List of Abbreviations

ANNs	Artificial Neural Networks
BANNs	Bootstrap-based Artificial Neural Networks
BMA	Bayesian Model Averaging
BMLPC	Bootstrapped Multiple-Layer Perceptron Neural Network Combination
CE	Coefficient of Efficiency
CN	Curve Number
DEM	Digital Elevation Model
DMIP	Distributed Model Intercomparison Project
ECMWF	European Centre for Medium-Range Weather Forecasts
ESRE	Environmental Systems Research Institute
ESRI	Environmental Systems Research Institute
ESP	Ensemble Streamflow Prediction
FFBP	Feed Forward Back Propagation
GA	Genetic Algorithm
GEP	Gene Expression Programming
GIS	Geographic Information System
GP	Genetic Programming
GLUE	Generalised Likelihood Uncertainty Estimation
HRUS	Hydrological Response Units
KGF	Kling and Gupta Efficiency
LOO	leave-one-out
LPM	Linear Perturbation Model
LRIS	Land Care Research Institute
LTFM	Linear Transfer Function Model
LVGFM	Linearly-Varying Gain Factor Model

MCS	Monte Carlo Simulation
MLR	Multiple Linear Regression
MLPNN	Multi-Layer Perceptron Neural Network
NAM	Nedbør-Afrstrømnings Model
NID	Neural Interpretation Diagram
NNM	Neural Network Method
OLS	Ordinary Least Squares
PET	Potential Evapotranspiration
RBFNN	Radial Basis Function Neural Network
RMSE	Root Mean Squared Error
RTMOCM	Real Time Model Output Combination Method
SA	Sensitivity Analysis
SCS	Soil Conservation Services
SLM	naïve Simple Linear Model
SMAR	Soil Moisture Accounting and Routing model
SNN	Simple Neural Network
SUFI	Sequential Uncertainty Fitting
SWAT	Soil and Water Assessment Tool
USDA	U.S. Department of Agriculture
WAM	Weighted Average Method
WMO	World Meteorological Organisation

Co-Authorship forms

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Co-Authorship Form

Graduate Centre
ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 5 is from the paper:

Multi-model approach using neural networks and symbolic regression to combine the estimate discharge of rainfall-runoff models. Summitted to the Journal of Hydrologic Engineering.

Nature of contribution by PhD candidate	Prepared the data, analysed data and results, simulated river flows using five rainfall-runoff models, developed multi-models using ANNs models and GEP model (DTREG solfware), tested and verified the models performance, wrote the paper
Extent of contribution by PhD candidate (%)	95%

CO-AUTHORS

Name	Nature of Contribution
Asaad Shamseldin	Supervised developments and assisted in proofreading
Keith Adams	Assisted in proofreading

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- in cases where the PhD candidate was the lead author of the work that the candidate wrote the text.

Name	Signature	Date
Asaad Shamseldin		11/12/2013
Keith Adams	Keith Adams	12 Dec 2013

Co-Authorship Form

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Graduate Centre
ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 6 is from the paper

The optimal number of rainfall-runoff models used in neural network combination system. Submited to the Journal of Hydrologic Engineering

Nature of contribution by PhD candidate	Prepared the data, analysed the data and results, simulated river flows using five rainfall-runoff models, developed multi-models, implemented the knowledge extraction techniques, wrote the paper
Extent of contribution by PhD candidate (%)	95%

CO-AUTHORS

Name	Nature of Contribution
Asaad Y. Shamseldin	Supervised developments and assisted in proofreading
Bruce W. Melville	Co-supervisor

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- in cases where the PhD candidate was the lead author of the work that the candidate wrote the text.

Name	Signature	Date
Asaad Shamseldin		11/12/2013
Bruce W. Melville		17/12/13

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Co-Authorship Form

Graduate Centre
ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 7 is from the paper

Uncertainty analysis in neural network multi-models of combining simulated river flows using bootstrap technique. Submited to the Journal of Hydrologic Engineering

Nature of contribution by PhD candidate	Prepared the data, analysed the data and results, simulated river flows using five rainfall-runoff models, developed multi-models, analysed model uncertanties using the bootstrap techniques, wrote the paper
Extent of contribution by PhD candidate (%)	95%

CO-AUTHORS

Name	Nature of Contribution
Asaad Y. Shamseldin	Supervised developments and assisted in proofreading
Bruce W. Melville	Co-supervisor

Certification by Co-Authors

The undersigned hereby certify that:
- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- in cases where the PhD candidate was the lead author of the work that the candidate wrote the text.

Name	Signature	Date
Asaad Shamseldin		11/12/2013
Bruce W. Melville		17/12/13

Table of contents

List of Tables

List of Figures

Chapter 1
Introduction

Numerous rainfall-runoff models, of varying degrees of sophistication and complexity, are available for river flow simulations (e.g. Kachroo and Liang, 1992; Ahsan and O'Connor, 1994; Arnold et al., 1998; Bell et al., 2001; DHI, 2007). Rainfall-runoff models help to provide an estimate of river flows when floods occur through a catchment. It can help in designing and planning flood control work. In each river flow simulation system, a single rainfall-runoff model is usually used. This single model may have been selected from among a number of competing alternative models based on model accuracy, the available data, availability of license software and having performed well in previous use in river flow forecasting and simulations (e.g. Kachroo et al., 1992; Ahsan and O'Connor, 1994; Tan and O'Connor, 1996; Arnold et al., 1998; Shamseldin and O'Connor, 1999; Xiong et al., 2001; Cao et al., 2006; Rahman et al., 2012; He et al., 2014).

No individual rainfall-runoff model is superior in providing river flow forecasts which are better for all types of catchments under all circumstances, than those of other competing models (Shamseldin et al., 2007). Consequently, many researchers have been working to improve model accuracy. For improving model accuracy, the first method can be to develop new rainfall-runoff models and the second method can be to modify existing models using the results of several models instead of one model (Fenicia et al., 2007). Each individual rainfall-runoff model provides information about the catchment processes involved in the rainfall-runoff transformation and this information may be different from each model. Thus, the motivation for this research is that an alternative approach for improving the modelling accuracy and reliability may be to combine the information from these different sources (e.g. Cavadias and Morin, 1986; Shamseldin et al., 1997 and 2007; Abrahart and See, 2002; Kim et al., 2006; Velázquez et al., 2011), rather than modifying the existing models and developing new models. The combination method, or multi-model approach, is an integration of the results obtained from competing models for improving modelling results that will be better than the result of the best individual model in the combination. A diagram illustrating the concept of the multi-model approach is shown in Figure 1.1. Moreover, the combination techniques have been commonly used in statistics, management, economics and meteorology (e.g. Blattberg and Hoch, 1990; Palm and Zellner, 1992; Armstrong, 2001; Zhang, 2003; Stock and Watson, 2004; Timmermann, 2005; Kücken et al., 2009; Shen et al., 2011; Demargne et al., 2013), since Bates and Granger (1969) have demonstrated that this approach can help improve forecast accuracy. In hydrology, this approach has increasingly gained popularity as an alternative for improving model accuracy (e.g. Shamseldin et al., 1997, 2007; See and Openshaw, 2000; See and Abrahart, 2001; Abrahart and See, 2002; Shamseldin and O'Connor, 2003; Georgakakos et al., 2004; Timmermann, 2005; Ajami et al., 2005; Kim et al., 2006; Zhang et al., 2009; Jeong and Kim, 2009; Exbrayat et al., 2011; Fernando et al., 2012; Liang et al., 2013).

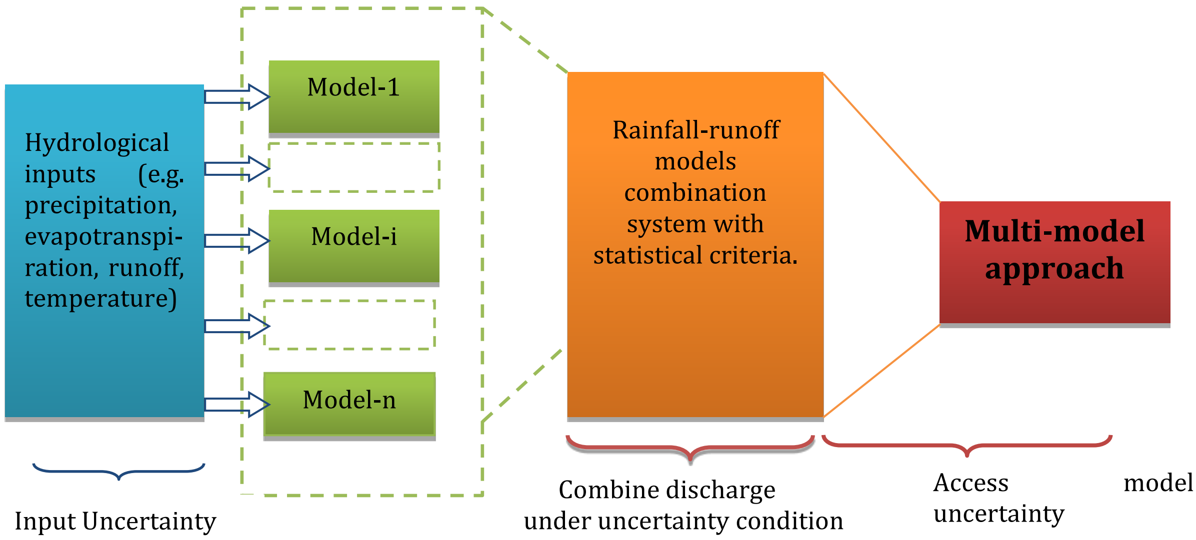


Figure 1.1: Schematic diagram of multi-model approach

Therefore, this study attempts to further improve and develop the multi-model approach. It addresses the issue *"How to optimally design the multi-model approach for improving the model accuracy of river flow simulation?"* However, the performance of the multi-model combination system is affected by the number of rainfall-runoff models used in the combination, due to the complexity of the multi-model, which increases with the number of individual models being used in the combination system (Velázquez et al., 2010). For example, if the number of individual rainfall-runoff models is increased, the complexity will increase in the multi-model combination system. However, the increase in complexity does not always guarantee a significant improvement in the performance of the multi-model combination system. To design the optimal multi-model, this study intends to investigate the effects of the number of rainfall-runoff models being used on the performance of the multi-model combination system. The next question in this study is: could we use simple methods in the multi model combination system? Previous studies of the multi-model approach in the context of combined rainfall-runoff models have demonstrated that the linear combination methods such as the simple average method (SAM) and weighted average method (WAM) (e.g. Timmermann, 2005; Ajami et al., 2006; Kim et al., 2006; Exbrayat et al., 2011) work fairly well but other more complex methods such as, non-linear neural networks methods and fuzzy based methods can produce better model results (e.g. Shamseldin et al., 1997; Xiong et al., 2001; Abrahart and See, 2002; Kim et al., 2006; Jeong and Kim, 2009). Then, the next question in this study is: can the additional complexity of the multi-model combinations be applied effectually in all the case studies (i.e. catchment)?

The uncertainties in the simulations of rainfall-runoff models are recognized as including too many factors such as model structure, model parameter and data (Refsgaard and Knudsen, 1996). However, the understanding of these uncertainties often remains incomplete (Krueger et al., 2010). At the present time, the reliability of the multi-model combination simulations is not satisfactory since all model predictions contain an element of uncertainty (Beven and Binley, 1992). Furthermore, it lacks important information in predicting uncertainty in multi-model combination systems (Velázquez et al., 2011). To answer this problem, there is a need to investigate and apply uncertainty analysis in the multi-model combination systems to provide the accuracy and reliability of the model outputs. Therefore, the proposed study assesses the uncertainties of developed multi-model combination simulations, to quantify the uncertainties of multi-model combination systems achieved by the combination technique.

1.1 Objectives of the research

The main objectives of this thesis were developed from the multi-model approach literature review. The aim of this research is to develop a multi-model approach for river flow simulations through the case studies of two contrasting catchments located in Thailand and New Zealand, respectively. There are three specific objectives of this research:

1. To compare the performance of two artificial neural networks (ANNs) combination methods: the multi-layer perceptron neural network (MLPNN) and the radial basis

function neural network (RBFNN), with a symbolic regression (Gene expression programming, GEP) in the multi-model combination systems.

2. To design the optimal number of rainfall-runoff models to be used in the multi-model combination systems.

3. To quantify the uncertainty and estimate the confidence intervals of multi-model simulations for providing accuracy and reliability of the model results.

Objective 1: To compare the performance of two ANNs combination methods with a symbolic regression (Gene expression programming, GEP) combination method in the multi-model combination systems.

The first objective of the research is a comparison of the performance of the non-linear combination methods such as the GEP and two ANNs methods: multi-layer perceptron neural network (MLPNN) and the radial basis function neural network (RBFNN) when used in the development of multi-models. When used as a rainfall-runoff model, GEP has been shown to produce better results than the ANNs model (Fernando et al., 2011; Roushangar et al., 2013). Their use in a combination method also differs from their use in a rainfall-runoff model. Shamseldin et al. 2007 found that the MLPNN was the appropriate ANN form for use in the context of combining simulated river flows. Therefore, this research attempts to investigate whether or not the use of GEP as a combination technique will lead to further improvement in the performance of multi-model combinations as well as, or better than, the use of MLPNN and RBFNN.

Objective 2: To design the optimal number of rainfall-runoff models to be used in the multi-model combination systems.

The second objective of the thesis is to design the optimal number of rainfall-runoff models to be used in the ANN combination systems. This research will investigate the optimal number of models which perform best in the ANN combination system. This study extends the work of Phukoetphim et al. (2013) by exploring whether or not the knowledge extraction techniques can be used to determine the number of optimal rainfall-runoff models in the ANN combination system.

Objective 3: To quantify the uncertainty and to estimate the confidence intervals of ANN multi-model simulations for providing the accuracy and reliability of the model results.

The third objective of the thesis is to quantify the uncertainty and to estimate the confidence intervals of the developed multi-model combinations. In a multi-model approach for river flow simulations, the ANN combination method has shown the potential for modelling results to perform better than any other combination methods and better than any single models (i.e. SAM, WAM, fuzzy based model) (Shamseldin et al., 1997 and 2007; Xiong et al., 2001; Abrahart and See, 2002; Jeong and Kim, 2009). Previous applications of ANN as a combination method in the context of rainfall-runoff modelling have not provided any technique of uncertainty analysis. To investigate this issue, the use of the bootstrap method was applied for quantifying the uncertainty associated with the ANN multi-model combinations, due to its benefits and capacity for uncertainty analysis in ANN models, according to previous applications (i.e. Abrahart 2003; Tiwari and Chatterjee 2010; Kasiviswanathan and Sudheer 2013).

1.2 Layout Organisation

Chapter 1 Introduction

This chapter first presents a brief introduction of the multi-model approach for this research. Following that, it specifies the objectives for this research. Finally, the structure of the thesis is presented.

Chapter 2 Literature Review

This chapter provides an overview of the background of the multi-model approach which is relevant to a multi-model approach for river flow forecasting. Then, it reviews the applications for improving the model accuracy which is related to the use of combination techniques in a multi-model combination. The literature review is dedicated to objective 1. The number of rainfall-runoff models used in multi-model combination systems is presented in addressing the second objective of the thesis. Then, Objective 3 is discussed in an attempt to review the uncertainty analysis of the ANN multi-model combination. The literature review revealed the research gaps which motivated the undertaking of this study. The summary and discussion are given at the end of the chapter.

Chapter 3 Study areas and data

Chapter 3 presents a brief overview of the selected study areas and data which were used in developing the multi-model approach in this research. Two case studies are considered. The first is the Mae Tuen River catchment located in Thailand and the other one is the Ohinemuri River catchment located in New Zealand. The overview includes the catchment description, data and sources of the selected catchments in this research. Finally, the summaries obtained from the study areas and data are discussed.

Chapter 4 Rainfall-runoff models and model efficiency criteria

This chapter first presents a description of the selected five rainfall-runoff models and the statistical methods applied in this research. These models are as follows: two empirical black-box models, namely, the linear perturbation model (LPM) and the linearly varying gain factor model (LVGFM); two conceptual models, namely, the soil moisture accounting and routing (SMAR) model and the Nedbør-Afrstrømnings model, (NAM)and finally a Physically-based model, namely, the soil and water assessment tool (SWAT). These models are used to simulate daily river flow through the two contrasting catchments (see Chapter 3). Then, the details of the evaluation criteria used for assessment of model efficiency are presented and are summarised at the end of the chapter.

Chapter 5 Multi-model approach using ANNs and GEP to combine the estimated discharge of rainfall-runoff models

This chapter addresses the first objective of the thesis. It presents the comparative performance of a symbolic regression (GEP) combination method with two neural networks combination methods: the multi-layer perceptron neuron network (MLPNN) and the radial basic function neural network (RBFNN) in the multi-model combination systems. A description of the combination methods used in this research work is presented. Then the chapter describes the methodology of the rainfall-runoff simulations and the analysis of the combination methods in multi-model combination systems of this research. Finally, conclusions obtained from the results are discussed.

Chapter 6 The optimal number of rainfall-runoff models used in the neural network combination system

This chapter addresses the second objective of the thesis. It presents general guidelines for the optimal number of rainfall-runoff models being used in ANN multi-model combination systems. Next, the knowledge extraction techniques: the Garson's algorithm method and the connection weight method, are addressed in this chapter. Following that, the optimal number of rainfall-runoff models is presented. Then, the results and analysis are discussed. At the end of the chapter, the conclusion and summary are given.

Chapter 7 Uncertainty analysis in artificial neural network multi-model combination systems

This chapter addresses the third objective of the thesis. It quantifies the uncertainty and estimates the confidence intervals of the developed ANN multi-model combinations. Firstly, it provides a brief description of the uncertainty analysis and the use of the bootstrap method. The bootstrap method was applied to analyse the uncertainty of the developed ANN multi-models through the case studies of the Mae Tuen River catchment, Thailand and the Ohinemuri River catchment, New Zealand, respectively. Then, the criteria for evaluation of model efficiency are explained. The summary and conclusions obtained from the results are discussed at the end of the chapter.

Chapter 8 Summary and conclusions

This chapter presents the summary and conclusions of the thesis. Firstly, it presents summaries and discussions on all the results of this study. Following that, the conclusions obtained from the results of this research are given. The chapter concludes with recommendations for future research directions.

Chapter 2

Literature Review

This chapter provides a review of the literature relevant to the study presented in this thesis. Firstly, it introduces the background of the multi-model approach which is relevant to its application in the context of rainfall-runoff modelling for river flow forecasting. Secondly, the combination methods and the number of rainfall-runoff models applied in multi-model combination systems are presented. Then, it discusses the uncertainty analysis of the ANN multi-model. Finally, the literature review reveals research gaps which motivated the current research.

2.1 Multi-model approach background

Bates and Granger (1969) published their application of combination techniques to economic forecasting. They presented the weighted average methods of combining two separate sets of forecasts. The results show that the combination of forecasts can outperform the individual forecasts. Since then, the advantages of the combination method for improving forecasts have been demonstrated in many fields (e.g. Armstrong, 1989; Palm and Zellner, 1992; Deutsch et al., 1994; Ridley, 1997; Armstrong, 2001; Timmermann, 2005; Atiya, 2008; Coulibaly, 2008; Velázquez et al., 2011; Demargne et al., 2013). Their results indicated that the technique of combination methods can lead to significantly reduced forecast error when compared with that of a single model. No single forecasting method proves to be the most accurate for every data time series, and the best forecasts are often produced by combining forecasting models.

In hydrology, the first combination techniques were investigated by Cavadias and Morin (1986). They applied the three different methods of Granger and Newbold (1977) to the combination of simulated discharges of ten hydrological models. Their results found that combining discharges improved performance by approximately 80% more than the individual simulated discharges. In the combination river flow forecasting system, Shamseldin et al., (1997) compared three combination methods: the simple average method (SAM), the weighted average method (WAM) and the artificial neural network (ANN) method to combine the results obtained from five rainfall-runoff models. The results based on the Nash-Sutcliffe criteria found that the combined outputs were more accurate than the best single individual model and the ANN combination method performed better than the other combination methods. Shamseldin and O'Connor (1999) later developed a real-time model output combination method (RTMOCM) and tested it using three rainfall-runoff models on five watersheds. Their results showed that the combined streamflow output was generally better than the individual models. See and Openshaw (2000) applied a hybrid multi-model approach for river flow forecasting. Four different approaches, namely the hybrid neural network, the simple rule-based fuzzy logic model, the ARMA model and the naive predictions (which use the current value as the forecast) were developed on a time series data from the River Ouse in northern England to provide a hybridized solution for six hours ahead flood prediction. Results show that their purposed approaches were superior to the other individual model developed on the same data set. Xiong et al. (2001) developed the RTMOCM further by introducing a novel concept combination method of the first-order Takagi-Sugeno fuzzy based system, comparing it with three other combination methods (i.e. SAM, WAM and ANN). Abrahart and See (2002) evaluated six alternative methods for river flow forecasting based on two contrasting catchments, namely the River Ouse located in northern England and the Upper River Wype located in Central Wales to improve the multi-model data fusion. These methods are the arithmetic-averaging, the probabilistic method, two different neural network operations and two different soft computing methodologies. They were applied to perform the data fusion. Each set of single model forecasts used in the fusion operation comprised a six-hour-ahead prediction. Their results found that all data fusion methodologies produced improvements and the multi-model data fusion operated better in overall terms in comparison to their individual modelling. Ajami et al. (2006) applied four multi-model combination techniques for streamflow forecasting, namely the simple model average (SMA), the multimodel superensemble (MMSE), modified multi-model superensemble (M3SE), and the weighted average method (WAM). Four model combination techniques were evaluated using the results from the Distributed Model Intercomparison Project (DMIP) for hourly streamflow forecasting. Results revealed that the multi-model approach provides superior to the current single-model simulation.

To improve the rainfall-runoff model performance, Kim et al. (2006) reviewed the combination methods which have been commonly applied in economic forecasting and examined their applicability to hydrologic forecasting. The combination methods were used to improve the accuracy of the existing ensemble streamflow prediction (ESP) forecasting system for forecasting the monthly inflow to the Daecheong Dam in Geum River, Korea. Their results revealed that the combination techniques improved the probabilistic forecasting accuracy of the existing ESP system.

In an ensemble forecast, Georgakakos et al. (2004) developed the multi-model river flow forecasting ensembles employing the simulations produced for the Distributed Model Intercomparison Project (DMIP). Results based on the root mean squared error (RMSE) confirmed that multi-model ensembles are more sophisticated and reliable than the single model ensemble. Ajami et al. (2006) later extended the work of Georgakakos et al. (2004) and Shamselding et al. (1997) by applying several multi-model combination techniques to the streamflow simulation results from the DMIP models. Their study revealed that the multi-model simulations are generally better than any single model simulations and the more sophisticated combination techniques may further improve simulation accuracy.

Since then, many studies have applied this technique (to take advantage of it for improving modelling results (e.g. Duan et al., 2007; Fenicia et al., 2007; Shamseldin et al., 2007; Vrugt and Robinson, 2007; Devineni et al., 2008; Weigel et al., 2008; Viney et al., 2009; Velázquez et al., 2010; Exbrayat et al., 2011; Evsukoff et al., 2012; Fernando et al., 2012; Liang, 2013).

2.2 Combination method

The early work of Bates and Granger (1969) demonstrated that combination techniques can help to improve forecast accuracy. More than 200 applications of the combination techniques were reviewed and summarized by Cleman (1989). To date, various combination methods have been applied in many fields (i.e. economics, statistics, business, management, science, industry and, meteorology). In hydrology, there are various linear, fuzzy based, Bayesian model averaging (BMA), non-linear neural network and symbolic regression methods, which have been used for producing the combined discharges. In a recent application, Dae and Kim (2009) have developed and reviewed useful guidelines for selecting an appropriate method for combining river forecasts.

2.2.1 Linear combination methods

Two linear combination methods: SAM (Shamseldin et al., 1997; Timmermann, 2005; Wang et al., 2005; Ajami et al., 2006; Kim et al., 2006; Goswami and O'Connor, 2007; Jeong and Kim, 2009) and WAM (Cavadias and Morin, 1986; Shamseldin et al., 1997; Shamseldin and O'Connor, 1999; Shamseldin and O' Connor, 2003; Coulibaly et al., 2005; Ajami et al., 2006; Goswami and O'Connor, 2007; Jeong and Kim, 2009; Exbrayat et al., 2011) are the most popular combination techniques used for river flow forecasting. They are also used as a benchmark for comparing the results with other combination methods (e.g. ANN methods, fuzzy based method, and regression methods).

SAM is the simplest method for combining the outputs by weight of the forecast outputs of the individual models. Its accuracy depends on the fact that the different models have the same level of performance results of each individual model, and the number of models involved. Many published studies show that SAM provides an alternative which can perform better than individual forecasts (Makridakis et al., 1982; Makridakis and Winkler, 1983; Cleman, 1989; Shamseldin et al., 1997, Timmermann, 2005).

WAM was first discussed by Bates and Granger (1969). It utilizes the multiple linear regression technique to combine the results obtained from different models, where each model has a different model weight. Cavadias and Morin (1986) in an early application applied the WAM to river flow simulations where they found that the combination method improved the performance of the simulated discharge results. Shameseldin and O'Connor (1999) developed a Real-Time Model Output Combination Method (RTMOCM) based on the structure of the Linear Transfer Function Model (LTFM) and the WAM for three rainfall-runoff models output combinations. Their results indicated that the combined model output of the RTMOCM were generally better than the individual rainfall-runoff models. Coulibaly et al. (2005) found that using WAM for combining three different models can significantly improve the accuracy of the daily reservoir inflow forecast.

2.2.2 Bayesian model averaging method

The Bayesian model averaging (BMA) method has recently been applied as an alternative for combining the forecast outputs. BMA is a technique of statistical postprocessing that reduces the overall model predictions by weighing each individual prediction based on their probabilistic likelihood measures. The better performing prediction receives higher weights than the worse predictions (Duan et al., 2007). Duan et al. (2007) applied the Bayesian averaging model to develop the probabilistic hydrological predictions from a nine-member ensemble of hydrological predictions. The results showed that the BMA model has the advantage of generating more skilful and equally reliable probabilistic predictions than original ensembles. Vrugt and Robinson (2007) applied the BMA method for probabilistic ensemble streamflow forecasting. The results demonstrated that BMA produces more accurate and reliable predictions than other individual watershed models. Recently, Liang et al. (2013) applied the use of BMA in ensemble hydrologic forecasting from two hydrological models, for the Dongwan Basin, China. The results showed that the multi-model ensemble hydrological forecast based on BMA can provide a robust forecast of flood events.

2.2.3 Fuzzy rule-based model

In the combining of models, See and Openshaw (2000) introduced a fuzzy rule-based model in the application of river flow forecasting. This method is based on fuzzy if-THEN rules which transform the individual model forecasts into the multi-model forecasts. Their results are based on the root mean squared error (RMSE) and show that the performances of fuzzy rule-based models were better than other individual model forecasts and integrated approaches. Later, Xiong et al. (2001) applied the fuzzy rule-based model (the first-order Takagi-Sugeno fuzzy system) as a combination method to produce the combination forecasts of five different conceptual rainfall-runoff models. The results demonstrated that the fuzzy rule based model was efficient in enhancing the flood forecasting accuracy. They recommended the use of the fuzzy rule-based model in the combination system for flood forecasting. Abrahart and See (2002) applied six data fusion strategies, including the fuzzy rule-based models, to combine data-driven and physically based hydrological models. These methods were used to produce forecast outputs in two contrasting catchments. The results indicated that the fuzzy rule-based multi-models were better suited to the estimation of flashier behaviour and their operations were better in overall terms, than their individual hydrological modelling.

2.2.4 Artificial neural networks model

Recently, ANN has become a very popular modelling tool in hydrological modelling, generally being used as a competing alternative to the nonlinear rainfall-runoff modelling (e.g. Hsu et al., 1995; Sajikumar and Thandaveswara, 1999; Tokar and Markus, 2000; Shamseldin et al., 2002; Ancil et al., 2004; Jeong and Kim, 2005; Kerem and Kisi, 2006; Ki, 2007; Wang et al., 2009; Izadifar and Elshorbagy, 2010; Wei et al., 2012; Singh et al., 2013). However, there are a small number of studies which intensively consider the development and application of combining simulated river flows (e.g. Shamseldin et al., 1997 and 2007; Xiong et al., 2001; See and Abrahart, 2001; Abrahart and See, 2002; Kim et al., 2006; Jean and Kim, 2009).

ANN was inspired by biological research; its origins are based on the human brain which consists of billions of neural cells that process information. It is a non-linear black-box model and its adaptability and ability to handle complex modelling problems make it very useful. ANNs can help to identify complex non-linear relationships between input and output and can provide rapid and reliable solutions. Their use in combination methods also differs from their other hydrological applications, as the ANNs work synergistically but not competitively with the integral models to produce better river flow simulation. There are various ANN types which can be used in river flow simulations.

Shamseldin et al. (1997) has shown the potential for modelling results improvements by using linear combination methods (i.e. SAM and WAM) and the non-linear method (i.e. ANN) for river flow forecasting. They found that the ANN combination method performed better than any other combination methods and better than any single rainfall-runoff models. Later, Xiong et al. (2001) found that the multi-model combination system which uses neural networks to produce the overall combined forecasts generally out-performed the linear and fuzzy based combination systems in forecasting accuracy and reliability. See and Abrahart (2001) applied two neural network data fusion approaches for combining four individual model outputs to produce a single final forecast, and found that these approaches provided a better solution than the individual models. Abrahart and See (2002) later applied six different model combination techniques: arithmetic-averaging, a probabilistic method, two different neural network operations and two different soft computing methodologies. Their results indicated that the ANN combination techniques provided the best solution for a stable regime, while the fuzzy probabilistic mechanism produced a superior output for flashier catchments with extreme events. Kim et al. (2006) developed the ANN ensemble monthly inflow forecasts, and found that the developed ANN model can significantly improve the model performance of the final predictive outputs and can reduce the increasing overall ensemble bias. Shamseldin et al. (2007) applied three ANNs combination methods (i.e. the simple neural network (SNN), the radial basic function neural network (RBFNN) and the multi-layer perceptron neural network (MLPNN) methods) for combining simulated river flows. They also found that the performance of all three combination methods was superior to other individual rainfall-runoff models. The results based on the Nash Sutcliffe model efficiency index showed that the MLPNN combination performed better than the other two combination methods. Jean and Kim (2009) found the ANN combination methods can remove the effect of bias in the overall ensemble forecasts.

2.2.3 Gene expression programming

In the context of hydrological modelling, gene expression programming (GEP) is a recent development in combination methods. The GEP has been recently introduced as a variant of Genetic Programming (GP) (Ferreira, 2001). It is an evolutionary algorithm like GP and Genetic Algorithms (GAs), used for performing symbolic regression to find a mathematical function that fits a set of data. Unlike traditional linear and non-linear regression, it does not require the form of the function to be specified in advance. GEP provides transparent models to perform symbolic regression (Ferreira, 2006). It is slightly different from other ANN models. It is not completely a black-box model and the relationship between the input (independent variables) and output (dependent variables) can be expressed in mathematical functions.

Fernando et al. (2009) employed the GEP combination method to develop a multi-model for the Chu River in Vietnam. Their results showed that the GEP multi-model produced superior results to those obtained using the individual models. Fernando et al. (2012) extended the work of Fernando et al. (2009) by applying the GEP method to combine the daily estimated outputs of four rainfall-runoff models in four different catchments. The results showed that the GEP combination method produced better results than those obtained from the individual models. These two studies suggest that the GEP method has considerable potential for combining the estimated discharges from different rainfall-runoff models. Fernando et al. (2012) recommended further research to compare the performance of GEP combination technique with other combination methods.

In the application of the comparison of GEP and ANN models, Aytex et al. (2008) developed a mathematical model for rainfall-runoff prediction based on GEP using the daily hydro-meteorological data of three rainfall stations and one streamflow station for Juniata River Basin in Pennsylvania State, USA. The GEP is compared with two different ANN techniques: the feed forward back propagation (FFBP) and the generalized regression neural network (GRNN) methods. Their results confirmed that the GEP can be proposed as an alternative to ANN models. Fernando et al. (2011) compared simulated river flows from the GEP model and ANN model using daily rainfall and runoff data of the Blue Nile catchment in East Africa. Their results demonstrated that the GEP model consistently outperformed the ANN model. Roushangar et al. (2013) compared the simulated runoff values from three different artificial intelligence approaches (i.e. GEP and ANNs) using daily stream flow data from the Vaniar River in Northwestern Iran. A comparison of the results demonstrated that the GEP model performed better than other ANN models in daily streamflow simulation.

2.3 Number of rainfall-runoff models applied in multi-model approach

The performance of the multi-model is affected by the number of models used in the combination. This is due to the complexity of the multi-model, which increases when the number of individual models being used increases in the multi-model combination systems (Velázquez et al., 2010). However, the selection of the number of rainfall-runoff models used in the multi-model combination system is a difficult task, which needs investigation.

Researchers in the field of meteorology have also investigated the impact of reducing the size of ensembles on weather forecasting. For example, Atger (1999) reported that the effect of a decrease in the ensemble size, on the precipitation forecasts issued by the European Centre for Medium-Range Weather Forecasts (ECMWF) is small. Verbunt et al. (2007) also found that the performance of an ensemble comprising 10 members of the ECMWF ensemble system was comparable with the full ensemble. Recently, Khan et al. (2013) reported that an ensemble of 8 members combining only the control forecasts from different ensemble systems was equally good in forecasting rainfall occurrence, as compared to the Grand ensemble comprised of all members from the participating weather ensembles.

However, there are only limited studies of work in hydrological modelling on the number of models to be used in the combination system. For example, Cavadias and Morin (1986) recommended combining the simulated discharges from two or more models as a means of improving performance. Ajami et al. (2006) tested the multi-model simulations to find the optimal number of models used in their multi-model. They found that the inclusion of at least four models is necessary for the multi-model to obtain consistently good results and that above that, five models would actually slightly worsen the results. Viney et al. (2009) compared predictions for one catchment, exploiting ten models and different types including lumped, semi-distributed, and fully distributed models. Their results based on the Nash-Sutcliffe criteria confirmed that the best performing ensemble is not necessarily the one that contains the best individual models, and some models which predicted well individually did not combine well with other models in their multi-models. Velázquez et al. (2010) recommended that further research was needed to apply different types of rainfall-runoff models (e.g. empirical black-box models, conceptual models and distributed physically based models) in the multi-model combination system which can achieve a greater improvement in accuracy and reliability of the model results. In a recent study, Phukoetphim et al. (2013) explored the knowledge extraction techniques from the artificial neural network (ANN) model to produce the combined outputs from four rainfall-runoff models. They recommended that the knowledge extraction techniques had considerable potential to be used for optimizing a combined rainfall-runoff model.

2.4 Uncertainty analysis of the ANN multi-model

In the multi-model combination system, the results of a number of competing rainfall-runoff models are used to produce combined outputs which can be more reliable than that obtained from the best individual model involved in the combination. However, while a number of these models combined may reduce the bias of the multi-model, it may also increase variance, as more parameters have to be estimated (e.g. Ajami et al., 2006; Viney et al., 2009; Velázquez et al., 2010 and 2011; Phukoetphim et al., 2013). This issue requires investigation and the application of uncertainty analysis in the multi-model combination system to provide accuracy and reliability of the combined outputs. The uncertainties in hydrological forecasts are attributable to many factors, such as model structure, poor estimation of the model parameters and errors in the input data (Refsgaard and Storm, 1996).

Uncertainty analysis is related to attempts to quantify the degree of confidence intervals in the model simulations and predictions, given the uncertainties in the model inputs (i.e. data and parameters). Various methods are available such as the Monte Carlo simulation (MCS), the Bayesian Model Averaging, (BMA) and the bootstrap methods, which can be used in uncertainty analysis in hydrological models (e.g. Beven and Binley, 1992; Yu et al., 2001; Wagener et al., 2003; Marshall et al., 2004; Kavetski et al., 2006; Han et al., 2007; Benke et al., 2008; Parasuraman and Elshorbagy, 2008; Renard et al., 2010; McMillan et al., 2011; Zhang et al., 2012; Kasiviswanathan and Sudheer, 2013; Wang et a., 2013). These techniques all have strengths and weaknesses and differ in their underlying assumptions in uncertainty analysis.

2.4.1 Bayesian Model Averaging (BMA)

In the literature on the multi-model approach in hydrological forecasts, there is little guidance regarding the assessment of their associated uncertainty. Recently, the BMA method has become a representative measure of uncertainty for multi-model ensemble forecasting. The BMA predictive probability distribution function (PDF) of a quantity of interest is the weighted average of the individual model PDFs, providing that the individual forecasts are bias-corrected. The weights assigned to each of the models reflect the particular model's relative contributions to the forecast skill over the training period (Raftery et al., 2005). Raftery et al. (2005) applied the BMA approach in weather forecasting to obtain calibrated and sharp predictive PDFs of future weather quantities from the output of ensembles. Results found that the BMA PDFs were much better calibrated than the ensemble itself and produced prediction intervals that were much sharper than those produced by sample climatology. Ajami et al. (2007) applied the BMA approach to improve the prediction skill and address model structural uncertainty, using multiple model outputs. Duan et al. (2007) applied the BMA model to estimate the uncertainty of hydrological model structures, and found that the BMA scheme has the advantage of generating more skilful and equally reliable probabilistic predictions than the original ensemble. They also recommended that BMA predictions employing multiple sets of weights are generally better than those using a single set of weights. In the application of hydrologic groundwater modelling, the BMA method has also been applied to uncertainty analysis (i.e. Neuman, 2003; Ye et al., 2004; Tsai, 2010). In a recent application, Liang et al. (2013) applied the BMA approach for providing quantitative evaluation of forecasting uncertainty (e.g. standard deviation and confidence interval), which calculated the estimation of the PDF of forecast variables.

2.4.2 Monte Carlo simulation (MCS)

MCS methods are stochastic techniques for specified probability distributions to compute the probability distribution of uncertainty in model outputs. These are based on running a certain number of model simulations, using a large random sample of the input variables and parameters. It helps to reduce the uncertainty analysis in rainfall-runoff models and allows the quantification of the model output uncertainty, resulting from uncertain model parameters (Shrestha, 2009).

Kuczera and Parent (1998) applied two MCS approaches for assessing parameter uncertainty in complex hydrological models. The first is the generalised likelihood uncertainty estimation (GLUE) framework and the second is the Metropolis algorithm.

Results demonstrated that the Monte Carlo-based approaches provided an advantage in dealing with parameter uncertainty in hydrological models. Khu and Werner (1999) applied MCS techniques to estimate the model uncertainty due to uncertain parameters. Their results showed that the method is more efficient and increases the feasibility of applying uncertainty analysis to computationally intensive simulation models. Yu et al. (2001) examined the uncertainty of model output caused by model calibration parameters, and found that the MSC was suitable for estimating the uncertainty of model outputs. Recently, Taibi et al. (2006) applied the MCS method to estimate the uncertainty in the input parameters on the simulated discharges. Their results found that the measured discharges were falling within the 95% confidence interval of the modelled discharge. Blasone et al. (2008) used adaptive Markov chain MC sampling within the the Generalised Likelihood Uncertainty Estimation (GLUE) methodology to improve the sampling of the high probability density region of the parameter space. Xiong and O'Connor (2008) modified the GLUE method to improve the efficiency of the GLUE prediction limits in enveloping the observed discharge. Krueger et al. (2010) demonstrated how model parameters as well as structural and data uncertainties can be accounted for explicitly and simultaneously, within the GLUE methodology. Their work demonstrated that discharge error estimates and by implication those of other evaluation data can serve as model independent benchmarks for testing model hypotheses. Majid et al. (2013) applied the MCS method to estimate the uncertainty of streamflow drought forecast and found that the MCS simulations of forecasted values lie within the 95% confidence intervals.

2.4.3Bootstrap method

Bootstrapping is a resampling technique with replacement of the number of samples, in order to quantify model uncertainty (Efron and Tibshirani, 1993). It is the simplest approach since it does not require complex computations of derivatives and Hessian-matrix inversions involved in linear methods or the Monte Carlo solutions of the integrals involved in the Bayesian approach (e.g. Dybowski and Roberts, 2001; Srivastav et al., 2007; Van Hinsbergen et al., 2009). This technique has been applied successfully in hydrological modelling (e.g. Abrahart, 2003; Tiwari and Chatterjee, 2010; Kasiviswanathan and Sudheer, 2013).

The bootstrap method has been applied in the ANN model development (Tiwari and Chatterjee, 2010). It is effective and easy to implement in practice to quantify the uncertainty analysis in comparison with the Bayesian approach (Sharma and Tiwari, 2009). Previous bootstrap results have shown that uncertainty ANN model calibrations (i.e. the training data) play a crucial role in the ANN model's performance during model predictions (i.e. Abrahart, 2003 and Han et al., 2007). The confidence intervals and prediction intervals from the bootstrapped ANN models can guide the ANN model to find the best ANN structure for synthetic flow generation (Jia and Culver, 2006).

In the application of hydrological modelling, Abrahart (2003) developed ANN models based on the bootstrap method to forecast discharge on the Upper River Wye in Central Wales. Results found that the bootstrap method provided marginal improvements in terms of greater accuracies and better global generalisations. Srivastav et al. (2007) reported that the bootstrap method effectively quantifies uncertainty in the ANN model outputs based on the hydrological model. Shu and Quarda (2007) applied ANN models for flood frequency analysis in the canonical physiographic space, and applied Bootstrap based artificial neural network (BANN) models for prediction of monthly runoff. Their results showed that the BANN models provide estimation superior to the original ANN models. Sharma and Tiwari (2009) applied the bootstrap based artificial neural network (BANN) analysis for prediction of monthly runoff in Upper Damodar Valley catchment. Their study recommended the BANN for better simulations of rainfall-runoff relationships. The bootstrap technique has shown capability of solving the problems of over-fitting and under-fitting during the training of ANN models for flood forecasting (Tiwari and Chatterjee, 2010).

Recently, Kasiviswanathan and Sudheer (2013) applied a bootstrap technique to estimate the values of the mean and standard deviation of ANN parameters, and to quantify the predictive uncertainty. They found that this method can effectively quantify the uncertainty bounds of ANN model outputs. Wang et al. (2013) developed the bootstrap-based wavelet neural networks (BWNNs) model to forecast the monthly water quality of Harbin in northeast China. They applied the bootstrapped method to assess the uncertainties from the model structure and input data for the WNNs model.

2.5 Research Gaps

The research gaps are summarised as follows;

(1) According to the literature, most research applied the SAM and WAM combination methods as a benchmark for comparing the results with other combination methods. The results also demonstrated that the non-linear combination methods such as ANNs, fuzzy based model and GEP methods outperform the most commonly used combination methods - SAM and WAM, and other individual rainfall-runoff models. When used as a rainfall-runoff model, the GEP model consistently outperformed the ANN models in daily streamflow simulations (e.g. Fernando et al., 2011; Roushanger et al., 2013). From the present author's knowledge based on the literature review, the comparison of the performance of ANNs and GEP has not been applied in the context of rainfall-runoff model combinations. This issue needs to be investigated in order to compare the performance of ANNs and GEP combination methods in the multi-model combination systems.

(2) Due to the complexity of the multi-model combination systems, adding the number of individual models and type of rainfall-runoff models used in the multi-model combination systems inhibits effectiveness. In developing multi-model performance, there is a need to design the optimal number of rainfall-runoff models to be used to improve performance in the multi-model combination system. The optimal number will therefore maintain a balance between complexity and performance in multi-model combination systems.

(3) The use of a multi-model approach will lead to significant production in forecast uncertainty. According to the literature, the ANN combination method

performed better than any other combination methods (i.e. SAM, WAM, fuzzy based model) and the best individual model in the combination. The use of ANN combination method has not provided any measure of forecast uncertainties in the developed multi-model combinations and also lacks significance in the prediction of uncertainty analysis. Thus, there is a need to investigate and apply the uncertainty analysis for quantifying the uncertainty in developed multi-model combination systems.

2.6 Motivations for the Thesis

This research was motived by a desire to fill the gaps listed below:

(1) To compare the performance of the GEP with two previously investigated ANNs (i.e., MLPNN and RBFN) by Shamseldin et al. (2007) in the multi model combination systems. This study aims to investigate whether or not the use of GEP will lead to further improvement in the performance of multi-model combinations as well as or better than the two previously investigated ANNs combination methods.

(2) To investigate the optimal number of models and the type of rainfall-runoff models, which best perform in multi-model combination systems for the case studies of the two contrasting catchments. To investigate this issue, the performance of the developed multi-models are assessed using statistical methods and scatter plots in this study.

(3) To quantify the uncertainty and to estimate the confidence intervals of the developed multi model combination systems for providing the accuracy and reliability of the model results. According to literature, the bootstrap method has been chosen, due to its benefits and capability for uncertainty analysis in ANN models.

2.7 Summary

This chapter first provides an overview of the background of the multi-model approach. Secondly, it reviews the research on the combination techniques for combining forecast outputs, thus addressing the first objective of the research. The review demonstrated that the non-linear combination methods (i.e. ANNs, fuzzy based and GEP methods) showed better performance than the most commonly used combination methods (i.e. SAM and WAM) and the best individual model. It also found that the comparison of the performance of GEP and ANNs in the multi-models has not been applied in previous applications of the multi-model approach. Thirdly, it reviewed the research on the number of rainfall-runoff models used in a multi-model combination system, addressing the second objective of the thesis. Results obtained from the review found that in terms of the number of rainfall-runoff models a group of at least four models is necessary for the combination system to obtain consistently good results and over five models would actually slightly worsen the results (Ajami et al., 2006). However, there is a need to draw up guidelines about the optimal number of rainfall-runoff models to be used in the multi-model combinations. Next, it reviews the research on the uncertainty analysis of the developed multi-model combination systems, addressing the third objective of the thesis. As the review, there has never provided any measure of forecast uncertainties in the developed multi-model combination systems.

Chapter 3
Study areas and Data

This chapter presents a brief overview of the study areas and data used in the thesis. The overview includes a description of the catchment, data and sources for the study areas. The study areas used are: (1) the Mae Tuen River catchment located in Thailand and (2) the Ohinemuri River catchment located in New Zealand. These were contrasting properties. The two catchments were used for testing the performance of each model, because each is different with respect to climate, topography, geology, land use, evapotranspiration and rainfall-runoff response. The selection of the catchments was based on two main reasons:

1) having sufficient data available (i.e. rainfall, evapotranspiration, maximum and minimum temperature and discharge) and,

2) having data available to input into each rainfall-runoff model for river flow simulations in multi-model combinations.

These selected catchments provide relatively good geographical coverage and represent different topographic features. The data available is obtained from different sources for the model simulations. The data needed for the development of the rainfall-runoff models for this research study includes:

- Hydrological data, which consists of rainfall, runoff discharge, temperature (maximum and minimum) and evapotranspiration.

- Physical data of the area such as watershed, slope, main stream length, soil data and land use.

ArcView GIS 9.3.1 from Environmental Systems Research Institute (ESRI) was used for pre-processing of spatial data in this research. The original DTM resolution used in this study was a 30 m x 30 m cell grid. DEM was used for catchment delineation and slope computation.

3.1 Catchment description

3.1.1Mae Tuen River catchment

The Mae Tuen River catchment is located in the Ping river basin of Northern, Thailand. It has an area of 502 km^2 (see Fig. 3.1). The Ping river basin is the largest of the eight river basins in Thailand. It has a catchment area of about 35,000 km^2 and extends over the provinces of Chiang Mai, Lamphun, Kamphaengphet, Tak and Nakhonsawan. It lies approximately between latitudes 15 °N – 20 °N and longitudes 98 °E – 100 °E. The Ping River flows downstream into the south to become the inflow for the Bhumiphol dam, which is a large dam with an active storage capacity of 9.7 billion m^3. The terrain of the basin is from undulating and rolling to steep in upland areas and flat along river floodplains. The landscape for the Ping River basin is characterised by more than 70% forests and the rest of the basin is terraced hillsides (Reda et al., 2013).

The topography of the Mae Tuen River catchment varies from 762 m above sea level in the south to 1822 m in the north (see Fig. 3.1; Land Development Department, Thailand). Land use types in the Mae Tuen River catchment include forest (82%), agriculture (16%), (rice is the major crop followed by orchards, maize, cassava, sugarcane, beans and other minor crops and pasture (1%). The remainder of the catchment is urban or rural-residential (see Fig. 3.2). Soil types within the catchment consist of clay, loam, orclay-loam (93%), rock (4.55%) and the rest of the area is urban (see Fig. 3.3). These soils are strongly favourable for agriculture and many have good infiltration (Land Development Department, Thailand).

Major flooding in October, 2011, is one example of flooding causing disaster to humans and the environment in this catchment (Reda et al., 2013). The climate regime of the catchment is humid, predominantly affected by the Monsoon. In general, during the rainy season from the middle of May to the end of October, most of the annual precipitation and particularly the heavy storms occur. The winter season, which starts in the middle of October and lasts to the middle of February, has very little precipitation. The summer season starts in the middle of February and ends in the middle of May and it is very dry. The average annual rainfall for the catchment varies between 1,020 mm and 1,225 mm and the mean temperature in summer is within the range of 28-30 °C, while the temperature in winter varies between 20 °C and 25 °C.

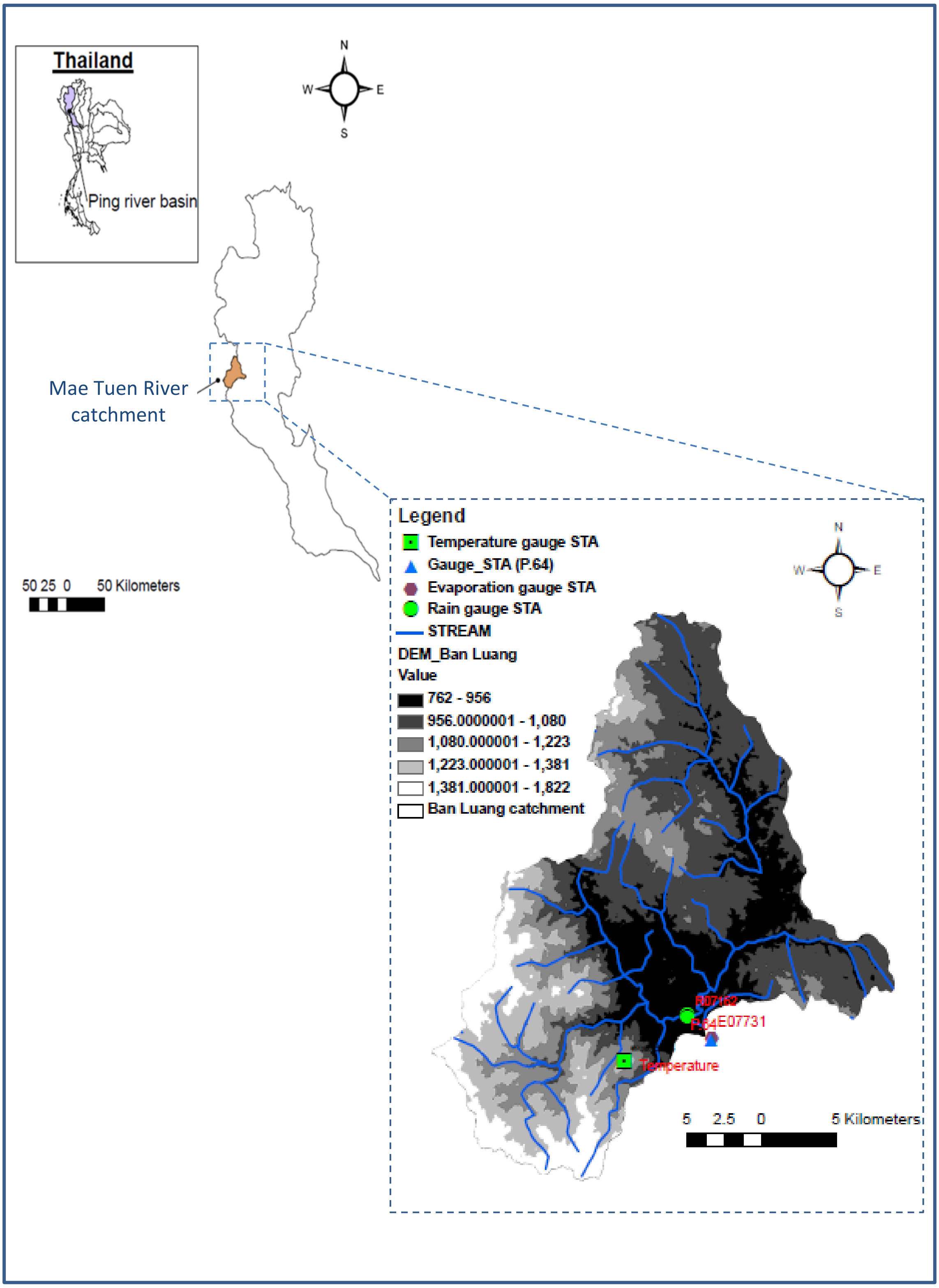


Figure 3.1: The Study area and its digital elevation model representing the topography in the area of the Mae Tuen River catchment, Thailand

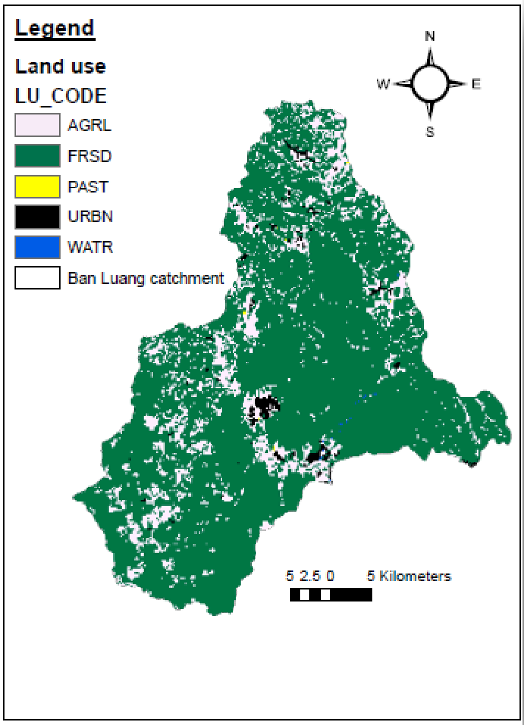


Figure 3.2: Land use, Mae Tuen River catchment

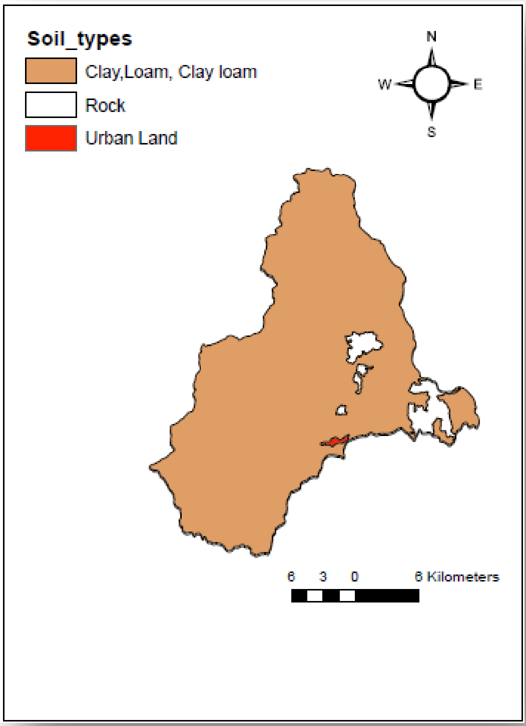


Figure 3.3: Soil types, Mae Tuen River catchment

3.1.2 Ohinemuri River catchment

The Ohinemuri River catchment is located in the Waihi basin on the Coromandel Peninsula, North Island of New Zealand (see Fig. 3.4). It has a drainage area of about 286 km^2. The Ohinemuri River is the major tributary of the Waihou River with its locationin the north-east of the town of Waihi as shown in Figure 3.4. It is a rapid responding or flashy river. The Ohinemuri River has a substantially greater peak flood than the Waihou River due to its geographical location with regular severe weather patterns and the very steep nature of its catchment. It flows westwards through the steep-sided Karangahake Gorge, exiting the ranges near the town of Paeroa on the low-lying flatlands of the Hauraki Plains. Therefore, the town of Paeroa is the most risk-prone area of flooding originating from the Ohinemuri River.

The topography of the Ohinemuri River catchment varies from 19 m above sea level at the stream flow gauging outlet station, to 849 m in the south part of the catchment (see Fig. 3.4). It lies at approximately -37.5 °N and 175.48 °E. The main land use within the Ohinemuri River catchment is mostly forest (49%); other land uses are pasture (48%), agriculture (1%), urban and rural residential (1.5%) and water (0.5%) (see Fig. 3.5). The predominant soil type in the Ohinemuri catchment is clay loam (97%); the remainder of the area is urban or rock (see Fig. 3.6).

The climate is influenced by predominantly westerly air masses of continuous cyclones and depressions. The climate regime of the catchment is a wet temperate climate with annual rainfall of nearly 2700 mm per year, with the higher western hills having the most rain. Its winters are cool and wet with the precipitation at its peak, the mean maximum temperature is within the range of 12-15 °C in July and the mean minimum temperature of 5 °C. Summers are drier with the mean maximum temperature of 25 °C and minimum temperature of 14.5 °C in February. There is a potential for drought occurring one year in ten.

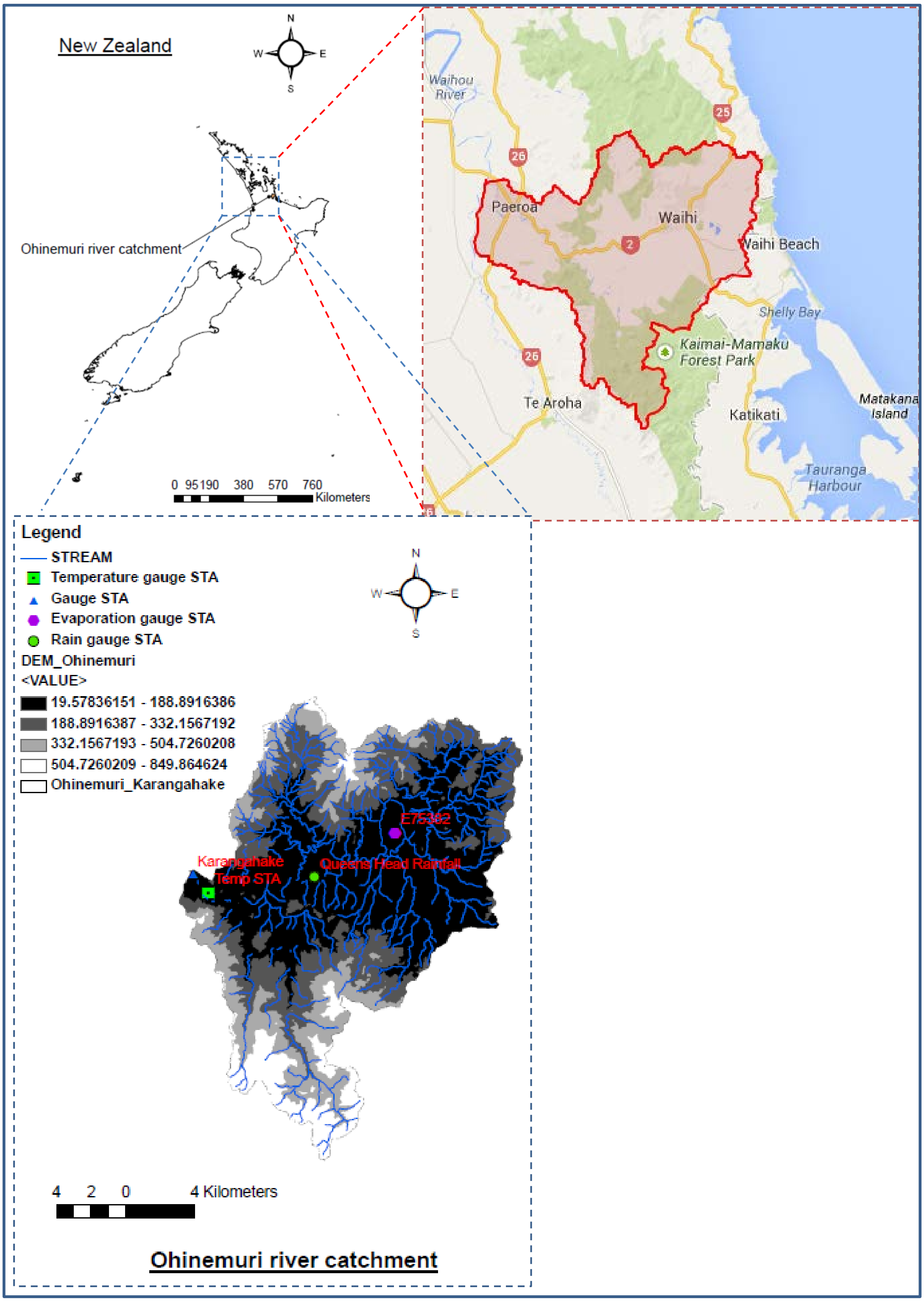


Figure 3.4: Location of the study area and its digital elevation model representing the topography in the area for the Ohinemuri River catchment, New Zealand

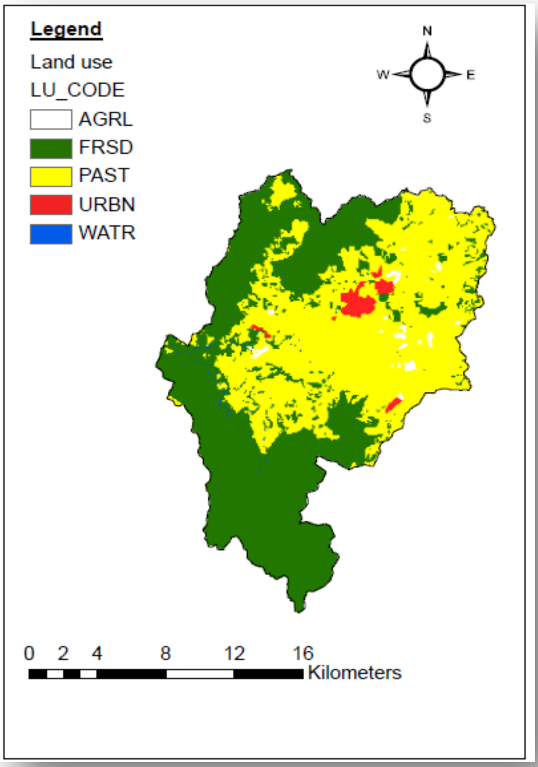


Figure 3.5: Land use 2010, Ohinemuri River catchment

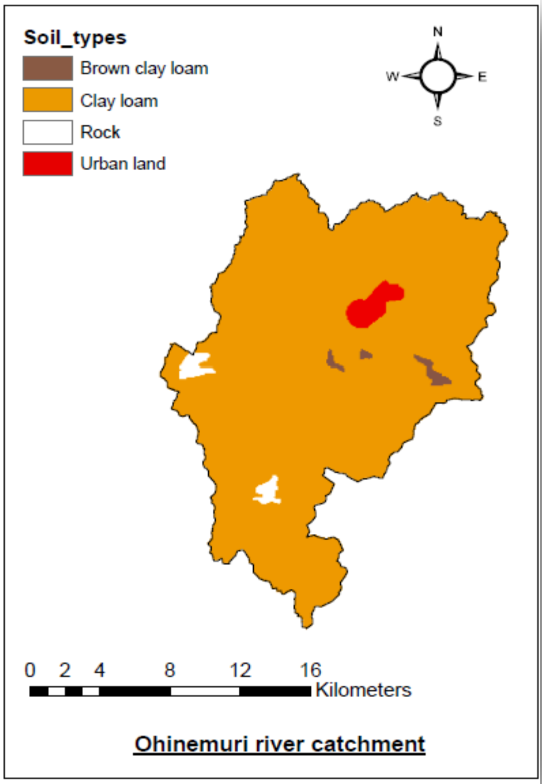


Figure 3.6: Soil types, Ohinemuri River catchment

3.2 Data and sources

The main data used in this study are (a) hydro-climatological data, such as rainfall, discharge, temperature (maximum and minimum), evapotranspiration, which is the most important input information for the rainfall-runoff model and (b) physical data such as a digital terrain model (DTM) of the catchment, soil data, and land use. The daily time series data is required as input data for each rainfall-runoff model to simulate daily runoffs of both catchments. The calibration and validation period for both catchments are shown in Table 3.1, and data sources are given in Table 3.2. Figures 3.7 and 3.8 show the long term means of monthly rainfall totals (mm), monthly evapotranspiration totals (mm), mean monthly discharge (m^3/s), and maximum and minimum monthly discharges for the Mae Tuen River catchment and the Ohinemuri River, respectively.

For the Mae Tuen River catchment, the hydro-climatological data such as rainfall, discharge, temperature (maximum and minimum), and evapotranspiration were obtained from the Thai Royal Irrigation Department and the Thai Meteorology Department, Thailand. The daily data (i.e. rainfall, discharge, temperature (maximum and minimum) and evapotranspiration) is available at the location of the gauging stations (see Fig. 3.1) from 1/04/1991 to 12/31/2002. These data sets were used for the calibration and validation of the rainfall-runoff models in the catchment (see Table 3.1). The observed discharge data obtained from the outlet of the gauging station, P.64 (see Fig. 3.1) was used for the calibration and validation process. The gauging station (P.64) is managed by the Thai Royal Irrigation Department. The high discharge occurs from August to September, with the maximum monthly mean of 1,201 m^3/s in August, and low discharge occurs from January to February (see Fig. 3.7). August to September is the month with the highest rainfall (see Fig 3.7). The second data group comprising DEM, soil data and land use was obtained from the Land Development Department, Thailand. DEM 30-m resolution topography data was used as a basis for the modelling process. The Geographic Information system (GIS) layer representing land use and soil types (see Figs 3.5 and 3.6) was based on an interpretation of aerial photographs.

For the Ohinemuri River catchment, the hydro-climatological data such as rainfall, temperature (maximum and minimum), and evapotranspiration were obtained from the NIWA website (http://www.niwa.co.nz), which are available to download online. The observed discharge data was obtained from the Environment Waikato Department, New Zealand. The observed discharge data obtained from the outlet of the Karangahake gauging station (see Fig. 3.2) was used for the calibration and validation process. The Karangahake gauging station is managed by the Environment Waikato Department. The total available period of daily data recorded (i.e. rainfall, discharge, temperature (maximum and minimum) and evapotranspiration) from January 1990 to October 1993. These data sets were used for the calibration and validation of each rainfall-runoff model in the Ohinemuri River catchment (see Table 3.1). The locations of the rainfall, discharge and temperature stations are shown in Figure 3.2. The highest discharge occurs from June to September (maximum monthly mean of 842 m^3/s in August), and low discharge occurs from January to March (see Fig. 3.4). The DEM, land use and soil data were obtained from Landcare Research Institute (LRIS) and the Environment Waikato Department.

The comparisons in climate information between Mae Tuen River and Ohinemuri River catchments are shown in Figures 3.9 to 3.11. As can be seen from Figures 3.9 to 3.11, there is a significant difference in the climates of both catchments. Figure 3.9 shows the mean daily maximum temperatures of both catchments. Overall, the temperature in Mae Tuen River is higher than the Ohinemuri River catchment. The highest temperature in Mae Tuen River catchment is about 31 °C from March to April, while in Ohinemuri River catchment it is about 23 °C during February. The high temperatures are usually accompanied by high humidity at Mae Tuen River catchment (see Figure 3.10). Mae Tuen River catchment's mean minimum temperature is warmer than Ohinemuri River catchment (see Fig. 3.10). Figure 3.11 shows the mean daily relative humidity and mean monthly rainfall of both catchments. Overall, Mae Tuen River's mean relative humidity is higher than the Ohinemuri River catchment during May to December, due to the high temperatures at Mae Tuen River catchment (see Fig. 3.10), and the Ohinermuri River catchment often receives more rain than the Mae Tuen River catchment.

Table 3.1: Description of Mae Tuen River and Ohinemuri River catchments

Catchment	Basin	Country	Area (km^2)	Climate	Mean Annual Rainfall (mm)	Mean Annual Evaporation(mm)	Mean Annual Discharge(m^3/s)	Calibration data	Validation data
Mae Tuen River	Ping	Thailand	501.79	Humid	1039	1147	7.02	1/4/1991 1/1/2001	31/12/2000 12/31/2002
Ohinemuri River	Waihi	New Zealand	285.39	Temperate	1645	1123	9.73	1/1/1990 31/12/1992	1/1/1993 31/08/1993

Table 3.2: Data and sources

Data	Country	Sources
Hydrology	Thailand	Thai Royal Irrigation Department (http://www.rid.go.th), and
		Thai Meteorology Department (http://www.tmd.go.th)
	New Zealand	Environment Waikato Department (http://www.waikatoregion.govt.nz), and
		NIWA (http://www.niwa.co.nz)
Physical	Thailand	Land Development Department (http://www.ldd.go.th)
	New Zealand	New Zealand Land Resource Inventory (http://www.landcareresearch.co.nz)
Digital Elevation Model (DEM)	Thailand	Land Development Department (http://www.ldd.go.th)
	New Zealand	Environment Waikato Department and Land Care Research Institute (LRIS)
Geoinformation System (GIS)	Thailand	Thai Royal Irrigation Department and the Thai Meteorology Department
	New Zealand	Environment Waikato Department and NIWA

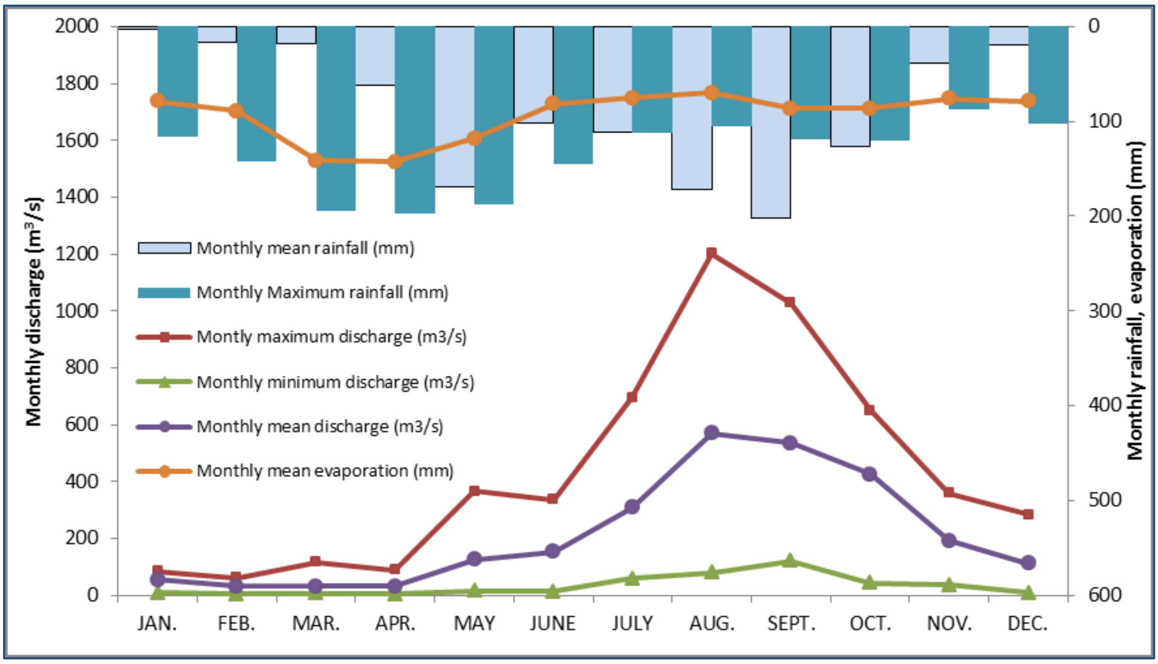


Figure 3.7: Monthly mean discharges (m^3/s), rainfall (mm), and evapotranspiration (mm), 1991 – 2002 at station P.64 gauge

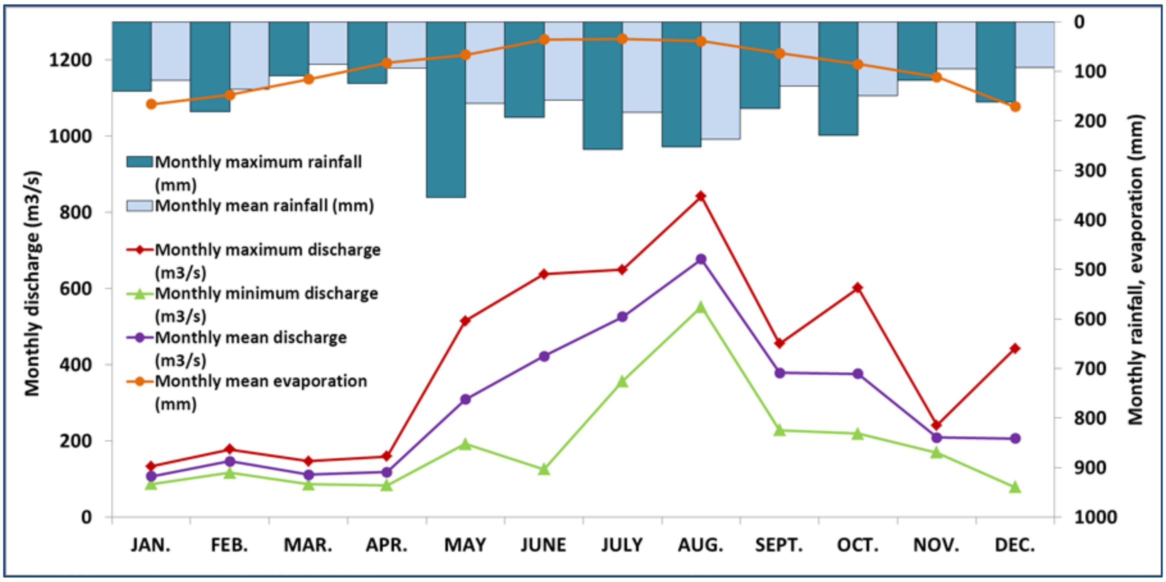


Figure 3.8: Monthly mean discharges (m^3/s), rainfall (mm), and evapotranspiration (mm), 1990 – 1993 at station Ohinemuri River gauge

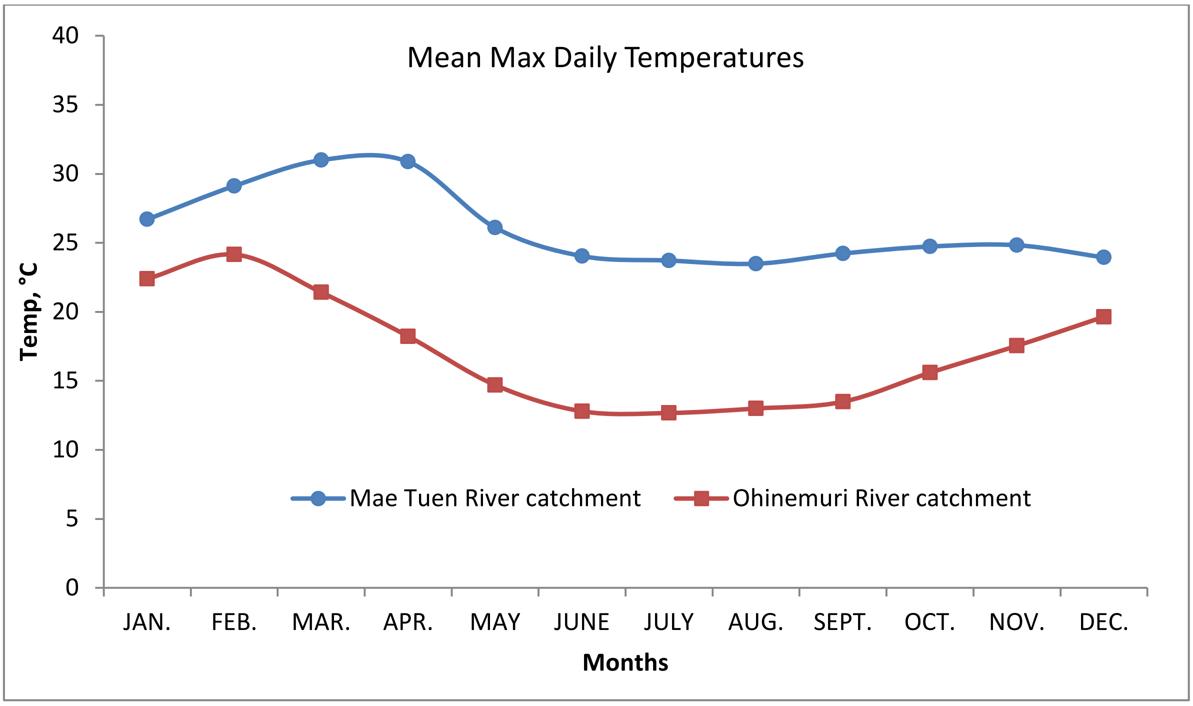


Figure 3.9: Mean daily maximum temperatures, Mae Tuen River and Ohinemuri River catchments

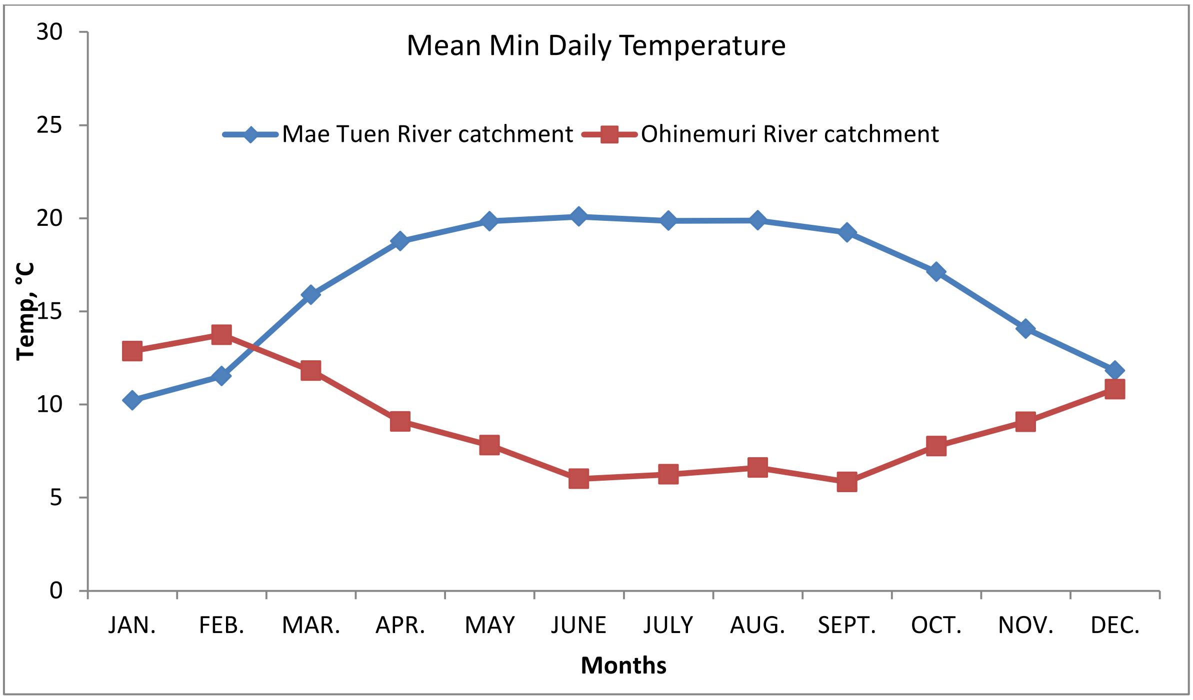


Figure 3.10: Mean daily minimum temperatures, Mae Tuen River and Ohinemuri River catchments

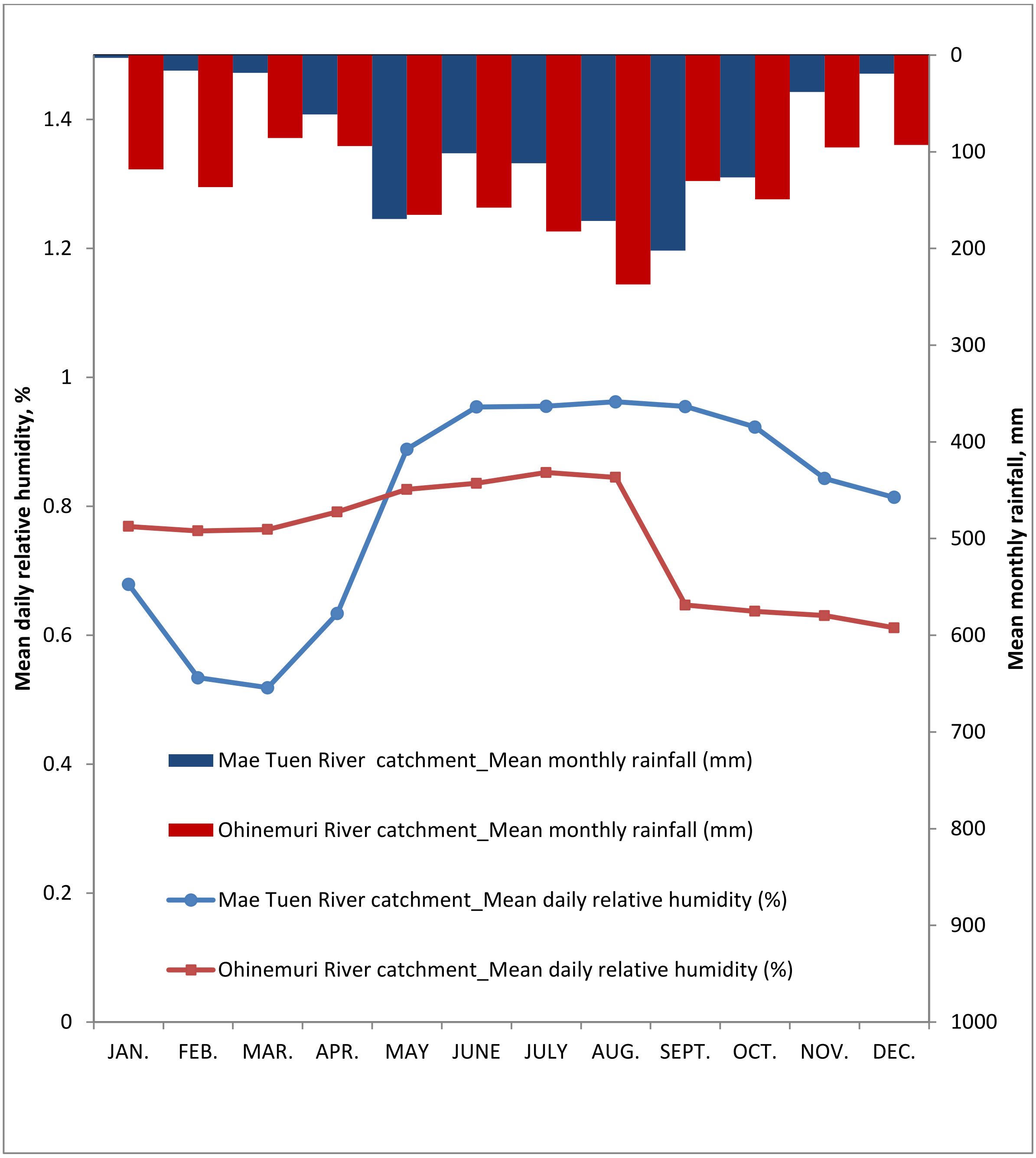


Figure 3.11: Mean daily relative humidity (%) and mean monthly rainfall (mm), Mae Tuen River and Ohinemuri River catchment

3.3 Summary

This chapter first gives a brief overview of the selected case studies with contrasting catchments for this study. The first case study is the Mae Tuen River catchment located in Thailand and other one is the Ohinemuri River catchment located in New Zealand. It describes their catchments including location of gauges (i.e. rainfall, evaporation, temperature, and streamflow), topography and catchment area, land use and soils. It also describes the hydrological features of the catchments in term of climate and hydrology. In this study, ArcView GIS 9.3.1 was used for pre-processing of spatial data and mapping. Then, it describes the nature and sources of different types of data used for both catchments. It also presents the graphs containing comparisons in climate information of both catchments, which demonstrate a significant difference in the climate of both catchments.

Chapter 4
Rainfall-runoff models and model efficiency criteria

This chapter of the thesis provides a brief description of the selected rainfall-runoff models to be used in the multi-model combination system and the evaluation of model performance. Firstly, this chapter starts with a brief description of the selected rainfall-runoff models, used in this research. The selected models were used to provide discharge estimates for the two catchments - one is located in Thailand and the other one is located in New Zealand. The details of the two catchments are presented in Chapter 3 of this thesis. Then, the details of model efficiency evaluation criteria used for the assessment of model efficiency are explained. Finally, the summary of this chapter is discussed.

4.1 Rainfall-runoff models

Rainfall-runoff models can provide information about the catchment processes involved in the rainfall-runoff transformation (see in Fig. 4.1). They can be used for many purposes such as river flow simulation, flood forecast, planning, design, operation and management of the water resource systems. Traditionally, rainfall-runoff modelling systems are classified into three main groups (Anderson and Burt, 1985);

- *Empirical black-box system or theoretical models*, in which little or no attempt is made to simulate the individual constituent hydrologic processes. In black-box models, a relationship is developed between the input time series (usually rainfall) and the output time series (usually discharges), without any consideration of the catchment elements or physical processes.

- *Lumped conceptual models*, which are used to simulate the most important hydrological mechanisms of the catchment response to rainfall, such as evapotranspiration, infiltration, interception and groundwater. In conceptual models, catchment elements are modelled as a number of interconnected storages, with physical processes represented by simple mathematical models.

- *Distributed physically-based models*, the modelling of catchment storages and processes are based on the laws of physics, which are usually represented by non-linear partial differential equations.

Conceptual models and empirical black-box models have often proved to be effective in the solution to a wide spectrum of important hydrological problems, such as river flow forecasting and the extension of hydrological records (O’Connor, 2005). Distributed models are not often used in flood forecasting because of the large data requirements, long processing time, and the large number of parameters. The distributed physically-based models are, however, well suited to solving problems such as predicting the effects of land use changes and pollution hazards (Beven, 1997).

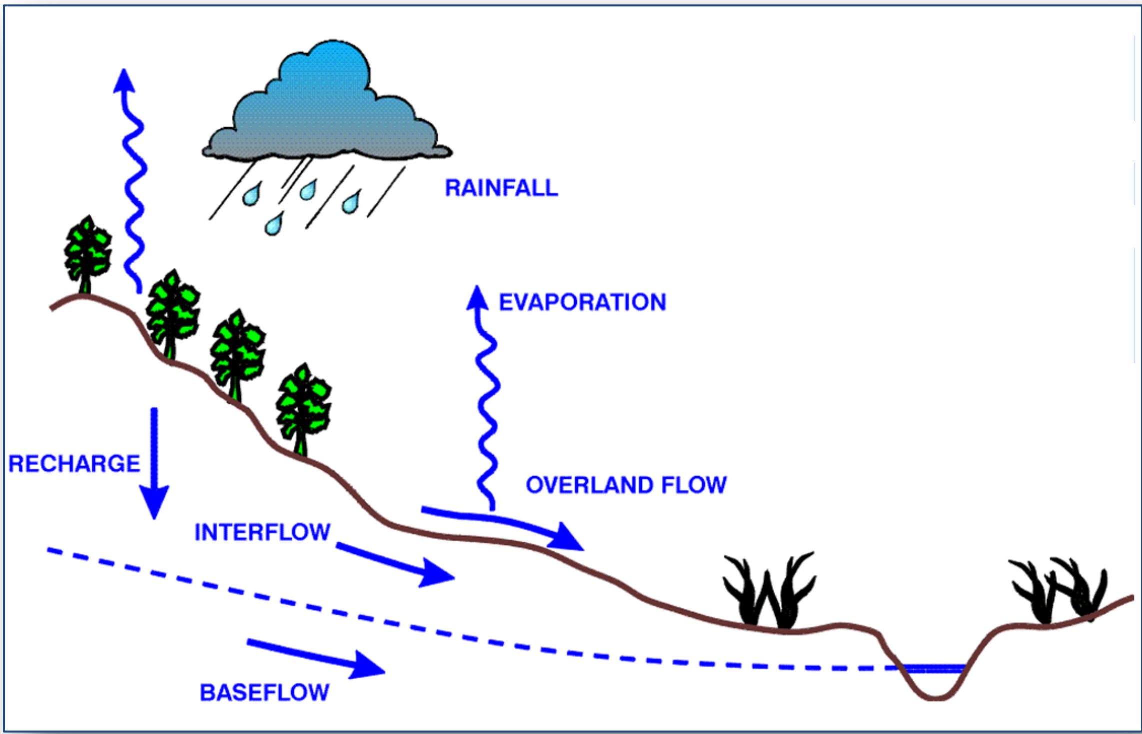


Figure 4.1: Schematic diagram of rainfall-runoff process

In terms of rainfall-runoff models, the question in this research for developing a multi-model approach is:

- *What rainfall-runoff model should we use to improve a multi-model combination system?*

To answer the above question, different types of rainfall-runoff models which could be used in river flow forecasting and could be efficiently combined with the combination method were used in this study. The selected models are generally good representatives of the wide spectrum of existing rainfall-runoff models. The rainfall-runoff models were selected for the purpose of this research work based on the following merits:

1) *The basis of model accuracy and familiarity with the model;*

2) *The available data;*

3) *Availability of license software;*

4) *Performance in previous use in river flow simulations (Shameseldin et al., 1997 and 2007; Refsgaard, 1997; Alansi et al., 2009; Rahman et al., 2012; Vu et al., 2012)*

5) *All the model parameters are obtained by automatic optimization strategy.*

According to the literature review in Chapter 2, multi-models using different types of models (.g. empirical black-box models, conceptual models and distributed physically based models) can achieve a greater improvement in accuracy, and at least four models are required to obtain consistent multi-model simulations. *For this study, three types (i.e., empirical black-box models, conceptual models and physically-based models) of five rainfall-runoff models were applied for the first time in multi-models in the context of rainfall-runoff model combinations.* The five rainfall-runoff models selected in this study are:

- Two empirical black-box models, namely: the linear perturbation model (LPM) (Nash and Barsi, 1983) and the linearly varying gain factor model (LVGFM) (Ahsan and O'Connor, 1994),

- Two conceptual models, namely: the soil moisture accounting and routing (SMAR) model (Tan and O'Connor, 1996) and the Nedbør-Afstrømnings model (NAM) (DHI, 2007), and

- A Physically-based model, namely the soil and water assessment tool (SWAT) (Arnold et al., 1998).

A brief description of the five selected rainfall-runoff models is described in the following section below.

4.1.1 The linear perturbation model (LPM)

In the context of rainfall-runoff modelling, the Linear Perturbation Model (LPM) was developed by Nash and Barsi (1983), and incorporates the seasonal behaviours of the catchment. They suggested that the use of LPM reduces dependence on linearity and increases the dependence of observed seasonal behaviour. It has been extensively applied in many hydrologic modelling studies (e.g. Kachroo et al., 1988, 1992; Kachroo, 1992; Linag et al., 1992; Shamseldin and O'Connor, 1997; Shamseldin, 1997; Shamseldin et al., 1997, 2007, Xiong et al., 2001) to produce the model efficiency for comparing the results of other types of rainfall-runoff models. Results showed that the LPM can perform significantly better in the case of catchments exhibiting marked seasonal behaviour. The model is based on the following two assumptions (Kachroo and Liang, 1992);

(1) If, in a particular year, each input function for each day is equal to its expected value of rainfall for that date, R_d, the output will also equal its expectation for that date, Q_d.

(2) the perturbations or departures from the data-expected input values, R_d are linearly related to the corresponding perturbations or departures from the data-expected output values, Q_d.

For a single input, the ratio between the inflow departure series of the LPM can be described by the equation as:

$$Q_t' = \sum_{j=1}^{m} R'_{t-j+1} h_j + e_t \tag{4.1}$$

where Q_t and R_t are the discharge and rainfall respectively at the t^{th} time-step, m is the memory length of the linear system, h_j is the series of discrete pulse response ordinates, $R_j^{'} = R_t - R_d$ and $Q_t^{'} = Q_t - Q_d$ are the respective departures of rainfall and discharge from their seasonal expectations, e_t is the error output term and for each day, d of the year = 1, 2, 3, ….., 365.

LPM input:

The LPM uses the information between the daily observed rainfall and discharge time series.

4.1.2 The linearly varying gain factor model (LVGFM)

The Linearly-Varying Gain Factor Model (LVGFM) was developed by Ahsan and O'Connor (1994) for the single-input to single-output case. It is exhaustively described in many hydrologic modelling studies (e.g. Shamseldin et al., 1997, 2007; Shamseldin and O'Connor, 1999; Xiong et al., 2001; Goswami et al., 2005; Fernando et al., 2012; He et al., 2014). LVGFM is based on the idea of the Simple Linear Model (SLM) (Nash and Foley, 1982) of a constant runoff coefficient (a gain factor, G). SLM assumes a linear time-invariant relationship between the total rainfall (R_t) and the total discharge (Q_t), which can be expressed after incorporating the model error term (e_t), by the equation (4.2). LVGFM assumes the amount of rainfall that transforms to runoff is a function of the state of the soil moisture. It assumes that the gain factor varies linearly with the selected index of the prevailing catchment wetness, without varying the shape (i.e. the weights) of the response function. Using a time varying gain factor G_i, the model output structure can be expressed by equation (4.3).

$$Q_t = \sum_{j=1}^{m} R_{t-j+1} h_j^{'} + e_t , \tag{4.2}$$

$$Q_t = G_i \sum_{j=1}^{m} R_{t-j+1} h_j^{'} + e_t , \text{ with } \quad \sum_{j=1}^{m} h_j^{'} = 1 \tag{4.3}$$

where Q_t and R_t are the discharge and rainfall respectively at the t^{th} time-step, m is the memory length of the linear system, $h_j^{'}$ is the jth discrete pulse response ordinate or weight.

The gain factor can be considered to vary linearly with the catchment wetness index (the soil moisture state), $Z(t)$ by assuming a linear relationship of the form:

$G(t) = a + bZ(t)$, where a and b are constants (4.4)

Although the antecedent precipitation index (API) provides a crude index of the catchment wetness index, Z(t), however Ahsan and O'Connor (1994) suggested that the value of *Z(t)* is obtained from the outputs of the SLM, operating as an auxiliary Model (See Fig. 4.2), using:

$$Z_{(t)} = \frac{\hat{G}}{\overline{Q}} \sum_{j=1}^{m} R_{t=j+1} + \hat{h}_j \tag{4.5}$$

where $\hat{G}$ and $\hat{h}_j$ are estimates of the gain factor and the pulse response ordinates respectively of the SLM and $\overline{Q}$ is the mean calibration discharge.

LVGFM input:

The LVGFM uses the information rainfall, runoff and evapotranspiration time series.

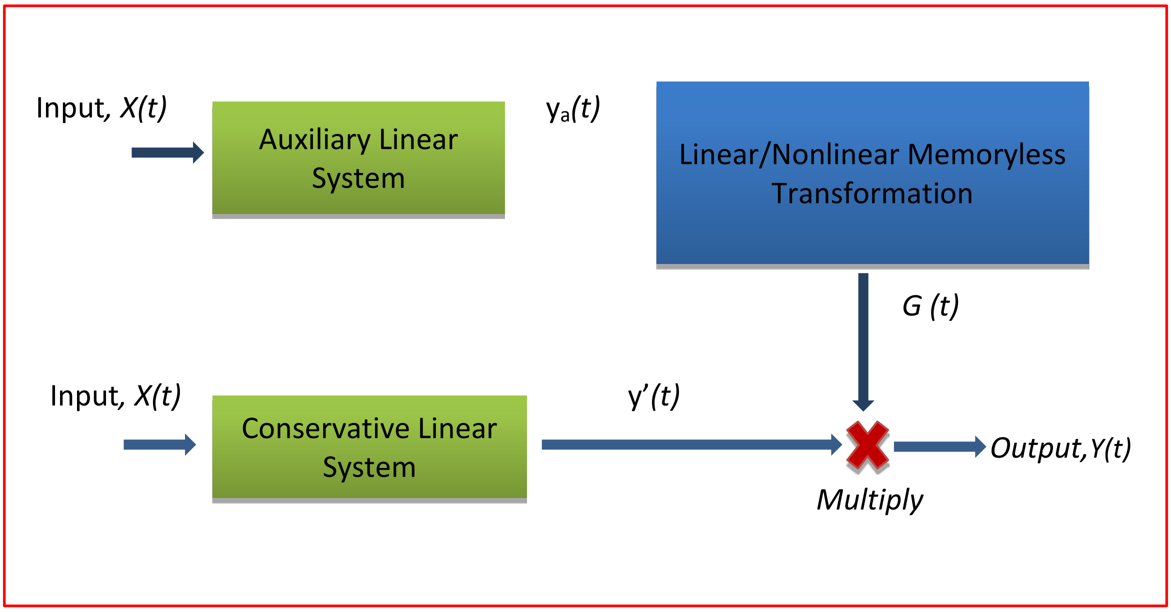


Figure 4.2: Schematic diagram of the Linearly Varying Gain Factor Model (Ahsan and O'Connor, 1994)

4.1.3 The soil moisture accounting and routing (SMAR) model

The Soil Moisture Accounting and Routing (SMAR) model is a development of the conceptual rainfall-runoff model originally introduced by Mandeville et al. (1970). Its water-balance component is based on "the Layers Water Balance Model" proposed by Nash and Sutcliffe (1969). The SMAR model structure consists of two parts, the first part is the water-balance (budget) part which at each time step keeps account of the balance between the rainfall, the evaporation, the runoff and the simulated soil storage. The second part is the routing part, which synthesizes the attenuation and the diffusive effects of the catchment by routing the different generated runoff components of the water balance part through linear storage systems.

The structure of the SMAR model is shown schematically in Figure 4.2 (Tan and O'Connor, 1996). Evaporation *E* occurs from the first layer at the potential rate and the second layer on exhaustion of the top layer at the remaining potential multiplied by a parameter *C* (whole value is less than unity). On the exhaustion of the second layer, evaporation from the third layer occurs at the potential rate multiplied by C^2 and so forth. However, if all soil layers were full and there was no subsequent rainfall, then a constant potential evaporation rate applied to the catchment would reduce the soil moisture in a roughly exponential manner. To estimate the potential evaporation depth, E_p, it is taken as the Pan evaporation depth or obtained from the Penman's equation, *E* multiplied by a conversion parameter *T* (less than unity), as denoted by $E_p = (T \times E)$. To determine the actual evaporation depths, E_a, it is assumed that the catchment is analogous to the vertical stack of horizontal soil layers of the total water storage depth, *Z* (mm). Each soil layer is taken as 25 mm except for the lowest, which may be less than 25 mm (Tan and O'Connor, 1996).

When the rainfall depth *R* exceeds the potential evaporation depth, E_p, a fraction (*H'*) of the excess contributes rainfall, as denoted by $X = (R\text{-}E_p)$ to generated direct runoff, *H* by producing the direct generated runoff component as denoted by $r_1.=H'.x.$ Normally, the fraction *H'* is taken as being proportional to the actual soil-moisture depth, *Z* in the top five layers, as denoted by *H'*=*H* x (*the actual soil moisture depth per 125mm of water),* but if *Z* is less than 125 mm, then *H'* is given by;

$$H^{'} = H \quad x \frac{the \quad actual \quad soil \quad moisture \quad depth \quad in \quad all \quad layers}{the \quad storage \quad capacity, \quad Z} \tag{4.6}$$

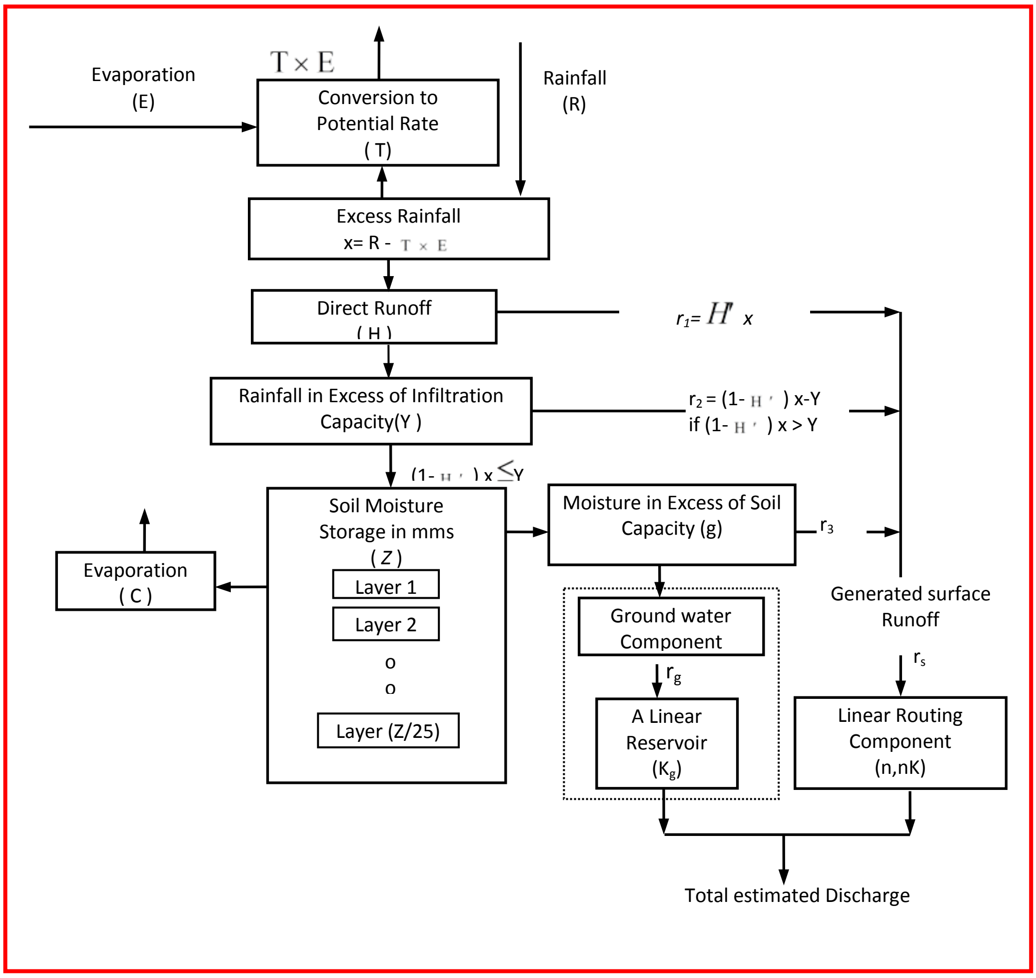
H is a parameter to be optimised. This is a modification of the original version of the SMAR model presented by O'Connell et al. (1970) in which fraction *H'* contributed to the generated runoff, *H* without any consideration of the actual soil moisture depth in the layers.

The remaining rainfall excess, as denoted by *(1-H').X* in excess of the infiltration capacity, *Y* also contributes to the generated runoff, r_2, denoted as $r_2 = \{(1-H').X\}-Y$, whereas the remainder (if sufficient) restores each layer to its field capacity, from the top layer downwards, until the rainfall is exhausted or until all the layers are at field capacity. Any still remaining surplus is further divided into two portions, the first is a ground water runoff component, r_g and other one is a generated runoff component, r_3 was controlled by a weight parameter, *g.* Thus, the total lumped generated runoff produced by the water balance component of the SMAR model *is* $r = (r_1+r_2+r_3)$. The model simulates four components of flow namely: (1) the direct overland flow, (2) the saturation overland flow, (3) the interflow and (4) the groundwater flow. A summary of the model parameters used in the SMAR model has 9 parameters that are described in Figure 4.3.

SMAR model input:

The inputs to the model for each time step are:

- precipitation
- evapotranspiration
- observed runoff



Parameter	Description
Z	The combined water storage depth capacity of the layers (mm)
T	A parameter (less than unity) that converts the given evaporation series to the model-estimated potential evaporation series.
C	The evaporation decay parameter, facilitating lower evaporation rates from the deeper soil moisture storage layers
H	The generated *'direct runoff'* coefficient
Y	The maximum infiltration capacity depth (mm)
n	The shape parameter of the Nash gamma function *'surface runoff'* routing element; a routing parameter
nK	The scale (lag) parameter of the Nash gamma function *'surface runoff'* routing element; a routing parameter
g	The weighting parameter, determining the amount of generated *'groundwater'* used as input to the *'groundwater'* routing element.
K_g	The storage coefficient of the *'groundwater'* (linear reservoir) routing element; a routing parameter

Figure 4.3: Schematic diagram of SMAR model (Tan and O'Connor, 1996)

4.1.4 The Nedbør-Afstrømnings Model (NAM)

The NAM mode is a lumped conceptual rainfall-runoff model which was originally developed in 1973 by the Department of Hydrodynamics and Water Resources at the Technical University of Denmark (DHI, 2007). Since then, the NAM model has become a well-proven engineering tool that has been applied to a number of catchments around the world, representing many different hydrological regimes and climatic conditions (Rahman et al., 2012).

The NAM model simulates the rainfall-runoff process by continuously accounting for the water content in four different and mutually interrelated storages that represent different physical elements of the catchment. The conceptual structure of the NAM model is shown in Figure 4.4. It is a simulation of the land phase of the hydrological cycle. The four storage zones represent the different physical composition units of the catchment in the vertical direction. These storages are snow storage, surface storage, lower or root zone storage and groundwater storages.

The basic data requirements for the NAM model include catchment area, initial conditions and time series of precipitation, potential evapotranspiration, temperature and observed discharge. Nevertheless, the temperature is required when snowmelt is included in the model. For the study of this research the snow melt parameters have been excluded, because the temperature in these study areas is almost never below 0°C (See Chapter 3). Hence, the parameter corresponding to this process is set to zero.

Based on the basic data requirements, the NAM model produces catchment runoff as well as information about other elements of the land phase of the hydrological cycle such as the temporal variation of the evapotranspiration, soil moisture event, groundwater recharge and groundwater levels. The resulting catchment runoff is divided into three components such as (1) overland flow, (2) interflow and (3) baseflow components. The basics of these components are described briefly below:

(1) Overland flow

When the amount of water in surface storage *U* exceeds the upper limit of the amount of water in the surface storage, U_{max}, the excess water P_N will enter the streams as overland flow and as well as to infiltration, *QOF* as denoted by equation below;

$$QOF = \begin{cases} CQOF \dfrac{L/L_{\max} - TOF}{1 - TOF} P_N & for\ L/L_{\max} > TOF \\ 0 & for\ L/L_{\max} \le TOF \end{cases} \tag{4.7}$$

where *CQOF* is the overland flow runoff coefficient ($(0 \le CQOF \le 1)$, *TOF* is the threshold value for overland flow $(0 \le CQOF \le 1)$, *L* is the soil moisture depth in the lower storage zone and L_{max} is the maximum of water content in the lower storage zone.

(2) Interflow

The interflow contribution, QIF, is assumed to be proportional to the amount of water in surface storage, *U,* and to vary linearly with the relative moisture content of the lower zone storage.

$$QIF = \begin{cases} (CKIF)^{-1} \dfrac{L/L_{\max} - TIF}{1 - TIF} U & for\ L/L_{\max} > TIF \\ 0 & for\ L/L_{\max} \le TIF \end{cases} \tag{4.8}$$

where *CKIF* is the time constant for interflow and *TIF* is the root zone threshold value for interflow ($(0 \le TIF \le 1)$.

(3) Baseflow

The baseflow, *BF* from the groundwater storage is estimated as the outflow from a linear reservoir with time constant, *CKBF*:

$$BF = \frac{GWL}{CKBF} \tag{4.9}$$

where GWL is the water depth of the groundwater storage.

The groundwater storage performs as a linear reservoir constantly draining to the stream as baseflow. The overland flow and interflow are routed through one linear reservoir before all of the catchment runoff components are added and routed through a final linear reservoir. In the study of this research, the basic NAM model was applied, including nine parameters to be determined by calibration. Table 4.1 briefly describes these model parameters to generate runoff, where these parameters are sufficient for most applications (e.g., Nlelsen and Hansen, 1973; Tingscachali and Gautan, 2000; Rahman et al., 2012).

NAM model input:

The inputs to the model, for each time step are:

- precipitation
- potential evapotranspiration
- observed discharge

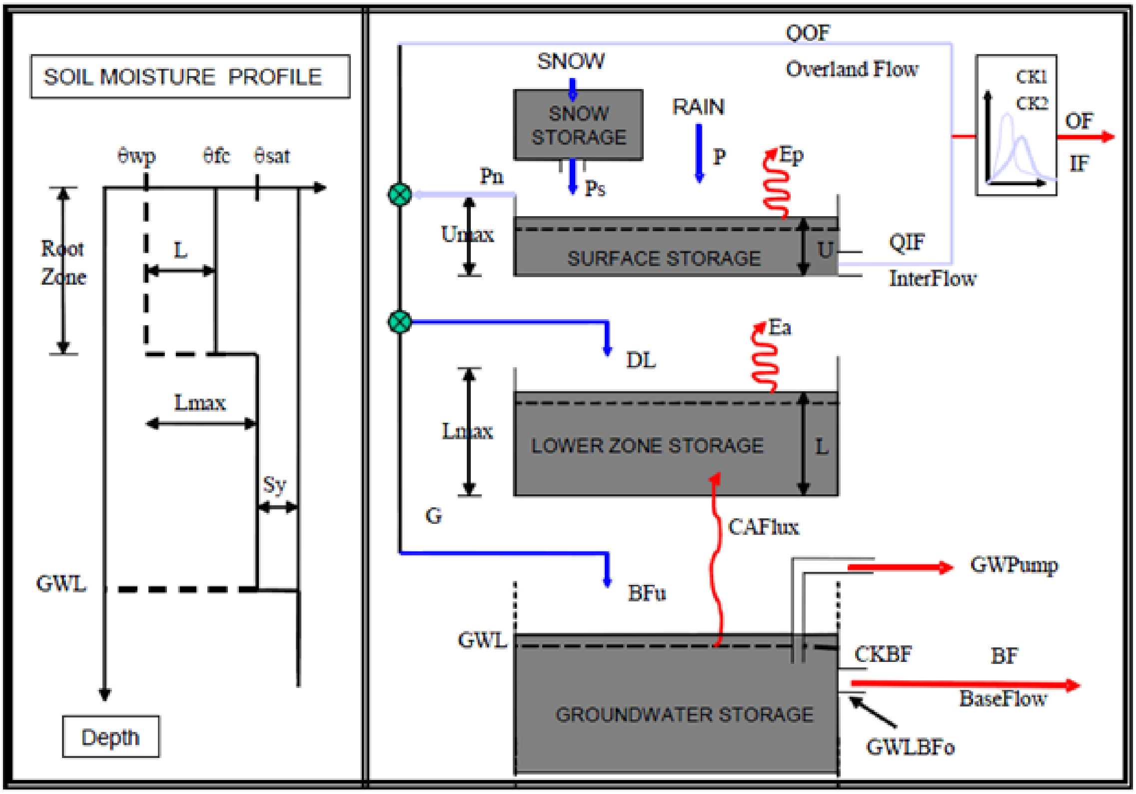


Figure 4.4: Structure of the NAM (DHI, 2007)

Table 4.1: NAM model parameter

Nam model parameter	Description	Unit	Parameter Boundaries
Umax	Maximum water content in surface storage	mm	10-25
Lmax	Maximum water content in root zone storage	mm	Umax=0.1Lmax
CQOF	Overland flow runoff coefficient	-	0-1
CKIF	Time constant for interflow	hour	500-1000
CK1,2	Time constant for routing interflow and overland flow	hour	3-48
TOF	Root zone threshold value for overland flow	-	0-0.7
TIF	Root zone threshold value for interflow	-	0-1
CKBF	Time constant for routing baseflowtime constant	hour	500-5000
TG	Root Zone threshold value for groundwater recharge	-	0-0.7

4.1.5 The SWAT model

The SWAT model is a semi-distributed conceptual model that operates on a daily time step (Arnold et al., 1998). It was developed at the U.S. Department of Agriculture (USDA). The SWAT model was developed to simulate the impact of management on water, sediment and agricultural chemical yields over long periods of time for complex catchments. Currently, the SWAT model has been extensively applied in water resource applications (e.g. Santhi et al., 2001; Cao et al., 2006; Schuol and Abbaspour, 2007; Keshta et al., 2009).

This model simulation requires specific information about weather, soil properties, topography, vegetation and land management practices occurring in the watershed. The simulated hydrological processes include precipitation, infiltration, surface runoff (i.e. the Soil Conservation Services (SCS), Curve Number (CN)), evapotranspiration, lateral flow and percolation. The model subdivides the watershed into several sub-watersheds, which are further divided into hydrological response units (HRUS) according to topography, land use and soil.

The model calculations are performed on a HRU basis and flow and water quality variables are routed from HRU to sub-basin and subsequently to the catchment outlet. The SWAT model simulates the hydrology of a watershed and is divided into two component systems. The first component is the land phase of the hydrologic cycle, as shown in Figure 4.5. This component controls the amount of water, sediment nutrient and pesticide loadings to the main channel in each sub-basin. The second component is the water or routing phase of the hydrologic cycle, which can be defined as the movement of water, sediments, etc. through the channel network of the watershed to the outlet (Neitsch et al., 2011).

The hydrological cycle, as simulated by SWAT, is based on the water balance equation;

$$SW_t = SW_0 + \sum_{i=1}^{t} (R_{day} - Q_{surf} - E_a - w_{seep} - Q_{lat} - Q_{gw}) \tag{4.10}$$

where SW_t is the final soil water content (mm H_2O), SW_0 is initial soil water content on day i (mm H_2O), t is the time (days), R_{day} is the amount of precipitation on day i (mm H_2O), Q_{surf} is the amount of surface runoff on day i (mm H_2O), E_a is the amount of evapotranspiration on day i (mm H_2O), w_{seep} is the amount of water entering the vadose zone from the soil profile on day i (mm H_2O), Q_{lat} is lateral flow from soil to channel and Q_{gw} is the amount of return flow on day *i* (mm H_2O).

The subdivision of the watershed enables the model to reflect differences in evapotranspiration for various crops and soils. The total runoffs are routed from each HRU to sub-watersheds of the catchment outlet. The water balance of each HRU in the watershed contains four storage volumes which are (1) snow, (2) the soil profile (unsaturated), the depth 0-2 m, (3) the shallow aquifer (unconfined), the depth 2-20 m, and (4) the deep aquifer (confined), the depth greater than 20 m (see Fig. 4.5). The soil profile can contain several layers. The soil-water processes include infiltration, percolation, evaporation, plant uptake, and lateral flow. Surface runoff is calculated using the SCS curve number method or the Green-Ampt infiltration equation. Many commonly used watershed models address the SCS curve number method to simulate runoff. For this study, the SCS curve number method was used to calculate surface runoff, the equation is defined as:

The soil conservation service (SCS);

$$Q = \frac{(R-0.2s)^2}{R+0.8s}, R > 0.2s$$

$$Q = 0.0, R \le 0.2s \tag{4.11}$$

where Q is the daily runoff (m^3/s), R is the daily rainfall (mm) and *s* is retention parameter (mm), which is related to the curve number, *CN* as denoted by;

$$s = 254\left(\frac{100}{CN} - 1\right) \tag{4.12}$$

Potential evaporation can be calculated using Hargreaves, Priestly-Taylor or Penman-Monteith method (Arnold et al., 1998). Outflow from a channel is adjusted for transmission losses, evaporation diversions and return flow. A comprehensive description of all the components in SWAT can be found in the literatures (e.g. Arnold and Allen, 1996; Arnold et al., 1998; Srinivasan et al., 1998; Neitsch et al., 2011).

SWAT model input:

- Topography data (i.e. Digital Elevation Model, DEM)
- Soil and Land use (maps and physical parameters)
- Weather data (i.e. precipitation, temperature (minimum and maximum))

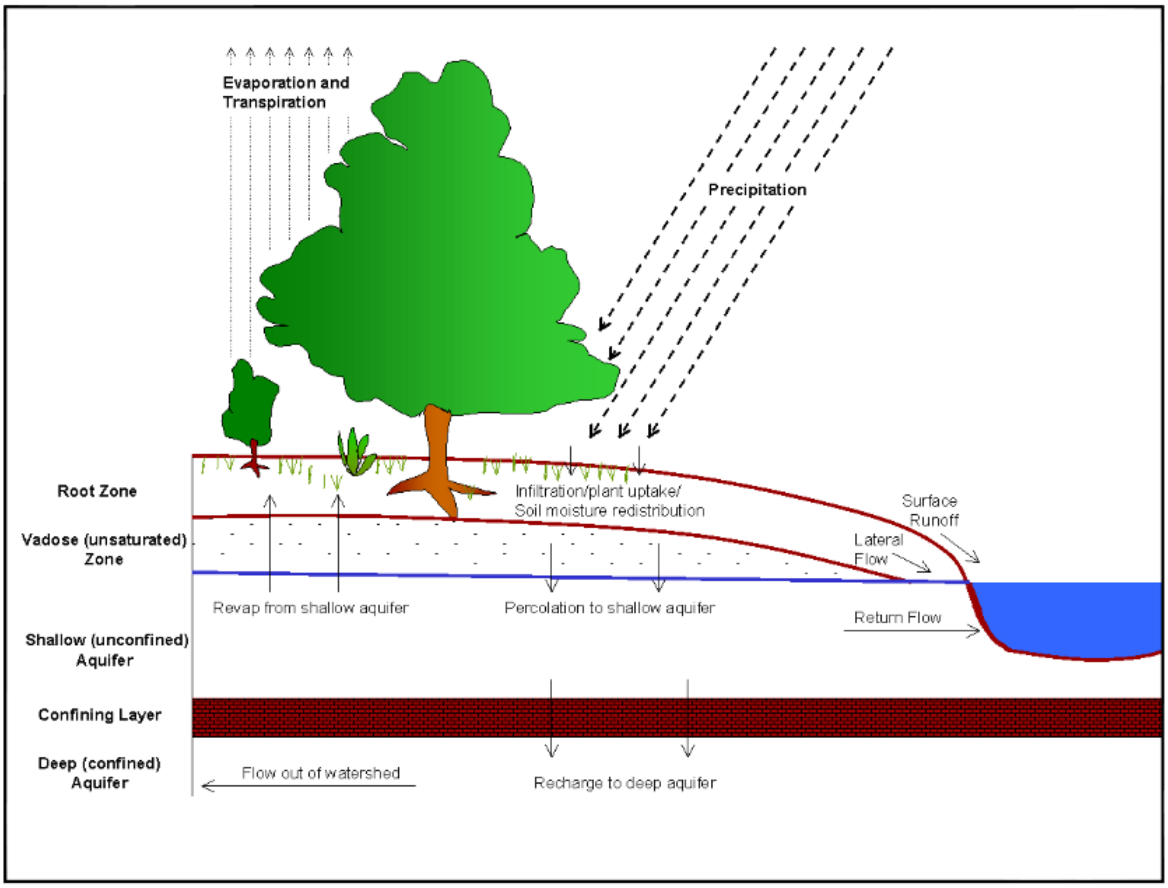


Figure 4.5: Schematic of the SWAT model representation of the hydrological process (Neitsch et al. 2009)

4.2 Evaluation of model performance

Legates and McCabe (1999) recommended that researchers should consider at least one goodness-of-fit and one absolute error measure (e.g. root mean square error, *RMSE*) for providing an accurate evaluation of the model's simulation abilities. To test the ability of the models, the statistical methods and graphical methods (i.e. the hydrograph plots) were used to evaluate the model performance in this study. Five statistical criteria were used including (1) the Coefficient of Efficiency, *CE* (Nash and Sutcliffe, 1970), (2) the Coefficient of Determination, R^2, (3) the Root Mean Squared Error, *RMSE*, (4) the percentage of deviation from observed runoff, *PBIAS*, and (5) the Kling and Gupta Efficiency, *KGE* (Gupta et al., 2009). These methods were used in assessing the relative performance of the selected five rainfall-runoff models and multi-model combination systems in this research. The graphical criteria, namely, the hydrograph plots such as the scatter and time series plot of the predicted flow and observed flow are used to judge the model performance.

The most widely used statistics reported for calibration and validation are R^2 and *CE* (ASCE, 1993; Duan et al., 2003). The R^2 provides an estimate of how well the variance of observed runoff values is replicated by the model predictions as where the values range between 0 to 1.0. The *CE* indicates how well the plot of the observed versus predicted values fits the 1:1 line. The *CE* ranges vary from $-\infty$ to 1; a CE value of 0 indicates the model performs no better than the average of the observed data and the CE value of 1 indicates a perfect match between the observed and predicted values. If R^2 and *CE* values are close to zero, the model prediction is considered unacceptable. In contrast, if these values approach one, the model predictions become highly accurate, and the negative values of *CE* indicate that the observed values are the better predictor than the predicted values.

The *RMSE* is the commonly used error index statistic (Chu and Shirmohammadi, 2004). *RMSE* is very sensitive to even small errors, which is good for comparing small differences in model performance. It indicates the absolute fit of the model to the data and how close the observed data points are to the model predicted values. It is a non-negative metric that has no upper bond and for a perfect model the *RMSE* value is close to zero. The *PBIAS* measures the average trend of the predicted values to be larger or smaller than observed values. The optimal value of *PBIAS* is 0.0, with low-magnitude values indicating accurate model prediction. The positive values of *PBIAS* indicate predicted values underestimation of observed values. The negative values of *PBIAS* indicate predicted values overestimation of observed values (Gupta et al. 1998).

The recently proposed *KGE* developed by Gupta et al (2009) is used to improve the Nash-Sutcliffe efficiency. This method can help reduce model calibration problems for river flow forecasting (Gupta et al. 2009). For a perfect model, the *KGE* values are close to one, which is similar to R^2 and *CE* as above. See Gupta et al. (2009) for further details of the *KGE* and its components. The equations used for calculating these statistics are shown below.

i. The Coefficient of Efficiency, *CE*

$$CE = 1 - \frac{\sum_{i=1}^{n}(S_i - Q_i)^2}{\sum_{i=1}^{n}(Q_i - \overline{Q})^2} \tag{4.13}$$

ii. Coefficient of Determination, R^2

$$R^2 = \left[\frac{\sum_{i=1}^{n}(Q_i - \overline{Q})(S_i - \overline{S})}{\sqrt{\sum_{i=1}^{n}(Q_i - \overline{Q})^2 \sum_{i=1}^{n}(S_i - \overline{S})^2}}\right]^2 \tag{4.14}$$

iii. Root Mean Square Error, *RMSE*

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(S_i - Q_i)^2}{n}} \tag{4.15}$$

iv. The percentage of deviation from observed runoff, *PBIAS*

$$PBIAS = \frac{\sum_{i=1}^{n}(Q_i - S_i)}{\sum_{i=1}^{n} Q_i} x100 \qquad (4.16)$$

v. The Kling and Gupta Efficiency, *KGE*

$$ED = \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \qquad , \alpha = \frac{\sigma_{S_i}}{\sigma_{Q_i}}, \beta = \frac{\bar{S}}{Q'}$$

$$KGE = 1 - ED$$

(4.17)

Where Q_i is the observed runoff value at time *i* *(m^3/s)*, S_i is the simulated runoff at time *i* (m^3/s), $\overline{Q}$ is the mean of the observed runoff data, $\overline{S}$ is the mean of the simulated runoff data and n is the number of data points.

4.3 Summary

In this chapter of the thesis, the five rainfall-runoff models which are used in producing the multi-model combination system and the evaluation of model performance have been presented. First the general description of the rainfall-runoff model and the selection rainfall-runoff models used for this study were presented. This was followed by a brief description of each selected rainfall-runoff model and the requirement for model input. For this study, three types of rainfall-runoff models (i.e. the empirical black-box models, conceptual models and physically-based model) were applied in multi-models for the first time in the context of rainfall-runoff model combinations. The selected models were used to provide discharge estimates for the case studies of two catchments (see Chapter 3) in this study. The chapter ends with the details of model efficiency evaluation criteria used for assessment of model efficiency.

Chapter 5

Multi-model approach using ANNs and GEP

This chapter presents the use of three combination methods: the Gene Expression Programming (GEP), the Multi-Layer Perceptron Neural Network (MLPNN) and the Radial Basis Function Neural Network (RBFNN) to develop multi-model combination systems. It is related to the first objective (see Chapter 1) of this thesis. The work in this chapter was motivated by a desire to investigate the performance of these three combination methods in the multi-model combination for this study (see Section 2.6 in Chapter 2). These three methods were used to combine the results from five different rainfall-runoff models to test the multi-model combination systems. The author's conclusion, based on the literature (see Chapter 2), is that the comparison of the performance of GEP and ANNs has not been applied in the context of combinations of rainfall-runoff models. Hence, this project addresses the issue for the first time.

The aims of the study in this chapter are: (1) to compare the performance of the GEP with two previously investigated ANNs: the MLPNN and RBFNN used by Shamseldin et al. (2007) in the multi-model combination systems, and (2) to investigate whether or not the use of GEP will lead to further improvement in the performance of multi-model combinations as well as, or

Note: Major research reported in this chapter is based on work submitted in the Journal of Hydrologic Engineering- *"Multi-model approach using neural networks and symbolic regression to combine the estimated discharges of rainfall-runoff models."*

better than, the two previously investigated ANNs (MLPNN and RBFNN) combination methods. This chapter starts by presenting the study areas, data and rainfall-runoff models used in the study. Following that, a brief description of the combination methods is presented. Then, the methodology of this research is discussed. After that, the details of model efficiency evaluation criteria used for assessment of model efficiency are presented. Finally, the summary obtained from the results and discussion is presented.

5.1 Study areas, data, and rainfall-runoff models

5.1.1 Study areas and data

Two contrasting catchments are used in this study to compare the performance of the GEP and the two ANNs combination methods (MLPNN and RBFNN) methods in the multi-model combination systems. These two catchments are the Mae Tuen River catchment located in Thailand and the Ohninemuri River catchment located in New Zealand. The two catchments differ with respect to climate, topography, geology, land use, and rainfall-runoff response. The details related to the study areas and data are already presented in Chapter 3 of this thesis.

5.1.2 Rainfall-runoff models

The comparative study presented in this research involves the use of three types of rainfall-runoff models, specifically, two empirical black-box models, two conceptual models and a semi-distributed physically based model. The two empirical black-box models selected for this study are the linear perturbation model (LPM) and the linear varying gain factor model (LVGFM). The two conceptual models selected in this study are the soil moisture accounting and routing (SMAR) model and the Nedbør-Afstrømnings Model (NAM). The semi-distributed physically based model selected is the soil and water assessment tool (SWAT). A brief description of these five rainfall-runoff models is already presented in Chapter 4 of this thesis. These five models are selected for the purpose of the investigation on the basis of model accuracy, familiarity with the model, the available data, availability of license software and having performed well in previous use in river flow simulations (Shameseldin et al. 1997 and 2007; Refsgaard 1997; Alansi et al. 2009; Rahman et al. 2012; Vu et al. 2012). These models are used to provide discharge estimates for two catchments - one is located in Thailand and the other is located in New Zealand.

5.2 The combination methods

This section of the chapter briefly describes three combination methods: (1) the multi-layer perceptron neural network (MLPNN), (2) the radial basis function neural network (RBFNN) and (3) the gene Expression programming (GEP). These three combination methods were used to combine the estimated discharges from the selected five rainfall runoff models to test the multi-model combination systems in Thailand and New Zealand catchments.

5.2.1 Multi-layer perceptron neural network (MLPNN)

The MLPNN is a feed-forward network and the most commonly used neural network type in hydrological applications. The MLPNN used in this study consists of three-layers, namely, an input layer, a hidden layer and an output layer (see Fig. 5.1), which each consist of several neurons. In general, one hidden layer is enough to handle almost all sorts of problems (Sherrod, 2003). The utilization of two hidden layers hardly contributes to any improvement to the model performance. This may also create the possibility of converging into local minima (Sherrod, 2003). In this study, the input layer has five neurons with the simulated runoff output from each of the rainfall-runoff models being assigned to one neuron. The five input neurons of each layer are connected to the neurons of the single hidden layer by weights, which produce the hidden output neuron. The transfer function in the hidden nodes of neural networks is usually a nonlinear function. The most widely used transfer function is the sigmoid function. It is used in this study in order to give the neural network the capacity of learning possible nonlinear functions. While in the output node, the linear transfer function is used instead in this study.

If we let *i, j,* and *k* represent the indices (number of neurons) of input, hidden and output layers, respectively, X_i represents the *ith* neuron of the input layer, H_j represents *jth* neuron of the hidden layer and represents the *kth* neuron of the output layer. The input neurons (X_i) of each layer are connected to the neurons of the single hidden layer by weights, which produce the hidden output neuron. The input-output transformation in each hidden neuron is achieved by a mathematical nonlinear transfer function, as defined by:

$$Y_j = f_j(\alpha_j + \sum_{i=1}^{n} w_{i,j} x_i) \tag{5.1}$$

where f_i is the activation function of hidden neurons, n is the number of neurons in the hidden layer, α_j is the threshold or bias of the hidden neuron, $w_{i,j}$ is the connection weight from the input node to the hidden node and n is the number of neurons in the input layer.

The output layer is the last layer, having a single neuron which produces the output neuron, Y_k , it is obtained in a similar way as the neurons in the hidden layer, using:

$$Y_k = f_k(\beta_k + \sum_{j=1}^{m} w_{j,k} Y_j) \tag{5.2}$$

where f_k is the activation function of output neurons, β_k is the threshold of output neurons, $w_{j,k}$ is the connection weight from hidden node to the output node and m is the number of neurons in the hidden layer.

The simulated model output, $\hat{Y}_k$ corresponding to the input vector, $X = [X_1, .., X_n]^T$ and the synaptic weights, $\hat{w}$ and a Gaussian white noise (model noise), ε , which can be represented as:

$$\hat{Y}_k = f(X; \hat{w}) + \varepsilon \tag{5.3}$$

where a hat denotes the simulated value. The function $f(x:\hat{w})$ will be referred to as a regression function.

Regression assumes that target *Y* is related to input vector *X* by stochastic and deterministic components. The stochastic component is the random fluctuation of *Y* about its mean $\mu_y(X)$:

$$Y = \mu_y(X) + \varepsilon(X) \tag{5.4}$$

If the 'true' functional relationship between $\mu_y(X)$ and *X* is defined by

$$\mu_y(X) = f(X : X_{true}) \tag{5.5}$$

where there is a set of parameters, regression is the attempt to estimate this relationship from a finite data set (a derivation or training set) by estimating the parameter values from the dataset.

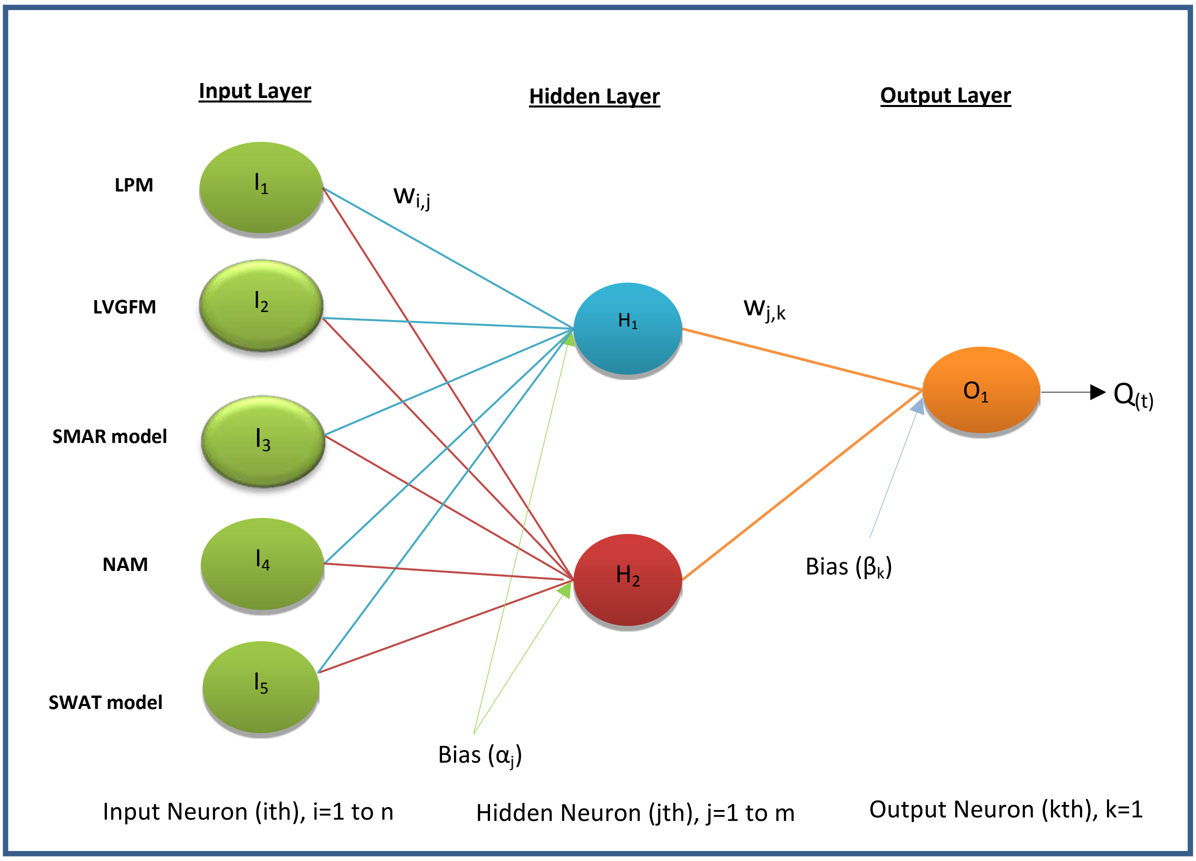


Figure 5.1: Network diagram showing the connection weights of MLPNN structure

5.3.2 Radial basis function neural network (RBFNN)

The RBFNN is another special type of neuron networks designed to solve the complex and common set of problems in hydrological applications. RBFNN is based on supervised learning and a feed forward neural network. It consists of an input layer, one hidden layer with a non-linear RBF activation function and a linear output layer (see Fig. 5.2). The input layer is a first layer with the number of neurons equal to a number of individual models. For the *jth* time period, the hidden layer transforms the data from the input space to the hidden space using a non-linear function, as there are no weights on the lines from the input nodes to the hidden nodes.

Each neuron in the hidden layer consists of a radial basis function (also called a kernel function) ϕ, which computes the distance between the input vector $X_i = \lfloor x_{1,j}, x_{2,j}, \ldots, x_{n,j} \rfloor$ and the neuron centre vector, c_j. The output of each hidden unit, H_l, is given by;

$$H_l(X_i) = \phi(\|X_i - c_j\|) \tag{5.6}$$

where $\|X_i - c_j\|$ is the Euclidean distance between the input and the hidden nodes. Different types of radial basis function could be used, but the most common is the Gaussian function; we consider that the radial basis function, as the Gaussian function, can be written as;

$$\phi_j(x_i) = \exp\left(-\frac{\|X_i - c_j\|^2}{2\sigma_j^2}\right) \tag{5.7}$$

where *x* is the training data and σ_j is the width of the Gaussian function. The final network output implement linear summation function as in an MLPNN, Y_k, or the output vector $Y_m = [Y_1(X), Y_2(X), ..., Y_n(X)]$, Y_m , which is a linear function, can be calculated as;

$$Y_m(X) = \sum_{j=1}^{n} w_{lm}\varphi_j(X_i) + \beta_m \tag{5.8}$$

where n is the number of neurons in the hidden layer, w_{lm} denotes the corresponding weights connecting the hidden neuron, l to the output neuron, m and β_m is the bias of output later.

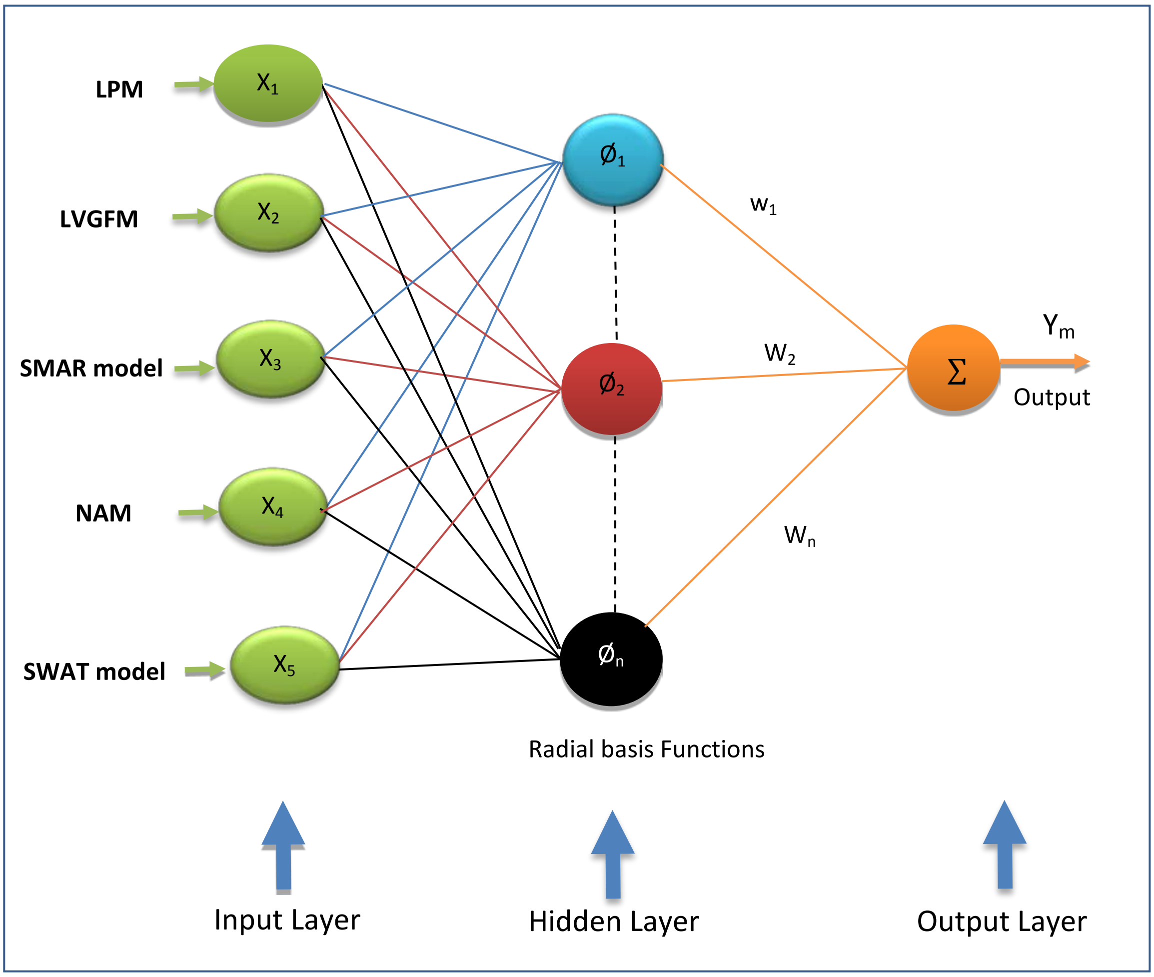


Figure 5.2: Network diagram showing the architecture of the radial basis function neural network

5.3.3 Gene Expression Programming (GEP)

Symbolic regression is the problem of fitting a function that involves finding a mathematical expression in symbolic mathematical form to provide the best relationship between dependent variables and independent variables. Gene Expression Programming (GFP) is an algorithm for performing symbolic regression by using a genetic and evolution algorithm to find a mathematical function that fits a set of data. Unlike traditional linear and non-linear regression, it does not require the form of the function to be specified in advance.

Symbolic regression was first proposed by Koza (1992) where the genetic algorithms (GAs) were first used to solve the symbolic regression problems. However, Koza (1992) proposed a powerful extension to GAs known as genetic programming (GP) to solve symbolic regression problems. GP is evolutionary search (optimization) techniques which are based on Darwin's evolution theory. GEP is like GAs and GP, a genetic algorithm as it uses populations of individuals, selects them according to fitness and introduces genetic variation using one or more genetic operators (Ferreira, 2006). The major differences between these three algorithms exist in the nature of their individuals. These are: the GAs are linear strings of fixed length (chromosomes); the GP are nonlinear entities of different sizes and shapes (parse trees), and the GEP are encoded as linear strings of fixed length (the genome or chromosomes) which are afterwards expressed as nonlinear entities of different sizes and shapes (i.e., simple diagram representations or expression trees). However, experiments have shown that GEP is 100 to 60,000 times faster than older GAs (Sherrod, 2003).

An original mathematical function of a symbolic regression application, (where a label *Y* is provided for the input *X* and a linear symbolic relationship of the form *Y*) can be expressed by

$$Y = a + b * X \tag{5.9}$$

where *X* is the independent variable, *Y* is the dependent variable, and *a* and *b* are parameters whose values are to be computed by the regression algorithm.

The nonparametric function of a symbolic regression is not known in advance, where the goal of the procedure is to find a mathematical expression that satisfactorily explains the relationship between the dependent and independent variables. It could be given in the form:

$$Y = f(\overline{X}) \tag{5.10}$$

where *Y* is a function and $\overline{X}$ is the input vector, $\overline{X} = \left[x_1, x_2, ..., x_{n,}\right]$

There are many potential forms of nonparametric functions including neural networks, polynomial constructs and decision trees. Symbolic regression is a subset of nonparametric functions that restricts the functions to be mathematical or logical expressions. GEP is a form of nonparametric function, which is used to instantaneously select the optimum set of mathematical expressions which involve the appropriate input variables. GEP tries to "evolve" a function to provide the value of "*Y*" based on input variables $X = (X_1, X_2, \ldots, X_n)$.

In this study, the development of multi-model combination systems is viewed as a symbolic regression which is solved using GEP. The multi-models developed are referred to in this study as the GEP-based multi-model combination. For this study, the GEP multi-models developed employs the simulated runoffs obtained from the selected five rainfall-runoff models to produce a multi-model combination system for the two contrasting catchments located in Thailand and New Zealand.

In the GEP process, several computer programs are evolved and selected on the basis of their fitness for solving that particular problem. The GEP operation used in solving the symbolic regression problem is shown in Figure 5.3. The operation can be summarized in the following steps (Ferreira, 2001);

- Generation of random potential solutions (chromosomes).
- Encode each solution as a computer program.
- Evaluate the fitness of each program.
- Select the best performing programs.
- Reproduction of new solutions using the best performing individual programs through genetic operators (i.e. replication, modification, transposition and recombination).
- The process is repeated until the stopping criterion is fulfilled.

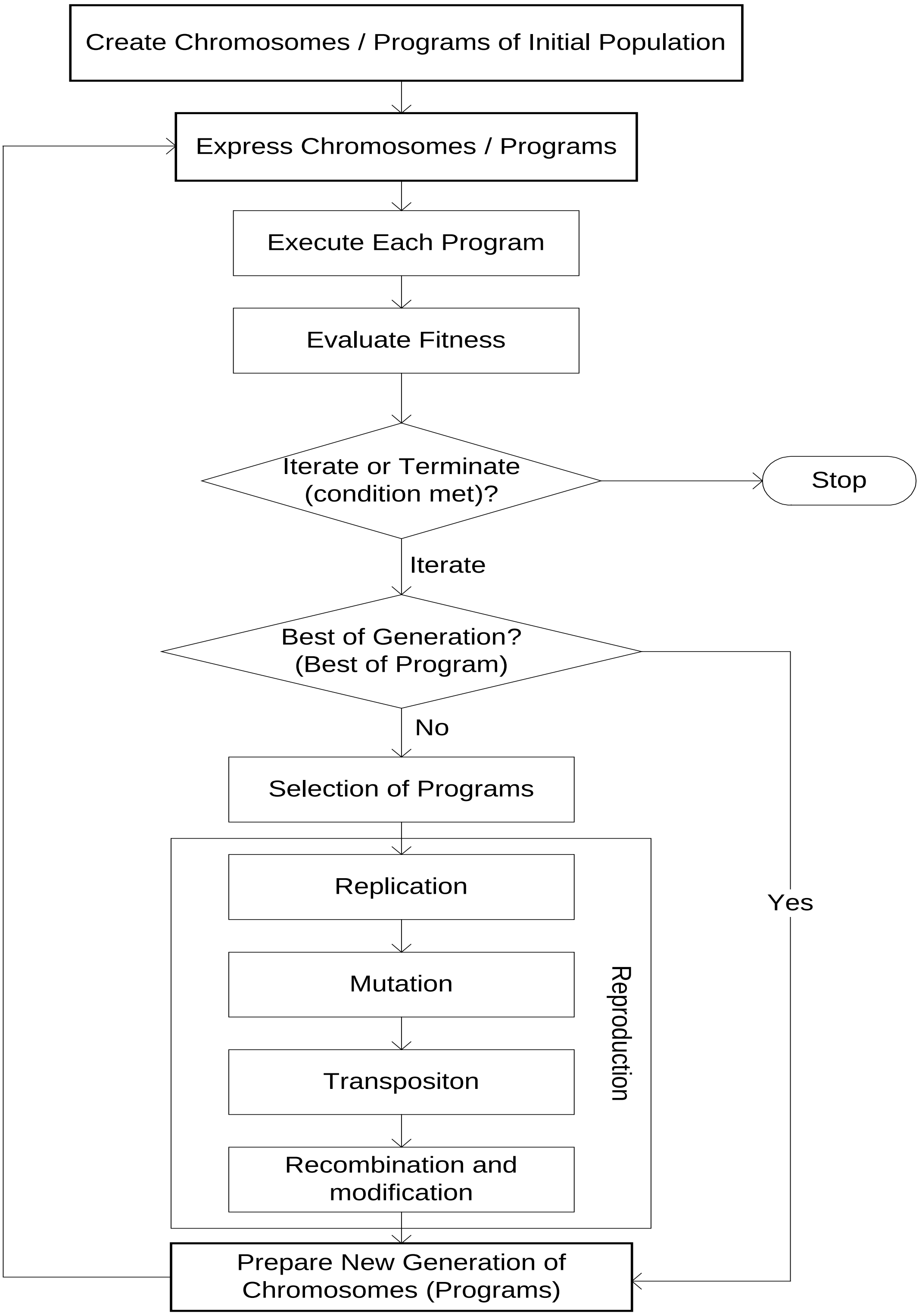


Figure 5.3: The flowchart of a gene expression algorithm (Ferreira, 2001).

5.4 Methodology

In this section, the application of the selected five rainfall-runoff models (i.e. LPM, LVGFM, SMAR model, NAM model and SWAT model) is presented. Each of the five rainfall-runoff models was applied to test the Mae Tuen River catchment located in Thailand and the Ohinemuri River catchment located in New Zealand, respectively. Each rainfall-runoff model was run with the daily data (i.e. rainfall, evapotranspiration, temperatures (maximum - minimum)) for each catchment a number of times to determine the optimum set of values of coefficients or the parameters for obtaining the model forms. Then, the applications of the developed multi-model combinations which were produced by the three combination methods - GEP, MLPNN and RBFNN, are presented.

5.4.1 Application of the selected five rainfall-runoff models

Different types of rainfall-runoff models are the LPM, the LVGFM, the SMAR model, the NAM model and the SWAT model, were applied to the daily data of two catchments in New Zealand and Thailand to test the multi-model combination systems in this study. They were used to simulate runoff at the outlet of two river flow gauging stations namely gauge P64 (see Fig. 3.1) and gauge Ohinemuri River (see Fig. 3.4), respectively. For model evaluation, the input data was split into two parts (see Table 3.2) for each rainfall-runoff model. The first part is 2/3rds of the available data, which was used for model calibration. The remaining 1/3rd was used for model verification (i.e. testing the consistency of the calibrated model on an independent set of data). All models were calibrated for each catchment to determine the optimum set of the parameter values.

The LPM used the seasonal expectation of daily observed rainfall and discharge time series to estimate the parameters of the model over the calibration period. The ordinary least squares (OLS) solutions were used for estimating the pulse response function except for the parametric forms where the parameters were optimised. In the application of the LVGFM, the catchment wetness index (the soil moisture state), $Z_{(t)}$ is determined from Eq. 4.4 (see Chapter 4), in which it was necessary first to fit the auxiliary SLM to the rainfall-runoff data for both catchments. The time varying gain factor G, used to calculate the soil moisture state, $Z_{(t)}$ in Eq. (4.4), is obtained by finding the sum of the least squares discrete pulse response ordinates of the SLM. The ordinary least squares (OLS) solution was used for estimation of the ordinates of the weighting function, and then the parameters of the results were optimised. Once the estimates of the parameters were obtained, the simulated flows could be obtained from the outputs of the calibrated auxiliary SLM. Then, the LVGFM was applied to the data of each catchment.

The SMAR model parameters were estimated using a simplex method of optimisation (Nelder and Mead, 1965). In this study, the SMAR model has nine parameters. There are five water balance parameters, T, H, Y, C and Z, plus the weighting parameter G for groundwater routing and the three routing parameters, N, NK and Kb giving a total of nine parameters in the SMAR model. The SMAR model was calibrated using daily discharge, rainfall and evaporation data. The NAM model set up requires initial conditions and time series of precipitation, temperature, evapotranspiration and stream discharge data. The calibration involves adjusting the coefficients for the exchange of water between storage units and the storage unit depth, so that model outputs match observed discharge values as closely as possible. The initial calibration was derived by using an auto-calibration tool provided within the model. The model was calibrated with the observed discharge by adjusting the response parameters. Its steps run through 4000 combinations of the different parameters, while trying to minimize the root mean square error and improve the water balance in order to match with the observed data.

The SWAT model version 2009 (Neitsch et al., 2011) was used in this study, which combines the SWAT model with the Arcview GIS interface version 9. To setup a SWAT model, the input data is a digital elevation model (DEM), soil types, land use and weather data. The three main steps of the SWAT model applied to set up a watershed simulation in this study are: (1) watershed delineation, (2) Hydrological Response Unit (HRU) analysis and, (3) weather data definition (see Figure 5.4). The first step is to partition both catchments into a number of sub-catchments. In particular, river networks of both catchments are automatically delineated from DEM by means of ArcSWAT interface. As a result, 31 sub-catchments and 29 sub-catchments respectively were generated from the Mae Tuen River catchment and the

Ohinemuri River catchment. For the second step, the HRU analysis used land use, soil, and slope for the input data. Once each layer was loaded, the delineated sub-basin maps, land use and soil maps were overlaid to determine the HRU features. The HRUs were portions of a sub-basin that possess unique land use and soil attributes. The SWAT simulates different land uses in each sub-basin. For a basic simulation, only a limited amount of data is required. In the final step, the precipitation and temperature (maximum and minimum) time series for each weather station were used for the weather data definition stage. Weather stations were assigned to each sub-catchment. The outputs from these steps were then used as inputs for the SWAT simulation (see Fig. 5.4). Figure 5.4 shows a flowchart of how GIS layers are integrated into ArcSWAT and prepared for a simulation of the SWAT model. The model was calibrated using the discharge data at the catchment outlet. The criterion used for calibrating the model was to minimize the differences between the simulated and observed discharges. The SWAT model calibration can be completed manually or by using the auto-calibration tool in SWAT (Van et al., 2005) or SWAT-CUP (Abbaspour et al., 2007).

For model calibration, the SWAT model parameters for simulated discharge were calculated using SWAT-CUP in this study. SWAT-CUP (SWAT Calibration Uncertainty Programs) is an interface that was developed for SWAT. Its main function is to calibrate SWAT and perform validation, sensitivity, and uncertainty analysis. In this study, sensitivity analysis was carried out using the sequential uncertainty fitting algorithm (SUFI-2) to obtain better estimates of the initial parameter ranges and to improve the effect of all parameters on observed discharge. More details of SWAT-CUP can be found in Abbaspour et al. (2007). In SUFI2, the parameter uncertainty accounts for all sources of uncertainties such as uncertainty in driving data variables (e.g., rainfall), the conceptual model, the parameters and the measured data. A schematic of the linkage between SWAT and SUFI2 is illustrated in Figure 5.5. Both watersheds were calibrated with established manual and auto-calibration methods. Once the model was calibrated, the simulated discharges were evaluated.

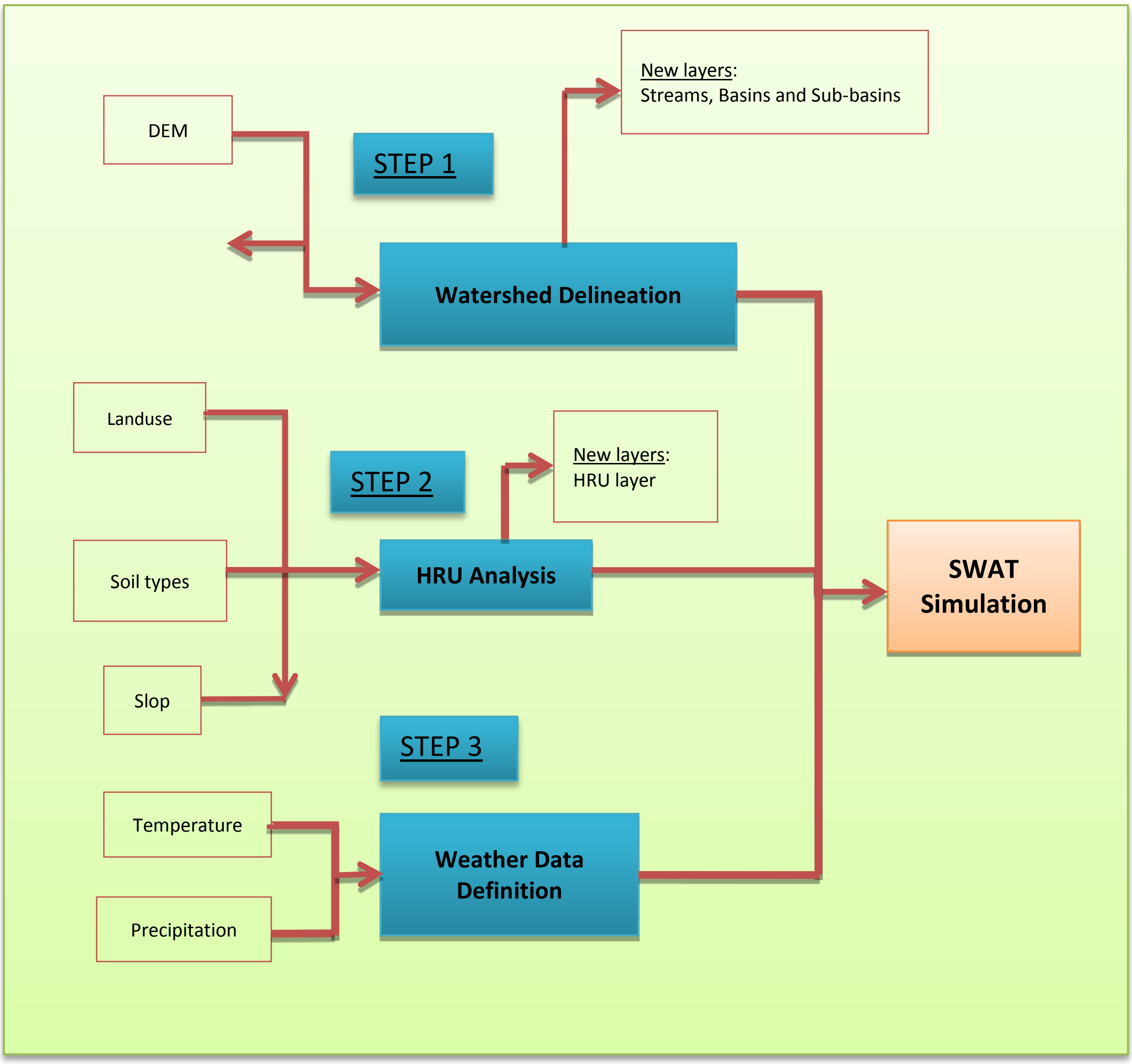


Figure 5.4: Schematic of SWAT model simulation

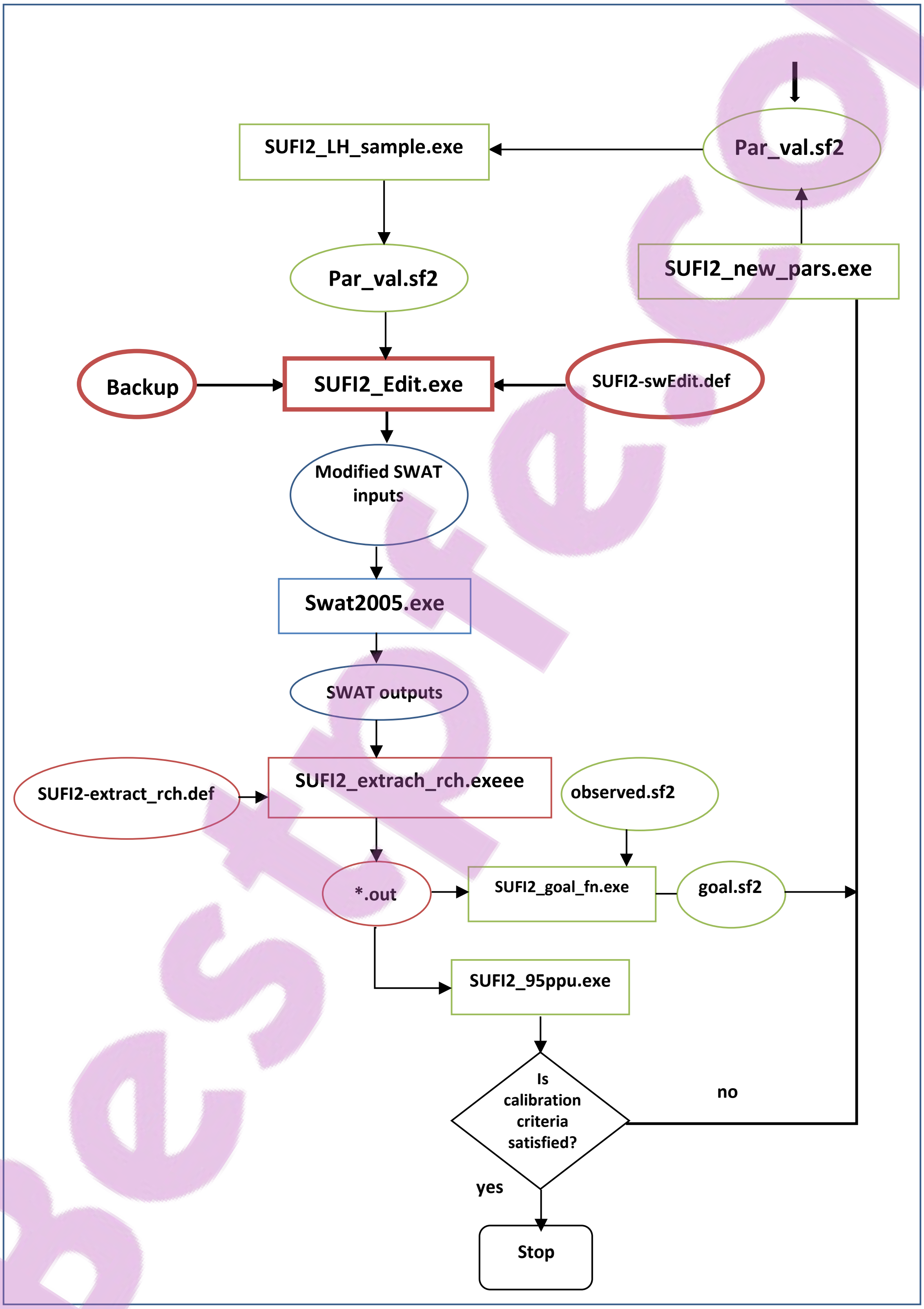

SUFI2_LH_sample.exe
Par_val.sf2
Par_val.sf2
SUFI2_new_pars.exe
Backup
SUFI2_Edit.exe
SUFI2-swEdit.def
Modified SWAT inputs
Swat2005.exe
SWAT outputs
SUFI2-extract_rch.def
SUFI2_extrach_rch.exeee
observed.sf2
*.out
SUFI2_goal_fn.exe
goal.sf2
SUFI2_95ppu.exe
Is calibration criteria satisfied?
no
yes
Stop

Figure 5.5: Showing the link between SWAT and SUFI2 (Abbaspour et al. 2007)

5.4.2 The application of three combination methods

The three combination methods, namely, MLPNN, RFBNN and GEP, were applied to produce the results from the selected five rainfall-runoff models. The DTREG (pronounced as D-T-Reg) which is a very recent and advanced predictive modelling software (Sherrod, 2003) was used to develop the MLPNN, RBFNN and GEP combination methods for this study.

5.4.2.1 The MLPNN

The determination of an optimal amount of neurons in the hidden layer is the most important aspect of the MLPNN architecture. However, using too many neurons may result in over-fitting data. If an inadequate number of neurons are used, the network will be unable to model complex data, so the resulting fit will be poor (Sherrod 2003). To date, there has been no exact solution to the question of how many neurons to use for the hidden layer (Stathakis, 2009). To build the MLPNN combination for both catchments, the numbers of neurons used in the hidden layer were varied systematically between 2 neurons and 50 neurons. The ANNs were trained in batch training using the scaled conjugate gradient method to find the optimal values of the weight and bias parameter values. To find the optimal number of hidden neurons, MLP used a step-wise search with increasing numbers of hidden neurons. It evaluated each one using cross validation. It is also used in order to avoid overtraining. The final architecture of the MLPNN model is optimally to have between 5 and 14 neurons in the hidden layer for both the Mae Tuen River and the Ohinemuri River catchments.

5.4.2.2 The RBLNN combination method

To build the RBFNN combination for both catchments, the optimum numbers of neurons in the hidden layer was determined by using an evolutionary method called Repeating Weighted Boosting Search (RWBS), developed by Sheng et al. (2005). This algorithm uses an evolutionary approach to determine the fitness centre points and spreads for each neuron. It also determines when to stop adding neurons to the network by monitoring the estimated leave-one-out (LOO) error and terminating when the LOO error increases due to over-fitting (Sherrod, 2003). The computation of the optimal weights between the neurons in the hidden layer and the summation layer is done using ridge regression. The training of RBFNN was initiated by determining the number of neurons in the hidden layer, the coordinates of the centre of each hidden layer and the radius (spread) of each RBFNN function in each dimension. Then, the weights were applied to the RBFNN function outputs as they are passed to the summation layer. The final architecture of the RBFNN model was developed having an optimal 8 and 7 neurons in the hidden layer for both Ban Laung and Ohinemuri catchments.

5.4.2.3 The GEP combination method

The GEP was used to find the mathematical function (see Eq. 5.10) relationship between the input variables (the individual rainfall-runoff model simulated runoffs) and the output (the combined runoffs). The simulated runoff of five rainfall-runoff models were model input parameters and the multi-model combination were model output parameters, so the mathematical function can be expressed in the form:

$$Q_t = f(Q_{LVM}, Q_{LVGFM}, Q_{SMAR}, Q_{Nam}, Q_{SWAT}) \tag{5.11}$$

where Q_t is the combined runoffs of the multi-model combination system, Q_{LVM} is the simulated runoff values obtained from LPM, Q_{LVGFM} is the simulated runoff values obtained from LVGFM, Q_{SMAR} is the simulated runoff values obtained from SMAR, Q_{NAM} is the simulated runoff values obtained from NAM and Q_{SWAT} is the simulated runoff values obtained from the SWAT model.

For both catchments, the GEP control parameters used in GEP model developments are shown in Table 5.1. DTREG software was used for the GEP algorithm (Sherrod 2003), so the mathematical functions obtained from the GEP model for the time period, t, of both catchments, produce the best fit for the data in the form of;

$$Q_t = ((((-0.6632368)Q_{SWAT}) + Q_{NAM}) - 3.059299) + Q_{LPM}$$

(5.12)

$$Q_t = (a\tan(Q_{NAM} - Q_{LVGFM})^3) + a\tan(a\tan(a\tan(Q_{SMAR} - (Q_{LVGFM} - Q_{LPM})))) + \\ a\tan(Q_{SMAR} - (Q_{LVGFM} - 4.4467)) + Q_{LVGFM}$$

(5.13)

equations (5.12) and (5.13) are the mathematical functions for the GEP- combination models of Mae Tuen River and Ohinemuri River catchments, respectively.

Table 5.1: The parameter setting for GEP

Control Parameter	Value
Population size	50
Max.tries for initial population	10000
Genes per chromosome	4
Gene head length	8
Max. generations	2000
Functions	+, -, *, /, Sqrt, Sin(x), Cos(x), Atan(x), Exp(x), Log(a), Square(a), Cube(a)
Mutation rate	0.044
Inversion rate	0.1
Transposition rate	0.1
One-point recombination rate	0.3
Two-point recombination rate	0.3

5.5 Evaluation of model performance

The selected five rainfall-runoff models and three combination methods (MLPNN, RBFNN and GEP) of combining the model output were all evaluated using (1) the Root Mean Squared Error, *RMSE*, (2) the Coefficient of Determination, R^2, (3) the percentage of deviation from observed runoff, *PBIAS*, (4) the Coefficient of Efficiency, *CE* (Nash and Sutcliffe, 1970) and (5) the Kling and Gupta Efficiency, *KGE* (Gupta et al., 2009). The description of these methods used for evaluating model performance is already provided in Chapter 4 in this thesis. The graphical criteria, namely, the hydrograph plots and scatter plots, were used in assessing the model performance.

5.6 Results and Discussion

5.6.1 Five rainfall-runoff models

The five rainfall-runoff simulation models: LPM, LVGFM, SMAR model, NAM and SWAT model and each of the three multi-models (i.e. MLPNN, RBFNN and GEP) performance measures for the Mae Tuen River and the Ohinemuri River catchments located in Thailand and New Zealand respectively, are presented in Table 5.2, which also shows the model efficiency values and their ranking.

In the calibration period of the Mae Tuen River catchment, the results indicated that the LPM outperformed other rainfall-runoff models which were characterised by strong seasonality (See Chapter 3). For the Ohinemuri River catchment, the LVGFM performs better than other individual rainfall-runoff models in terms of *R^2, RMSE, CE and KGE*. In the case of the verification period of the Ohinemuri River catchment, it was found that the results were variable; the SWAT model performs better than other rainfall-runoff models in terms of *R^2.* However, the LPM performs better than other individual rainfall-runoff models in term of *RMSE, CE, KGE and PBIAS.* For the Mae Tuen River catchment, the NAM model performs better than other individual rainfall-runoff models in terms of *R^2, RMSE, CE and KGE,* except only in term of *PBIAS, the LPM* outperformed other individual rainfall-runoff models. Regarding the two black-box models including LPM and LVGFM, which involve only the parameters of the daily observed rainfall and discharge time series and rely on the ordinary least squares to produce the simulation, results show better outputs than the distributed physically-based model (SWAT model), which involved many parameters based on the complex law of physic elements generally expressed as systems of non-linear partial differential equations.

5.6.2 The developed multi-model combinations

In the calibration period, the results in Table 5.2 show that the performance of GEP is generally the best in terms of *R^2, RMSE, CE* and *KGE* for the Ohinemuri River catchment. However, only the RBFNN outperformed other individual simulation models and the other two combination methods in term of PBIAS.

For the Mae Tuen River catchment, the RBFNN performed better than other individual simulation models and other two combination methods, in terms *of R^2, RMSE, CE and KGE*. For the Ohinemuri River catchment, in the verification period, the RBFNN performs better than other single models and other combination methods in terms of *RMSE, CE* and *PBIAS*. Nevertheless, the MLPNN was the best in terms of *R^2, RMSE, CE* and *KGE* for the Mae Tuen River catchment.

In this study, GEP was used to perform symbolic regression functions and develop a multi-model combination system. The mathematical functions obtained from the GEP model are shown in equations 5.12 and 5.13 for the GEP-multi-model combinations of the Mae Tuen River and the Ohinemuri River catchments, respectively. The final GEP-combination model of Mae Tuen River catchment shows that the Q_{SWAT}, Q_{NAM} and Q_{LPM} were used as input to produce a mathematical function (see Eq. 5.12), while the Q_{LVGFM} and Q_{SMAR} have not been included. It is shown that the Q_{LVGFM} and Q_{SMAR} are not significant in the GEP-combination model.

In the case of the Mae Tuen River catchment, the Q_{NAM}, Q_{LVGFM}, Q_{SMAR}, Q_{LPM} were used as input to produce a mathematical function (see Eq. 5.13) except only Q_{SWAT} was not used in the GEP combination model. Clearly, the equation suggests that the Q_{LVGFM} is significant in the GEP-combination model, due to it being more frequently used than any other input models. However, the results in equations 5.12 and 5.13 show that the GEP combination method has the advantage over other neural network combination methods, as the combination function can be expressed as a simple mathematical function.

5.6.3 Comparison of model performance

Comparison of the performance of five individual rainfall-runoff models (i.e. LPM, LVGFM, SMAR model, NAM model and SWAT model) with the other three combination methods (i.e. GEP, MLPNN and RBFNN) is based on a statistical analysis.

Table 5.2 indicates that the *KGE* provides an overall measure of the performance of five rainfall-runoff models and three combination methods, outperforming the traditional CE at the all-time series of both catchments. Figures 5.6 and 5.7 show the observed and estimated peak hydrographs from some selected peaks in both the Mae Tuen River catchment and the Ohinemuri River catchment. The results in Figures 5.6 and 5.7 show that the combination can reproduce the flood peak hydrographs better than those of the individual rainfall-runoff models. Figures 5.8 and 5.9 show that all combined models performances during low and medium flows are reliable and during high flow are negligible. The R^2 values in Figure 5.8

show that in the RBFNN multi-model there is a stronger linear relationship between the combined models and observed flow of the calibration period than in other combined models. Figure 5.9 demonstrates that the GEP multi-model has higher correlation coefficients of the calibration period than do the other combined models. However, the RBFNN multi-model exhibits a stronger linear relationship between the combined model and observed flow than in the other combined models of the verification period. The results in Figures 5.8 and 5.9 (show OR indicate OR demonstrate) that the combined model values have scatter points around the best fit line in calibration and verification periods. Moriasi et al. (2007) suggest that any R^2 greater than 0.5 for daily flow comparison is an acceptable threshold for hydrology calibration. The R^2 values of the MLPNN, RBFNN and GEP combination methods are greater than 0.50, which were considered satisfactory. Thus, these combination methods are capable of the multi-model approach for river flow simulations.

Table 5.2: Calibration and verification results from five models and combined models for Mae Tuen River catchment, Thailand and Ohinemuri River catchment, New Zealand

Model	Mae Tuen River, Chiang Mai,TH (Area=501.79km^2)										Ohinemuri River, NZ (Area=285.39km^2)									
	R^2	Rank	RMSE	Rank	CE	Rank	KGE	Rank	PBIAS %	Rank	R^2	Rank	RMSE	Rank	CE	Rank	KGE	Rank	PBIAS %	Rank
Calibration																				
LPM	0.55	4	6.93	4	0.55	4	0.64	4	**-0.05**	**1**	0.71	7	6.59	7	0.71	7	0.78	3	0.56	2
LVGFM	0.52	5	7.13	5	0.52	5	0.62	5	2.56	6	0.81	4	5.39	4	0.81	4	0.86	2	1.89	5
SMAR model	0.50	7	7.29	7	0.50	7	0.61	7	-1.04	2	0.75	6	6.41	6	0.73	5	0.77	4	16.53	6
NAM	0.51	6	7.23	6	0.51	6	0.61	7	3.78	7	0.76	5	6.33	5	0.74	6	0.77	4	17.47	7
SWAT model	0.41	8	8.72	8	0.29	8	0.62	5	12.79	8	0.68	8	9.26	8	0.43	8	0.30	5	48.33	8
Combination models																				
MLPNN	0.62	2	6.41	2	0.62	2	0.70	2	2.28	4	0.84	2	4.96	2	0.84	2	0.86	2	-1.08	3
RBFNN	**0.64**	**1**	**6.21**	**1**	**0.64**	**1**	**0.72**	**1**	2.35	5	0.83	3	5.14	3	0.83	3	**0.88**	**1**	**-0.50**	**1**
GEP	0.58	3	6.69	3	0.58	3	0.67	3	1.20	3	**0.86**	**1**	**4.69**	**1**	**0.85**	**1**	**0.88**	**1**	-1.51	4
Verification																				
LPM	0.58	5	7.59	6	0.57	6	0.60	6	**-6.49**	**1**	0.64	4	6.14	3	0.61	3	**0.78**	**1**	5.39	2
LVGFM	0.47	7	8.43	7	0.47	7	0.57	7	-7.38	2	0.55	6	6.71	5	0.54	5	0.64	4	12.96	4
SMAR model	0.62	4	7.21	5	0.61	5	0.68	4	-14.55	7	0.57	7	6.80	6	0.53	6	0.58	6	25.64	6
NAM	0.66	3	6.88	3	0.65	3	0.68	4	-17.72	8	0.65	3	6.21	4	0.61	4	0.60	4	23.04	5
SWAT model	0.50	6	9.10	8	0.38	8	0.68	3	-11.19	6	**0.67**	**1**	7.02	7	0.49	7	0.38	7	39.45	7
Combination models																				
MLPNN	**0.73**	**1**	**6.11**	**1**	**0.72**	**1**	**0.74**	**1**	-10.14	5	0.63	5	6.08	2	0.62	2	0.66	3	7.36	3
RBFNN	0.67	2	6.65	2	0.67	2	0.72	2	-8.09	4	0.66	2	**5.77**	**1**	**0.66**	**1**	0.70	2	**2.83**	**1**
GEP	0.62	4	7.13	4	0.62	4	0.67	5	-8.02	3	0.65	3	6.21	4	0.61	4	0.60	5	23.04	5

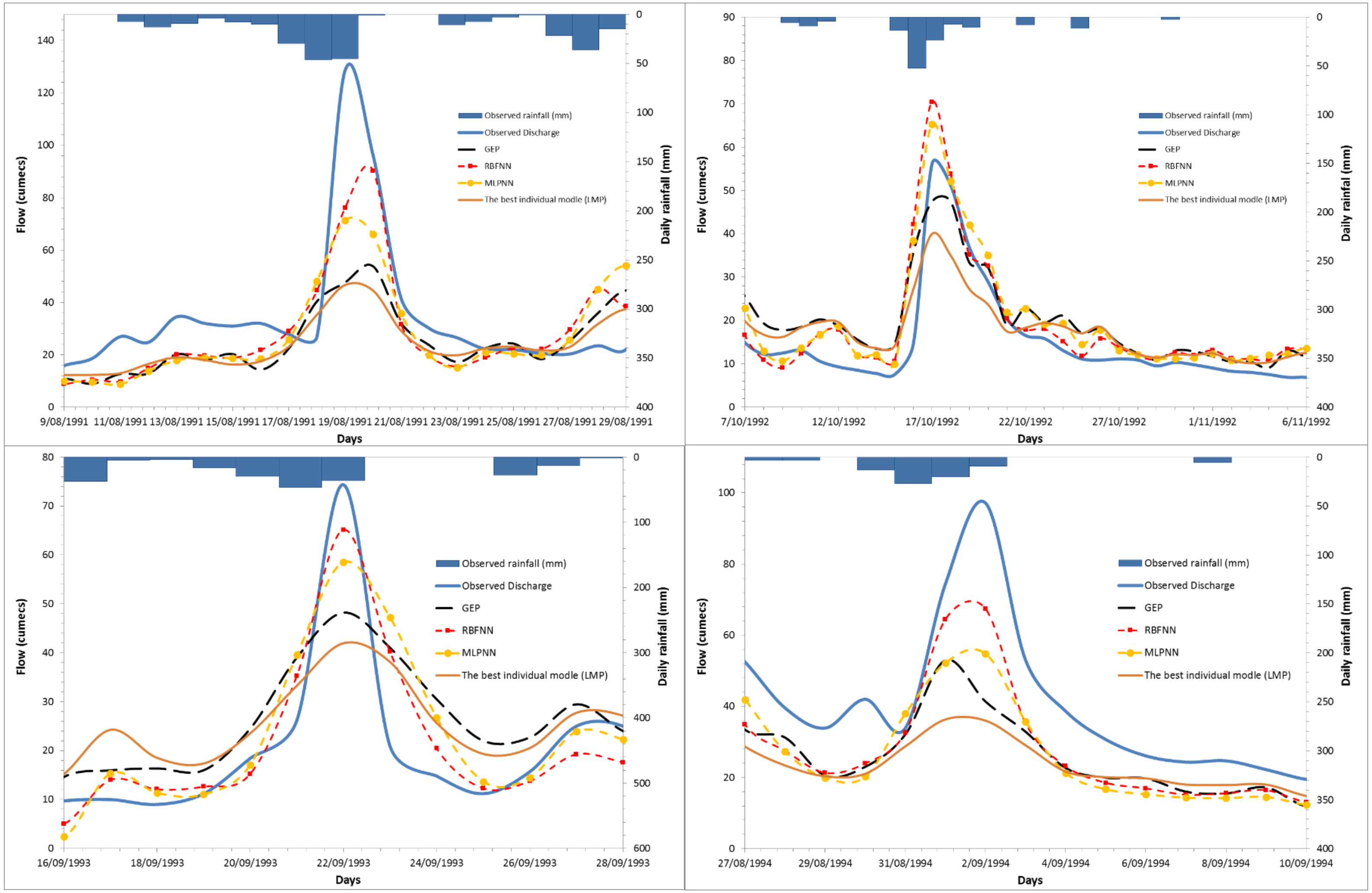


Figure 5.6: Comparison of the observed and simulated flood hydrographs of combined model of Mae Tuen River catchment

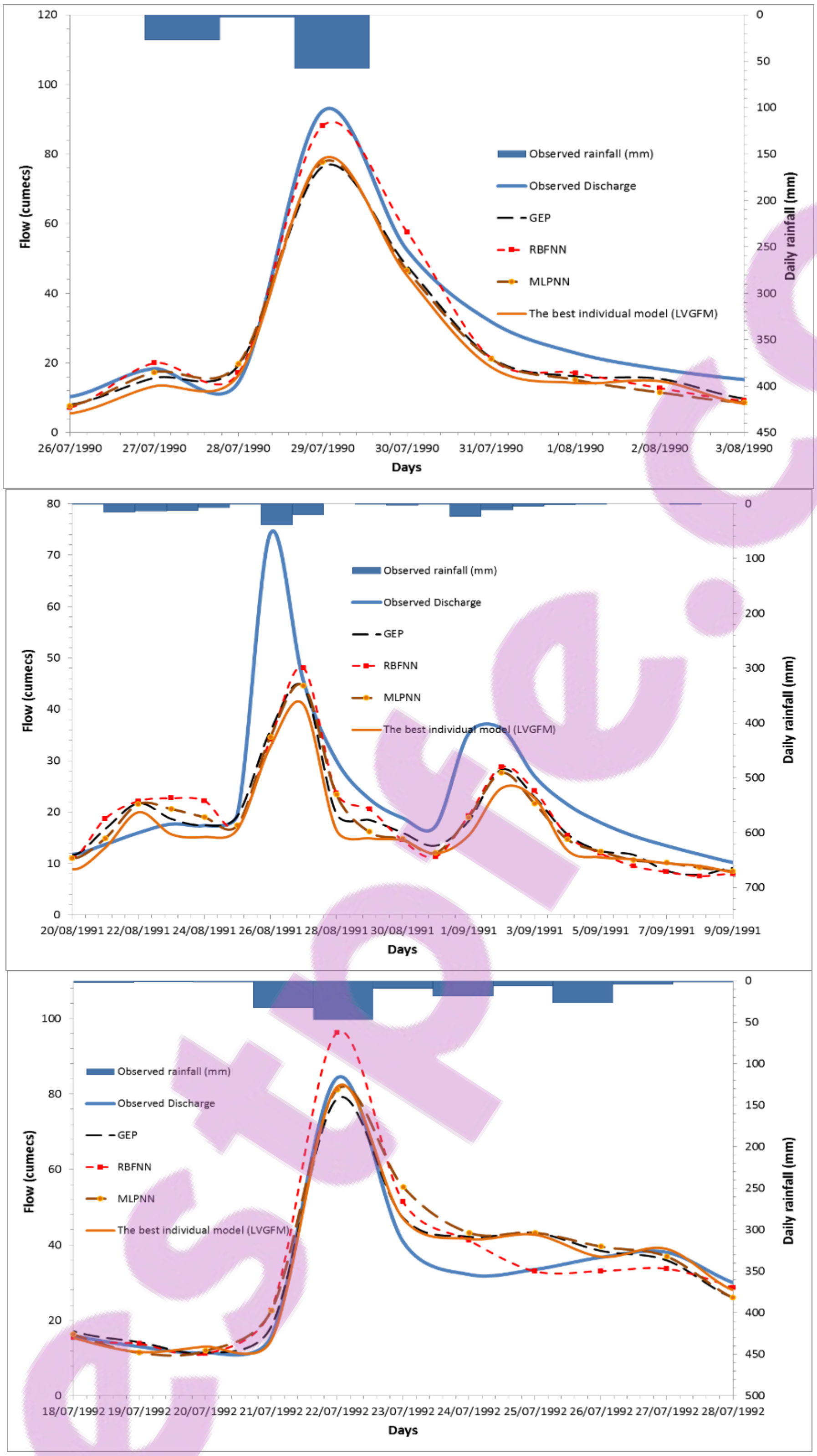


Figure 5.7: Comparison of the observed and simulated flood hydrographs of the combined model of Ohinemuri River catchment

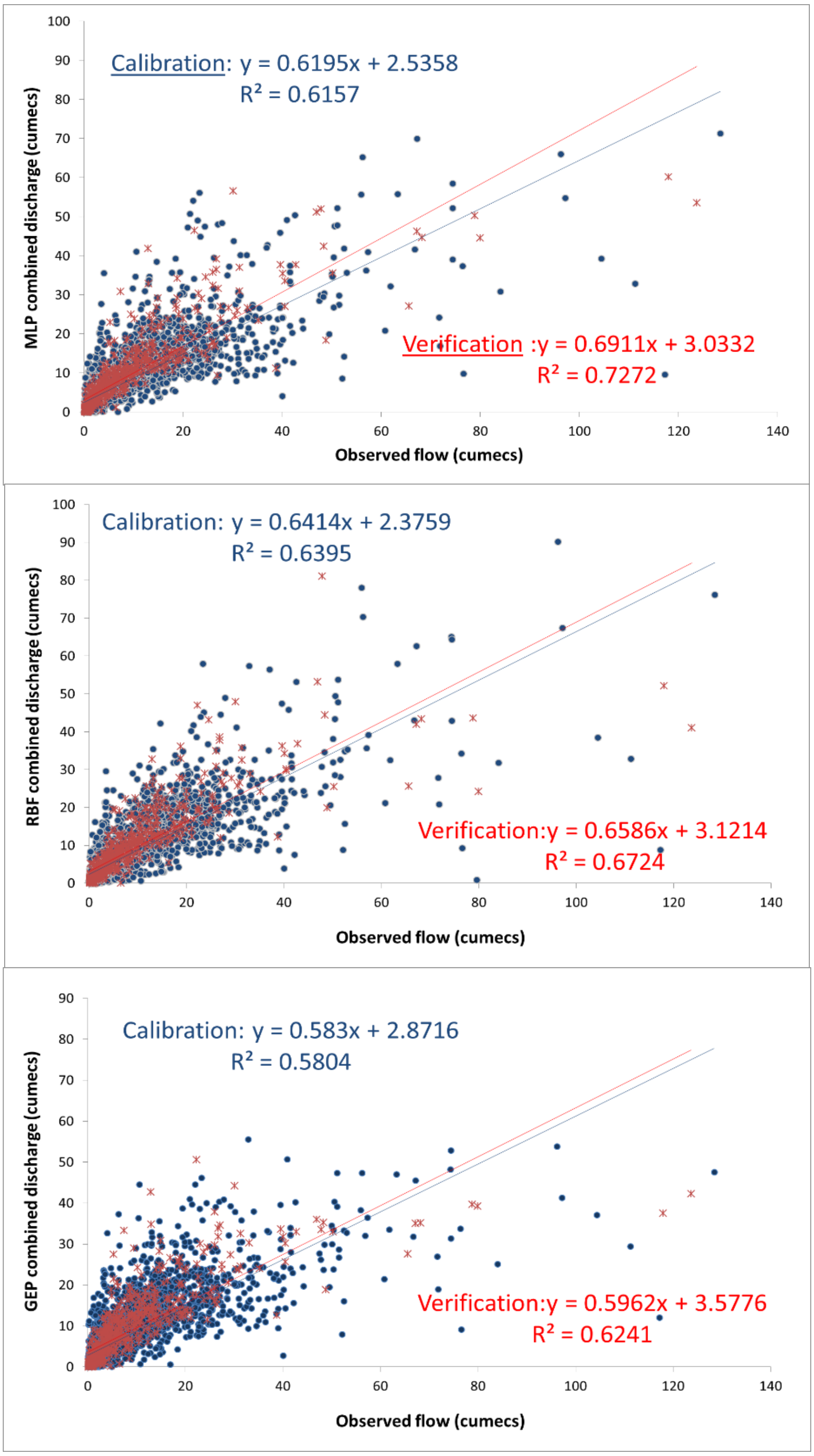


Figure 5.8: Scatter plots of observed discharge versus three combined models (i.e. RBF, MLP and GEP) at Mae Tuen River catchment during the time series, 4/01/1991 to 31/12/2002

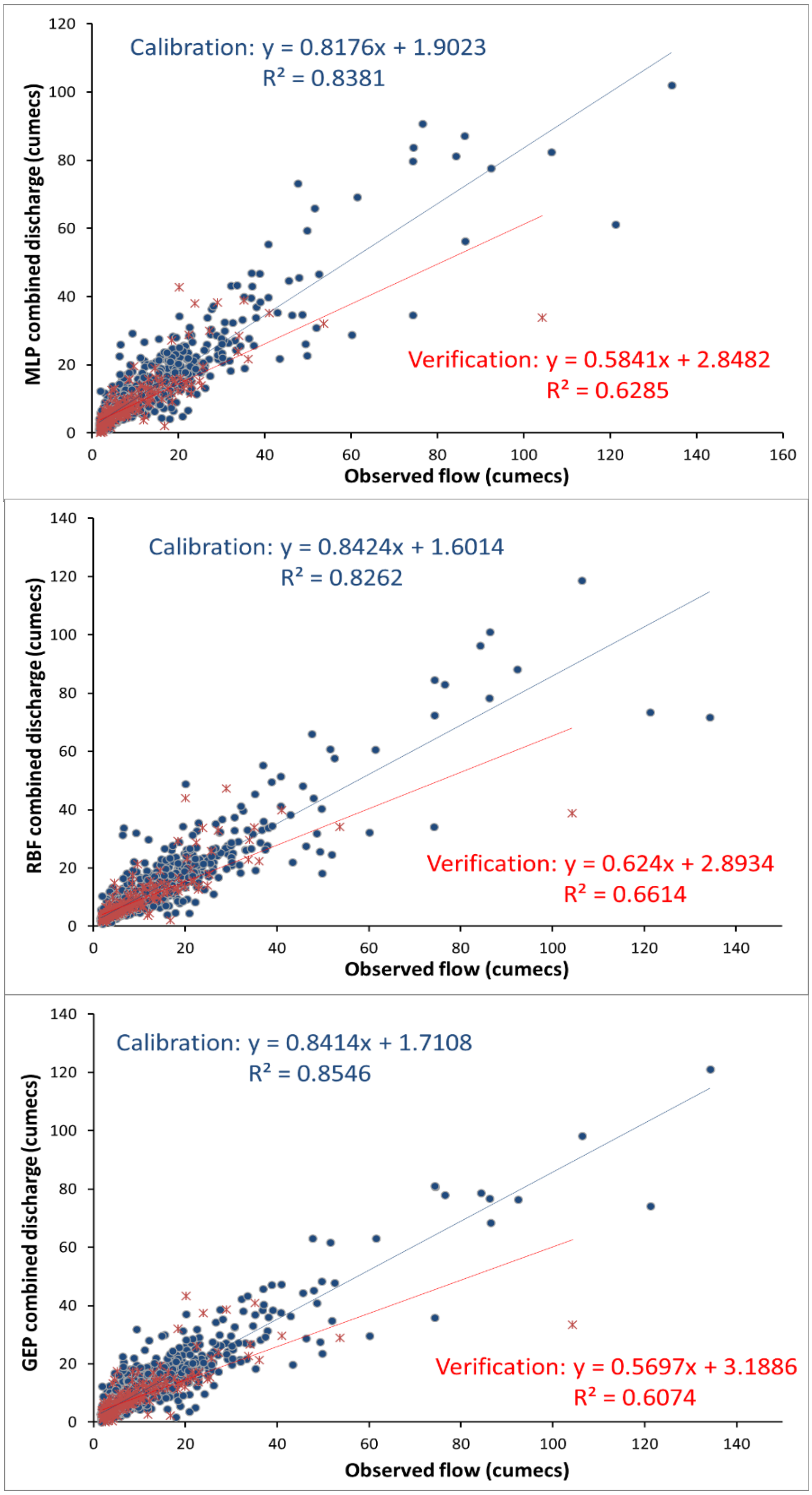


Figure 5.9: Scatter plots of observed discharge versus three combined models (i.e. RBF, MLP and GEP) at Ohinemuri River catchment during the time series, 1/01/1990 to 31/08/1993

5.8 Summary

The study presented in this chapter addressed the first objective of the thesis. This study aims to investigate whether or not the use of GEP will lead to further improvement in multi-model combination systems as well as or better than the two previously investigated -MLPNNN and RBFNN combination methods by Shamseldin et al. (2007). The chapter presents the comparative performance of three combination methods - the GEP, MLPNN and RBFNN in the multi-model combination systems, using the daily time series of two different catchments in Thailand and New Zealand. The combination techniques involve the use of five rainfall-runoff models, specifically, two black-box models (LPM and LVGFM), two conceptual models (SMAR model and NAM) and a semi-distributed conceptual model (the SWAT model) to produce the multi-model combinations for testing both catchments. Results show that the combined model is superior to all five individual models in calibration period and verification period. Overall, results of the developed multi-model combinations indicated that the use of GEP as a combination technique shows better performance in the multi-model combination systems than the use of MLPNN and RBFNN for the catchment in New Zealand. However, the use of RBFNN outperforms both combination methods (i.e. MLPNN and GEP) for the catchment in Thailand.

Chapter 6

The optimal number of rainfall-runoff models used in ANN combinations

The study in this chapter was motivated by a desire to investigate the optimal number of models to be used in the multi-model combination systems when applied to the case studies of two contrasting catchments located in Thailand and New Zealand, respectively (See Section 2.6 in Chapter 2), the initial part of this chapter highlights these basic requirements. Following that, the knowledge extraction techniques from the trained ANN multi-models is presented. It also addresses two techniques: (1) the Garson's algorithm method and (2) the connection weight approach. Then, it presents the application of the optimal number of rainfall-runoff models in the developed ANN multi-model combination systems and presents the detail of statistical methods used for assessment of the model performance. Finally, the summary obtained from results is discussed.

Note: Major research reported in this chapter is based on work submitted in the Hydrological Sciences Journal- "*The optimal number of rainfall-runoff models used in the neural network combination system.*"

6.1 Introduction

Phukoetphim et al. (2013) found that the knowledge extraction techniques had considerable potential for optimizing combined rainfall-runoff models in the multi-model combination system. The proposed study aims to extend the work of Phukoetphim et al. (2013) by exploring whether or not the knowledge extraction techniques can be used to determine the number of optimal rainfall-runoff models in the ANN combination system. ANN is a black-box, so it is not possible to monitor the input and output relationship and the interpretation of its weights, as knowledge may not be extracted from the ANN (Phukoetphim et al., 2013). The knowledge extraction technique can help in understanding the overall influence and contribution of connection weights and input variables in the ANN model (Olden et al., 2002). This chapter investigates the optimal number of models, which perform best in the ANN combination system. The research attempts to explore the interpretation of connection weights and the input parameters contribution in the trained ANN. To my knowledge based on the literature of combining simulated river flows, this technique has never been addressed in the ANN combination.

In this study, the MLPNN was used to combine the results obtained from the selected five rainfall-runoff models for testing the multi-model combination systems in Thailand and New Zealand catchments with contrasting properties. The five selected rainfall-runoff models are two empirical black-box models, two conceptual models and a semi-distributed physically based model. The two empirical black-box models selected for this study are the linear perturbation model (LPM) (Kachroo et al., 1988) and the linear varying gain factor model (LVGFM) (Ahsan and O'Connor, 1994). The two conceptual models selected in this study are the soil moisture accounting and routing (SMAR) model (Mandeville et al., 1970) and the Nedbør-Afstrømnings Model (NAM) (DHI, 2007). The semi-distributed physically based model selected is a soil and water assessment tool (SWAT) (Arnold et al., 1998).

Details regarding the study areas and data are presented in Chapter 2, the selected five rainfall-runoff models and evaluation of model performance are presented in Chapter 3 and the MLPNN combination method is already described in Chapter 4.

6.2 Knowledge extraction from artificial neural network

The major drawback of ANNs is that little is known about what is happening inside the ANN system. It is difficult to control the ANN parameters within the system because ANN works like a black-box, all dependencies (between parameters and responses) are hidden within the neural network structures. Thus, it is not possible to monitor these parameters, as knowledge may not be extracted from the neural network. The objective of ANN calibration or training is to find an optimal estimate of connection weights that best fit between the observed data and the simulation output (Kingston et al., 2006).

The optimal estimate of connection weights depends upon the input variables used in model training. Therefore, researchers have been trying to extract knowledge from ANN. Gallant (1988) was the first to use knowledge extraction from ANN based on the connection weight method in an ecological study. There are many different tools and techniques for extracting knowledge from a trained neural network model used in other fields (e.g. Andrews et al., 1995; Gallant, 1988; Özesmi and Özesmi, 1999; Olden and Jackson, 2002a; Olden et al., 2004; Weckman et al., 2009; Chapman and Purse, 2011) but there are limited studies in the field of hydrology (e.g. Wilby et al., 2003; Sudheer, 2005; Kalteh, 2008; Jain and Kumar, 2009). Their results found that the knowledge extraction technique can help us understand the overall influence and contribution of connection weights and input variables in the ANN model. To date, only one application has applied these techniques in the context of combination river flow forecasting systems (Phukoetphim et al., 2013).

Garson (1991) proposed a Garson's algorithm of partitioning the neural network weights to determine the relative importance of each input variable in the network. This method is used to eliminate connections whose weights do not significantly influence the network output, thereby illuminating the significant interactions being modeled. Later, Garson's algorithm became popular and was used in ecology for rule extraction for trained ANNs. Olden et al. (2004) applied nine different methodologies for assessing variable contributions in ANN by using simulated data with known correlative properties. Their results indicated that the connection weight approach provides the best overall method for accurately quantifying variable importance in ANN. To estimate monthly runoff, Kalteh (2008) applied a Neural Interpretation Diagram (NID) and Garson's algorithm to obtain a quantitative estimate of the contribution of each input variable on the output of a trained MLP. These approaches were used to assess the importance of input variables on the output. The results showed that the input variable can be considered to be significant with a 95% confidence level. Many studies have described different methodologies for assessing the importance of input variables in neural networks. The sensitivity analysis helps to evaluate the significance of the selection of input variables to the network prediction (Lek et al., 1996). Sudheer (2005) applied the sensitivity analysis to extract the knowledge embedded in trained ANN river flow models in order to identify the strength of the relationship between individual input variables and the outputs. The results suggest that the input variables to the network can be further pruned to improve outputs by using this analysis.

In order to understand the simulation and model process, it is imperative to analyse these weights and extract information regarding the contribution of input variables on the final network output. The interrelationship between the input variables significantly affects the contribution towards the final network output. Thus, this study focuses on exploring the interpretation of connection weight and input parameter contribution in the trained MLPNN by using two methods. These methods are (1) Garson's algorithm and (2) the connection weight method. The proposed study applied these techniques to provide an explanation for the significance of the relationship between the individual rainfall-runoff model and the MLPNN combination systems. It may help to optimally estimate the number of input variables (rainfall-runoff models) to be used in the MLPNN combination systems. These techniques are described in the coming sections.

6.2.1 Garson's algorithm

Garson (1991) introduced a method of partitioning the neural network weights in order to determine the relative importance of each input variable in the trained neural network. It uses the absolute value of connection weights to calculate the variable contribution with each input neuron, to the output neurons. The network diagram of connection weights used in the application of Garson's algorithm is shown in Figure 5.1 (see Chapter 5). The relative contribution (*RC*) of each input neuron to the output neuron via each hidden neuron is calculated as the product of the input-hidden connection weight and the hidden-output connection weight as follows by:

$$RC_{ji} = (w_{i,j})x(w_{j,m}) \tag{6.1}$$

The relative importance (RI) of each input variable (*i*) is then calculated by the following equation:

$$RI_i = \sum_{j=1}^{n} \frac{|RC_{ij}|}{\sum_{i=1}^{N} |RC_{ij}|} \tag{6.2}$$

6.2.2 Connection weight approach

This method calculates the product of the raw input-hidden and hidden-output connection weights between each input neuron and output neuron and sums the products across all the hidden neurons (Olden and Jackson, 2002). The information contained in each network weight can be used to analyse the importance of network inputs. In this study, this method is selected to determine the contributing importance of each simulated rainfall-runoff model. The relative contribution (*RC*) of each input neuron to the output neuron via each hidden neuron (see Fig. 5.1) can then be calculated in a similar way in equation (6.1). Thus, the relative importance (*RI*) of each input variable (*i*) in predicting the output is as follows:

$$RI_i = \sum_{j=1}^{n}(RC_{ij})$$
(6.3)

Garson's algorithm and connection weight approach methods calculate the product of the raw input-hidden and hidden-output connection weights. The weights and bias are obtained as per the trained MLPNN model are shown in Tables 6.1 and 6.2.

Table 6.1: Connection weights for combined discharge output of five rainfall-runoff models for the case studies of Mae Tuen River catchment, Thailand and Ohinemuri River catchment, New Zealand

	Mae Tuen River cathment		Ohinemuri River catchment	
Name	Hidden Neuron 1, H_1	Hidden Neuron 2, H_2	Hidden Neuron 1, H_1	Hidden Neuron 2, H_2
LPM	$W_{1,1}$=-0.031	$W_{2,1}$=-0.009	$W_{1,1}$=-0.003	$W_{2,1}$=0.010
LVGFM	$W_{1,2}$=-0.016	$W_{2,2}$=0.015	$W_{1,2}$=0.029	$W_{2,2}$=0.057
SMAR model	$W_{1,3}$=-0.007	$W_{2,3}$=-0.010	$W_{1,3}$=-0.032	$W_{2,3}$=-0.018
NAM model	$W_{1,4}$=-0.026	$W_{2,4}$=0.003	$W_{1,4}$=-0.010	$W_{2,4}$=-0.006
SWAT model	$W_{1,5}$=0.018	$W_{2,5}$=0.008	$W_{1,5}$=-0.016	$W_{2,5}$=-0.015
Combined discharge	$W_{m,1}$=-0.866	$W_{m,2}$=-0.288	$W_{m,1}$=-0.761	$W_{m,2}$=0.799

Table 6.2: Bias for combined discharge output of five rainfall-runoff models for the case studies of Mae Tuen River catchment, Thailand and Ohinemuri River catchment, New Zealand

	Mae Tuen River catchment	Ohinemuri River catchment
Neuron	Bias (optimized)	
α1	0.657	-2.069
α2	-0.321	-0.537
β	0.970	0.234

6.3 The optimal number of rainfall-runoff models

The MLPNN combination method was applied to combine the results obtained from the five selected rainfall-runoff models for producing the multi-model combination systems for this study. The five selected models are two empirical black-box models, the LPM and the LVGFM; two conceptual models, the SMAR model and the NAM, and a semi-distributed physically based model, the SWAT model (see Chapter 4). These models are applied to the daily data to simulate runoffs from two different catchments in New Zealand and Thailand, respectively (see Chapter 3).

In this study, the MLPNN multi-model has been trained on the neural network topologies (see Chapter 5), after which, two methods were followed: (1) Garson's algorithm method and (2) the connection weight approach. Both were applied for extracting knowledge from the weight trained MLPNN combined models to determine the influence of the input variables on the network outputs. These methods were used to estimate the ranking importance of the input variables, which model was to be used in the MLPNN multi-model combination systems. However, the determination of the optimal number of model performances involves the use of different numbers of models in producing the combined forecasts. To design the optimal number of models to be used in the multi-model combination, the selected models used in the trained MLPNN multi-model are based on the ranking (deciding which is most important to network output), calculated by Garson's algorithm and the connection weight approach, respectively (see Table 6.3). To investigate this issue, the performance of differently developed MLPNN combined models is assessed using the commonest evaluation criteria used in hydrology: root mean squared error (*RMSE*), the coefficient of efficiency (*CE*) and the % of deviation from observed runoff, *PBIAS* (e.g. ASCE, 1993; Duan et al., 2003). A description of these methods used for evaluating model performance is already provided in Chapter 4. The scatter plot was used in assessing model performances in this study.

Table 6.3: The importance of each model for the case studies Mae Tuen River catchment and Ohinemuri River catchment

Model	Mae Tuen River catchment			
	Garson's Algorithm		Connection Weight Approach	
	Importance	Ranking	Importance	Ranking
LPM	0.26	**1**	-0.86	3
LVGFM	0.25	2	2.43	**1**
SMAR model	0.14	5	-1.30	5
NAM	0.17	4	0.87	2
SWAT model	0.18	3	0.86	4
Model	Ohinemuri River catchment			
	Garson's Algorithm		Connection Weight Approach	
	Importance	Ranking	Importance	Ranking
LPM	0.06	5	0.45	2
LVGFM	0.43	**1**	1.11	**1**
SMAR model	0.26	2	0.37	3
NAM	0.09	4	0.10	4
SWAT model	0.16	3	-0.02	5

6.4 Results and discussions

This study applied two methods: (1) Garson's algorithm method and (2) the connection weight approach, for extracting knowledge from the trained MLPNN combination models. These methods were applied to estimate the ranking importance of the input variables, estimating which model is to be used in the MLPNN combination system. Table 6.3 shows the importance of each input variable (each rainfall-runoff model) from the trained MLPNN calculated by the Garson's algorithm and the connection weight approach methods. Results in Table 6.3 show that these methods can be used to identify the best set of input-output parameters for the development of simulation procedures using the MLPNN method for rainfall-runoff combination systems.

For the Mae Tuen River catchment, the results calculated by Garson's algorithm method show that the LPM is the most important input model followed by the LVGM, the SMAT model, the NAM and the SMAR model, respectively. In the case of the Ohinemuri River catchment, the LVGFM is the most important model followed by the SMAR model, the SWAT model, the NAM and the LPM, respectively. For both catchments, the results calculated by the connection weight method indicate that the LPM was the most important model.

The performance of the different developed MLPNN combination systems was assessed using the coefficient of efficiency, *CE*, root mean square error, *RMSE*, and the percentage of deviation from observed runoff, *PBIAS* for both catchments. These statistical performance criteria were considered for the comparison of the models. Table 6.4 shows the summary statistics efficiency values and their ranking of the developed MLPNN multi-models for both catchments. These models were developed using different combinations of model inputs, which are based on the ranking (the most important input to network outputs) calculated by Garson's algorithm and the connection weight approach (see Table 6.3).

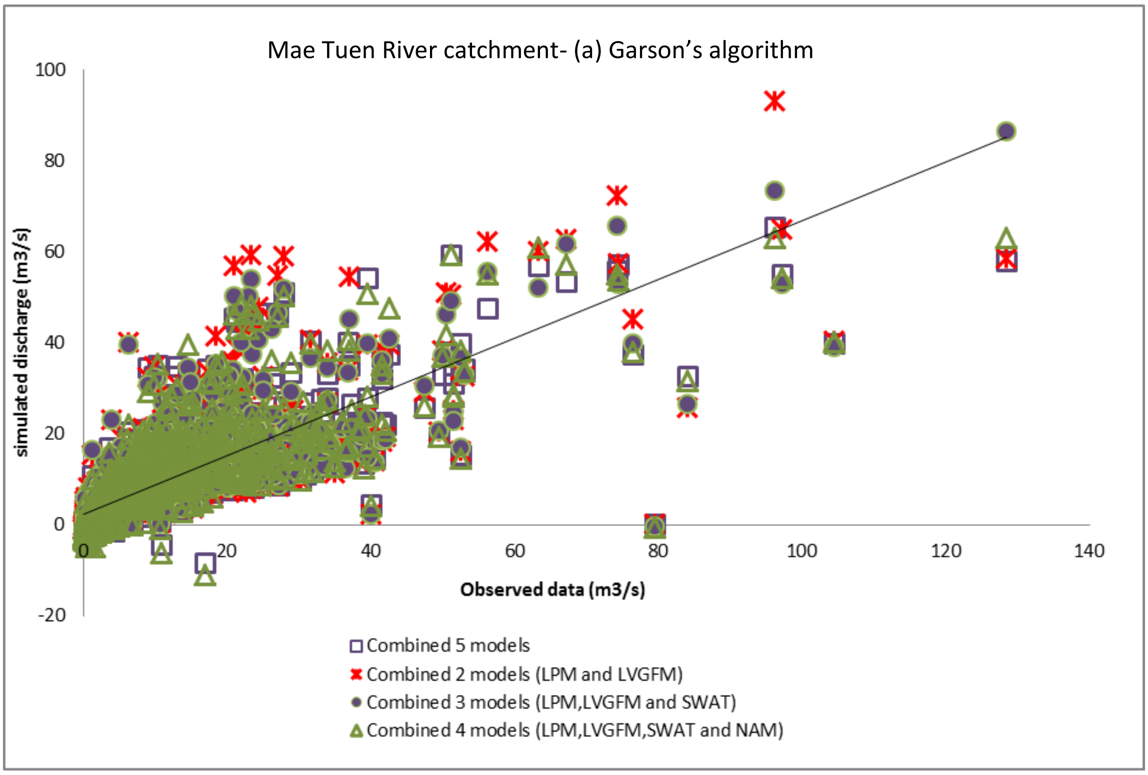
For the Mae Tuen River catchment, results in Table 6.4 show that in terms of *CE* and *RMSE* values, the combination of four models, namely, the LVGFM, the SMAR model, the SWAT model and the NAM) are the best performers in the MLPNN combination system. The results based on *PBIAS* values show that the combination of three models, namely, the LVGFM, the SMAR model and the SWAT model have the best performances in the MLPNN combination system for the calibration period. In the verification period, the combination of two models, namely, the LVGFM and the SMAR model constitute the best performance model in the MLPNN combination system. Overall results show that the combination of four models, namely, the LVGFM, the SMAR model, the SWAT model and the NAM are the optimal number of models calculated by Garson's algorithm in the calibration period. In the verification period, the combination of four models, namely, the LVGFM, the NAM, the LPM and the SWAT model calculated by connection weight method are the optimal number of models.

For the Ohinemuri River catchment, the results in Table 6.4 show that, in terms of *CE* and *RMSE* values, the combination of three models, namely, the LVGFM, the SMAR model and the SWAT model, are the best performing of the MLPNN combination, in the calibration period. However, in terms of *PBIAS* values, the combination of three models, namely, the LVGFM, the SMAR model and the SWAT model, constitute the best performance in the MLPNN combination. In terms of *CE* and *RMSE* values, the combination of three models, namely, the LVGFM, the SMAR model and the SWAT model are the best models in the MLPNN combination for verification periods. In terms of PBIAS values, the combination of four models, namely, the LVGFM, the SMAR model, the SWAT model and the NAM are the best performing of the MLPNN combination. Overall, results show that the combination of three models, namely, the LVGFM, the SMAR model and the SWAT model constitutes the optimal number of models in the MLPNN combination as calculated by the connection weight approach. The combination of three models, namely, the LVGFM, the SMAR model and SWAT model is the optimal number of models in the MLPNN combination as calculated by Garson's algorithm method.

Figures 6.1 and 6.2 present the distribution of two variables between the observed discharge values and the developed MLPNN multi-model simulated runoff values for the Mae Tuen River catchment and the Ohinemuri River catchment, respectively. A scatter plot was used to check a number of aspects of the distribution of two variables, which these relationships would show in a straight line. The results in Figures 6.1 and 6.2 show that the distribution of the developed MLPNN multi-model simulated runoff values and the observed discharge values are strongly related.

Table 6.4: Performance of the developed MLPNN combination systems for the Mae Tuen River catchment and the Ohinemuri River catchment

(a)Mae Tuen River catchment	Calibration						Verification					
	CE	Rank	RMSE	Rank	PBIAS%	Rank	CE	Rank	RMSE	Rank	PBIAS%	Rank
(1) Garson's Algorithm Numbers of model used in MLPNN												
2 models (LPM and LVGFM)	0.589	4	6.624	4	0.977	2	0.609	4	7.230	5	**-6.100**	**1**
3 models (LPM,LVGFM and SWAT)	0.593	3	6.588	3	**0.481**	**1**	0.603	5	7.283	6	-6.217	2
4 models (LPM+LVGFM+SWAT+NAM)	**0.622**	**1**	**6.353**	**1**	1.743	3	**0.695**	**1**	**6.389**	**1**	-10.620	4
(2) Connection Weight Approach Numbers of models used in MLPNN												
2 models (LVGFM and NAM)	0.575	6	6.737	6	2.125	4	0.693	2	6.406	2	-12.341	6
3 models (LVGFM,NAM and LPM)	0.585	5	6.654	5	2.136	5	0.657	3	6.776	4	-10.146	4
4 models (LVGFM,NAM,LPM and SWAT)	**0.622**	**1**	**6.353**	**1**	1.743	3	**0.695**	**1**	**6.389**	**1**	-10.620	5
The selected 5 models	0.616	2	6.406	2	2.656	6	0.693	2	6.408	3	-9.203	3
(b)Ohinemuri catchment												
(1) Garson's Algorithm Numbers of model used in MLPNN												
2 models (LVGFM and SWAR)	0.483	5	8.846	5	-1.087	7	0.290	5	11.220	5	3.052	4
3 models (LVGFM,SMAR and SWAT)	0.482	6	8.853	6	**-0.567**	**1**	**0.242**	**1**	**11.009**	**1**	-0.266	2
4 models (LVGFM,SWAR,SWAT and NAM)	0.448	7	9.136	7	-0.862	2	0.257	2	11.079	2	**0.004**	**1**
(2) Connection Weight Approach Numbers of model use in MLPNN												
2 models (LVGFM and LPM)	0.505	2	8.650	2	-0.891	3	0.322	7	11.361	7	2.824	5
3 models (LVGFM,LPM and SWAR)	**0.525**	**1**	**8.475**	**1**	-0.971	4	0.275	3	11.156	3	3.855	6
4 models (LVGFM,LPM,SWAR and NAM)	0.500	4	8.694	4	-0.997	5	0.312	6	11.315	6	2.471	4
The selected 5 models	0.501	3	8.682	3	-1.053	6	0.280	4	11.179	4	2.140	3



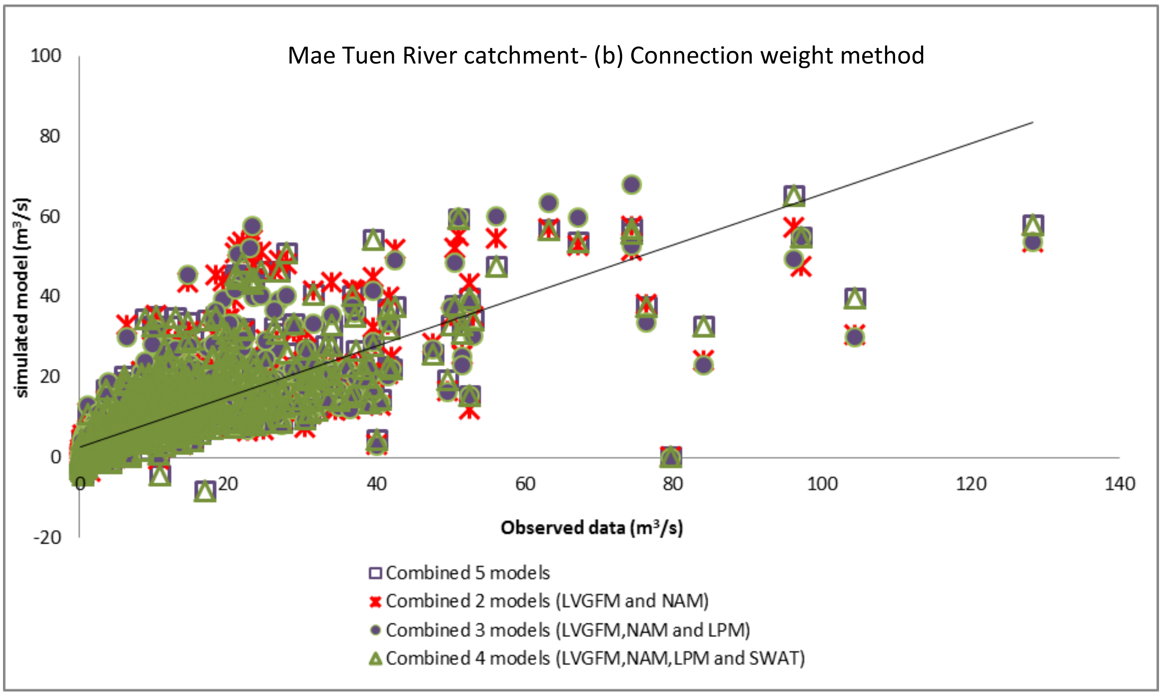
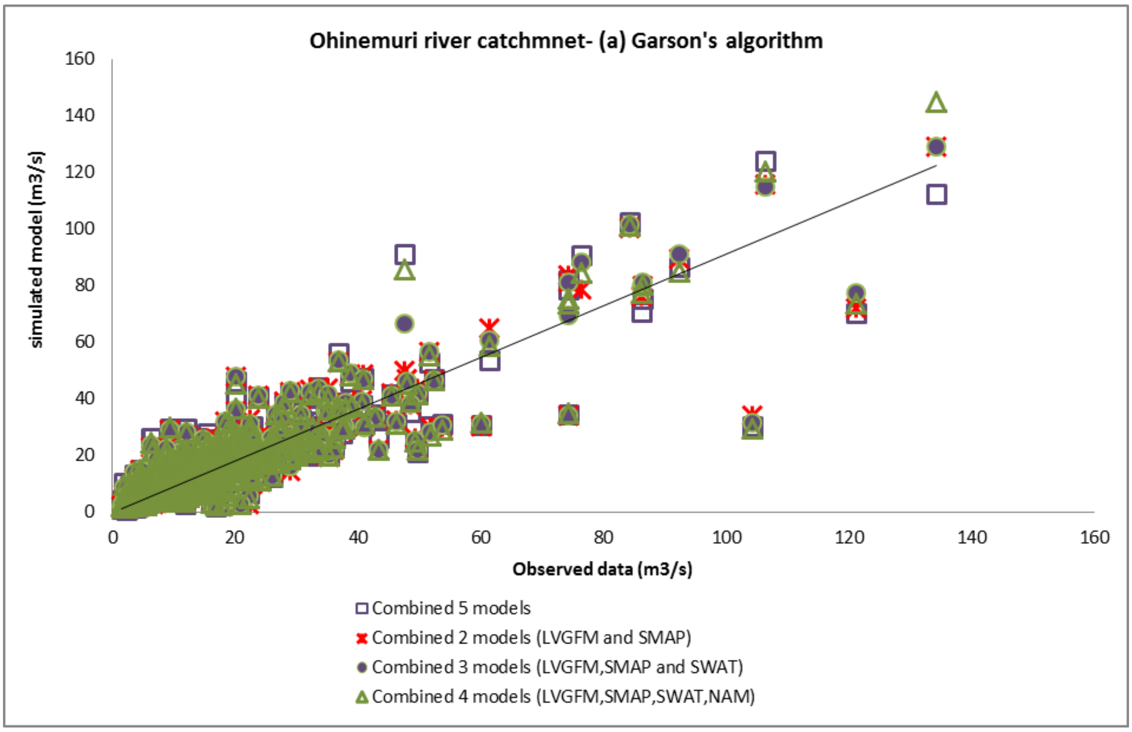


Figure 6.1: Scatter plot of simulated discharge between two variables for Mae Tuen River catchment: (a) Garson's algorithm and (b) Connection weight appraoch



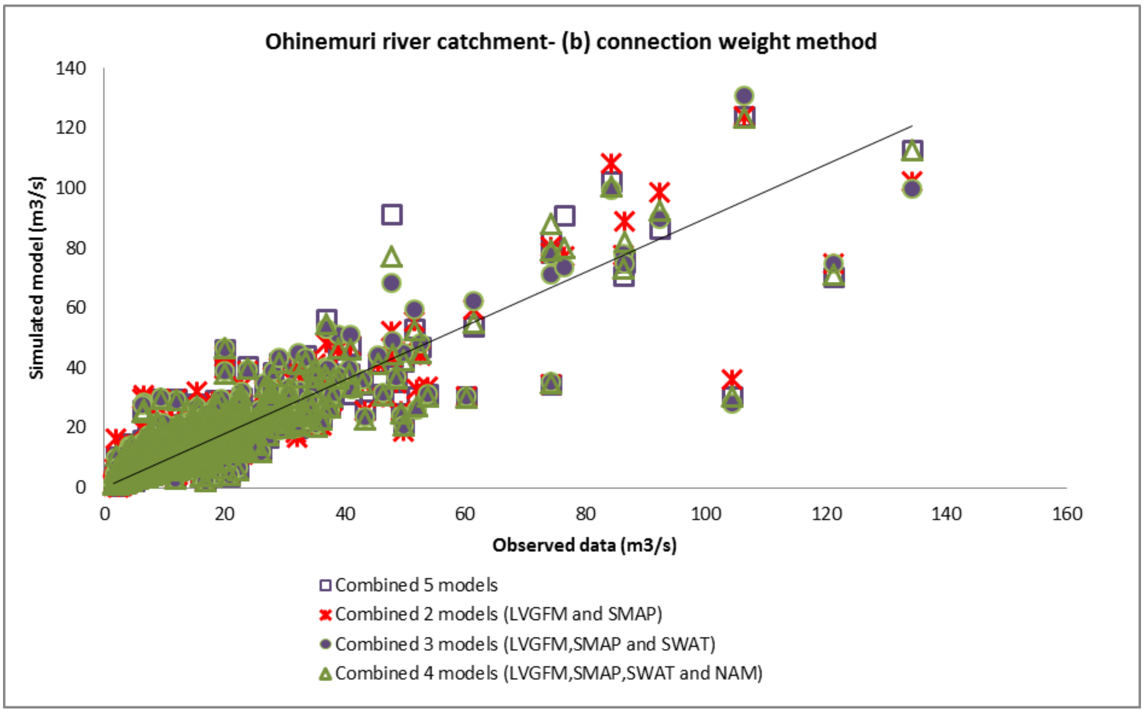


Figure 6.2: Scatter plot of simulated discharge between two variables for Ohinemuri River catchment: (a) Garson's algorithm and (b) Connection weight approach

6.5 Summary

The study presented in this chapter investigates the optimal numbers of models applied in the multi-model combination system rather than applying the new approach for multi-model combination. It addresses the second objective of the thesis, to design the optimal number of rainfall-runoff models to be used to increase performance in the ANN combination systems. This study explores the interpretation of connection weight and input parameter contribution in the trained MLPNN multi-model by using two approaches: Garson’s algorithm and the connection weight approach. These methods were applied for extracting knowledge from the MLPNN, which is used to combine the results from five different competing rainfall-runoff models for the case studies of two different catchments in Thailand and New Zealand. Overall, results demonstrated that these approaches can be used to determine the relative importance of each simulated rainfall-runoff model, assessing which perform best in the multi-model combination system.

Chapter 7

Uncertainty analysis in ANN multi-model combination systems

This chapter was motivated by a desire to address the research gaps in section 2.6. (see Chapter 2). It addresses the third objective of the thesis (see Chapter 1), to analyse the uncertainty and estimate the confidence interval of developed multi-model combination systems. The chapter begins by briefly describing the uncertainty analysis of multi-model combination systems. Following that, it presents a bootstrap method that is applied for quantifying the uncertainty associated with the multi-model combinations. Then, the details of model efficiency evaluation criteria used for assessment of model efficiency are presented. Finally, the summary obtained from the results is discussed.

From the literature (see Chapter 2), the results showed that the ANN combination methods produced better performance than other combination methods (i.e., SAM, WAM, fuzzy based method). This study proposes to use the ANN model to quantify uncertainty and estimate the confidence interval of the developed multi-model combination system. To the author's best knowledge based on the literature (see Chapter 2), previous applications of the ANN combination method have not provided any measure of forecast uncertainties in the developed multi-model combinations.

Note: Major research reported in this chapter is based on work submitted in the Journal of Hydrologic Engineering- *"Uncertainty analysis in neural network multi-models of combining simulated river flows using bootstrap technique".*

Thus, this study is the first to address a preliminary investigation of the uncertainty analysis in the ANN multi-model combination.

7.1 Uncertainty analysis

The reliability of the model predictions is still not satisfactory since the model prediction contains an element of the uncertainty (Beven and Binley, 1992). The uncertainty analysis can help provide insights into the level of confidence and reliability on the model outputs used in practice. However, one of the most important problems in using the ANN model is uncertainties in its predictions (Aires, 2004). The uncertainties in the results of the ANN model are of two types associated with the predictions. The first is uncertainty from the model structure and the second is the inherent uncertainty in training datasets (Heskes, 1997). The reliability of ANN model predictions can be enhanced through providing by quantifying uncertainty in model prediction (e.g., Aires 2004; Shrestha et al. 2009; Tiwari and Chatterjee 2010).

In this study, the uncertainty analysis assesses the effects of parameter uncertainties on the uncertainties of the ANN multi-model combination results. The model results generated from the simulation can be represented as probability distributions (or histograms), tolerance zones and confidence intervals. In practical regression problems, there are two types of prediction that need to be obtained in correspondence for a given input *X*. First is the mean $\mu_y(X)$ of a regression function, $f(x:\hat{\omega})$ (see Eq. 5.5) and the second is the target valued *Y* associated with *X* as given by equation 5.4. It is important to associate these estimates with their corresponding measure of confidence intervals of the model outputs. Thus, the study aims to assess the uncertainty and estimate the confidence interval of the developed multi-model simulations.

According to the literature (see Chapter 2), there are only a few studies in hydrologic modelling regarding the uncertainty analysis of the ANN hydrological models. In this study, the bootstrap method is chosen for the proposed study due to its benefits and capability of performing uncertainties analysis in ANN models according to the literature presented in Section 2.4 (see Chapter 2). In this study, the bootstrap method was applied to quantify the uncertainty of the developed ANN multi-model combinations through the case studies of two different catchments, namely (1) the Mae Tuen River catchment located in Thailand and (2) the Ohinemuri River catchment located in New Zealand. In this study, the MLPNN as a non-linear combination method was used to combine the results obtained from the selected five rainfall-runoff models to produce multi-model combination systems. The five selected models are two black-box models, (1) the linear perturbation model (LPM) and (2) the linearly varying gain factor model (LVGFM), and two conceptual models, (1) the soil moisture accounting and routing (SMAR) model and (2) the Nedbør-Afstrømnings model (NAM). In addition a distributed physically-based model is included - the soil and water assessment tool (SWAT) model. These models were run at a daily time step using the data inputs (i.e. rainfall, runoff, temperature (maximum and minimum), evaporation) to simulate river flows at the outlet of two river flow gauging stations; namely gauge P64 (see Fig. 3.1) and gauge Ohinemuri River (see Fig. 3.4) , respectively. For model evaluation, the input data was split into two parts (see Table 3.1). The first part is 2/3rds of the available data, which was used as the calibration period. The remaining 1/3rd was used as the verification period. All models were calibrated to each catchment to determine the optimum set of the parameter values. The application of these five rainfall-runoff models is discussed in Chapter 5.

The descriptions and details of the study areas and data, the MLPNN combination method and the selected five rainfall-runoff models, are already presented in Chapter 2, Chapter 3 and Chapter 4, respectively in this thesis.

7.1.1 Bootstrapped MLPNN combination

The bootstrap method, also called the sampling method was introduced by Efron (1979). Efron (1979) recognized that a bootstrap method is a more extensive use of nonparametric statistics for estimating statistical error. It is used to estimate the uncertainties of ANNs with different model structures and to construct the confidence intervals for model predictions (Efron and Tibshirani, 1993). In this study, the uncertainty analysis is considered a function of the system parameters. The bootstrapped MLPNN combination (BMLPC) was used to quantify the uncertainty (variance) in the parameters (weights between neurons) of the trained MLPNN multi-models in this study. Then, the statistical properties of the parameter variation, which is a measure of uncertainty, were addressed for estimating the standard error of the estimated MLPNN combination simulations, and their corresponding confidence intervals.

The following is a detailed description to estimate the variance denoted as σ_B^2 in the weights and the output of the network over the whole trained MLPNN multi-model using bootstrap technique. In the BMLPC data sets are defined as, $S_i = \{(X_i, Y_i) | i = 1,...,N\}$, where (X_i, Y_i) represents the *ith* pair of *X*, input (the individual rainfall-runoff model simulated runoffs) and *Y*, output (the combined river flows) samples for training the network (i.e., determining the weights). A number of bootstrap samples (*B*) are drawn with replacements from a pair of datasets randomly at a particular time in the bootstrap re-sampling. Each re-sampling delivers the data set which can be represented as $S^{*b} = \{(X_i, Y_i)^{*b} | i = 1,...,N\}$, $b = 1,2,3,...B$, due to the sampling with replacement, some of N(X_i, Y_i) - pairs from S may be sampled more than once, while others may not be selected at all. In BMLPC, from each re-sampled data set, S^{*b} is used as a data set for training a different MLPNN to give a regression function defined as $f(X; \hat{w}_b)$, where $\hat{w}_b$ is thereby obtained from network weight values. Then, in correspondence with a new input *X*, the BMLPC estimate of the *B* regression function $f_B(X; \hat{w}_b)$ is denoted as:

$$f_B(X : \hat{w}_b) = \sum_{b=1}^{B} \frac{f(X; \hat{w}_b)}{B}, b = 1,2,..., B \tag{7.1}$$

The BMLPC estimate of the bias is denoted as $\hat{\phi}_B = f_B(X;\hat{w}_b) - f(X;\hat{w})$ and the variance of the outputs of each BMLPC is the estimate of the model uncertainty variance, associated with the input X as given by:

$$\sigma_B^2(X) = \frac{1}{B-1}\sum_{b=1}^{B}[f(X;\hat{w}_b) - f_B(X;\hat{w})]^2 \tag{7.2}$$

The BMLPC estimate of standard error (SE_B) of $f(X:\hat{w}_b)$, which is a function of $f(X:\hat{W}_b)$, is given by (Efron and Tibshirani, 1993);

$$SE_B(f(X;\hat{w}_b) = \sqrt{\frac{1}{B}[\sum_{b=1}^{B} f(X;\hat{w}_b) - f_B(X;\hat{w})]^2} \tag{7.3}$$

The estimated uncertainty interval is expected to contain a specified portion of observation from the validation period dataset. Assuming that target values, *Y* (the combined river flows) follow a normal distribution, the confidence intervals are defined as follows:

$$f(X;\hat{w}) \pm t_{confidence} SE_B(f(X;\hat{w}_b) \tag{7.4}$$

7.2 Modelling performance evaluation

The coefficient of efficiency, CE (Nash and Sutcliffe, 1970), root mean square error (RMSE), and the percentage of deviation from observed runoff, PBIAS, were used to evaluate the accuracy of the BMLPC and MLPNN multi-model. The details and equations of CE, RMSE and PBIAS are already given in Section 4.2 (see Chapter 4).

7.3 Results and discussions

For the experimental design of uncertainty analysis in the MLPNN multi-model of both catchments, the use of the bootstrap technique was applied to build the BMLPC models. Each BMLPC model operation was based on a continuous process of data selection and parameter adjustment using the random samples with replacement from the available data sets (i.e. the individual rainfall-runoff model simulated runoffs). The data sets are divided into two data sets: the training data set and testing data set. These spilt sampling procedures used WMO, 1992 for reference. In this procedure, 2/3rds of the total available data sets were used for training the networks, and the remaining 1/3rd data sets were used for testing the networks. The remaining data set is used for testing the final result in order to confirm the actual predictive capability of the network.

To develop the MLPNN combination for both catchments (see Chapter 5), the numbers of neurons used in the hidden layer were varied systematically between 2 neurons and 20 neurons. All networks were trained in batch training using the scaled conjugate gradient method to find the optimal values of the weight and bias parameter values. To assess the optimal number of hidden neurons, MLP used a step-wise search with increasing numbers of hidden neurons. It evaluated each one using cross-validation in order to avoid overtraining. The final architecture of the MLPNN model developed is optimal, having 2 neurons in the hidden layer for both catchments (see Fig. 5.1).

The BMLPC models were developed using the same architecture of the MLPNN multi-model in order to maintain consistency for both catchments. The summary of statistics of performance measured by the BMLPC models and the MLPNN multi-models of both catchments is described in Table 7.1. Results in Table 7.1 show that the standard deviation, the skewness, the mean and median values are relatively close to each model value for both catchments, except only the standard deviation values vary for the Ohinemuri River catchment. The results demonstrate that the estimated relationships of each model result in similar network performance. The performance of BMLPC models and MLPNN multi-models has been evaluated using three statistical methods: the CE, the *RMSE* and the *PBIAS*. Table 7.2 shows the statistical results of the trained MLPNN multi-models and the BMLPC models for both catchments. For both catchments, the BMLPC models outperformed the MLPNN multi-models in terms of *RMSE, CE* and *PBIAS*, respectively. The results show that the BMLPC model has better trained results than the MLPNN multi-model, which can be attributed to the bootstrap technique.

To quantify the uncertainty in the trained MLPNN parameters (see Fig. 5.1), the summary of the connection weight values of BMLPC models for both catchments have been plotted as the box plots (see Figs. 7.1 and 7.2). Box plots display the range and distribution of the connection weight values of the BMLPC model along a number line. For the Mae Tuen River catchment, results in Figure 7.1 show that the connection weight values (the input-hidden weights) of wI_2,H_1 connected from the LVGFM to H_1, wI_3,H_1 connected from the SMAR model to H_1, wI_5,H_1 connected from the SWAT model to H_1, wI_2,H_2 connected from the LVGFM model to H_2, wI_3,H_2 connected from the SMAR model to H_2 and wI_5,H_2 connected from the SWAT model to $H_{2,}$ do not have much variation as their range is very small (see Fig. 5.1). On the other hand, the connection weight values of wI_1,H_1, wI_4,H_1, wI_1,H_2, and wI_4,H_2 have high variation values, which are the weight connections from the LPM to H_1 (wI_1,H_1), the NAM to H_1 (wI_4,H_1), the LPM to H_2 (wI_1,H_2) and the NAM to H_2 (wI_4,H_2), respectively.

The results in Figure 7.1 show that the connection weight values of wI_1,H_1 and wI_1,H_2 connected from the LPM and the connection weight values of wI_4,H_1 and wI_4,H_2 connected from the NAM, are more significant in producing the trained MLPNN multi-models, in which these inputs require larger weight values. In addition, results in Figure 7.1 show that the mean value of the hidden-output connection weights between the hidden node 1 (H_1) and the model outputs is relatively high when compared to the hidden node 2 (H_2). It is demonstrated that the behaviour of H_1 performs a major role in the trained MLPNN combination system.

For the Ohinemuri River catchment, the results in Figure 7.2 show that the connection weight values of wI_1,H_1, wI_3,H_1, wI_2,H_2 and wI_5,H_2 do not have much variation. These are the weight connections from the LPM to H_1 (wI_1,H_1), the SMAR model to H_1 (wI_3,H_1), the LVGFM to H_2 (wI_2,H_2) and the SWAT model to H_2 (wI_5,H_2), respectively. The connection weight values of wI_2,H_1 connected from the LVGFM, wI_4,H_1 connected from the NAM model and wI_5,H_1 connected from the SWAT model are more significant in producing the trained MLPNN multi-models. The mean value of the hidden-output connection weights between the H_1 and the model outputs is relatively high when compared to weights from the H_2. The results also show that the behaviour of the H_1 performs a more major role in the model outputs than the H_2.

Analysis of the results of both catchments shows that the bootstrap method can help in understanding the behaviour of the system parameters (i.e. the connection weights) and the input variables (i.e. the individual rainfall-runoff model simulated runoffs) in the multi-model combination system. It represents the significant input variables to the model outputs and it can also be used to reduce the complexity of rainfall-runoff models combination systems by showing which individual models should be included in the MLPNN multi-model combination system.

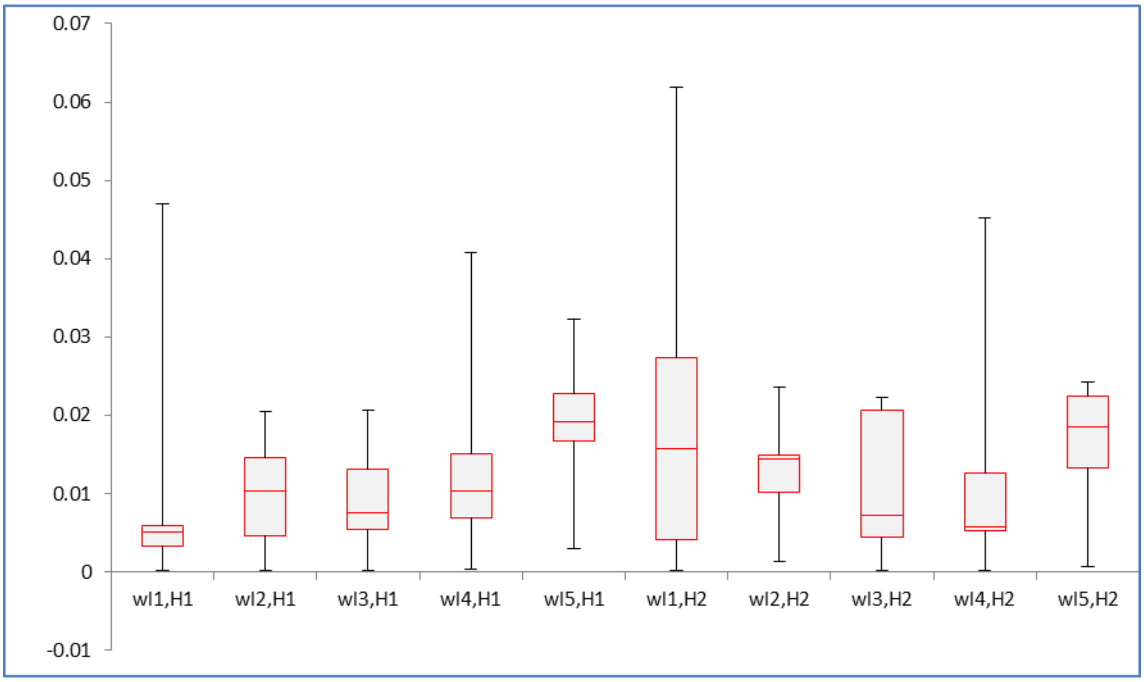
As mentioned earlier, the bootstrap method is a very powerful tool for variance estimation based on the actual data. The study used bootstrap confidence intervals to infer the trained MLPNN combination system significance level of the effects. The estimate of the confidence interval provides some idea of how uncertain we are about the estimated or unknown parameter. The most widely used confidence intervals are the 95% and 99% confidence intervals, which have 0.95 and 0.99 probabilities of containing the parameter respectively. The 95% confidence intervals of the trained MLPNN combination system were evaluated in this study, as shown in Figures 7.3 and 7.4 for the Mae Tuen River catchment and the Ohinemuri River catchment, respectively. To access the uncertainty in the trained MLPNN combination system, Figures 7.3 and 7.4 provide the 95 percent confidence intervals of the trained MLPNN combination system for the Mae Tuen River catchment and the Ohinemuri River catchment, respectively. Figure 7.3 confirms that the MLPNN multi-model is able to compute any of the hydrograph peak flows except for a few. Figure 7.4 shows that some simulated discharge outputs fail to capture the hydrograph peak flow. It is worth mentioning that the actual value of the hydrograph peak flow does not fall in the prediction confidence interval for the MLPNN combination simulations.

Table 7.1: Summary of statistic of performance measures of BMLPC models and MLPNN multi-models for Mae Tuen River catchment and Ohinemuri River catchment.

	Standard Deviation	Skewness	Mean	Median
(a) Mae Tuen River catchment				
BMLPC	8.222	2.543	6.548	3.992
MLPNN multi-model	8.234	2.226	7.100	4.584
(b) Ohinemuri River catchment				
BMLPC	12.612	4.386	10.074	6.089
MLPNN multi-model	11.016	4.324	9.587	6.053

Table 7.2: The evaluation performance results for the BMLPC models and the trained MLPNN multi-models for the Mae Tuen River catchment and the Ohinemuri River catchment.

Catchment	Model	RMSE	CE	PBIAS%
Mae Tuen River,	BMLPC	**5.932**	**0.663**	**0.071**
Thailand	MLPNN multi-model	6.406	0.631	0.569
Ohinemuri River,	BMLPC	**3.934**	**0.911**	**-0.163**
New Zealand	MLPNN multi-model	5.087	0.818	-0.186


0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
wI1,H1
wI2,H1
wI3,H1
wI4,H1
wI5,H1
wI1,H2
wI2,H2
wI3,H2
wI4,H2
wI5,H2

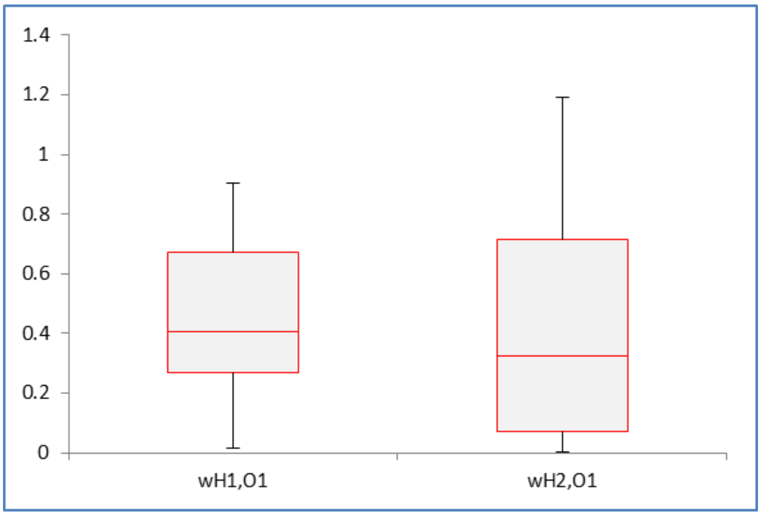
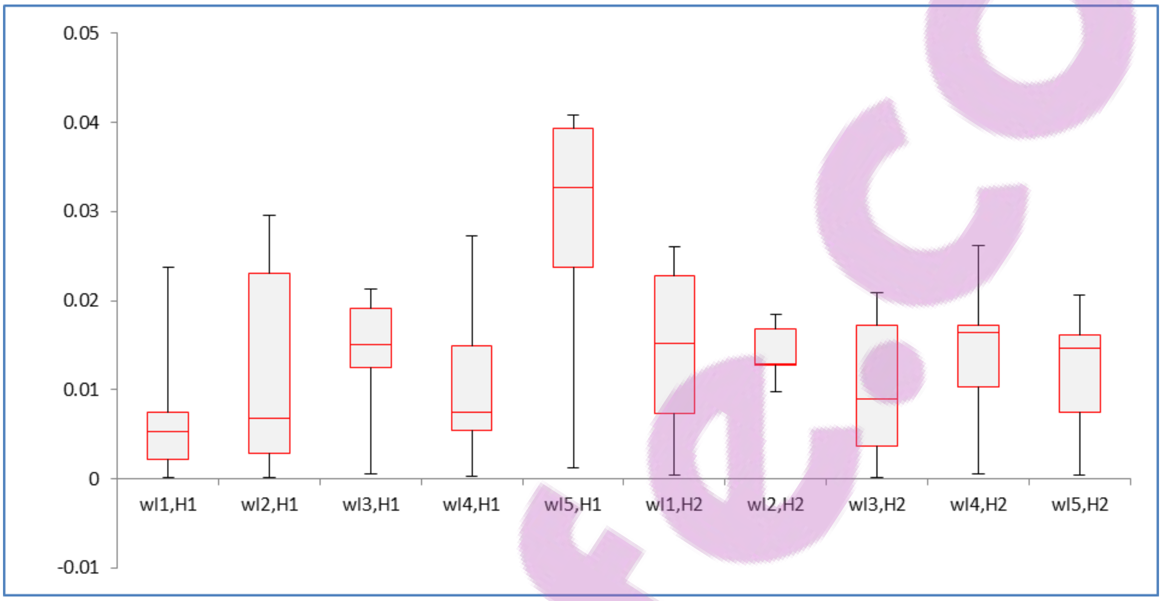

1.4
1.2
1
0.8
0.6
0.4
0.2
0
wH1,O1
wH2,O1

Figure 7.1: The connection weight values of the BMLPC for the trained MLPNN combination system for the Mae Tuen River catchment



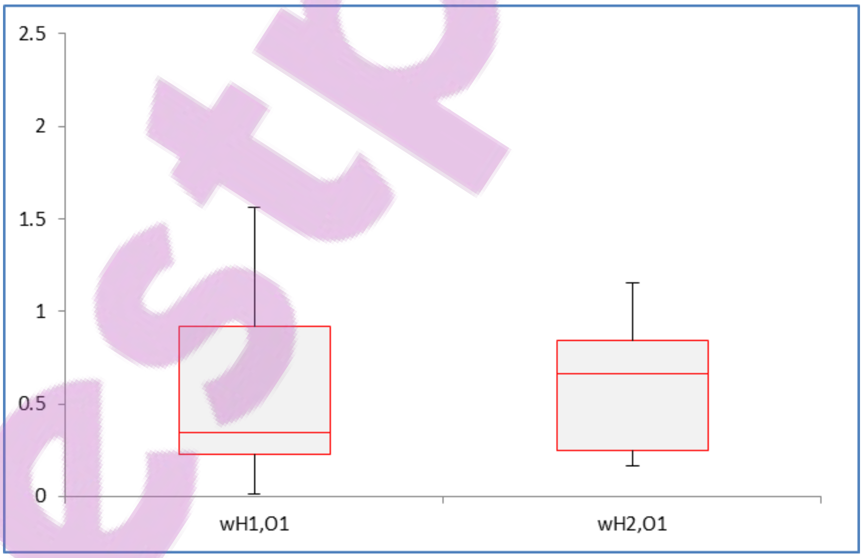


Figure 7.2: The connection weight values of the BMLPC for the trained MLPNN combination system for the Ohinemuri River catchment

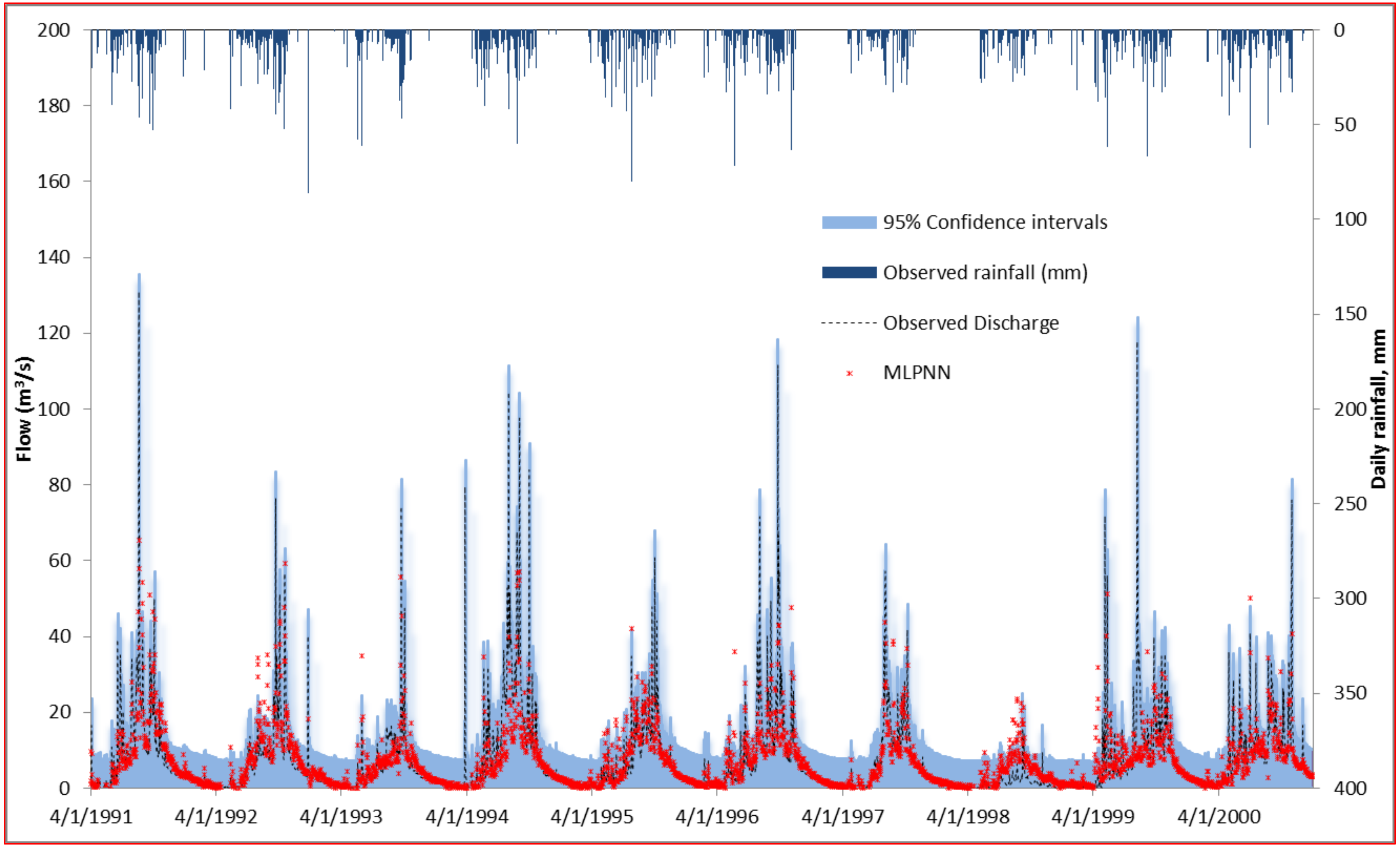


Figure 7.3: The 95% confidence intervals of the trained MLPNN combination system for the Mae Tuen River catchment

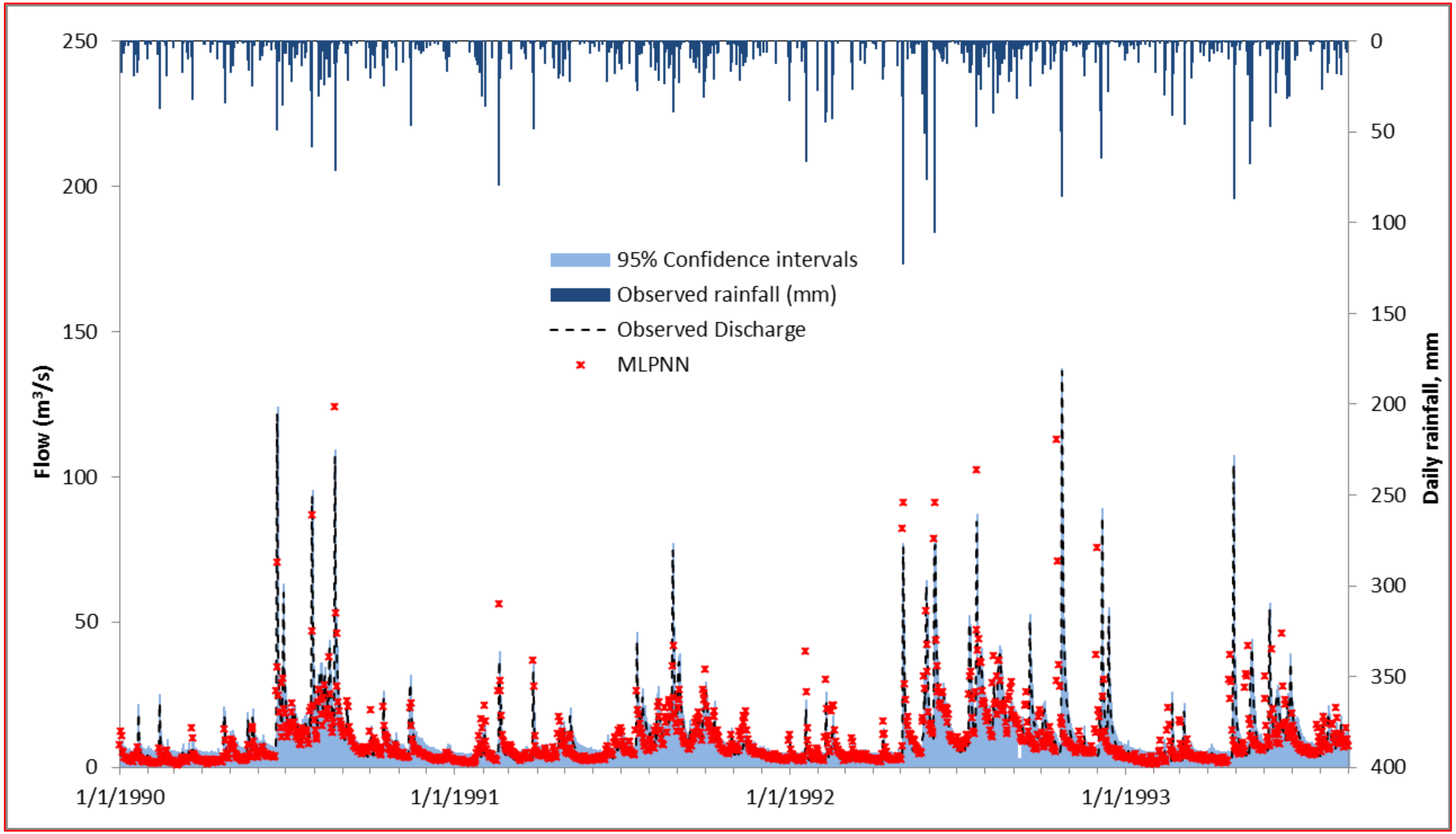


Figure 7.4: The 95% confidence intervals of the trained MLPNN combination system for the Ohinemuri River catchment.

7.4 Summary

This chapter addressed the third objective of the thesis. The study investigates and applies uncertainty analysis for quantifying the uncertainty in the developed multi-model combination systems. The objectives in this study are to analyse the uncertainty and estimate the confidence interval of the developed multi-model combination systems. This investigation presents a bootstrap method of behaviour uncertainty analysis for the MLPNN combination simulations. The bootstrap method has been applied to quantify the uncertainties of the developed MLPNN multi-model due to the parameters estimation and model prediction. This method is demonstrated through two case studies of two contrasting catchments located in Thailand and New Zealand, respectively. The results show that the bootstrap method is very effective for quantifying the uncertainty associated with the output of the trained MLPNN multi-model.

Chapter 8
Summary, Conclusion and Future Work

The main aim of this thesis is to develop the multi-model approach for river flow simulations. The development of the approach is tested using two contrasting catchments. The first catchment is the Mae Tuen River catchment located in Thailand and the second is the Ohinemuri River catchment located in New Zealand.

The main issues in the specific area of research explained in this thesis were identified on the basis of a review of the relevant literature (see Chapter 2) and the research objectives of this thesis were outlined in Chapter 1. Applications of methods and presentation and discussion of results obtained in this research have been presented in previous chapters. This research has applied the multi-model approach for river flow in Thailand and New Zealand catchments, which can lead to improving the modelling accuracy and reliability rather than modifying the existing models and developing new models.

The study has provided general guidance about the use of a combination river flow simulation system in two contrasting catchments (Thailand and New Zealand) and the optimal number of models to be used in the system. In addition, this study has contributed to the literature by presenting a novel application of uncertainties of simulated flows for the applied multi-model approach. The results indicate an improved multi-model approach and outlined approach to improve the reliability of the simulated flows, which is extremely beneficial to many users. The developed multi-model combination river flow can be used with other models to develop flood risk map and building planning regulations. The use of the improver river flow modelling system would really enhance the reliability of the developed maps and planning regulations.

In the first section of this chapter, the conclusions reached from the results of this research are provided. The chapter firstly presents separately the conclusions drawn from each of objectives in section 1.1 of the thesis. Finally, recommendations for possible future directions in the specific area of research of this thesis are presented.

8.1 Multi-model approach using ANNs and GEP

The first objective of the thesis was to compare the performance of the MLPNN, the RBFNN and the GEP combination methods in the multi-model combination system. This study presents a comparison between the GEP, the MLPNN and the RBFNN combination methods in multi-model combination systems of two different contrasting catchments. These combination methods involve the use of three different types of rainfall-runoff models: specifically, two black-box models – the LPM and the LVGFM, two conceptual models – the SMAR model and the NAM and a semi-distributed physically based model - the SWAT model to produce the multi-model combination system. These five rainfall-runoff models were applied to the daily data of each catchment for river flow simulations.

Overall, results found that the LPM outperformed other rainfall-runoff models which were characterised by strong seasonality in the calibration period, and the NAM performs better than other individual models in the verification period of the Mae Tuen River catchment. For the Ohinemuri River catchment, in the calibration period, results showed that the LVGFM has the best performance. However, for the verification period, results of the selected five models are quite inconsistent. The results also show that the SWAT model has a worse performance than other individual rainfall-runoff models which provide a significant input in producing the combined results in the multi-model combination system of both catchments. The other individual rainfall-runoff models required fewer parameters to produce the daily river flow simulation. This suggests that a complex model does not necessarily enhance the model performance; nevertheless the worst performance might be a significant input in producing the combined river flow simulations.

The question arises in this study as to whether or not a complex model with a poorer performance might make a significant input in producing the combined outputs in the multi-model combination system. This issue was investigated in Chapter 6 of the thesis. However, overall results show that the multi-model approach using different types (i.e. the empirical black-box models, the conceptual models and the semi-distributed physically based models) of rainfall-runoff models can achieve even greater improvement in accuracy and reliability of river flow simulations.

In the application of multi-models, comparison of the results reveals that the GEP performs better than neural network methods in the case of the catchment located in New Zealand. Nevertheless, the RBFNN method outperforms the GEP and the MLPNN combination method in the case of the catchment located in Thailand. The results in this study contrast with Shamseldin et al.'s (2007) results. They found that the MLPNN was identified as the appropriate ANN form for use in the context of combining outputs. Moreover, the results in Chapter 5 show that the GEP multi-model combination has the advantage over MLPNN and RBFNN multi-model combinations, in that the method can be expressed as a simple mathematical function. However, it is not conclusive as to which combination method produces better results than other methods in the multi-model combination system. The results suggest the selection of the best combination method to be used in conjunction with the multi-model approach may depend on the catchment type.

There are a few instances of overestimation and underestimation of simulated flow hydrographs (see Figs. 5.6 and 5.7) due to the uncertainty of the multi-model simulation and the calibrated model parameters of the selected five rainfall-runoff models applied in the multi-model combination system. As discussed in Chapter 4, the auto-calibration is chosen in this study for finding optimal parameter values. However, selecting the optimal values can significantly impact on the accuracy of model results.

8.2 The optimal number of rainfall-runoff models used in ANN combinations

According to the literature (see Chapter 2), the complexity of the multi-model combination system increases with the increase in the number of individual models. It will affect the performance of the multi-model combination system. The study investigates the optimal number of models and types of rainfall-runoff models to be used to improve performance in multi-model combination systems of two contrasting catchments located in Thailand and New Zealand. The optimal number will therefore maintain a balance between complexity and performance in multi-model combination systems. To investigate this problem, the performances of the developed MLPNN multi-model combinations and the selected five rainfall-runoff models (i.e. the LPM, the LVGFM, the SMAR model, the NAM and the SWAT model) were assessed using statistical methods which are commonly used in hydrology, namely the *RMSE,* the R^2*,* the *CE,* the *PBIAS* and the *KGE*. The graphical criteria involving the hydrograph plots and scatter plots were used in assessing the model performances.

The knowledge extraction techniques, namely the Garson's algorithm and the connection weight approach methods were applied in this research. The first method, Garson's algorithm, represents the connection weights using the absolute value, to components associated with each input neuron to the output neuron. The second method, the connection weight approach, uses the connection weights to calculate the variable contribution of each input neuron to the output neuron. For the Mae Tuen River catchment, results calculated by Garson's algorithm and the connection weight approach methods found that the LPM and LVGFM outputs are the most important in the multi-model combination system. For the Ohinemuri River catchment, results calculated by Garson's algorithm and connection weight approach methods found that the LVGFM is the most important input model in the MLPNN multi-model combination. The results calculated by the connection weight approach method indicate that the LVGFM output was the most important model in the MLPNN multi-model combinations of both catchments. Overall, results demonstrate that the knowledge extraction techniques had considerable potential for optimizing combined rainfall-runoff models in multi-model combination systems. It can also be used to reduce the complexity of multi-models by eliminating the least significant contributing input rainfall-runoff models. Two approaches (Garson's algorithm and the connection weight methods) can be used to identify the best set of input-output parameters for the development of simulation procedures using the MLPNN multi-model combinations. They can also be used to determine the relative importance of each simulated rainfall-runoff model, which was the model to be used in the trained MLPNN multi-model combination systems. The results in this study also illustrate that the developed MLPNN multi-models are not a black-box system, and that these approaches can help to explain the physical effects of input parameters on the model network.

For the Mae Tuen River catchment, the results showed that the combined four rainfall-runoff model outputs, namely, the LVGFM, the SMAR model, the SWAT model and the NAM model have the best performance in the multi-model combination system. For the Ohinemuri River catchment, results showed that the combination of three rainfall-runoff models (i.e. LVGFM, LPM and SWAR model) performs best in the MLPNN multi-model combination system in the calibration period. The combination of three rainfall-runoff models (i.e. LVGFM, SMAR model and SWAT model) also has the best performance in the MLPNN multi-model combination system in the verification periods. The results in Chapter 5 found that the SWAT model performed worse than other individual rainfall-runoff models of both catchments. However, when the SWAT model was used in the MLPNN multi-model combination system, results showed that the SWAT model, while producing a poorer performance than other models, nevertheless has a significant input in producing the combined outputs in the multi-model combination system. This is a strong justification for the argument that the worst performance of the complex model that involves many parameters based on complex laws of physical elements may improve the performance of multi-model combination systems. Overall, the analysis found that the optimal number of rainfall-runoff models which best perform in a MLPNN multi-model combination system depends on the selection of numbers of rainfall-runoff models to be used in the multi-model combination system.

8.3 Uncertainty analysis in ANN multi-model combination systems

The third objective of the thesis was to quantify the uncertainty and estimate the confidence intervals of the developed multi-model combination systems. Chapter 6 is dedicated to examining the behaviour of uncertainty analysis in the developed ANN multi-model combinations applied to two contrasting catchments located in Thailand and New Zealand. The bootstrap method was used to quantify the uncertainties of multi-model combination systems achieved by the developed MLPNN model for this research.

Results in Chapter 6 found that the BMLPNN models provide the greatest improvements in estimating the generalization errors of the trained MLPNN multi-models for simulating runoffs, at each time step. The bootstrap method can help our understanding of the behaviour of the system parameters (i.e. the connection weights) and the input variables (i.e. the individual rainfall-runoff model simulated runoffs) in the multi-model combination system. It can also be used to reduce the complexity of a rainfall-runoff models combination system by eliminating the least significant contributor to input variables, while the bootstrap confidence intervals can be the indicator with which to evaluate uncertainties of the developed multi-model simulations.

Results of two case studies (see Figs. 7.3 and 7.4) indicate that the proposed method can effectively quantify the uncertainty bounds of the MLPNN multi-model outputs. However, some simulated discharge outputs fail to capture the hydrograph peak flow for the catchment in New Zealand. Overall results of the proposed study show that the Bootstrap method can provide the greatest improvements to estimating the generalization errors of the developed multi-model approach for river flow simulations of each different catchment characteristic. This method is very effective for quantifying the uncertainty and estimating the confidence intervals of the developed multi-model combination systems.

8.4 Future Research Directions

This section provides the recommendation for future research directions in this research. These recommendations would further develop the multi-model approach for enhancing the accuracy of river flow simulations. Therefore, these recommendations are listed below:

1. Further investigation is needed on the individual rainfall-runoff models used in the multi-model combination systems, in which these models can operate at hourly or shorter time steps (flood events). These models need to be applied to different catchment types in order to establish guidelines about their use in different situations.

2. Building a multi-model combination system with more complex rainfall-runoff modes, or even the use different types of models, need to be investigated to improve a multi-model combination.

3. There are a number of linear and non-linear based combination methods, which can be used for producing the combined runoffs. However, there are only a limited number of studies which extensively consider the development and

applications of combination methods in the context of rainfall-runoff modelling. Further work needs to be done on improving a multi-model approach through use of more combination techniques.

4. More comprehensible techniques of knowledge extraction (i.e. TREPAN algorithm, C45 rule algorithm) for the trained neural networks model require investigation.

5. It is recommended here that more approaches and tools should be explored that can account for the uncertainty in an individual model's results and multi-model combination simulation's results.

6. How to reduce the uncertainty analysis in the developed multi-model combination systems? This question still need to be investigated further.

7. Future study is needed to improve a multi-model approach for enhancing the accuracy of river flow forecasting by considering uncertainty in future hydrology events (e.g. impacts of climate change). Such an approach is extremely beneficial to many users.

References

Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., Srinivasan, R. (2007). Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. Journal of Hydrology, 333(2–4), 413-430.

Abebe, A. J., Solomatine, D. P., and Venneker, R. (2000). Application of adaptive fuzzy rule-based models for reconstruction of missing precipitation events. Hydrological Sciences Journal, 45(3), 425-436.

Abrahart, R. J. (2003). Neural network rainfall-runoff forecasting based on continuous resampling. Journal of Hydroinformatics, 5, 51-61.

Abrahart, R. J., and See, L. (2002). Multi-model data fusion for river flow forecasting: an evaluation of six alternative methods based on two contrasting catchments. Hydrological Earth System Sciences, 6(4), 655-670.

Ahsan, M., and O'Connor, K. M. (1994). A simple non-linear rainfall-runoff model with a variable gain factor. Journal of Hydrology, 155(1-2), 151-183.

Aires, F. (2004). Neural network uncertainty assessment using Bayesian statistics with application to remote sensing: 1. Network weights. Journal of Geophysical Research: Atmospheres, 109(D10), D10303.

Ajami, N. K., Duan, Q., Gao, X., and Sorooshian, S. (2006). Multimodel combination techniques for analysis of hydrological simulations: Application to distributed model intercomparison project results. Journal of Hydrometeorology, 7(4), 755-768.

Ajami, N. K., Duan, Q., and Sorooshian, S. (2007). An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resources Research, 43(1), W01403.

Alansi, A. W., Amin, M. S. M., Abdul Halim, G., Shafri, H. Z. M., and Aimrun, W. (2009). Validation of SWAT model for stream flow simulation and forecasting in Upper Bernam humid tropical river basin, Malaysia. Hydrological Earth System Sciences. Discuss., 6(6), 7581-7609.

Anctil, F., Michel, C., Perrin, C., and Andréassian, V. (2004). A soil moisture index as an auxiliary ANN input for stream flow forecasting. Journal of Hydrology, 286(1-4), 155-167.

Anderson, M. G., and Burt, T. P. (1985). Hydrological Forecasting. Chichester, New York: John Wiley and Sons, Ltd.

Andrews, R., Diederich, J., and Tickle, A. B. (1995). Survey and critique of techniques for extracting rules from trained artificial neural networks. Knowledge-Based Systems, 8(6), 373-389.

Armstrong, J. S. (1989). Combining Forecasts: The End of the Beginning or the Beginning of the End? International J. Forecasting, 5, 585-588.

Armstrong, J. S. (2001). Combining forecasts Principles of Forecasting: A Handbook for Researchers and Practitioners, J. Scott Armstrong (ed.).

Arnold, J. G., and Allen, P. M. (1996). Estimating hydrologic budgets for three Illinois watersheds. Journal of Hydrology, 176(1–4), 57-77.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Williams, J. R. (1998). Large Area Hydrologic Modeling and Assessment Part I: Model Development 1. JAWRA Journal of the American Water Resources Association, 34(1), 73-89.

ASCE. (1993). Criteria for Evaluation of Watershed Models. Journal of Irrigation and Drainage Engineering, 119(3), 429-442.

Atger, F. (1999). Verification of intense precipitation forecasts from single models and ensemble prediction systems. Nonlin. Processes Geophys., 8(6), 401-417.

Atiya, A. F. (2008). Forecast Combinations of Computational Intelligence and Linear Models for the NN5 Time Series Forecasting competition. International J. Forecasting.

Aytex, A., Asce, M., and Alp, M. (2008). An application of artificial intelligence for rainfall-runoff modeling. Journal of Earth System Science, 117(2), 145-155.

Bates, J. M., and Granger, C. W. J. (1969). The combination of forecasts. Journal of the Operational Research Society (20), 451-468.

Bell, V. A., Carrington, D. S., and Moore, R. J. (2001). Comparison of rainfall-runoff models for flood forecasting, Part2: Calibration and Evaluation of models: Environmental Agency R&D Report W242.

Benke, K. K., Lowell, K. E., and Hamilton, A. J. (2008). Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model. Mathematical and Computer Modelling, 47(11–12), 1134-1149.

Beven, K. (1997). TOPMODEL: a critique. Hydrological Processes, 11(9), 1069-1085.

Beven, K., and Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. Hydrological Processes, 6(3), 279-298.

Blasone, R.-S., Vrugt, J. A., Madsen, H., Rosbjerg, D., Robinson, B. A., and Zyvoloski, G. A. (2008). Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. Advances in Water Resources, 31(4), 630-648.

Blattberg, R. C., and Hoch, S. J. (1990). Database Models and Managerial Intuition: 50% Model + 50% Manager. Management Science, 36(8), 887-899.

Cao, W., Bowden, W. B., Davie, T., and Fenemor, A. (2006). Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. Hydrological Processes, 20(5), 1057-1073.

Cavadias, G., and Morin, G. (1986). The Combination of Simulated Discharges of Hydrological Models. Nordic Hydrology, 17, 21-32.

Chapman, D. S., and Purse, B. V. (2011). Community versus single-species distribution models for British plants. Journal of Biogeography, 38(8), 1524-1535.

Chu, T. W., and Shirmohammadi, A. (2004). Evaluation of the SWAT model's hydrology component in the piedmont physiographic region of Maryland. Trans. ASAE, 47(4), 1057-1073.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography

International J. Forecasting, 5, 559-583.

Coulibaly, P. (2008). Multi-model approach to hydrologic impact of climate change From Headwaters to the Ocean (pp. 249-255): CRC Press.

Coulibaly, P., Hache, M., Fortin, V., and Bobee, B. (2005). Improving Daily Reservoir Inflow Forecasts with Model Combination. Journal of Hydrologic, 10(2), 91-99.

Dawson, C. W., and Wilby, R. L. (2001). Hydrological modelling using artificial neural networks. Progress in Physical Geography, 25(1), 80-108.

Demargne, J., Wu, L., Regonda, S., Brown, J., Lee, H., He, M., Seo, D. J., Hartman, R., Herr, H. D., Fresch, M., Schaake, J., Zhu, Y. (2013). The Science of NOAA's Operational Hydrologic Ensemble Forecast Service. Bulletin of the American Meteorological Society.

Deutsch, M., Granger, C. W. J., and Teräsvirta, T. (1994). The combination of forecasts using changing weights. International Journal of Forecasting, 10(1), 47-57.

Devineni, N., Sankarasubramanian, A., and Ghosh, S. (2008). Multimodel ensembles of streamflow forecasts: Role of predictor state in developing optimal combinations. Water Resources Research, 44(9), W09404.

DHI. (2007). MIKE 11 - A modelling system for rivers and channels - Reference Manual. DHI software 2007.

Duan, Q., Ajami, N. K., Gao, X., and Sorooshian, S. (2007). Multi-model ensemble hydrologic prediction using Bayesian model averaging. Advances in Water Resources, 30(5), 1371-1386.

Duan, Q., Gupta, H. V., Sorooshian, S., Rousseau, A. N., and Turcotte, R. (2003). Calibration of Watershed Models (Vol. 6). Washington, DC: AGU.

Dybowski, R., Roberts, S. J., Dybowski, R., and Roberts., S. J. (2001). Confidence intervals and prediction intervals for feedforward neural networks. New York: Cambridge Univ. Press.

Efron, B. (1979). Bootstrap methods- Another Look at the Jackknife. Annals of Statistics, 7(1), 1-26.

Efron, B., and Tibshirani, R. J. (1993). An Introduction to the Bootstrap. New York: Chapman and Hall.

El Ouazzani Taibi, A., Zhang, G. P., and Elfeki, A. (2006). Hydrologic responses of the Zwalm catchment using the REW model: incorporating uncertainty of soil properties. Hydrology Earth System. Sciences. Discuss., 3(1), 69-114.

Evsukoff, A. G., Cataldi, M., and de Lima, B. S. L. P. (2012). A multi-model approach for long-term runoff modeling using rainfall forecasts. Expert Systems with Applications, 39(5), 4938-4946.

Exbrayat, J. F., Viney, N. R., Seibert, J., Frede, H. G., and Breuer, L. (2011). Multi-model data fusion as a tool for PUB: example in a Swedish mesoscale catchment. Advanced Geoscience. 29, 43-50.

Fenicia, F., Solomatine, D. P., Savenije, H. H. G., and Matgen, P. (2007). Soft combination of local models in a multi-objective framework. Hydrology Earth System Sciences. Discuss., 4(1), 91-123.

Fernando, A., Shamseldin, A., and Abrahart, R. (2012). Use of Gene Expression Programming for Multimodel Combination of Rainfall-Runoff Models. Journal of Hydrologic Engineering, 17(9), 975-985.

Fernando, A. K., Shamseldin, A. Y., and Abrahart, R. J. (2011). Comparison of two data-driven approaches for daily river flow forecasting. Paper presented at the 19th International Congress on Modelling and Simulation.

Fernando, Shamseldin, A. Y., and Abrahart, R. J. (2009). Using gene expression programming to develop a combined runoff estimate model from conventional rainfall-runoff model outputs. Paper presented at the 18th World IMACS / MODSIM Congress, .

Ferreira, C. (2001). Gene Expression Programming: A New Adaptive Algorithm for Solving Problems. Complex Systems, 13(2), 87-129.

Ferreira, C. (2006). Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence. Germany: Springer-Verlag.

Gallant, S. I. (1988). Connectionist expert systems. Commun of ACM, 31(2), 152-169.

Gallant, S. I. (1988). Connectionist Expert Systems. Communications of the ACM, 31(2), 152-169.

Garson, G. D. (1991). Interpreting neural-network connection weights. AI Expert, 6(4), 46-51.

Georgakakos, K. P., Seo, D. J., Gupta, H., Schaake, J., and Butts, M. B. (2004). Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. Journal of Hydrology, 298(1-4), 222-241.

Goswami, M., and O'Connor, K. M. (2007). Real-time flow forecasting in the absence of quantitative precipitation forecasts: A multi-model approach. Journal of Hydrology, 334(1-2), 125-140.

Goswami, M., O'Connor, K. M., Bhattarai, K. P., and Shamseldin, A. Y. (2005). Assessing the performance of eight real-time updating models and procedures for the Brosna River. Hydrology and Earth System Sciences, 9(4), 394-411.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology, 377(1–2), 80-91.

Gupta, H. V., Sorooshian, S., and Yapo, P. O. (1998). Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resources Research, 34(4), 751-763.

Han, D., Kwong, T., and Li, S. (2007). Uncertainties in real-time flood forecasting with neural networks. Hydrological Processes, 21(2), 223-228.

He, Z., Zhao, W., and Liu, H. (2014). Comparing the performance of empirical black-box models for river flow forecasting in the Heihe River Basin, Northwestern China. Hydrological Processes, 28(1), 1-7.

Heskes, T. (1997). Practical confidence and prediction intervals Advances in Neural Information Processing Systems 9 (pp. 176-182): MIT Press.

Hsu, K.-l., Gupta, H. V., and Sorooshian, S. (1995). Artificial Neural Network Modeling of the Rainfall-Runoff Process. Water Resources Research, 31(10), 2517-2530.

Izadifar, Z., and Elshorbagy, A. (2010). Prediction of hourly actual evapotranspiration using neural networks, genetic programming, and statistical models. Hydrological Processes, 24(23), 3413-3425.

Jain, A., and Kumar, S. (2009). Dissection of trained neural network hydrologic models for knowledge extraction. Water Resource. Research, 45(7), W07420.

Jeong, D. I., and Kim, Y.O. (2009). Combining single-value streamflow forecasts – A review and guidelines for selecting techniques. Journal of Hydrology, 377(3–4), 284-299.

Jia, Y., and Culver, T. B. (2006). Bootstrapped artificial neural networks for synthetic flow generation with a small data sample. Journal of Hydrology, 331(3–4), 580-590.

Kachroo, R. K., and Liang, G. C. (1992). River flow forecasting. Part 2. Algebraic development of linear modelling techniques. Journal Hydrology, 133, 17-40.

Kachroo, R. K., Liang, G. C., and O'Connor, K. M. (1988). Application of the linear perturbation model (LPM) to flood routing on the Mekong River. Hydrological Sciences Journal/Journal des Sciences Hydrologiques, 33(2), 193-214.

Kachroo, R. K., and Natale, L. (1992). Non-linear modelling of the rainfall-runoff transformation. Journal of Hydrology, 135(1–4), 341-369.

Kachroo, R. K., Sea, C.H., Warsi, M.S., Jemenez, H., and Saxena, R. P. (1992). River flow forecasting. Part 3. Applications of linear techniques in modelling rainfall-runoff transformations. Journal of Hydrology, 133, 41-97.

Kalteh, A. M. (2008a). Rainfall-runoff modelling using artificial neural networks (ANNa): modelling and understanding. Caspian Journal of Environmental Sciences, 6, 53-58.

Kalteh, A. M. (2008b). Rainfall-runoff modelling using artificial neural networks (ANNs): modelling and understanding. Caspian Journal of Environmental Sciences, 6(1), 53-58.

Kasiviswanathan, K. S., and Sudheer, K. P. (2013). Quantification of the predictive uncertainty of artificial neural network based river flow forecast models. Stochastic Environmental Research and Risk Assessment, 27(1), 137-146.

Kavetski, D., Kuczera, G., and Franks, S. W. (2006). Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. Water Resources Resources, 42(3), W03407.

Kerem Cigizoglu, H., and Kisi, Ö. (2006). Methods to improve the neural network performance in suspended sediment estimation. Journal of Hydrology, 317(3-4), 221-238.

Keshta, N., Elshorbagy, A., and Carey, S. (2009). A generic system dynamics model for simulating and evaluating the hydrological performance of reconstructed watersheds. Hydrology Earth System Sciences, 13(6), 865-881.

Khan, M., Shamseldin, A., and Melville, B. (2014). Impact of Ensemble Size on Forecasting Occurrence of Rainfall Using TIGGE Precipitation Forecasts. Journal of Hydrologic Engineering, 19(4), 732-738.

Khu, S. T. and Werner, M. G. F., (1999). Reduction of Monte-Carlo simulation runs for uncertainty estimation in hydrological modelling. Hydrol. Earth Syst. Sci., 7(5), 680-692

Ki, O. (2007). Streamflow Forecasting Using Different Artificial Neural Network Algorithms. Journal of Hydrologic Engineering, 12(5), 532-539.

Kieran, O. Ä., (2005). River flow forecasting River Basin Modelling for Flood Risk Mitigation (pp. 197-213): Taylor & Francis.

Kim, Y. O., Jeong, D., and Ko, I. H. (2006). Combining rainfall-runoff model outputs for improving ensemble streamflow prediction. Journal of Hydrologic Engineering, 11(6), 578-588.

Kingston, G. B., Maier, H. R., and Lambert, M. F. (2006). A probabilistic method for assisting knowledge extraction from artificial neural networks used for hydrological prediction. Mathematical and Computer Modelling, 44(5–6), 499-512.

Koza, J. R. (1992). Genetic Programming: On the Programming of Computers by Means of Natural Selection: MIT Press, Cambridge, Massachusetts.

Krueger, T., Freer, J., Quinton, J. N., Macleod, C. J. A., Bilotta, G. S., Brazier, R. E., Butler, P., Haygarth, P. M. (2010). Ensemble evaluation of hydrological model hypotheses. Water Resources. Research, 46(7), W07516.

Kücken, M., Gerstengarbe, F. W., and Orlowsky, B. (2009). A Combination of Cluster Analysis and Kappa Statistics for the Evaluation of Climate Model Results. Journal of Applied Meteorology and Climatology, 48(9), 1757-1765.

Kuczera, G. and Parent, E., (1998). Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. Journal of Hydrology, 211(1-4), 69-85.

Legates, D. R., and McCabe, G. J. (1999). Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation. Water Resources Research, 35(1), 233-241.

Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., and Aulagnier, S. (1996). Application of neural networks to modelling nonlinear relationships in ecology. Ecological Modelling, 90(1), 39-52.

Liang, G. C., Kachroo, R.K., Kang, W., and Yu, X. Y. (1992). River flow forecasting. Part 4. Applications of linear modelling techniques for flow routing on large catchments. Journal of Hydrology, 133, 99-140.

Liang, Z., Wang, D., Guo, Y., Zhang, Y., and Dai, R. (2013). Application of Bayesian Model Averaging Approach to Multimodel Ensemble Hydrologic Forecasting. Journal of Hydrologic Engineering, 18(11), 1426-1436.

Majid, D., Bahram, S., Farzin, N.S., Ashkan, F., and Roohollah, N. (2013). Uncertainty analysis of streamflow drought forecast using artificial neural networks and Monte-Carlo simulation. International Journal of Climatology, 34(4), 1169-1180.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., et al. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. Journal of Forecasting, 1(2), 111-153.

Makridakis, S., and Winkler, R. L. (1983). Averages of Forecasts: Some empirical results. Management Science, 29(7), 987-996.

Mandeville, A. N., O'Connell, P. E., Sutcliffe, J. V., and Nash, J. E. (1970). River flow forecasting through conceptual models part III - The Ray catchment at Grendon Underwood. Journal of Hydrology, 11(2), 109-128.

Marshall, L., Nott, D., and Sharma, A. (2004). A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. Water Resources Research, 40(2), W02501.

Mason, J. C., Price, R. K., and Tem'Me, A. (1996). A neural network model of rainfall-runoff using radial basis functions. Journal of Hydraulic Research, 34(4), 537-548.

McMillan, H., Jackson, B., Clark, M., Kavetski, D., and Woods, R. (2011). Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. Journal of Hydrology, 400(1-2), 83-94.

Moriasi, D. N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., and Veith, T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, American Society of Agricultural and Biological Engineers, 50(3):885-900.

Nash, J. E., and Barsi, B. I. (1983). A hybrid model for flow forecasting on large catchments. Journal of Hydrology, 65(1–3), 125-137.

Nash, J. E., and Foley, J. J. (1982). Linear models of rainfall-runoff systems. In: 'Rainfall-Runoff Relationship', Proceedings of the International Symposium on Rainfall-Runoff modelling. Mississippi State University, May 1981, USA. Edited by V.P. Singh, Water Resources Publications pp.: 51-66, 1982

Nash, J. E., and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models: Part 1 - A discussion of principles. Hydrology, 10, 282-290.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., and Wiliams, J. R. (2011). Soil and Water Assessment Tool Theoretical Documentation Version 2009. Texas, USA.: Texas Water Resources Institute Technical Report No.406.

Neuman, S. P. (2003). Maximum likelihood Bayesian averaging of uncertain model predictions. Stochastic Environmental Research and Risk Assessment, 17(5), 291-305.

NIelsen, S. A., and Hansen, E. (1973). Numerical Simulation of the Rainfall-runoff Process on a Daily Basis. Nordic Hydrology, 4, 171-190.

O'Connell, P. E., Nash, J. E., and Farrell, J. P. (1970). River flow forecasting through conceptual models part II - The Brosna catchment at Ferbane. Journal of Hydrology, 10(4), 317-329.

Olden, J. D., and Jackson, D. A. (2002). Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. Ecological Modelling, 154(1-2), 135-150.

Olden, J. D., Joy, M. K., and Death, R. G. (2004). An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. Ecological Modelling, 178(3–4), 389-397.

Özesmi, S. L., and Özesmi, U. (1999). An artificial neural network approach to spatial habitat modelling with interspecific interaction. Ecological Modelling, 116(1), 15-31.

Palm, F. C., and Zellner, A. (1992). To combine or not to combine? Issues of combining forecasts. Journal of Forecasting, 11, 687-701.

Parasuraman, K., and Elshorbagy, A. (2008). Toward improving the reliability of hydrologic prediction: Model structure uncertainty and its quantification using ensemble-based genetic programming framework. Water Resources Research, 44(12), W12406.

Phukoetphim, P., Shamseldin, A., and Melville, B. (2013). Knowledge extraction from artificial neural network(s CHECK. IMIGHT HAVE THIS WRONG) for rainfall-runoff models combination system. Journal of Hydrologic Engineering, in process.

Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M. (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. Monthly Weather Review, 133(5), 1155-1174.

Rahman, M., Goel, N., and Arya, D. (2012). Development of the Jamuneswari Flood Forecasting System: Case Study in Bangladesh. Journal of Hydrologic Engineering, 17(10), 1123-1140.

Reda, A. G., Tripathi, N. K., Soni, P., Tipdecho, T., and Phalke, A. (2013). Temporal climate trend of Ping Basin of Thailand and implications for Mekong Region. Earth Science and Climatic Change, 4(4).

Refsgaard, J. C. (1997). Parameterisation, calibration and validation of distributed hydrological models. Journal of Hydrology, 198(1–4), 69-97.

Refsgaard, J. C. (1997). Validation and Intercomparison of Different Updating Procedures for Real-Time Forecasting. Nordic Hydrology, 28, 65-84.

Refsgaard, J. C., and Knudsen, J. (1996). Operational Validation and Intercomparison of Different Types of Hydrological Models. Water Resources Research, 32(7), 2189-2202.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W. (2010). Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. Water Resources Research, 46(5), W05521.

Rezaeianzadeh, M., Stein, A., Tabari, H., Abghari, H., Jalalkamali, N., Hosseinipour, E. Z., Singh, V. P. (2013). Assessment of a conceptual hydrological model and artificial neural networks for daily outflows forecasting. International Journal of Environmental Science and Technology, 10(6), 1181-1192.

Ridley, D. (1997). Optimal weights for combining antithetic forecasts. Computers and Industrial Engineering, 32(2), 371-381.

Roushangar, K., Mehrabani, F. V., and Alami, M. T. (2013). Forecasting daily stream flows of Vaniar River using genetic programming and neural networks approaches. Journal of Civil Engineering and Urbanism, 3(4), 197-200.

Sajikumar, N., and Thandaveswara, B. S. (1999). A non-linear rainfall–runoff model using an artificial neural network. Journal of Hydrology, 216(1–2), 32-55.

Santhi, C., Arnold, J. G., Williams, J. R., Dugas, W. A., Srinivasan, R., and Hauck, L. M. (2001). Validation of the Swat model on a large river basin with point and nonpoint sources1. JAWRA Journal of the American Water Resources Association, 37(5), 1169-1188.

Schuol, J., Abbaspour, K. C., Srinivasan, R., and Yang, H. (2008). Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model. Journal of Hydrology, 352(1–2), 30-49.

See, L., and Abrahart, R. J. (2001). Multi-model data fusion for hydrological forecasting. Computers and Geosciences, 27(8), 987-994.

See, L., and Openshaw, S. (2000). A hybrid multi-model approach to river level forecasting. Hydrological Sciences Journal, 45(4), 523-536.

Senthil Kumar, A. R., Sudheer, K. P., Jain, S. K., and Agarwal, P. K. (2005). Rainfall-runoff modelling using artificial neural networks: comparison of network types. Hydrological Processes, 19(6), 1277-1291.

Shamseldin, A. Y., Nasr, A. E., and O'Connor, K. M. (2002). Comparsion of different forms of the multi-layer feed-forward neural network method used for river flow forecasting. Hydrology and Earth System Sciences, 6(4), 671-684.

Shamseldin, A. Y., and O'Connor, K. M. (1999). A real-time combination method for the outputs of different rainfall-runoff models. Hydrological Sciences Journal, 44(6), 895-912.

Shamseldin, A. Y., and O'Connor, K. M. (2003). A "consensus" real-time river flow forecasting model for the Blue Nile River. IAHS-AISH Publication (281), 82-89.

Shamseldin, A. Y., O'Connor, K. M., and Liang, G. C. (1997). Methods for combining the outputs of different rainfall-runoff models. Journal of Hydrology, 197(1-4), 203-229.

Shamseldin, A. Y., O'Connor, K. M., and Nasr, A. E. (2007). A comparative study of three neural network forecast combination methods for simulated river flows of different rainfall—runoff models. Hydrological Sciences Journal, 52(5), 896-916.

Sharma, S.K. and Tiwari, K.N. (2009). Bootstrap based artificial neural network (BANN) analysis for hierarchical prediction of monthly runoff in Upper Damodar Valley Catchment. Journal of Hydrology, 374 (1), 209-222.

Shen, S., Li, G., and Song, H. (2011). Combination forecasts of International tourism demand. Annals of Tourism Research, 38(1), 72-89.

Sheng, C., Xunxian, W., and Harris, C. J. (2005). Experiments with repeating weighted boosting search for optimization signal processing applications. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 35(4), 682-693.

Sherrod, P. H. (2003). DTREG predictive modeling software, User Guide, www.dtreg.com.

Shiri, J., and Kisi, Ö. (2010). Comparison of genetic programming with neuro-fuzzy systems for predicting short-term water table depth fluctuations. Computers & Geosciences, In Press, Corrected Proof.

Shrestha, D. L. (2009). Uncertainty Analysis in Rainfall-Runoff Modelling: Application of Machine Learning Techniques. Delft University of Technology.

Shrestha, D. L., Kayastha, N., and Solomatine, D. P. (2009). A novel approach to parameter uncertainty analysis of hydrological models using neural networks. Hydrol. Earth Syst. Sci. Discuss., 6(2), 1677-1706.

Shu, C., and Quarda, T.B.M.J. (2007). Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space. Water Resources Research 43 (7), W07438.

Singh, A., Imtiyaz, M., Isaac, R., and Denis, D. (2013). Comparison of Artificial Neural Network Models for Sediment Yield Prediction at Single Gauging Station of Watershed in Eastern India. Journal of Hydrologic Engineering, 18(1), 115-120.

Singh, P., and Deo, M. C. (2007). Suitability of different neural networks in daily flow forecasting. Applied Soft Computing, 7(3), 968-978.

Srinivasan, R., Ramanarayanan, T. S., Arnold, J. G., and Bednarz, S. T. (1998). Large Area Hydrologic Modeling and Assessment Part II: Model Application 1. JAWRA Journal of the American Water Resources Association, 34(1), 91-101.

Srivastav, R. K., Sudheer, K. P., and Chaubey, I. (2007). A simplified approach to quantifying predictive and parametric uncertainty in artificial neural network hydrologic models. Water Resources Research, 43(10), W10407.

Stathakis, D. (2009). How many hidden layers and nodes? International Journal of Remote Sensing, 30, 2133-2147.

Stock, J. H., and Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. Journal of Forecasting, 23(6), 405-430.

Sudheer, K. P. (2005). Knowledge Extraction from Trained Neural Network River Flow Models. Journal of Hydrologic Engineering, 10(4), 264-269.

Tan, B. Q., and O'Connor, K. M. (1996). Application of an empirical infiltration equation in the SMAR conceptual model. Journal of Hydrology, 185(1–4), 275-295.

Timmermann, A. G. (2005). Forecast Combinations. Centre for Economic Policy Research, 90–98 Goswell Rd, London EC1V 7RR, UK.

Tingsanchali, T., and Gautam, M. R. (2000). Application of tank, NAM, ARMA and neural network models to flood forecasting. Hydrological Processes, 14(14), 2473-2487.

Tiwari, M. K., and Chatterjee, C. (2010). Uncertainty assessment and ensemble flood forecasting using bootstrap based artificial neural networks (BANNs). Journal of Hydrology, 382(1–4), 20-33.

Tokar, A. S., and Markus, M. (2000). Precipitation-Runoff Modeling Using Artificial Neural Networks and Conceptual Models. Journal of Hydrologic Engineering, 5(2), 156-161.

Tsai, F. C. (2010). Bayesian model averaging assessment on groundwater management under model structure uncertainty. Stochastic Environmental Research and Risk Assessment, 24(6), 845-861.

Van Griensven, A., and Bauwens., W. (2005). Application and evaluation of ESWAT on the Dender basin and Wister Lake basin. Hydrologic Processes, 19(3), 827-838.

Van Hinsbergen, C. P. I., Van Lint, J. W. C., and Van Zuylen, H. J. (2009). Bayesian committee of neural networks to predict travel times with confidence intervals. Transportation Research Part C: Emerging Technologies, 17(5), 498-509.

Velázquez, J. A., Anctil, F., and Perrin, C. (2010). Performance and reliability of multimodel hydrological ensemble simulations based on seventeen lumped models and a thousand catchments. Hydrology Earth System Sciences, 14(11), 2303-2317.

Velázquez, J. A., Anctil, F., Ramos, M. H., and Perrin, C. (2011). Can a multi-model approach improve hydrological ensemble forecasting? A study on 29 French catchments using 16 hydrological model structures. Adv. Geosci., 29, 33-42.

Verbunt, M., Walser, A., Gurtz, J., Montani, A., and Schär, C. (2007). Probabilistic Flood Forecasting with a Limited-Area Ensemble Prediction System: Selected Case Studies. Journal of Hydrometeorology, 8(4), 897-909.

Vernieuwe, H., Georgieva, O., De Baets, B., Pauwels, V. R. N., Verhoest, N. E. C., and De Troch, F. P. (2005). Comparison of data-driven Takagi–Sugeno models of rainfall–discharge dynamics. Journal of Hydrology, 302(1–4), 173-186.

Viney, N. R., Bormann, H., Breuer, L., Bronstert, A., Croke, B. F. W., Frede, H., Gräff, T. Hubrechts, L., Huisman, J. A., Jakeman, A. J., Kite, G. W., Lanini, J. Leavesley, G., Lettenmaier, D. P., Lindström, G., Seibert, J., Sivapalan, M., Willems, P. (2009). Assessing the impact of land use change on hydrology by ensemble modelling (LUCHEM) II: Ensemble combinations and predictions. Advances in Water Resources, 32(2), 147-158.

Viney, N. R., Vaze, J., Chiew, F. H. S., Perraud, J., Post, D. A., and Teng, J. (2009). Comparison of multi-model and multi-donor ensembles for regionalisation of runoff generation using five lumped rainfall-runoff models. Paper presented at the 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation.

Vrugt, J. A., and Robinson, B. A. (2007). Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging. Water Resources Research, 43(1), W01411.

Vu, M. T., Raghavan, S. V., and Liong, S. Y. (2012). SWAT use of gridded observations for simulating runoff – a Vietnam river basin study. Hydrology and Earth System Sciences, 16(8), 2801-2811.

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., and Gupta, H. V. (2003). Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis. Hydrological Processes, 17(2), 455-476.

Wang, F., Zhou, B., Xu, J., Song, L., and Wang, X. (2009). Application of neural network and MODIS 250 m imagery for estimating suspended sediments concentration in Hangzhou Bay, China. Environmental Geology, 56(6), 1093-1101.

Wang, W., Gelder, P. V., and Vrijling, J. K. (2005). Improving daily stream flow forecasts by combining ARMA and ANN models. Paper presented at the International conference on innovation advance and implementation of flood forecasting technology.

Wang, Y., Zheng, T., Zhao, Y., Jiang, J., Wang, Y., Guo, L., Wang, P. (2013). Monthly water quality forecasting and uncertainty assessment via bootstrapped wavelet neural networks under missing data for Harbin, China. Environmental Science and Pollution Research, 1-15.

Weckman, G., Millie, D. F., Ganduri, C., Rangwala, M.,Young, W., Rinder, M., Fahnenstiel, G. L. (2009). Knowledge Extraction from the Neural 'Black Box' in Ecological Monitoring. Journal of Industrial and Systems Engineering, 3(1), 38-55.

Wei, S., Song, J., and Khan, N. I. (2012). Simulating and predicting river discharge time series using a wavelet-neural network hybrid modelling approach. Hydrological Processes, 26(2), 281-296.

Weigel, A. P., Liniger M.A., and Appenzeller, C. (2008). Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? Quarterly Journal of the royal meteorological society, 134, 241–260.

Wilby, R. L., Abrahart, R. J., and Dawson, C. W. (2003). Detection of conceptual model rainfall—runoff processes inside an artificial neural network. Hydrological Sciences Journal, 48(2), 163-181.

World Meteorological Organization (WMO), 1992. Simulated real-time intercomparison of hydrological models. Oper. Hydrol. Rep. 38, WMO No. 779, Geneva.

Xiong, L., and O'Connor, K. M. (2008). An empirical method to improve the prediction limits of the GLUE methodology in rainfall–runoff modeling. Journal of Hydrology, 349(1–2), 115-124.

Xiong, L., Shamseldin, A. Y., and O'Connor, K. M. (2001). A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system. Journal of Hydrology, 245(1-4), 196-217.

Ye, M., Neuman, S. P., and Meyer, P. D. (2004). Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff. Water Resources Research, 40(5), W05113.

Yu, P.-S., Yang, T.-C., and Chen, S.-J. (2001). Comparison of uncertainty analysis methods for a distributed rainfall–runoff model. Journal of Hydrology, 244(1–2), 43-59.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159-175.

Zhang, L., Zhao, W., He, Z., and Liu, H. (2009). Application of the Takagi–Sugeno fuzzy system for combination forecasting of river flow in semiarid mountain regions. Hydrological Processes, 23(10), 1430-1436.

Zhang, X., and Zhao, K. (2012). Bayesian Neural Networks for Uncertainty Analysis of Hydrologic Modeling: A Comparison of Two Schemes. Water Resources Management, 26(8), 2365-2382.