainfall-runoff models can be detected as not flexible and they require many parameters for calibration.

Beven (2001) reported that the ungauged catchment problem is one of the real challenges for hydrological modellers in the twenty-first century. Furthermore, the traditional method of investigation and the collection of data in the field involving the installation and maintenance of a network of instruments tend to be costly. Furthermore, some of these models are expensive, and of limited applicability. The availability of rainfall-runoff data is important for the model calibration process. Rainfall-runoff modelling for sites where there are no discharge data is a very much more difficult problem. However, it is considered that the main limitation in the development of a design flood hydrograph estimation procedure lies in the availability of rainfall and streamflow data, rather than any inherent limitations in the techniques used to develop the procedure. However, discharge data are available at only a small number of sites in any region. In this respect the problem is that there are very few major floods for which reliable rainfall and streamflow data are available, particularly on small catchments. Any relationships developed are therefore based on data from relatively small storms, and hence the flood estimates are made from extrapolated relationships. Even more often, physical measurements of the pertinent quantities are very difficult and expensive especially in a virgin rural area. That is reasons why many catchments in many countries in the world are not installed the measurement instruments. These difficulties lead us to explore the use of neural networks as a way of obtaining models based on experimental measurements. In terms of hydrologic applications, this modelling tool is still in its

nascent stages. An attractive feature of this model is their ability to extract the relationship between the inputs and outputs of a process, without the physics being explicitly provided to them. The goal is to create a model for predicting runoff from a gauged or ungauged catchment. For long term runoff modelling, use a continuous model rather than a single-event model.

Rainfall-runoff modelling software's and guideline from USA, Australia and United Kingdom are required as reference for understanding and development of hydrologic model in Malaysia. Those models and guidelines to study the modelling technique, hydrologic problems, management and design of urban or rural watershed system. Since the present software and guidelines are based on the compilation of the practice of urban stormwater management of USA, United Kingdom and Australia, hence it is important for us to develop our own. Furthermore, various well-known currently available rainfall-runoff models such as HEC-HMS, MIKE-11, SWMM, etc. have been successfully applied in many problems and watersheds. However, the existing popular rainfall-runoff models can be detected as not flexible and they require too many parameters for calibration. Obviously, the models have their own weaknesses, especially in the calibration processes and the ability to adopt the non-linearity of processes. However, there are also many areas where today's tools are lacking the features and functions needed to build these applications effectively (Wasserman, 2000). Furthermore, the software's are not robust and performed by selective calibration. The rapid development of modern Malaysia, the demand of water resources utility has also increased, and therefore, time has already come to develop new techniques to overcome the problems regarding the hydrology and water resources design and management. In this context, one of the main potential areas of application of rainfall-runoff models is the prediction and forecasting of streamflow. An alternative approach to predicting suggested in recent years is the neural network method, inspired by the functioning of the human brain and nervous systems. Artificial neural networks are able to determine the relationship between input data and corresponding output data. When presented with simultaneous input-output observations, artificial neural network adjust their connection

weights (model parameters), and discover the rules governing the association between input and output variables.

1.3 Study Objectives

The research is focused on the application of the neural networks method on the rainfall-runoff modelling. Comparison between neural networks and other methods is made.

The overall objective of the present study is developing mathematical models that are able to provide accurate and reliable runoff estimates from the historical data of rainfall-runoff of selected catchments area. To address the performance of various rainfall-runoff models applied in Malaysian environment, the following specific objectives are made:

- (i) To develop rainfall-runoff model using artificial neural network (ANN) methods, based on the Multilayer Perceptron (MLP) model and Radial Basis Function (RBF) computation techniques.
- (ii) To examine and quantify the predicting accuracy of neural networks models using multiple inputs and output series.
- (iii) To evaluate and compare the neural networks and multiple linear regression (MLR) models for daily flow prediction only.
- (iv) To compare and evaluate the performance of the neural networks models against XP-SWMM and HEC-HMS models for daily and hourly predictions.

1.4 Research Approach and Scope of Work

The present study was undertaken to develop daily and hourly rainfall-runoff models using the ANNs method that can possibly be used to provide reliable and accurate estimates of runoff based on rainfall as input variable. The ANN models used are the MLP and RBF. It is believed that the ANN is able to overcome the non-linear relationship between rainfalls against runoff. The ANN methods of computation are MLP and RBF. Calibration methods (algorithm) apply for MLP is back-propagation and the transfer function used is tangent sigmoid (tansig). Meanwhile, calibration methods apply for RBF is Generalized Regression Neural Network (GRNN) and the transfer function used is Gaussian for hidden units.

The modelling work was carried out using five years period of daily data and ten years period of hourly data consisting the rainfall and runoff records from selected catchments in Peninsular of Malaysia. There are four catchments being selected for analysis and modelling. Those stations have sufficient length of records and fairly good quality of data. Those are Sungai Bekok (Johor, Malaysia), Sungai Ketil (Kedah, Malaysia), Sungai Klang (Kuala Lumpur, Malaysia), and Sungai Slim (Perak, Malaysia) catchments. Those sites were selected to demonstrate the development and application of ANN, multiple linear regression (MLR), XP-SWMM and HEC-HMS models. It is emphasized that the MLR model is only applied to model the daily rainfall-runoff for those catchments. The data required to carry out this study are catchment physical data, rainfall and river (at catchments outlet). The data of all these gauges is recorded and maintained by Department of Drainage and Irrigation (DID) Malaysia.

This study is subjected to the following limitations:

- (i) Analyses treat the catchment as one single catchment. No sub-division of catchment is carried out.
- (ii) It is assumed that the HEC-HMS and XP-SWMM can be applied to a big catchment without sub-division.

- (iii) The available observed data for analysis are rainfall, runoff or streamflow, evapotranspiration, and size of the catchment area. Other data or parameters such as time of concentration, runoff coefficient and infiltration loss coefficient in the HEC-HMS and XP-SWMM will be estimated.

1.5 Significance of the Study

The relationship, or the operation of transforming the input (rainfall) into the output (runoff), is implied uniquely by any corresponding input-output pair. This relationship can be abstracted and used to find the output for any arbitrary input or, the input corresponding to any given output, though, in practice, in analysing systems which are not exactly linear time variant, or where the data are subject to errors. Problems may arise both in identifying the operation or in computing an input corresponding to a given output function of time (Singh, 1982). Overton and Meadows (1976) defined mathematical model as, “a quantitative expression of a process or phenomenon one is observing, analyzing, or predicting”. Meanwhile, Woolhiser and Brakensiek (1982) defined mathematical model as, “a symbolic, usually mathematical representation of an idealized situation that has the important structural properties of the real system. Mathematical models that require precise knowledge of all the contributing variables, a trained artificial intelligence such as neural networks can estimate process behaviour even with incomplete information. It is a proven fact that neural networks have a strong generalization ability, which means that once they have been properly trained, they are able to provide accurate results even for cases they have never seen before (Hecht-Nielsen, 1991; Haykin, 1994). This generalization capability provides an understanding of how the runoff hydrograph system can respond under different rainfall and catchments characteristics.