

**School of Civil and Mechanical Engineering
Department of Civil Engineering**

**River Flow Forecasting Using an Integrated Approach of Wavelet Multi-
Resolution Analysis and Computational Intelligence Techniques**

Honey Badrzadeh

**This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University**

October 2014

DECLARATION

To the best of my knowledge and belief, this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature:



Honey Badrzadeh

Date: September 2014

ABSTRACT

Reliable river flow forecasting is a key element in achieving sustainable water resources and environmental management. Accurate short term and long term river flow forecasts are particularly essential for the design of hydraulic structures, flood and drought analysis, irrigation scheduling, reservoir operation and environmental planning. Due to stochastic characteristics of hydrological events, forecasting the future condition of surface water is always associated with uncertainty. A large number of modelling techniques, ranging from physically-based to data-driven approach, have been studied to alleviate this uncertainty. As a result of technological advances in the recent years, computational intelligence approaches (CI) have become increasingly popular in hydrological modelling. Compared to conceptual and physics-based methods, CI models require minimum observation data to simulate complex hydrological processes.

This thesis focuses on improving the accuracy and reliability of river flow forecasting. Developing hybrid CI models, wavelet multi-resolution analysis is applied in conjunction with computational intelligence techniques. Two promising data-driven approaches of artificial neural networks (ANN) and adaptive neuro-fuzzy inference system (ANFIS) are adopted. Various types of ANN, ANFIS and hybrid wavelet models, are developed. Historical data of four Australian rivers, with different characteristics, are employed to investigate different applications of proposed approach in river flow forecasting.

Firstly, the impact of multivariate input selection on daily river flow forecasting is investigated when both rainfall and river flow historical time series are applied as inputs. Back propagation feed forward neural networks (BPFF), ANFIS with fuzzy C-mean clustering (FCM), hybrid wavelet neural networks (WNN) and wavelet neuro-fuzzy (WNFC) model are developed and applied for forecasting the flow of two different rivers of Harvey and Avon River in Western Australia. Application of different mother wavelet of Haar, Daubechies and Coiflet and different level of decomposition are studied.

Secondly, different CI models are applied for short, mid and long term river flow forecasting. Different input combinations (forward stepwise selection) and signal processing techniques (Coiflet, Haar and Daubechies discrete wavelets) are applied on mean daily, weekly and monthly river flow time series of Ellen Brook River in Western Australia. Preprocessed data are applied as the input of multi-layer back propagation neural networks and adaptive neuro-fuzzy inference system with grid partitioning.

Thirdly, the application of different CI models for forecasting multi-step ahead of daily river flow is studied and improved. Artificial neural networks, adaptive neuro-fuzzy inference system with subtractive clustering and their associated wavelet hybrid models (WNN and WNFS, respectively) are applied for 1, 2, 3, 4 and 5 days ahead forecasting in Harvey River, Western Australia. Daubechies and Symlet wavelets are used to decompose river flow and rainfall time series to different levels.

Finally, developed models are applied to real time river flow forecasting for the purpose of timely flood warning. ANN, ANFIS with grid partitioning and their hybrid models, in conjunction with discrete wavelet transform, are applied for 1, 6, 12, 24, 36 and 48 hour ahead river flow forecasting. Casino gauging station of Richmond River, NSW, Australia, which is highly prone to flooding, is considered as the case study. The accuracy of forecasting is further improved when an upstream river flow data (Wiangaree station), are employed as additional input.

In each case study, optimum structure of different CI models is determined and the best fitted model among all is selected. The outcomes of this study confirm the robustness of CI models in river flow forecasting. Considering highly nonlinear and non-stationary characteristics of river flow time series, wavelet analysis significantly improved forecasting reliability in the proposed hybrid models, especially for longer lead time and higher step ahead forecasting. Moreover, hybrid models are highly outperform classical CI models in forecasting sudden extreme events. The outcome of this study will assist hydrologists and decision makers in forecasting river flows and sustainable planning and management of water resources.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincere gratitude to my supervisor, Dr. Rnajan Sarukkalige, for his continuous guidance and encouragement throughout this study. Without his advice and active support, this project would not have been possible. My appreciation is also extended to my co-supervisor Prof. Amithirigala Jayawardena for his enormous knowledge, invaluable guidance and advice.

I would like to acknowledge Department of Civil Engineering of Curtin University for providing all sorts of support in conducting this research. I also would like to give my special thanks to Prof. Hamid Nikraz, former head of the department and member of thesis committee, for his endless support and encouragement.

Last but not least, I would like to thank my family members, especially my father and brother, my dearest friends, Bahar Baniahmad and Kaveh Espandar, for their permanent encouragement, understanding and generous support during my PhD study. I would like to extend my special thanks to my fiancé, Dr. Meysam Banimahd, for his ongoing support, valuable comments and editing the thesis.

LIST OF PUBLICATION

Articles in Journals

- Badrzadeh, H., Sarukkalige, R., & Jayawardena, A. W. (2013). Impact of multi-resolution analysis of artificial intelligence models inputs on multi-step ahead river flow forecasting. *Journal of Hydrology*, 507(0), 75-85.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2013.10.017>
- Badrzadeh, H., Sarukkalige, R. & Jayawardena, A. W. (2014). Improving ANN-based short and long term seasonal river flow forecasting with signal processing techniques, *River research and application journal*, doi: 10.1002/rra.2865.
- Badrzadeh, H., Sarukkalige, R. & Jayawardena, A. W. (2015). A wavelet neuro-fuzzy computational model for stream flow forecasting, *Nonlinear processes in geophysics*, Under review.

Articles in peer reviewed conference proceedings

- Badrzadeh, H. & Sarukkalige, R. (2012) River flow forecasting using an integrated approach of wavelet analysis and artificial neural networks, in *Proceedings of the 34th Hydrology & Water Resources Symposium*, Nov 19-22 2012, pp. 1571-1578. Sydney, NSW: Engineers Australia.
- Badrzadeh, H., Sarukkalige, R. & Jayawardena, A. W. (2012) Combined wavelet-neural network model for intermittent stream flow prediction, in *Vimonsatit, V. and Singh, A. and Yazdani, S., Research, Development, and Practice in Structural Engineering and Construction*, Nov 28-Dec 2 2012, pp. 769-774. Perth, Western Australia.
- Badrzadeh, H. & Sarukkalige, R. (2014) Improving fuzzy-based model for seasonal river flow forecasting, in *Proceedings of the 35th Hydrology & Water Resources Symposium*, Feb 24-27 2014, pp. 994-1001. Perth, WA: Engineers Australia.

TABLE OF CONTENTS

Declaration	ii
Abstract	iii
Acknowledgements	v
List of Publication	vi
Table of Contents	vii
List of Figures	xi
List of Tables.....	xv
List of Acronyms.....	xvii
List of Notations.....	xix
Chapter 1: Research Overview	1
1.1 Background	1
1.2 Motivation	2
1.3 Thesis objective and scope.....	3
1.4 Structure of the thesis.....	4
Chapter 2: A Review on River Flow Forecasting Methods.....	8
2.1 Introduction	8
2.2 Physically-based models	9
2.3 Conceptual models.....	12
2.4 Data driven models	17
2.4.1 Classical data driven approach.....	18
2.4.2 Computational intelligence approach	18
Chapter 3: Computational Intelligence Approach	21
3.1 Introduction	21
3.2 Artificial Neural Networks.....	21
3.2.1 Introduction	21

3.2.2	Neuron modeling and activation functions	24
3.2.3	Neural network architecture	29
3.2.3.1	Feed-forward multilayer perceptron ANN	29
3.2.4	Neural network learning	31
3.2.4.1	Backpropagation algorithm	31
3.3	Neuro-fuzzy modelling	35
3.3.1	Introduction	35
3.3.2	Fuzzy logic	37
3.3.3	Fuzzy inference systems	40
3.3.4	Adaptive neuro-fuzzy inference system.....	41
3.3.5	Input space partitioning.....	45
3.3.5.1	Grid partitioning	45
3.3.5.2	Scatter partitioning (Clustering)	45
3.4	Wavelet multi-resolution Analysis	49
3.4.1	Introduction	49
3.4.2	Fourier transform.....	51
3.4.3	Short-time fourier transform	51
3.4.4	Continuous wavelet transform	52
3.4.5	Discrete wavelet transform	53
3.4.6	Mother wavelets	54
3.4.7	Time series decomposition by wavelet.....	57
3.5	Summary	59
 Chapter 4: Structure of Proposed Hybrid Models		60
4.1	Introduction	60
4.2	Wavelet neural networks.....	60
4.2.1	Neural networks sub-model	61
4.2.2	Wavelet sub-model	64
4.3	Wavelet neuro-fuzzy with grid partitioning	66
4.3.1	ANFIS sub-model with grid partitioning.....	66
4.3.2	Wavelet sub-model	68
4.4	Wavelet neuro-fuzzy with clustering.....	69
4.4.1	Hybrid wavelet neuro-fuzzy model with subtractive clustering	70

4.4.2 Hybrid wavelet neuro-fuzzy model with fuzzy C-mean clustering.....	71
4.5 Performance criteria.....	73
4.6 Summary	76

Chapter 5: Daily River Flow Forecasting Using Multivariate Inputs 77

5.1 Introduction	77
5.2 Case studies	78
5.3 Application of ANN.....	83
5.4 Application of ANFIS.....	85
5.5 Improving the efficiency with hybrid models	86
5.6 Conclusion.....	93

Chapter 6: Short Term and Long Term River Flow Forecasting... 93

6.1 Introduction	93
6.2 Study area and data used.....	94
6.3 Input selection for models.....	95
6.4 Results and discussion	100
6.4.1 Performance of ANN-based models in river flow forecasting.....	100
6.4.2 Performance of Fuzzy-based models in river flow forecasting.....	111
6.5 Conclusion.....	122

Chapter 7: Multi-Step Ahead River Flow Forecasting..... 123

7.1 Introduction	123
7.2 Study area and data used.....	124
7.3 Results and discussion	125
7.3.1 Application of ANN	125
7.3.2 Improving the efficiency of ANN with WNN	128
7.3.3 Application of ANFIS.....	138
7.3.4 Improving the efficiency of ANFIS with WNF	139
7.3.5 Model comparison	144
7.4 Conclusion.....	148

Chapter 8: Real Time Runoff Flow Forecasting for Flood Risk Management	150
8.1 Introduction	150
8.2 Study area and data used	151
8.3 Analysis and results	154
8.3.1 ANN-based models.....	154
8.3.2 Fuzzy-based models.....	159
8.3.3 Models comparison.....	165
8.4 Conclusion.....	169
 Chapter 9: Conclusions and Future Work.....	170
9.1 Conclusion.....	170
9.2 Recommendation for future works	173
 References	175

LIST OF FIGURES

Figure 1. 1	Main structure of the thesis.....	7
Figure 2. 1	Schematic of MIKE SHE distributed model structure (Graham and Butts, 2005)..	11
Figure 2. 2	Schematic of storage system in conceptual model.....	13
Figure 2. 3	Learning system in black box data driven approaches.....	17
Figure 3. 1	Schematic of a neuron function system.	25
Figure 3. 2	Sigmoid activation function with different steepness parameter.....	26
Figure 3. 3	Tan-Sigmoid activation function with different steepness parameter.....	26
Figure 3. 4	Schematic diagram of a three-layer feed-forward neural network.	30
Figure 3. 5	Structure of neural networks with back propagation training algorithm.....	32
Figure 3. 6	Levenberg-Marquardt algorithm shifts from the steepest descent to the Gauss-Newton method of decreasing the value of μ	34
Figure 3. 7	The basic structure of FIS.....	40
Figure 3. 8	Reasoning mechanism of a Sugeno fuzzy model with two inputs and rules.	43
Figure 3. 9	Equivalent ANFIS architecture for two-input first-order TSK fuzzy model with two rules.....	43
Figure 3. 10	Grid partitioning of two inputs into 9 fuzzy rules.....	46
Figure 3. 11	(a) Shifted and (b) Scaled wavelet illustration.	53
Figure 3. 12	Daubechies family wavelet.....	55
Figure 3. 13	Haar mother wavelet function.....	56
Figure 3. 14	Mexican Hat, Morlet, Coiflet1 and Symlet2 mother wavelets.	56
Figure 3. 15	Diagram of multi-resolution analysis of signal.....	57
Figure 3. 16	Two noisy signals and their (a) Daubechies3 (b) Coiflet1 wavelet coefficients.....	58
Figure 4. 1	Structure of the proposed hybrid WNN model for N step ahead forecasting.....	61
Figure 4. 2	Feed-forward neural network sub-model flow chart.....	63
Figure 4. 3	hybrid wavelet neural networks model flow chart.....	65
Figure 4. 4	Structure of Wavelet Neuro-Fuzzy hybrid model for N step ahead forecasting.	66
Figure 4. 5	Adaptive neuro-fuzzy with grid partitioning sub-model flow chart.....	67
Figure 4. 6	Flow chart of hybrid wavelet neuro-fuzzy model with grid partitioning.	69
Figure 4. 7	Flow chart of hybrid wavelet neuro-fuzzy model with subtractive clustering.....	71
Figure 4. 8	Flow chart of hybrid wavelet neuro-fuzzy model with FCM clustering.....	72
Figure 5. 1	Map of Northam weir and Dingo road station location in Western Australia (Bureau of meteorology, 2013)	80
Figure 5. 2	Daily river flow time series at the Dingo road station in the Harvey River, Western Australia (1976-2011).....	81
Figure 5. 3	Daily rainfall time series near the Dingo road station in the Harvey River, Western Australia (1976-2011).....	81

Figure 5. 4 Daily river flow time series at the Northam weir station in the Avon River, Western Australia (1978-2010).....	82
Figure 5. 5 Daily rainfall time series near the Northam weir station in the Avon River, Western Australia (1978-2010).....	82
Figure 5. 6 (a) Harvy and (b) Avon River mean daily river flow hydrographs in selected years.	83
Figure 5. 7 Scatter plots between Dingo road station observed and modelled daily river flow: (a) ANN single flow input; (b) ANN with multivariate input.	85
Figure 5. 8 Scatter plots between Northam weir station observed and modelled daily river flow: (a) ANN with single input; (b) ANN with multivariate input.....	85
Figure 5. 9 Scatter plots between Dingo road station observed and modelled daily river flow with: (a) ANN1; (b) Hybrid WNN12; (c) ANFIS1; (d) Hybrid WNFC4.....	89
Figure 5. 10 Comparison of the Dingo road observed and predicted daily river flow with WNN12.....	91
Figure 5. 11 Comparison of the Dingo road observed and predicted daily river flow with WNN12 in the validation set (2007-2011).	91
Figure 5. 12 Comparison of the Northam weir observed and predicted daily river flow with WNN9	92
Figure 5. 13 of the Northam weir observed and predicted daily river flow with WNN9 in the validation set (2006-2010).	92
Figure 6. 1 Location of Ellen Brook catchment in the Western Australia.....	94
Figure 6. 2 (a) Daily; (b) Weekly and (c) Monthly river flow time series at the Railway Parade station on the Ellen Brook River, Western Australia (1977-2010).....	97
Figure 6. 3 ACF of Ellen Brook River daily, weekly and monthly flow time series (1977-2010).....	98
Figure 6. 4 Nash-Sutcliffe coefficient of efficiency of (a) training and (b) validation set, for different BPNN and WNN models.....	104
Figure 6. 5 Ellen Brook weekly river flow time series and its wavelet coefficients with Coif1 wavelet.	105
Figure 6. 6 Scatter plots of observed and forecasted river flow with the best fitted BPNN and WNN models for daily, weekly and monthly forecasting.	106
Figure 6. 7 Comparing observed versus modeled monthly river flow with best fitted BPNN model.....	108
Figure 6. 8 Comparing observed versus modeled monthly river flow with best fitted WNN model.....	108
Figure 6. 9 Comparing observed versus modeled weekly river flow with best fitted BPNN model.....	109
Figure 6. 10 Comparing observed versus modeled weekly river flow with best fitted WNN model.....	109
Figure 6. 11 Best fitted hybrid neuro-fuzzy model (WNFG-M2) structure for monthly forecasting.....	111

Figure 6. 12 Nash-Sutcliffe coefficient of efficiency of (a) training and (b) validation set, for different ANFIS and WNFG models.	114
Figure 6. 13 Ellen Brook daily river flow signal and its wavelet coefficients with db5 wavelet.	116
Figure 6. 14 Scatter plots of observed and forecasted river flow with the best fitted ANFIS and WNFG models for daily, weekly and monthly forecasting.	117
Figure 6. 15 Comparing observed versus modeled weekly river flow with best fitted ANFIS model.....	120
Figure 6. 16 Comparing observed versus modeled weekly river flow with best fitted WNFG model.....	120
Figure 6. 17 Comparing observed versus modeled monthly river flow with best fitted ANFIS model.....	121
Figure 6. 18 Comparing observed versus modeled monthly river flow with best fitted WNFG model.....	121
Figure 7. 1 ACF of Harvey River daily flow and rainfall time seires.....	124
Figure 7. 2 Scatter plots of observed and ANN forecasted flow for different lead time.	128
Figure 7. 3 Daily river flow time series and its db5 wavelet coefficients with four level of resolution.	129
Figure 7. 4 Different hybrid WNN model efficiency for different lead-time (L) in training and validation set.	135
Figure 7. 5 Best fitted ANN and WNN model efficiency (RMSE) variation over the lead time in (a) training; (b) verification set.....	136
Figure 7. 6 Comparison of the observed and modeled river flow for 5-day ahead with ANN5-5 and WNND5-14 models (1972-2011).	137
Figure 7. 7 Scatter plots of observed and ANFIS forecasted flow for different lead time.	139
Figure 7. 8 Hybrid WNF1-1 model structure, generated with subtractive clustering approach.	140
Figure 7. 9 Different hybrid WNF model efficiency for different lead-time (L) in training and validation set.....	142
Figure 7. 10 ANFIS and WNF models' efficiency (RMSE) variation over the lead time in (a) training; (b) verification set.	142
Figure 7. 11 Comparison of the observed and modeled river flow for 5-day ahead with ANFIS5-2 and WNF5-2 models (1972-2011).....	143
Figure 7. 12 Variation of different models' performance (R^2) over the lead time.	144
Figure 7. 13 Comparison of the observed and modeled river flow for 5-day ahead with best fitted WNN and WNF models (2006-2011).....	145
Figure 7. 14 Scatter plots of five-day ahead forecasting of the best fit ANN, WNN, ANFIS and WNF model.	147
Figure 7. 15 Highest observed river flow in historical time series and its estimation with developed models.	148

Figure 8. 1 Richmond River catchment and its flood plain (http://australiasevereweather.com/floods , Bath, 2014).....	152
Figure 8. 2 Hourly river flow and rainfall time series of Casino station for (a) training and (b) validation set.	153
Figure 8. 3 Hourly river flow and rainfall time series of Wiangaree station for (a) training and (b) validation set.	153
Figure 8. 4 Nash-Sutcliffe coefficient of (a) training and (b) validation set of ANN-based models.	157
Figure 8. 5 Root mean square error of (a) training and (b) validation set of ANN-based models.	157
Figure 8. 6 Scatter plots of observed and simulated river flow for 24 hour lead time with different ANN-based models.	158
Figure 8. 7 Comparing observed flood versus ANN-based modeled hourly river flow (24 hour ahead forecasts).	159
Figure 8. 8 Comparing four highest observed peak flow versus ANN-based modeled values (24 hour ahead forecasts).	159
Figure 8. 9 Nash-Sutcliffe coefficient of (a) training (b) validation set of fuzzy-based models.	163
Figure 8. 10 Root mean square error of (a) training and (b) validation set of fuzzy-based models.	163
Figure 8. 11 Scatter plots of observed and simulated river flow for 24 hour lead time with different fuzzy-based models.	164
Figure 8. 12 Comparing observed flood versus fuzzy-based modeled hourly river flow (24 hour ahead forecasts).	165
Figure 8. 13 Comparing four highest observed peak flow versus fuzzy-based modeled values (24 hour ahead forecasts).	165
Figure 8. 14 Scatter plots of observed and simulated river flow for 12 hour lead time with best fitted ANN, WNN, ANFIS and WNF models.	167
Figure 8. 15 Scatter plots of observed and simulated river flow for 24 hour lead time with best fitted ANN, WNN, ANFIS and WNF models.	167
Figure 8. 16 Scatter plots of observed and simulated river flow for 36 hour lead time with best fitted ANN, WNN, ANFIS and WNF models.	168
Figure 8. 17 Scatter plots of observed and simulated river flow for 48 hour lead time with best fitted ANN, WNN, ANFIS and WNF models.	168

LIST OF TABLES

Table 3. 1 Different types of linear activation functions.	28
Table 3. 2 Triangular, Trapezoidal, Gaussian, Generalised bell and Sigmoidal membership functions.....	39
Table 5. 1 Daily statistical parameters of stream flow and rainfall data sets of the Dingo road and Northam weir stations.	79
Table 5. 2 ANN models structure and performance.....	84
Table 5. 3 ANFIS models' performance.	86
Table 5. 4 Hybrid models' structure and performance for Dingo road station daily river flow forecast.....	87
Table 5. 5 Hybrid models' structure and performance for Northam weir station daily river flow forecast.....	88
Table 6. 1 Statistical parameters of Ellen Brook river flow data sets of the Railway parade station.....	95
Table 6. 2 Input selection for different BPNN models.....	98
Table 6. 3 Input selection for different ANFIS models.....	99
Table 6. 4 Input pre-processing type for hybrid WNN models.	100
Table 6. 5 Input pre-processing type for hybrid WNFG models.....	100
Table 6. 6 BPNN and WNN models structure and performance for daily river flow forecasting.	102
Table 6. 7 BPNN and WNN models structure and performance for weekly flow forecasting..	102
Table 6. 8 BPNN and WNN models structure and performance for monthly flow forecasting.	103
Table 6. 9 Accuracy of developed ANN-based models in simulating daily, weekly and monthly extreme flow values.....	110
Table 6. 10 ANFIS and WNFG models structure and performance for daily river flow forecasting.....	112
Table 6. 11 ANFIS and WNFG models structure and performance for weekly river flow forecasting.....	113
Table 6. 12 ANFIS and WNFG models structure and performance for monthly river flow forecasting.....	113
Table 6. 13 Accuracy of developed Fuzzy-based models in simulating daily, weekly and monthly extreme flow values.....	119
Table 7. 1 ACF of Harvey River daily flow and rainfall time series (1972-2011).	124
Table 7. 2 Different input combinations.	125
Table 7. 3 ANN structure and performance for different lead time.	126
Table 7. 4 Different WNN model's structure and performance for 1 day ahead lead-time.....	130
Table 7. 5 Different WNN model's structure and performance for 2 day ahead lead-time.....	131

Table 7. 6	Different WNN model's structure and performance for 3 day ahead lead-time.....	132
Table 7. 7	Different WNN model's structure and performance for 4 day ahead lead-time.....	133
Table 7. 8	Different WNN model's structure and performance for 5 day ahead lead-time.....	134
Table 7. 9	ANFIS performance for different lead time.	138
Table 7. 10	WNF models' structure and performance for different lead time.....	141
Table 7. 11	Best fitted models performances for different lead time.	146
Table 7. 12	Accuracy of different model in simulating five-day ahead extreme flow values.	147
Table 8. 1	Statistical parameters of Richmond River flow data sets of the Casino and Wiangaree stations.....	152
Table 8. 2	ANN models structure and performance using only Casino station data.	154
Table 8. 3	ANN models structure and performance using Casino and Wiangaree stations data.	154
Table 8. 4	Hybrid WNN models structure and performance using only Casino station data. .	155
Table 8. 5	Hybrid models structure and performance using Casino and Wiangaree stations data.	156
Table 8. 6	ANFIS models structure and performance using only Casino station data.	160
Table 8. 7	ANFIS models structure and performance using Casino and Wiangaree stations data.	161
Table 8. 8	WNF models structure and performance using only Casino station data.....	162
Table 8. 9	WNF models structure and performance using Casino and Wiangaree stations data.	162

LIST OF ACRONYMS

A	Approximation
ACF	Autocorrelation function
ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural networks
AR	Autoregressive
ARIMA	Autoregressive integrated moving average
ARMA	Autoregressive moving average
ARX	Autoregressive exogenous
B	Bias
BP	Back propagation algorithm
BPF	Back propagation feed-forward
BPNN	Back propagation neural networks
CI	Computational intelligence
Coif	Coiflets wavelet
CWT	Continuous wavelet transform
D	Details
d	Agreement index
db	Daubechies wavelet
DWT	Discrete wavelet transform
EPA	Environmental protection agency
FCM	Fuzzy C-mean clustering
FFNN	Feed-forward neural networks
FIS	Fuzzy inference systems
FL	Fuzzy logic
FT	Fourier transform
HSPF	Hydrologic simulation program FORTRAN
I	Identity matrix
IHDM	Institute of hydrology distributed model
LM	Levenberg-Marquardt algorithm
LSE	Least square error
MA	Moving average
MAE	Mean absolute error
MF	Membership function

MFIS	Mamdani fuzzy inference systems
MLP	Multilayer perceptron
MLR	Multiple linear regression
MRA	Multi-resolution analysis
MSE	Mean square error
NSE	Nash-Sutcliffe coefficient of efficiency
PB	Percent bias
PE	Potential error
R	Coefficient of correlation
R ²	Coefficient of determination
RBF	Radial basis function
RMAE	Relative mean absolute error
RMSE	Root mean square error
R-R	Rainfall-runoff
RSR	Root mean square error-standard deviation ration
SAC-SMA	Sacramento soil moisture accounting
SHE	Système Hydrologique Européenne (European Hydrological System)
SMAR	Soil moisture accounting and routing
SMSC	Soil moisture storage capacity
SOFM	Self organizing feature map
SOM	Self organized map
STDV	Standard deviation
STFT	Short-time Fourier transform
SWM	Standford watershed model
SWMM	Storm water management hydrological model
SWWA	South West of Western Australia
Sym	Symlet wavelet
TAR	Threshold autoregressive
TSK	Takagi-Sugeno-Kang fuzzy
Var	Variance
WNF	Wavelet neuro-fuzzy
WNFC	Wavelet neuro-fuzzy with C-mean clustering
WNFG	Wavelet neuro-fuzzy with grid partitioning
WNFS	Wavelet neuro-fuzzy with subtractive clustering
WNN	Wavelet neural networks

LIST OF NOTATIONS

α	Weighting exponent
ε_t	White noise
φ_p	Autoregressive coefficient (order p)
$\Psi_{(\alpha,\beta)}$	Mother wavelet
θ_q	Moving average coefficient (order q)
$\mu_R(x)$	Membership of an element x
∇V	Gradient
$\nabla^2 V(w)$	Hessian matrix
a_i, b_i, c_i	Premise parameters
p_i, q_i, r_i	Consequent parameters (polynomial parameters)
b_j	Threshold value (bias) of j^{th} neuron
C	Number of clusters
D_i	Density measure of i^{th} cluster centre
e_n	Model error in n^{th} iteration
f_h	Hidden layer activation function
f_o	Output layer activation function
I_t	Stream flow time series
$J(w)$	Jacobian matrix
K	Number of inputs
m	Number of membership function
m_{ij}	Membership degree of j^{th} variable in data set for i^{th} cluster
N	Length of time series
$O_{k,i}$	Output of i^{th} node in k^{th} layer
Q_t	River flow time series
Q_{dt}	Daily river flow time series
Q_w	Weekly river flow time series
Q_{mt}	Monthly river flow time series
Q_{sim}	Simulated river flow
Q_{obs}	Observed river flow
Q_{DWT}	Wavelet coefficient of river flow time series
r	The steepness parameter

r_b	Positive constant (1.5 * Neighbourhood radius)
R_t	Rainfall time series
R_{DWT}	Wavelet coefficient of rainfall time series
v_i	Cluster centre of the i^{th} cluster
$W(\alpha, \beta)$	Scaled (α -scaled) and shifted (β -translation) wavelet coefficients
$W(\alpha, \beta)^2$	Wavelet spectrum
W_{ij}	Connection weight from the i^{th} neuron to the j^{th} neuron
$W(t)$	Window function
x	An input vector
x_n	Normalized value of x
y^l	Output of the l^{th} layer

Chapter 1

Research Overview

1.1 BACKGROUND

Water demands are increasingly growing due to population growth and irrigation and industrial developments. Surface water availability is likely to decrease as a reason of global warming, urbanizations and excessive groundwater extractions. On the other hand, in various regions around the world, extreme weather conditions resulting floods, droughts and heat waves.

Understandably, reliable information on current and future water availability is essential to properly manage the limited water resources and flood mitigation. Authorities in water sector cannot allocate water resources optimally for water demands like agricultural, industrial, domestic, hydropower generation and environmental maintenance, unless they are equipped with a reliable forecasting of river flow. Accurate forecasting of river flow, as the main part of the available water resources, is also a key element in drought analysis and design of water related infrastructures. Therefore, improving the accuracy and reliability of river flow forecasting is an ongoing research. Researchers are keen to develop and investigate various types of hydrological forecasting approaches to attain better management of scarce water resources and minimize the risk of any potential flooding.

1.2 MOTIVATION

Like many countries around the world, Australia is increasingly facing water scarcity. Many parts of Australia continent are in drought. South Western Australia, in particular, is suffering from extended dry period since 1975. Climate change projections for Southern part of the continent, more populated part of Australia, indicates reduction in total rainfall and water supply (Charters and Williams, 2006). At the same time, climate change causes more frequent rainfall events with higher intensity which increases the risk of flooding (Bates et al., 2008). Undoubtedly, effective water governance policies will become critical to cope with water crises.

Forecasting future surface water availability is a key element in assisting decision makers in water resources planning and management (Nash and Sutcliffe, 1970; Nayak et al., 2005; Sene, 2010; Piotrowski and Napiorkowski, 2011; Zeng et al., 2012). Forecasting water availability is always associated with large uncertainties and complexities. For example, determining the rate of runoff generated by rainfall and its routing is a very complex matter as an extensive number of parameters are involved in this process. A significant amount of research has been carried out to improve the accuracy of forecasting ranging from physically-based to data-driven approaches. Therefore, improving the accuracy of forecasting is a continuing research field as each hydrological forecasting approach has its own characteristics and limitations. This study focuses on developing river flow forecasting model with minimum parameter requirements (ungauged catchments) and maximum accuracy for long term as well as extreme event forecasting. For this research, computational intelligence (CI) approach is selected because of their cost-efficiency, accuracy and robustness.

In order to boost the forecasting performance, some hybrid approaches have been proposed recently (Sivakumar and Berndtsson, 2010). One of the recent developments in river flow forecasting is based on coupling computational intelligence models with wavelet analysis. River flow historical data are non-stationary time series with a wide range of frequency components. By applying wavelet transform, river flow complex time series can be decomposed into its major sub-components (Zhou et al., 2008). However, a comprehensive literature review

confirms the lack of research on wavelet neuro-fuzzy techniques with a subtractive clustering method for river flow simulation and forecasting. Furthermore, only a limited number of river discharge time series have been used for verification of the wavelet neural network based models. More data from different areas with different characteristics would be required to conclusively prove the advantages of this hybrid approach (Wei et al., 2012). Available research in the literature mainly focuses on the forecasting river flow by using only river flow discharge time series. Forecasting could be improved by adding other hydrological time series and variables which affect river flow (Adamowski and Sun, 2010; Pramanid et al., 2011). In addition, very few researchers explored the application of hybrid models on seasonal river flow forecasting and lead times of more than one day, but less than one month (Wu et al., 2009; Nournani et al., 2013).

Taking these considerations into account, this study aims at improving seasonal, short term, long term and real time river flow forecasting by various classical and hybrid computational intelligence approaches with different structures and input selections. By providing more accurate tools, the ultimate scope of this research is to assist decision makers in sustainable water resources planning, flood protection, mitigation of contamination or licensing of exploitations.

1.3 THESIS OBJECTIVE AND SCOPE

The main purpose of this research is to develop highly efficient, reliable and accurate data driven model for river flow forecasting. Each of the forecasting approaches has its own advantages and disadvantages and there is no perfect model or modelling technique to guarantee precise future long term prediction. Reviewing current available river flow forecasting and rainfall-runoff methods, computational intelligence techniques were found as a powerful approach for modelling complex hydrological process. In this study different type and structure of artificial neural networks (ANN), adaptive neuro-fuzzy inference system (ANFIS) and hybrid wavelet models will be developed. Comparing the performance of models, the best

fitted model for reaching the most accurate results in different study areas will be determined. In summary, the main objectives of this study are;

- Developing highly efficient model for accurate river flow forecasting by investigating and comparing the performance of artificial neural networks and adaptive neuro-fuzzy inference system approaches.
- Applying different methods for initiating fuzzy inference system (FIS) structure in ANFIS modelling, including grid partitioning, subtractive clustering and C-mean clustering (FCM).
- Finding the optimum structure and most effective training algorithm of neural network for river flow forecasting.
- Investigating the impact of wavelet multi-resolution analysis of CI model inputs on forecasting accuracy. Explore the performance of hybrid wavelet models by decomposing data series into the low and high frequency signals with different type of discrete wavelet transforms and into different level of decomposition.
- Developing and validating different computational intelligence techniques for real time, short term, long term and multi-step ahead prediction of stream flow. Also, determining and validating the best fitted CI model structure for seasonal river flow forecasting.

1.4 STRUCTURE OF THE THESIS

This thesis is designed in four main parts of introduction, methodology, applications and conclusion which are expanded in 9 chapters. Figure 1.1 depicts the structure of the thesis. Following is a brief description of each chapter;

Chapter One- This chapter mainly identifies the problem and the main reasons of conducting this research. It highlights the core objectives of the study and thesis outlines.

Chapter Two- This chapter introduces various types of river flow forecasting and rainfall-runoff models. Physically-based, conceptual and data-driven approaches are reviewed. Methodologies behind most popular models are briefly explained. Advantages and drawbacks of different approaches are identified.

Chapter Three- Three different CI approaches, namely, artificial neural networks, fuzzy modelling and wavelet analysis (as a part of hybrid models) are discussed in details in this chapter. The structure of feed-forward neural networks with back propagation training algorithm, adaptive neuro-fuzzy inference system with grid partitioning, subtractive and C-mean clustering is described. The application of wavelet multi-resolution analysis in signal decomposition is also presented. In addition to theoretical description of approaches, a review on their background and applications in hydrology is also provided.

Chapter Four- This chapter presents the structure of developed models, including four hybrid models of wavelet neural networks, wavelet neuro-fuzzy with grid partitioning, subtractive clustering and C-mean clustering. Optimum performance criteria are also selected for achieving most efficient models, especially for extreme event forecasting.

Chapter Five- In this chapter, application of developed models with multivariate inputs for daily river flow forecasting is investigated. Rainfall time series are added as an additional input. Two different rivers from Western Australia (Harvey and Avon Rivers) are selected as case studies. The best structure of ANN and ANFIS with C-mean clustering, alone and in conjunction with Daubechies and Haar mother wavelet (WNN, WNFC), are determined. The effect of adding an additional input is discussed.

Chapter Six- This chapter investigates the application of developed models in both short and long term river flow forecasting. Different input combinations (forward stepwise selection) and signal processing techniques (Coiflet, Haar and Daubechies mother wavelets) are applied on multi-layer back propagation neural networks (WNN) and adaptive neuro-fuzzy inference system with grid partitioning (WNFG). The data of the Railway parade station on Ellen Brook River, Western Australia, is used as a case study. Daily, weekly and monthly river flow forecasting is conducted. The impacts of right selection of the inputs and pre-processing the raw data with wavelet are showcased in this section.

Chapter Seven- In this chapter the accuracy of multi step ahead daily river flow forecasting is improved by applying Daubechies and Symlet multi-resolution analysis on ANN and ANFIS models' input. A novel approach of hybrid wavelet neuro-fuzzy with subtractive clustering is introduced for river flow forecasting. Overall 215 different models for various lead-times of 1 to 5 days ahead, with different input combinations (forward stepwise time series, multivariate input and wavelet coefficients) were developed for forecasting daily river flow of the Dingo road station on Harvey River, Western Australia. Highly satisfactory results achieved as the forecasting accuracy significantly improved for longer lead time and extreme event simulation.

Chapter Eight- In this chapter the application of developed models for timely flood warning is investigated. Feed-forward ANN, adaptive neuro-fuzzy with grid partitioning alone and in conjunction with Daubechies discrete wavelet transform (db3) are applied for forecasting 1, 6, 12, 24, 36 and 48 hour ahead of river flow. Hourly rainfall and river flow data of two stations on Richmond River in NSW, Australia, which is highly prone to flooding, are used. Highly reliable results are achieved for forecasting up to 24 hour ahead of flooding event, especially when an upstream flow time series added as the model input.

Chapter Nine- Summary of research outcomes and general conclusions are presented in this chapter. The recommendations for future studies are also provided in this chapter.

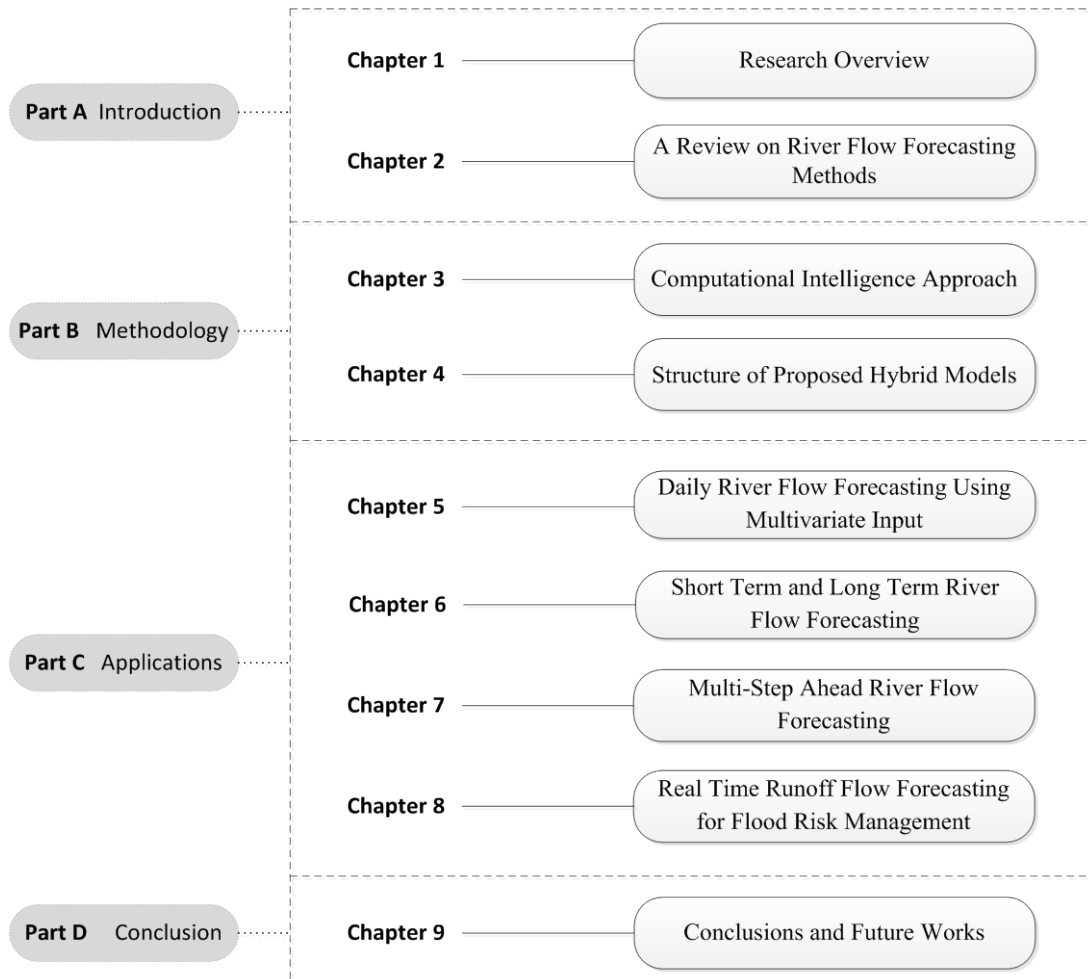


Figure 1. 1 Main structure of the thesis

Chapter 2

A Review on River Flow Forecasting Methods

2.1 INTRODUCTION

As a consequence of issues like water increased demands and climate change, the need for accurate river flow forecasting has grown rapidly in the past decades. Knowing future conditions of surface water resources is one of the key elements for an appropriate risk-based and sustainable water resources planning.

The application of river flow forecasting could be categorized into two main types. The first application is short term river flow forecasting to predict sudden extreme conditions such as flooding (Werner, et al., 2005; O'Connor, 2006). Being prepared a day or even a few hours before such an event could assist hazard adaptation which can reduce costs and save lives (Carpernter, et al., 1999). The second application is long term forecasting for the purpose of sustainable water resources management. Knowing the quantity of future surface water resources is required for determining optimum reservoir operations, irrigation allocations, groundwater extraction regulation and demands supply planning (Valenca, et al., 2005; Ghanbarpour et al, 2009; Sudheer, et al., 2014).

There are various types of river flow forecasting and rainfall-runoff (R-R) techniques ranging from deterministic to stochastic models (Clarke, 1973). The oldest and still the most widely-used rainfall-runoff approach is based on the rational formula (Mulaney, 1845), which estimates runoff rate from rainfall intensity and the catchment area. Technological advances have made a significant impact on

hydrology science in the last centuries. A growing number of scientific theories and mathematical techniques have been developed for measurements, modelling and forecasting of hydrological phenomena. Selecting the best approach for forecasting depends on the purpose of the modelling and available historical spatial and temporal data in the river catchment to simulate complex non-linear hydrological process. In general, there are three main types of forecasting models, namely, physically-based, conceptual and data-driven models (Dawson and Wilby, 2001; Sene, 2010). The following sections provide a brief introduction to different types of river flow forecasting models.

2.2 PHYSICALLY-BASED MODELS

Physically-based models, known also as “distributed” or “deterministic” models, simulate the complex hydrological process in the catchment mathematically. These models consist of nonlinear partial differential equations which spatially represent the physical process of runoff generation in a catchment. They improve our understanding of hydrological system by representing interaction of the spatial-temporal variables. The drawback of deterministic models is that they are very costly and time consuming (Chau, et al., 2005). They require a large amount of data, such as catchment characteristics and meteorological parameters to represent sub-surface and surface runoff generation and routing. For solving of the complex equations of the hydrological process, numerical solutions like finite element, finite difference, boundary integral and integral finite difference must be implemented (Gosain, et al., 2009).

Several physically-based distributed models have been developed and applied in hydrological forecasting. One of the pioneering physically-based models is European Hydrological System - *Système Hydrologique Européenne* (SHE). SHE has been developed by three European institutions, namely SOGREAH (France), Danish hydraulic institute and UK institute of hydrology (Beven, et al., 1980). SHE is a distributed physically-based model which simulates water movement in the hydrological cycle by applying a grid-based finite difference method. Partial equations of mass, energy conservation or momentum are derived based on the

spatially distributed data of catchment parameters, precipitations and catchment hydrological response in the orthogonal grid network (Abbott et al., 1986). Catchment parameters are assumed constant within each grid but could be different from other grids. Based on SHE model, an integrated hydrological modelling system of MIKE SHE has been further developed by DHI water and environment (Refsgaard and Storm, 1995). MIKE SHE represents hydrological process, including evapotranspiration, surface flow, unsaturated flow, sub surface, channel flow and their interactions (Butts et al., 2004). Figure 2.1 illustrates the schematic of MIKE SHE model and its numerical solutions for different hydrological process.

Another well known physically-based model is the Institute of Hydrology Distributed Model (IHDM) (Beven, 1985). This model uses two-dimensional finite element approach. Compared to SHE model, it needs less computational time and parameters as it does not forecast the hydrological response of every point in the catchment. Another example of such simplified model is the popular TOPMODEL (Beven and Kirkby, 1979). This model assumes that the hydraulic gradient of subsurface saturated zone is similar to the local surface slope. It also considers similar hydrological response for the points with same topographic index and thereby eliminates the need for calculations in every point of the watershed. This model also minimizes the number of parameters by simplifying surface flow and unsaturated zone routing algorithms. O'Connor (2006) argues that these kinds of model are not truly physically-based model as they actually apply conceptual model to each grid of the watershed. Many other physically-based models have been developed and applied in various case studies. Some of the most widespread among all are as follows;

ECOMAG model is developed by Motovilov et al. (1999) and consists of hydrological, geochemical and biological process in daily time scale. HYDROTE distributed model is developed in 2001 (Fortin et al., 2001a, b). This model is GIS compatible and its hydrological unit is a small vertical homogenous unit. Downer and Ogden (2004) are developed fully distributed GSSHA model by improving the older two-dimensional model of CASC2D (Julien and Sagharian, 1991). The main improvement was in discharge prediction, when runoff is not produced by Hortonian process. In 2004, MODHMS model with the ability of three-dimensional subsurface

modelling and two-dimensional surface modelling was developed (Panday and Huyakorn, 2004). This model is capable of simulating complex surface and groundwater interactions (Donn et al., 2012).

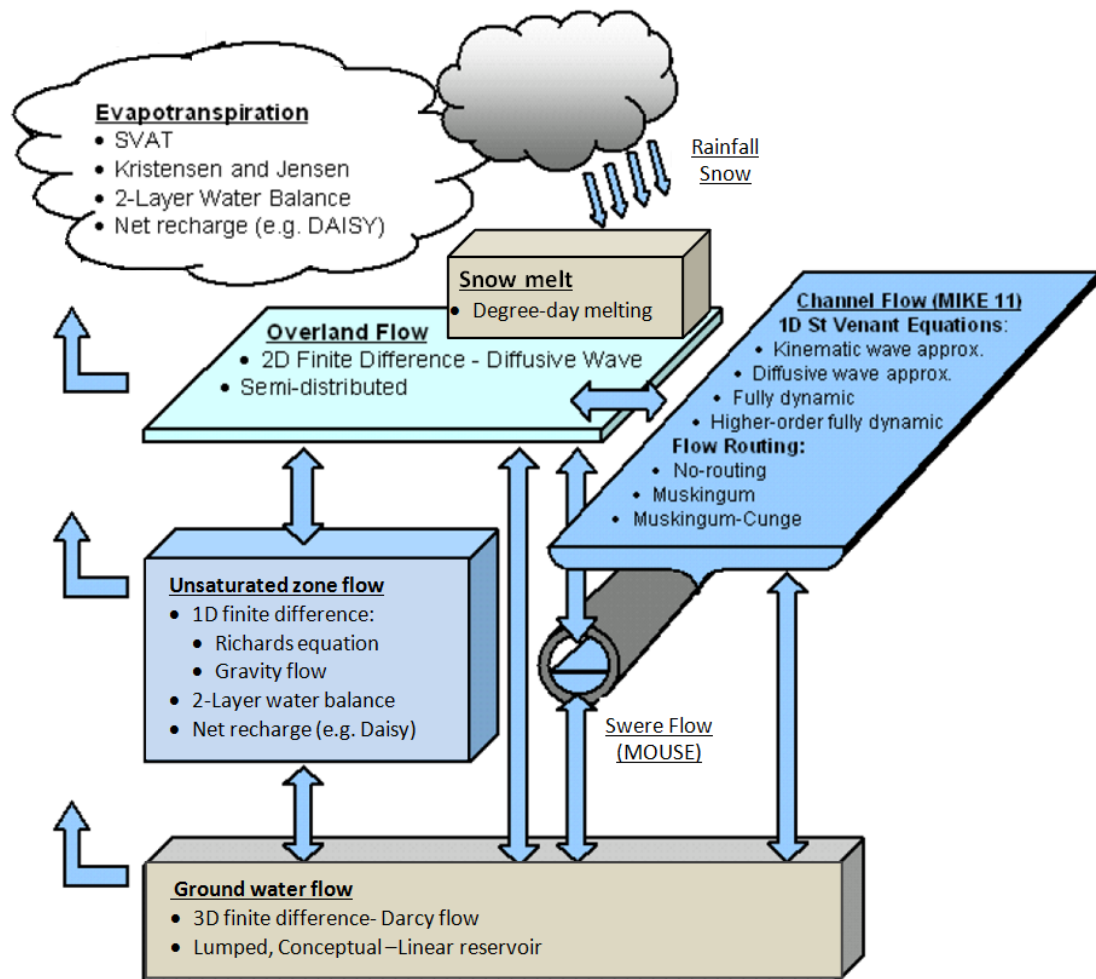


Figure 2. 1 Schematic of MIKE SHE distributed model structure (Graham and Butts, 2005).

Although physically-based models are more sophisticated than the other types of models, they are not applicable and accurate enough for flood forecasting due their complexity and extensive data demands. The main drawbacks of the physically-based models are as follows;

- They are not the exact representation of the hydrological process as it is very difficult to measure and understand catchment parameters such as soil parameters and determine their variation over the time (Liu, et al., 2011).
- There are difficulties in solving catchment descriptive equations. Even applying various available numerical techniques may not lead to convergence of solutions due to complexity of nonlinear partial differential equations.
- They are not cost-effective. Considerable costs are involved in setting up these models including measuring an extensive set of parameters from the field, appropriate softwares and training time.
- They are not suitable for large catchments due to their high-resolution data requirement.
- The accuracy of the model depends on grid size. Most of hydrological data are measured in points and could be homogenous in small scale while grid scale often covers a much bigger area.
- Due to time-consuming nature of complex numerical simulations, physically-based models may not be suitable for real-time flood forecasting.
- Physically-based forecasts are subject to high level of uncertainty as there are many possible sources of error in calibrating the model (Huang and Liang, 2006).

In conclusion, physically-based models can be considered as a powerful tool for providing spatial information of the hydrological parameters within the catchment. Their outcomes would be beneficial for solving many water management problems such as assessing water storage within the catchment rather than river flow forecasting (O'Connor, 2006).

2.3 CONCEPTUAL MODELS

Conceptual models, also called gray-box models, are process-based models too. They formulate physical process of hydrological cycle by most influential elements like

rainfall, evaporation losses and the soil moisture. In fact, they are simplified representation of the hydrologic system. Conceptual R-R models predominantly consist of a number of linked conceptual store buckets and the mathematical relationship between these storages (also called reservoirs) in order to maintain mass balance. Figure 2.2 illustrates the schematic of the typical conceptual storage and the way they are connected to each other in the hydrological cycle.

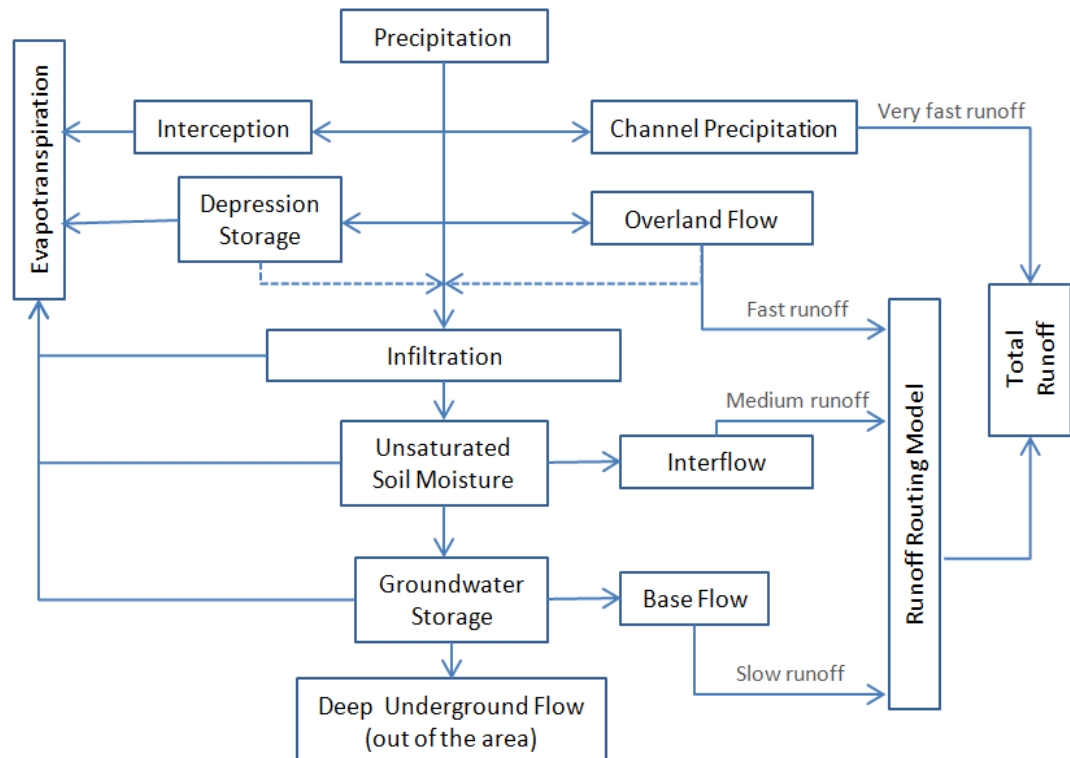


Figure 2. 2 Schematic of storage system in conceptual model

Based on the simulation duration, conceptual models can be classified into event-based or continuous models (Jayawardena, 2014). Event-based model simulates only one single rainfall-runoff event by given initial conditions, while continuous model covers extended period of time (Berthet, et al., 2009). Furthermore, conceptual models can be categorized to lumped and semi-distributed models (Todini, 1988). Most of the conceptual models are lumped, which catchment is considered as a single uniform unit (Refsgaard, 1997). Instead of incorporating the spatial variation of hydrological, hydrogeological and meteorological parameters, their average value will be employed in an input-output system.

Following is a brief overview of most widely used conceptual models.

Stanford Watershed Model (U.S.A) - Stanford watershed model (SWM) is one of the earliest conceptual models, developed in Stanford University (Linsley and Crawford, 1960). SWM is a lumped model which is capable of continuous simulating of runoff based on the continuity equation, using daily and hourly precipitation. In 1966, the basic SWM model is further improved (SWM- IV) by adding more parameters and routing techniques (Crawford and Linsley, 1966). This model requires up to 35 parameters for calibrating modelled evapotranspiration, infiltration, interception, overland and inter flow. Adding components of water quality, concept of SWM model transformed into wide spread Hydrologic Simulation Program FORTRAN (HSPF) model by a US environmental protection agency (EPA) and documented by Johanson et al. (1980).

Tank Model (Japan) - Tank model is another pioneering conceptual model developed by Sugawar (1961). Tank model is a simple lumped, continuous model, consist of four storage tanks, laid in vertically parallel series. The top tank is fed by precipitation and has a side outlet which corresponding surface runoff and a bottom outlet lead into the next tank, representing the infiltration. Evaporation is first subtracted from this tank and then from other tanks in downward order. Second and third tanks have similar outlets which their side outlets provide intermediate and sub-base runoff, respectively. The last tank has only the side outlet providing base flow. Total runoff would be the sum of all these runoff. The top tanks can have two side outlets for modelling the flood. For calibrating the model, a set of outlets and storages coefficients need to be determined. Despite model simple structure, the behaviour of model is highly dependent on storage conditions and similar precipitation may lead to a significantly different runoff (Podger, 2004). The tank model simulates R-R process in daily scale.

SMAR (Ireland) - The soil moisture accounting and routing (SMAR) is a daily lumped model, introduced by O'Connell et al. (1970). The first version of SMAR, which is also known as Layers Model, extensively improved during years of testing (Kachroo, 1992; Tuteja and Cunnane, 1999). SMAR model divides the soil to different horizontal layers with a strict soil moisture capacity and applies two main procedures of water balance and routing in sequence. The water balance component which maintains the balance between rainfall, evaporation, runoff and soil moisture storage in different layers, has five parameters to calibrate. The routing component has four parameters and calculates the generated runoff in the catchment outlet by applying the classic Gamma distribution model (Nash, 1959), given total runoff from the balance component.

Sacramento Model (U.S.A) – Sacramento soil moisture accounting (SAC-SMA) is another lumped continuous R-R model (Burnash et al., 1973). SAC-SMA model efficiency is highly related to the length and quality of available data. It needs long term mean daily rainfall, evaporation, air temperature and stream flow data for river flow forecasting. SAC-SMA model structure has 5 stores, two upper zone (tension and free water) and three lower zones (tension, primary free and supplementary free water). Evapotranspiration is removed from tension stores and runoff is released from free stores (surface runoff, inter flow and base flow). At first the upper zone receives the rainfall and next, water evaporates or moves to the lower stores based on defined movement rules. SAC-SMA model needs 16 parameters to be calibrated to represent catchment water balance process.

Xinjiang Model (China) - Xinjiang model is a semi-distributed conceptual model which is highly efficient in humid and semi-humid regions. This model is developed in 1973 in China and published in 1977 (Zhao, 1977). In this model, evapotranspiration is the controlling factor, as runoff is generated when soil moisture exceeds the field moisture capacity. Therefore, rainfall first feed the soil moisture deficit, then the subsequent precipitation will become runoff. Xinjiang model divides the soil to three layers of upper, lower and deeper. Generated runoff from these three layers are immediate, surface and groundwater runoff,

respectively. As a semi-distributed model, Xinanjiang applies a parabolic curve to consider spatial distribution of the soil moisture storage capacity (SMSC) over the catchment. The basic version of the model has been further modified by introducing a double parabolic curve (Jayawardena and Zhou, 2000). The modified Xinanjiang model has to calibrate 11 parameters including Muskingum routing parameters.

The literature of conceptual models is very vast. Almost all large hydrological research centres around the world have developed and applied their own conceptual model for hydrological forecasting. Currently, SMAR (O'Connell et al., 1970), Sacramento (Burnash et al., 1973), SimHyd (Chiew et al., 2002), GR4J (Perrin et al., 2003), AWBM (Boughton, 2004) and IHACRES (Croke et al., 2002) are the most popular conceptual models used in Australia (Vaze, et al., 2012). Compare to physically-based mode, these models are more popular, easier to develop and require fewer parameters for calibrating the catchment. However, conceptual models have some limitations as summarized below:

- In a lumped model, catchment is considered as a homogeneous unit by utilizing the average value of spatially heterogeneous parameters. Taking the average values of the catchment characteristics for simulating various hydrological process can significantly affect model accuracy.
- Developed model is not applicable for any other catchments, as model parameters optimized based on the unique characteristics of the selected catchment (catchment size and type, climate, topography, geology, vegetation and soil type).
- The model calibrates its parameters based on available historical rainfall-runoff events and may not be suitable for forecasting different rainfall-runoff trends in future.
- Event-based models are unable to be applied to ungauged catchments as an extensive amount of data is required for model calibration.

- Semi-distributed conceptual models have similar limitations of physically-based models, including extensive data requirement and using relatively inaccurate catchment parameters due to measurement difficulties.
- Many assumptions need to be made for simulating a complex process by a simplified model.

2.4 DATA DRIVEN MODELS

Another alternative for hydrological modelling is to apply data driven (also called black box) techniques on hydrological time series. Unlike process-based models, these models require very limited understanding of the hydrological system and mainly rely on the quality of the available data. Data driven models find the relation between inputs (river flow and/or rainfall time series) and output (runoff) without considering the underlying hydrological process. Figure 2.3 depicts the learning system in data driven method. These methods can be categorized in two main types of classical and computational intelligence approaches.

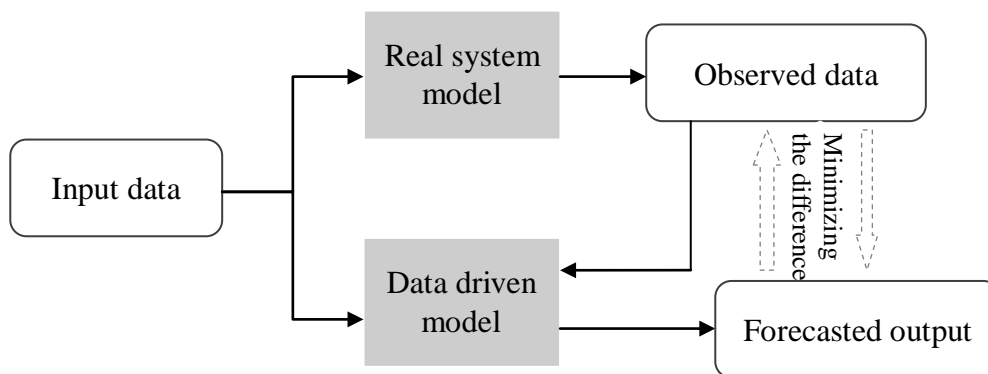


Figure 2. 3 Learning system in black box data driven approaches.

2.4.1 Classical data driven approach

The classical data driven models are generally regression models. Autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), seasonal ARIMA, autoregressive exogenous (ARX), threshold autoregressive (TAR) and multiple linear regression (MLR) are the most popular regression models (Wang, 2006). Among them, ARIMA has been the most frequently used method for river flow forecasting that is first introduced by Box and Jenkins (1970). ARIMA is an extended type of ARMA, which has two main components of autoregressive and moving average as following;

$$\begin{array}{c}
 \text{AR component} \qquad \qquad \qquad \text{MA component} \\
 \underbrace{\hspace{10em}} \qquad \qquad \qquad \underbrace{\hspace{10em}} \\
 ARMA(p, q) = Z_t = (\varphi_1 Z_{t-1} + \dots + \varphi_p Z_{t-p}) + (\varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}) \quad (2.1)
 \end{array}$$

where p is the order of autoregressive, q is the order of moving average, t is the time step (e.g. 12 for monthly modelling), ε_t is a white noise and φ and θ are the AR and MA coefficients, respectively. The past events are processed by AR component and the summation of forecasting error is presented by MA component.

These traditional techniques usually assume that a signal is stationary and can be described by a set of linear equations. Therefore, they are not reliable for achieving accurate river flow forecasting as river flow time series is highly nonlinear and nonstationary (Martins et al., 2011).

2.4.2 Computational intelligence approach

In the last two decades, computational intelligence (CI) approaches have been increasingly substituted regression models and applied in many hydrological forecasting. CI models are capable of recognizing complex non-linear relationships between input and output data sets. A number of different types of CI methods which are successfully applied in hydrological forecasting is as follows;

- Artificial neural networks (multi-layered perceptron, radial basis function, recurrent, product unit)
- Fuzzy rule-based systems
- Adaptive neuro-fuzzy inference system
- Support vector machines
- Chaos theory and dynamic systems
- Hybrid wavelet models
- Genetic algorithm/programming
- Swarm intelligence optimization (ant colonies, fish schooling, bee algorithm)

Given the complexity of rainfall-runoff process, computational intelligence methods are generally very powerful tool for river flow forecasting. Although CI models do not provide detailed information on hydrological process (black box type models) and require high quality historical time series, they are highly reliable and accurate. Following is a summary of CI models' advantages over physically-based and conceptual models for river flow forecasting application:

- Unlike physically-based models, CI models do not require a large number of hydrological and geological parameters for representing the catchment behaviour. CI models are able to achieve accurate forecasts by applying high quality river flow time series (long historical records) as the single input .
- CI based models are self-trained. The input-output relationship is formulated automatically based on historical data in a catchment. Therefore, understanding the complex interaction between hydrological and geological process is not necessary for developing the model.
- They are able to train the model with multiple effective inputs like meteorological parameters. Therefore, future climate changes could be considered in the CI modelling process.
- Contrary to conceptual, semi-distributed or even distributed physically-based models, no assumptions or estimations need to be taken for formulation and calibrating the catchment.

- Developed CI models are also easily applicable to different case studies with different catchment characteristics as they extract all necessary information from time series analysis.
- These models can be cost-effective as in-field measurements or gauging station maintenance would be reduced.
- Computational intelligence are the most efficient models for infilling of missing rainfall and river flow data to be used in river flow forecasting or any other hydrological applications.
- These models are the best option for modelling ungauged catchments when there is no other feasible solution for modeling. They are able to simulate the catchment by using effective inputs such as upstream data or data from other catchments with similar characteristics (Dawson et al., 2006; Besaw et al., 2010).

Despite the numerous advantages of data driven approaches, they also have some limitations. The main drawbacks of CI methods could be categorized as followings;

- These models require high quality historical data as the simulation is based on the previous trends. Accurate river flow forecasting with short period of river flow recoding is not achievable unless there are some other effective inputs data with good quality are available.
- Unlike process-based methods, they do not provide insight into the underlying hydrological processes in the catchment.

In this study, a number of CI based approaches are developed for river flow forecasting, using artificial neural networks, adaptive Neuro-fuzzy inference system and hybrid wavelet-CI techniques. More details on these CI approaches, are given in Chapter three.

Chapter 3

Computational Intelligence Approach

3.1 INTRODUCTION

In this study, computational intelligence (CI) approach is chosen for river flow forecasting. CI models are capable of simulating and forecasting hydrological events based on available historical data. They require very limited knowledge on complex rainfall-runoff process and huge catchment and meteorological parameters involved in this process.

This chapter briefly introduces the concept of artificial neural networks, fuzzy modelling and wavelet analysis. The methodology of specific types of CI approaches, applied in the developed models, is explained in more details.

3.2 ARTIFICIAL NEURAL NETWORKS

3.2.1 Introduction

Artificial neural networks (ANN) are generally computational models, inspired by the operations of biological neural system. Artificial neural networks are parallel distributed processing networks that are modelled after cortical structures of the brain. Artificial neural networks have flexible structures that are capable of identifying complex nonlinear relationships between input and output data sets (Adamowski and Sun, 2010). It can be used to forecast future output values from

given input data set by minimizing the error between the predicted and actual outputs.

The concept of Artificial Neural Networks (ANN) was first introduced by Warren McCulloch and Walter Pitts in 1943. They published the fundamentals of neural computing by proposed a neural model with binary neuron and a fixed threshold (McCulloch and Pitts, 1943). The initial concept of ANN was described algorithmically for the first time by Rosenblatt. He introduced perceptron algorithm for supervised learning of ANN input–output system (Rosenblatt, 1958). His work was the basis of feed-forward multi-layered neural networks development. The theory, algorithms and application of artificial neural networks have made significant progress since 1980s. In 1982, the self organized map (SOM) algorithm was introduced by Kohonen (Kohonen, 1982). Kohonen neural networks became widely known after he presented learning rule of unsupervised self organizing feature map (SOFM) in his book in 1988 (Kohonen, 1988). In 1986, the backpropagation training algorithm for training multilayer perceptron neural networks was first introduced (Rumelhart et al., 1986), which grounded significant growth in ANN applications. Broomhead and Lowe (1988) introduced radial basis function neural networks, as an alternative to the multilayer perceptrons.

Artificial neural networks are currently being used in different fields such as finance, medicine and a wide range of engineering applications. The startup period of studying ANNs' application in hydrology occurred throughout the 1990s. The study carried out by Daniell (1991) could be referred as the first paper on neural network application in hydrologic modeling. This study listed ten potential applications of neural networks in hydrology and water resources while it illustrated two examples of ANN applications itself. Since then, the application of ANN in hydrology and water resources modelling has attracted a lot of attention. Maier and Dandy (2000), the ASCE task committee (2000a, b) and more recently Maier et al. (2010) published comprehensive reviews of ANN applications in hydrology.

Different types of ANNs have been used in hydrological modeling like radial basis function (RBF) (Fernando and Jayawardena, 1998; Moradkhani et al., 2004; Nor et al., 2007; Partal, 2009; Lin and Wu, 2011), bayesian neural networks (Kingston et al., 2005; Khan and Coulibaly, 2006; Jiang et al., 2012) and feed-forward multilayer

perception (MLP), which is the most popular neural network paradigm in hydrological forecasting (Fernando and Jayawardena, 1998; ASCE task committee, 2000b; Dawson and Wilby, 2001; Kim and Barros, 2001; Sivakumar et al., 2002; Cigizoglu, 2003a; Kim and Valdes, 2003; Kumar et al., 2005; Srinivasulu and Jain, 2006; Dawson et al., 2006; Nayebi et al., 2006; Machado et al., 2011; Weilin et al., 2011).

Artificial neural networks are known as one of the promising techniques for river flow forecasting (Dibike and Solomatine, 2001; Chiang et al., 2004). Many studies have been carried out to investigate ANN applications in river flow forecasting in comparison with traditional linear and conceptual methods. Karunanithi et al. (1994) compared ANN and autoregressive moving average (ARMA) performance for daily and hourly river flow forecasting in the Pyung Chang River, Korea. They found ANN is a more accurate predictor, especially for high river flows. They noted that ANNs are more robust for simulation of noisy data compared to ARMA. Tawfik et al. (1997) applied a simple three-layer back propagation neural network with linear transfer function for forecasting discharge rate at two gauging locations on the Nile River. They showed that the ANN approach is more accurate than commonly used techniques for most of the cases considered. Abrahart and See (2000) compared the forecasting power of neural network and autoregressive moving average models for river flow prediction in two contrasting catchments. They concluded that ANN models were less demanding and faster, while their accuracy were similar and sometimes better in comparison to ARMA models. Imrie et al. (2000), presented a methodology for training ANN that generalised well on the new data. They used backpropagation and cascade-correlation learning architecture for training the network. They revealed that a function with a similar shape to the cubic polynomial function might be necessary for ANNs to predict extreme values. Dibike and Solomatine (2001) concluded that ANN can exhibit a comparable or even better performance than a calibrated conceptual model such as Sugawara-IHE tank model. They also stated that the back propagation neural networks (BPNN) networks showed slightly better performance than radial base function networks (RBF) for their case study. Birikundavyi et al. (2002) achieved excellent results of up to 5 days ahead river flow forecasting of Mistassibi River in Quebec, Canada using ANN. They also showed that the ANN result outperform the PREVIS conceptual model

and classic autoregressive model coupled with a Kalman filter. Cigizoglu (2003b), used ANN for predicting monthly river flow of a station with short record length in Turkey. This study indicated that using ARMA model to generate synthetic monthly flow and applying this series as the training sets of ANN, significantly improves prediction results. Moradkhani et al. (2004), compared the result of different ANN models with different structures for the Salt River stream flow forecasting. They concluded that the selection of a training set is crucial in the ANN modelling. Machado et al. (2011) developed an ANN model for monthly forecasting of Jangada River flow, Brazil. Comparing the results with the IPHMEN conceptual model showed that the ANN was superior in reproducing the observed flows.

Many studies have shown that adding other meteorological data, such as rainfall, temperature, soil moisture, evaporation and evapotranspiration can improve the accuracy of neural network based river flow forecasting (Poff et al., 1996; Kim and Barros, 2001; Nayebi et al., 2006; Anctil et al., 2008; Adamowski and Sun, 2010; Pramanik et al., 2011).

3.2.2 Neuron Modeling and Activation functions

Neural networks are networks of interconnected neurons (nodes), which are the fundamental unit of ANN. Neurons are able to receive and transmit signals from one to another. The basic neuron model as a binary threshold processing unit, is introduced by McCulloch and Pitts (1943). The most common structure of a neuron is illustrated in Figure 3.1.

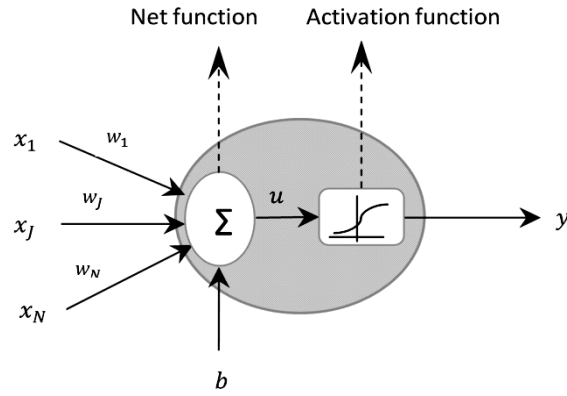


Figure 3. 1 Schematic of a neuron function system.

Each neuron has a number of inputs and outputs. A neuron computes an output by applying net and activation function (transfer function) on inputs. First, net function sum weighted inputs (u), then output is computed based on activation function ($y = f(u)$). The net function is usually linear, as follows;

$$u = \sum_{i=1}^N x_i w_i + b \quad (3.1)$$

where x is an input vector, w_i is the connection weight from the i th neuron in the input layer; b is the threshold value or the bias of the neuron. There are various types of activation functions including Sigmoid (Logistic), Hyperbolic tangent (Tansigmoid), Inverse tangent, Threshold, Gaussian radial basis and Linear, while the first two are the most commonly used in the hydrological modelling (Dawson and Wilby, 2001).

One of the most applied activation functions for most applications is Sigmoid. The output of this function is bounded into the range of zero and one, for inputs of minus to plus infinity, which is considered as the desirable characteristics of this function. Sigmoid activation function is expressed mathematically as shown in following equation;

$$y = f(u) = \frac{1}{1 + \exp(-ru)} \quad (3.2)$$

where r is the steepness parameter. As r decreases the shape of function alters from S-shape to the linear shape. Figure 3.2 shows the shape of this function with different steepness parameters.

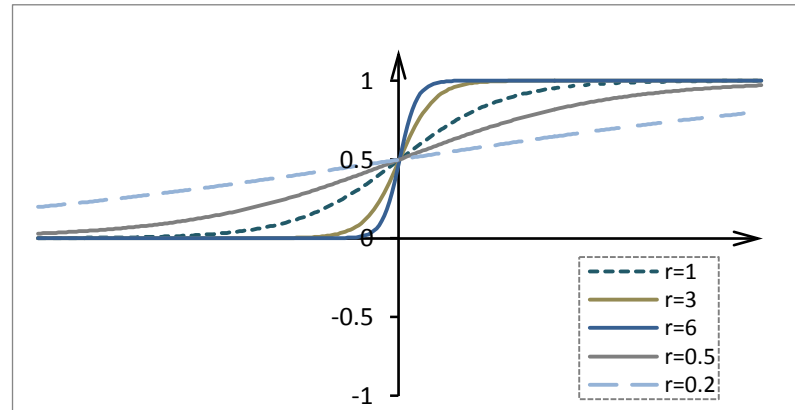


Figure 3.2 Sigmoid activation function with different steepness parameter.

Sigmoid hyperbolic tangent (Tan-sigmoid) transfer function belongs to sigmoid family and bounded between -1 and 1 (Figure 3.3). Tan-sigmoid activation function formula is as follows;

$$f(u) = \tanh(u) = \frac{\exp(ru) - \exp(-ru)}{\exp(ru) + \exp(-ru)} \quad (3.3)$$

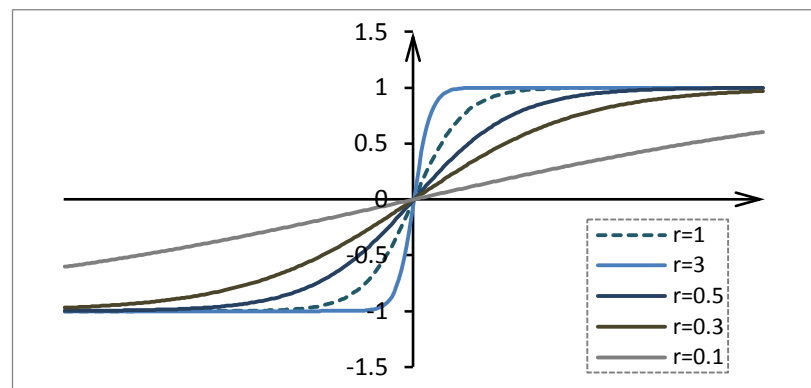


Figure 3.3 Tan-Sigmoid activation function with different steepness parameter.

The shape of this function is similar to the Sigmoid function with different boundaries. Having the bipolar output (range between ± 1) could be beneficial in certain networks with negative output values. According to Kalman and Kwasny (1992), using the tan-sigmoid as activation function achieves the best result for the ANN with backpropagation algorithm.

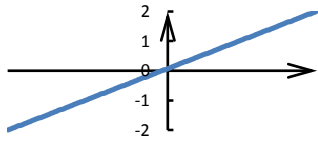
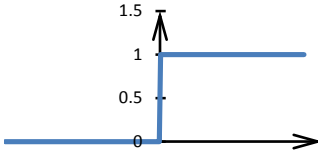
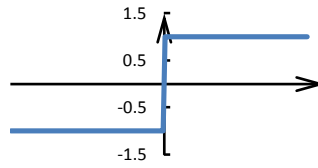
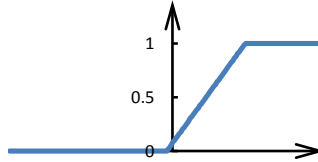
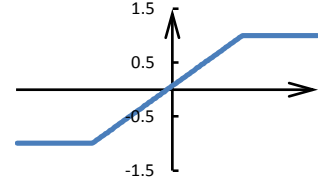
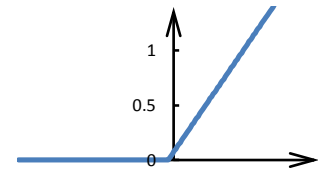
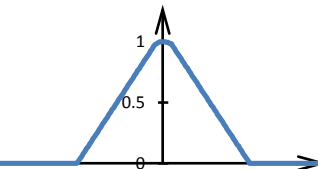
None of the sigmoid functions reach their theoretical minimum or maximum. For example, neurons that use the tan-sigmoid function could be considered fully activated around ± 0.9 . Therefore, the extremes of ± 1 can be used as inputs to the network, but it is ineffective to train the network to achieve extreme values in outputs.

Linear function is a simple activation function which does not limit the output range. This function is usually useful for the output layer of the ANN. The linear activation function formula is as follows;

$$f(u) = au + b \quad (3.4)$$

There are various types of linear activation functions. Table 3.1 illustrates the most popular linear transfer function.

Table 3. 1 Different types of linear activation functions.

Unbounded linear activation function	$f(u) = u$	
Hardlimiter activation function	$f(u) = \begin{cases} 0 & \text{if } u < 0 \\ 1 & \text{if } u \geq 0 \end{cases}$	
Symmetric hardlimiter function	$f(u) = \begin{cases} -1 & \text{if } u \leq 0 \\ 1 & \text{if } u > 0 \end{cases}$	
Saturating activation function	$f(u) = \begin{cases} 0 & \text{if } u < 0 \\ u & \text{if } 0 \leq u \leq 1 \\ 1 & \text{if } u > 1 \end{cases}$	
Symmetric saturating function	$f(u) = \begin{cases} -1 & \text{if } u < -1 \\ u & \text{if } -1 \leq u \leq 1 \\ 1 & \text{if } u > 1 \end{cases}$	
Positive linear activation function	$f(u) = \begin{cases} 0 & \text{if } u < 0 \\ u & \text{if } u > 0 \end{cases}$	
Triangular activation function	$f(u) = \begin{cases} 0 & \text{if } u > 1 \\ 1 + u & \text{if } -1 \leq u \leq 0 \\ 1 - u & \text{if } 0 \leq u \leq 1 \\ 0 & \text{if } u < -1 \end{cases}$	

3.2.3 Neural network architecture

Neural network architecture is defined based on the way neurons are connected to each other, which determines how computations proceed. About 30 different neural networks architectures have been created, developed and used so far (Krycha and Wagner, 1999).

Artificial neural networks are typically composed of different layers of neurons: an input layer, one or more hidden layer and an output layer. The number of neurons in the input and output layer depends on the problem and the number of hidden layers and the number of neurons in hidden layers should be specified. In practice, having a single hidden layer with enough neurons, usually leads to an accurate approximation needed (Lippmann, 1987; Cybenko, 1989). Having a greater number of hidden neurons, gives the network flexibility to solve more complex problems, while having too many neurons may cause overfitting problem. Different approaches have been introduced to reach an optimum number of neurons. One of the most promising solutions to achieve this number is trial and error procedure (Shamseldin, 1997).

In terms of the pattern of connections between the layers, ANN can be designed in feed-forward or recurrent form. Recurrent neural networks are mainly used when there are temporal patterns in the data. Feed-forward neural networks are the most common neural networks in use, so much so that some users identify the phrase “neural networks”, only feed-forward networks (Mehrotra et al., 1997). There are different type of feed-forward neural networks such as multilayer perceptron (MLP) and the radial basis function (RBF). The most popular neural network paradigm in hydrology is the multilayer feed-forward neural networks (Fernando and Jayawardena, 1998; ASCE task committee, 2000b; Dawson1 and Wilby, 2001; Kumar et al., 2005; Firat, 2008; Weilin et al., 2011), which is also a used in this study.

3.2.3.1 Feed-forward multilayer perceptron ANN

Feed-forward multilayer perceptron neural networks (MLP), are composed of several layers of neurons. The connections between neurons (information flow) are in one

direction, from the input layer, through hidden layers and to the output layer. Figure 3.4 shows a schematic diagram of three-layer feed-forward neural network (with one hidden layer).

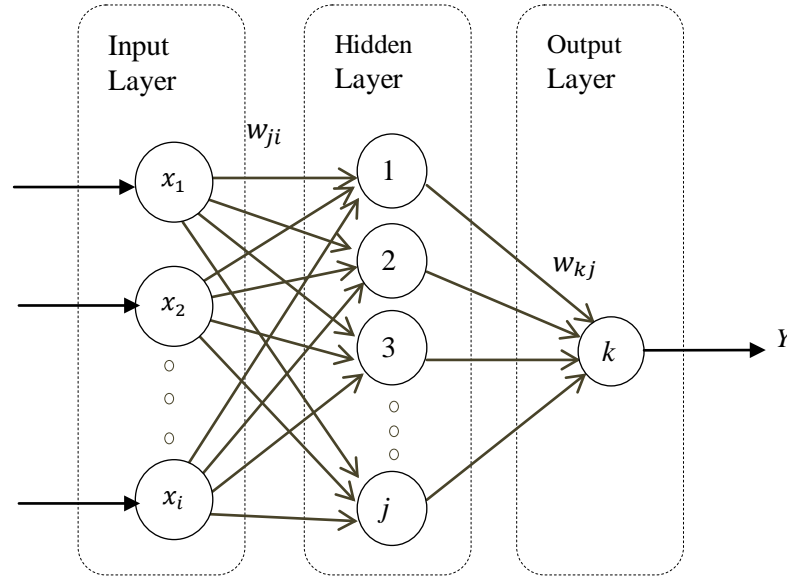


Figure 3. 4 Schematic diagram of a three-layer feed-forward neural

As shown in Figure 3.4, in MLP's there are no connections between neurons in the same layer or any feedback connections between layers. The output of neurons in each layer, are applied as the inputs to the next layer. The final output of this network could be achieved by the following equation:

$$Y = f_o \left[\sum_j w_{kj} \cdot f_h \left(\sum_i w_{ji} x_i + b_j \right) + b_k \right] \quad (3.5)$$

where x is an input vector, w_{ji} is the connection weight from the i^{th} neuron in the input layer to the j^{th} neuron in the hidden layer; b_j is the threshold value or bias of j^{th} hidden neuron; w_{kj} is the connection weight from the j^{th} neuron in the hidden layer to the k^{th} neuron in the output layer; b_k is bias of k^{th} output neuron and f_h and f_o are the activation function for hidden and output layer, respectively.

3.2.4 Neural network learning

Learning is the fundamental characteristic of neural networks. In the learning process, learning or training algorithm is updating network parameters to achieve a desired model performance based on a set of training data. The learning algorithm adjusts the weights and biases of the network to minimize the error between computed and observed output. There are three main classifications for ANN learning: supervised, unsupervised and reinforcement. The most commonly used learning paradigm, among all, is supervised learning neural networks. Unsupervised learning cannot trained neural networks to reach a target outputs and they are mainly efficient for pattern classification purposes. While, supervised learning algorithms require both inputs and associated output (target output) for training the network and they are very suitable for solving time series forecasting problems.

3.2.4.1 Backpropagation algorithm

The most widely used learning algorithm for training the neural networks is the backpropagation algorithm (Zhang and Barrion, 2006; Nawi et al., 2013). Backpropagation algorithm (BP) is a supervised algorithm which adjusts the connection weights and biases in the backward direction. It is an optimization procedure based on gradient descent to minimize the total error between the desired and actual outputs.

The learning process of the backpropagation algorithm consists of two parts, forward and backward propagation. Figure 3.5 illustrates the backpropagation training of a feed-forward neural networks.

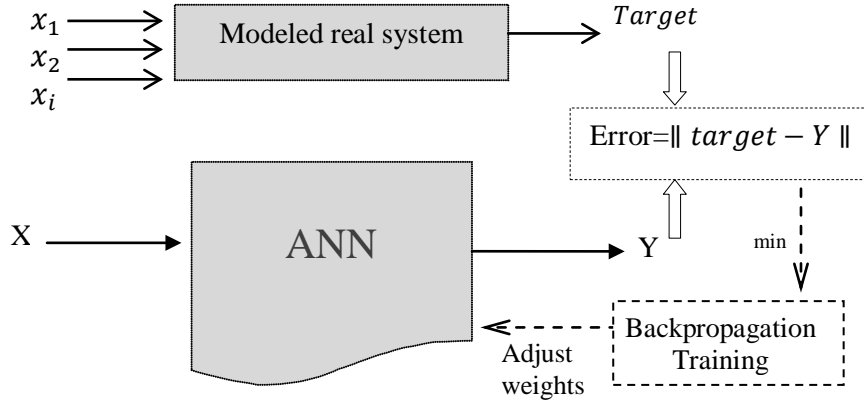


Figure 3.5 Structure of neural networks with back propagation training algorithm.

First, training pattern is propagated in a forward direction from input to output layer. As discussed previously, for the P layer network the output of each layer will be as follow;

$$y^0 = x \quad (3.6)$$

$$y^{l+1} = f^{l+1}(w^{l+1}y^l + b^{l+1}) \quad l = 0, 1, \dots, P - 1 \quad (3.7)$$

where x is the input vector, l is the layer number, f is the transfer function, w is connection weight, b is the threshold value and y^p would be the network output. Then the computed output is compared with the observed target to determine the error value. For each output node the computed output (y^p) is compared with its corresponding target value (t) to determine the difference.

$$e_n = t_n - y_n^p \quad (3.8)$$

The error function is based on the least squared errors as it can penalize large deviations more than small ones. The error function for this network is:

$$V = \frac{1}{2} \sum_{n=1}^N (t_n - y_n^p)^T (t_n - y_n^p) = \frac{1}{2} \sum_{n=1}^N e_n^T e_n \quad (3.9)$$

These errors are used to adjust the connection weight layer by layer in a backward direction, from the output layer to the input layer.

A number of training algorithms have been developed for error back propagation learning. In this study gradient descent, gradient descent with the adaptive learning rule, Bayesian regularization and Levenberg-Marquardt algorithm (LM) is used to train the network. LM algorithm has the fastest convergence among other algorithms and it is able to obtain lowest mean square error in many cases (Cigizoglu and Kisi, 2005; Beal et al., 2012; Lam et al., 2012). LM is a combination of steepest descent and the Gauss-Newton method. The one step weight update equation uses Newton's method. This equation minimizes the error function (V) with respect to the parameter vector w , as follows;

$$\Delta w = -[\nabla^2 V(w)]^{-1} \nabla V(w) \quad (3.10)$$

where $\nabla^2 V$ is the hessian matrix and ∇V is the gradient, then it can be shown as follows;

$$\nabla^2 V(w) = J^T(w)J(w) + S(w) \quad (3.11)$$

$$\nabla V(w) = J^T(w)e(w) \quad (3.12)$$

where the $J(w)$ is the Jacobian matrix which can be computed as:

$$J(w) = \begin{bmatrix} \frac{\partial e_1(w)}{\partial w_1} & \dots & \frac{\partial e_1(w)}{\partial w_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial e_N(w)}{\partial w_1} & \dots & \frac{\partial e_N(w)}{\partial w_n} \end{bmatrix} \quad (3.13)$$

And $S(w)$ is;

$$S(w) = \sum_{i=1}^N e_i(w) \cdot \nabla^2 e_i(w) \quad (3.14)$$

According to the Gauss-Newton method $S(w) \approx 0$, and the one step weight update equation becomes:

$$\Delta w = -[J^T(w)J(w)]^{-1}J(w)e(w) \quad (3.15)$$

Applying the Levenberg-Marquardt modification the Gauss-newton method, the equation becomes:

$$\Delta w = -[J^T(w)J(w) + \mu I]^{-1}J(w)e(w) \quad (3.16)$$

where I is the identity matrix and $\mu > 0$ which is modified by some factor (β) in each epoch. For large values of μ the algorithm becomes steepest descent with the step of $(1/\mu)$ which is the standard backpropagation, and for the small values of μ the algorithm becomes Gauss-Newton method (Figure 3.6).

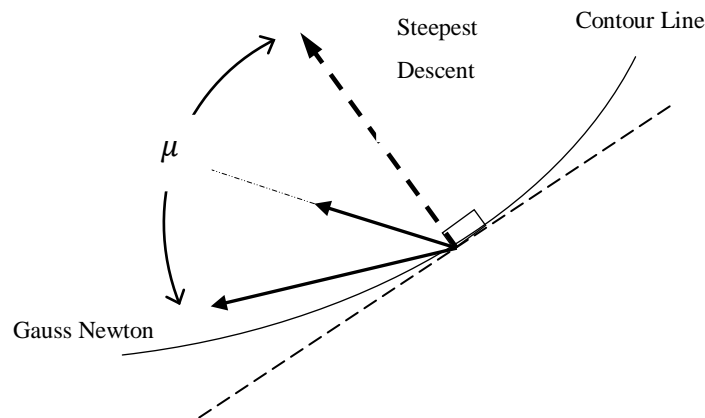


Figure 3. 6 Levenberg-Marquardt algorithm shifts from the steepest descent to the Gauss-Newton method of decreasing the value of μ .

In the next step again the network is trained with the new weight:

$$W_{Epoch+1} = W_{Epoch} + \Delta w \quad (3.17)$$

The sum of squares errors ($V(w)$) is recomputed using new weights, if $V(w)$ is increased, μ is multiplied by factor β and if it is reduced μ is divided by β . This procedure will be iterated again and again until the $V(w)$ has been reduced to some error goal, which then the algorithm is assumed to be converged.

3.3 NEURO-FUZZY MODELLING

3.3.1 INTRODUCTION

The concept of fuzzy logic (FL), was originally proposed by Zadeh (1965). In the fuzzy logic, linguistic variables are often used rather than numerical values in order to facilitate the expression of rules and facts. In the last two decades fuzzy logic modelling has been applied to various fields of engineering problems, including earth and sciences field (Demicco and Klir, 2004). There are also a number of studies, which have investigated FL application in hydrological forecasting, in particular.

Liong et al. (2000) used a fuzzy reasoning method with various shapes of membership functions to predict the water level in Dhaka, Bangladesh. Although the prediction accuracy of this method was lower than that of neural networks, the level of accuracy was still acceptable. Hundecha et al. (2001), developed a fuzzy rule-based routine to simulate different processes involved in runoff generation and compared the result with HBV semi-distributed conceptual model for Neckar River catchment in Germany. In this research the fuzzy logic-based model was found to reproduce the observed discharges well, although it overestimated the peak flows. Sen and Altunkaynak (2006), presented various uncertainties embedded methods for determining the runoff coefficient, and rainfall-runoff formulation including statistics, probability, perturbation and fuzzy system modelling. They applied these methods on two stations in Istanbul, Turkey and concluded that fuzzy logic approach yields the least relative error among other alternative runoff calculation methods. Turan and Yurdusev (2009), compared feed-forward back propagation neural networks, generalized regression neural networks and fuzzy logic for estimation of unmeasured data using the data of the four runoff gauge stations on the Birs River in Switzerland. The feed-forward back propagation neural networks (FFNN) algorithm was selected over other models for Souhieres station. They concluded that the best method should be sought based on the flow values and the basin characteristics. Wang and Altunkaynak (2012), conducted a comparative case study between storm water management hydrological model (SWMM) and a fuzzy logic model for total runoff prediction in the Cascina Scala watershed in Italy. They found predicted total runoffs from either the SWMM or fuzzy logic are reasonably well according to the

performance criteria. For large rainfall events, the fuzzy logic model generally outperforms the SWMM. However, the SWMM can produce the time varying hydrograph which cannot be generated by a fuzzy model. Jayawardena et al. (2014), demonstrated the robustness of fuzzy modelling for daily and 6-hourly discharge prediction. They used three different types of fuzzy inference systems, namely, Mamdani, Larsen and TSK for river flow prediction in 4 different rivers located in different climatological regions. Although all approaches were found efficient, TSK model slightly outperformed the other two.

The adaptive neuro-fuzzy inference system (ANFIS) models, which consist of both ANN and fuzzy logic methods, was first introduced by Jang (1993). The neuro-fuzzy inference system combines the advantages of both fuzzy logic and neural networks techniques, the learning ability of a neural network and the interpretable manner of the fuzzy logic in representing knowledge (Jang and Gulley, 1995). Neuro-fuzzy modelling has been successfully used in many hydrological studies. Nayak et al. (2005) used a neuro-fuzzy hybrid approach for short term river flow prediction. They showcased the application of model by applying it on the Baitaraini River in India. It was found that ANFIS outperforms traditional ARMA models. Also, comparing to ANN, it saved considerable computational time. Kisi (2006) investigated the abilities of neuro-fuzzy techniques, with various combinations of inputs, to improve the accuracy of daily evaporation estimation. It was found that the neuro-fuzzy computing techniques could be employed successfully in modelling evaporation process. Keskin et al. (2006), used ANFIS model for monthly river flow prediction in Dim stream in Turkey. They employed AR(2) model to generate synthetic monthly flows and used this data for prediction with ANFIS. They compared the results with the ANFIS performance when only a limited number of the observed flows were employed and shown that extended data series improved the prediction performance significantly in both low and high flows. Chang and Chang (2006), used a neuro-fuzzy hybrid approach to construct a water level forecasting system during flood periods. The ANFIS model provided accurate and reliable water level prediction for next three time steps. Aqil et al. (2007) carried out a study based on the comparative analysis of neural network and fuzzy systems for predicting the runoff in the Cilalawi River basin in Indonesia. Three different models of Levenberg-Marquardt-FFNN, Bayesian regularization-FFNN and ANFIS model were applied. The neuro-fuzzy

model had proved better generalization capabilities and Bayesian regularization algorithm appeared to be the worst. Kermani and Teshnehlab (2008) investigated the potential of neuro-fuzzy system with a Sugeno inference engine, and using different numbers of membership functions for daily river flow prediction. They achieved reliable results which were superior to the results of conventional autoregressive methods. Firat et al. (2009), applied two types of fuzzy inference systems, ANFIS and Mamdani fuzzy inference systems (MFIS), for predicting municipal water consumption time series. The results demonstrated that the ANFIS method is superior to the MFIS method in predicting monthly water consumption time series. Talei et al. (2010) compared fifteen ANFIS models, with different selection of input data, for event-based runoff forecasting. They determined the best input selection for long term, short term and overall discharge forecast.

The application of fuzzy modelling in hydrology is not as widespread as other computational intelligence approaches like neural networks. The reason could be the complex structure of if-then rules in fuzzy input-output modelling as the number of rules exponentially propagates with the number of inputs (Jacquin and Shamseldin 2009).

3.3.2 FUZZY LOGIC

Fuzzy logic focuses on linguistic variables in natural language to reduce and explain the system complexity. Fuzzy logic uses fuzzy set theory as a major tool.

Unlike classical sets, fuzzy sets don't have a crisp boundary and deal with reasoning that is approximate rather than precise. While classical sets either belong to a set or not, fuzzy sets can have partial membership. In a classical set, the membership of an element x in a classical set R is defined as follows;

$$\mu_R(x) = \begin{cases} 1 & \text{if } x \in R \\ 0 & \text{if } x \notin R \end{cases} \quad (3.18)$$

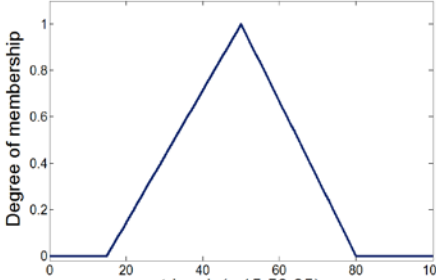
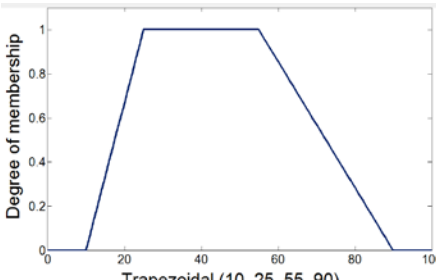
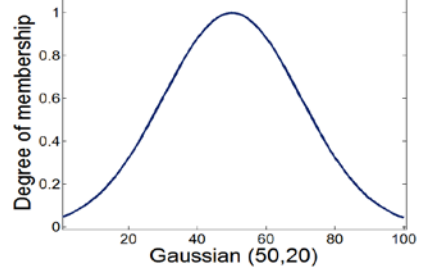
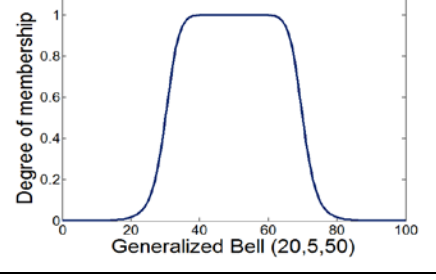
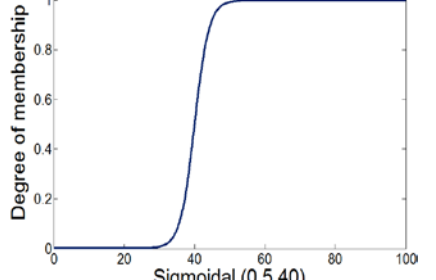
In contrast to a classical set, a fuzzy set can express the degree which an element belongs to a set. This concept is defined by degree of memberships that can take

values between 0 and 1, which a degree of zero means that the value is not belong to the set and a degree of one means that the value is completely representative of the set:

$$\mu_R(x): X \rightarrow [0,1] \quad (3.19)$$

where $\mu_R(x)$ is called the membership function (MF) of the fuzzy set R. Based on the problem, different type of fuzzy membership functions can be used. The most common types of membership functions are triangular, trapezoidal, gaussian, generalised bell and sigmoid functions. These membership functions are defined in Table 3.2. The shape of these membership functions has also been illustrated using the MATLAB fuzzy logic toolbox and applying sample parameters. There is no clear approach for selecting the optimum membership type or number. In this study the optimum structure of FIS is achieved by trial and error.

Table 3.2 Triangular, Trapezoidal, Gaussian, Generalised bell and Sigmoidal membership functions.

<p>Triangular membership function</p>	$\text{triangle}(x; a, b, c) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{c-x}{c-b}, & b \leq x \leq c \\ 0, & c \leq x \end{cases}$	
<p>Trapezoidal membership function</p>	$\text{trapezoid}(x; a, b, c, d) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & b \leq x \leq c \\ \frac{d-x}{d-c}, & c \leq x \leq d \\ 0, & d \leq x \end{cases}$	
<p>Gaussian membership function</p>	$\text{gaussian}(x; c, \sigma) = e^{-1/2(\frac{x-c}{\sigma})^2}$	
<p>Generalised bell membership function:</p>	$\text{gbell}(x; a, b, c) = \frac{1}{1 + (x-c)/a ^{2b}}$	
<p>Sigmoidal membership function</p>	$\text{sig}(x; a, c) = \frac{1}{1 + \exp[-a(x-c)]}$	

3.3.3 FUZZY INFERENCE SYSTEMS

The most important modelling tool based on fuzzy set theory is fuzzy inference systems (FIS). Fuzzy inference systems is a knowledge base system in which the information of input and output data is converted into linguistically interpretable if-then fuzzy rules. FIS has different applications such as data classification, decision analysis, expert systems, time series predictions, robotics and pattern recognition. The basic structure of FIS consists of three conceptual steps, including fuzzification, fuzzy inference process and defuzzification (Figure 3.7).

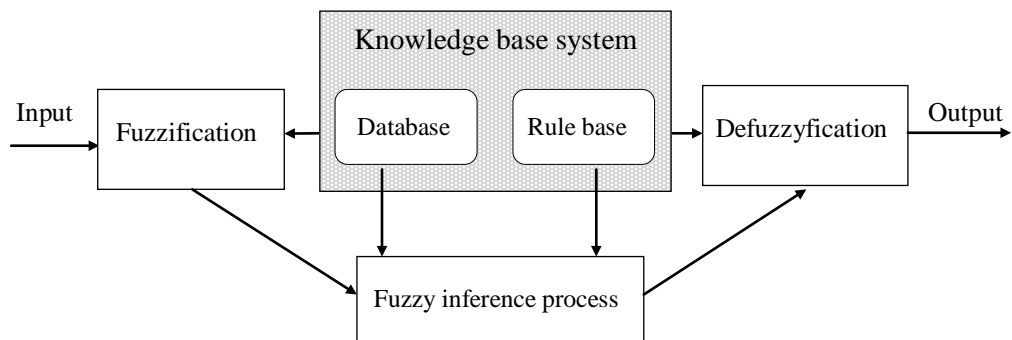


Figure 3.7 The basic structure of FIS.

In this process firstly classical or crisp data are converted into fuzzy data by defining the associated membership functions (MFs). Then MFs are connected with the fuzzy rules to drive the fuzzy output.

There are four main types of fuzzy inference system: Mamdani, Sugeno, Tsukamoto and Larsen fuzzy inference system, which the first two are the most widely used. The Mamdani FIS (Mamdani and Assilian, 1975), contains three stages of fuzzification, fuzzy rule inference process and defuzzification as mentioned above. The Sugeno fuzzy model, also known as TSK fuzzy model, first introduced by Takagi, Sugeno and Kang (Takagi and Sugeno, 1985; Sugeno and Kang, 1988). TSK is more efficient in optimization and adaptive techniques and its output membership functions could be either linear (first-order) or constant (zero-order) in respect to values of input. TSK type had the advantage of not requiring to go through de-

fuzzification. Typical rules in a TSK model with two inputs and one output have the form:

$$\text{Rule 1 } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ then } z_1 = f_1(x, y), \quad (3.20)$$

$$\text{Rule 2 } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ then } z_2 = f_2(x, y), \quad (3.21)$$

where x and y are inputs, A and B are fuzzy sets which are linguistic terms with MFs, and z_1 and z_2 are crisp output functions. The overall output is obtained by weighted average. For example, for an AND rule, the output is:

$$Z = \frac{z_1 * w_1 + z_2 * w_2}{w_1 + w_2} \quad (3.22)$$

where the firing strength is:

$$w_1 = \mu A_1(x) \wedge \mu B_1(y) \quad (3.23)$$

$$w_2 = \mu A_2(x) \wedge \mu B_2(y) \quad (3.24)$$

Since the overall output is a crisp set, there is no need for de-fuzzification required by this approach.

3.3.4 ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM

A neuro-fuzzy system integrates fuzzy inference systems and neural networks which have the potential to capture the benefits of both methods. Fuzzy systems have the advantages of describing the fuzzy rules and being interpretable, which make it possible to represent the real world process and identify the reason of particular value in the fuzzy system output. In the other hand, fuzzy systems need an expert knowledge or instructions to define fuzzy rules and tuning the parameters of fuzzy systems (e.g. membership functions parameters). By increasing the complexity of the process, developing fuzzy rules and membership functions become more difficult

and sometimes impossible. In the neural networks approaches, the opposite situation can be observed. Neural networks are not able to explain the behavior of the system based on prior knowledge, but they are trainable which gives them the ability of tuning their structures from input-output data. Considering these facts, using a hybrid model of fuzzy and neural networks eliminates these problems. However, ANFIS has more computational complexity restrictions than ANN.

Neoru-fuzzy systems, have recently attracted many researcher's interest (Firat 2008; Googhari and Lee, 2011). The neuro-fuzzy modeling has the natural language description of fuzzy systems and learning capability of neural network but the drawback is their highly constrained learning process and their complexity compared to the neural networks.

One of the most popular integrated systems is adaptive neuro-fuzzy inference system (ANFIS) which has shown promising results in modelling nonlinear time series. ANFIS was first introduced by Jang (1993). In ANFIS a feed-forward network finds fuzzy if-then rules for reaching optimal model. Specifically the ANFIS system of interest here is functionally equivalent to the TSK first-order fuzzy model (Jang et al., 1997). A common rule set for a first-order TSK fuzzy model is the following:

$$\text{If } x \text{ is } A_n \text{ and } y \text{ is } B_n, \text{ Then } z = f_n(x, y) = p_n \times x + q_n \times y + r_n \quad (3.25)$$

where A and B are fuzzy sets in the antecedent; p_n , q_n and r_n are polynomial parameters of n^{th} rule (also called the consequent parameters) and z is a crisp function in the consequent.

Figure 3.8 illustrates the fuzzy reasoning mechanism of a two-input first-order Sugeno fuzzy model with two rules. The weights are usually obtained by multiplication of membership grades.

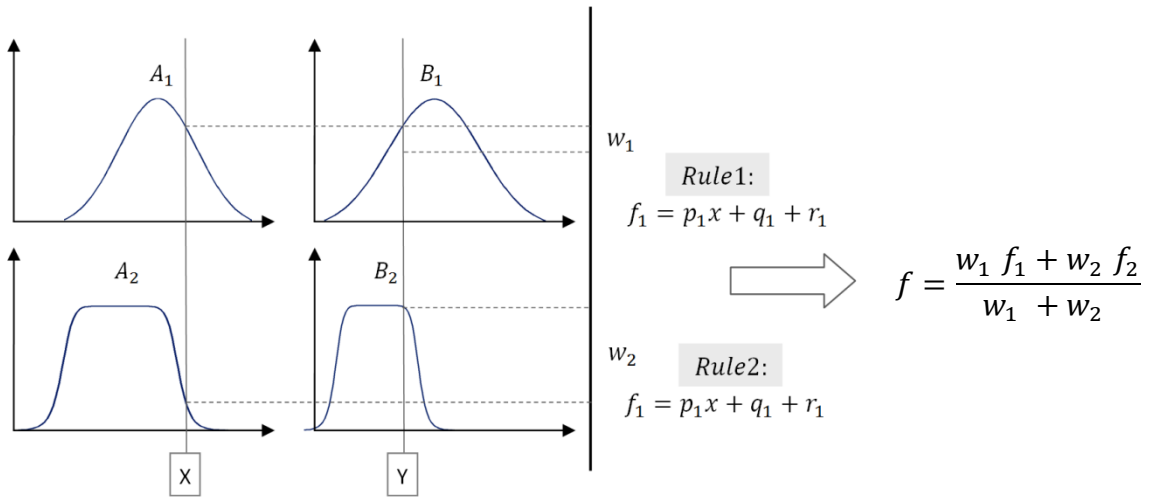


Figure 3. 8 Reasoning mechanism of a Sugeno fuzzy model with two inputs and rules.

The ANFIS structure for the TSK first-order fuzzy model consists of five layers to facilitate the learning process. Figure 3.9 illustrates the ANFIS structure where two inputs (x and y) with two fuzzy sets with linguistic labels (A and B) and one output are considered.

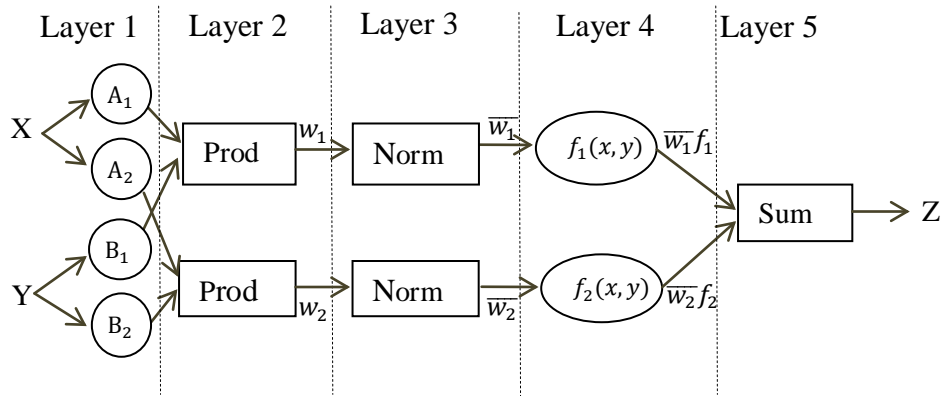


Figure 3. 9 Equivalent ANFIS architecture for two-input first-order TSK fuzzy model with two rules.

The first layer implements a fuzzification process and determine the membership grade of a fuzzy set:

$$O_{1,i} = \mu A_i(x) \quad \text{for } i = 1,2 \quad (3.26)$$

$$O_{1,i} = \mu B_{i-2}(y) \quad \text{for } i = 3,4 \quad (3.27)$$

Based on the problem, different membership functions can be applied. For instance, if the membership function of i^{th} node is a generalized bell function, the output of i^{th} node in the first layer defines as :

$$O_{1,i} = \mu A_i(x) = \frac{1}{1 + |(x - c)/a|^{2b}} \quad (3.28)$$

where $\{a_i, b_i, c_i\}$ are premise parameters that change the shape of the membership function.

The second layer generates the firing strengths (weights). This layer implements fuzzy AND operation by applying T-norm operators (usually multiplication) to the incoming signals in every neuron

$$O_{2,i} = w_i = \mu A_i(x) \times \mu B_i(y), \quad i = 1,2 \quad (3.29)$$

The third layer normalizes the firing strengths from the previous layer:

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1,2 \quad (3.30)$$

The fourth layer calculates rule outputs based on the consequent parameters:

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i) \quad (3.31)$$

Finally the fifth layer computes the overall outputs by summing all the incoming signals from layer 4:

$$O_{5,i} = \sum_i \bar{w}_i f_i = \frac{w_1 f_1 + w_2 f_2}{w_1 + w_2} \quad (3.32)$$

$$z = (\bar{w}_1x)p_1 + (\bar{w}_1y)q_1 + (\bar{w}_1)r_1 + (\bar{w}_2x)p_2 + (\bar{w}_2y)q_2 + (\bar{w}_2)r_2 \quad (3.33)$$

where the $O_{k,i}$ is the output of i^{th} node in k^{th} layer. The most substantial component of ANFIS is the rules which are defined by premise parameters (a_i, b_i, c_i) and the consequent parameters (p_i, q_i, r_i). The best values of the parameters for providing rules that would idealistically model the target system, are reached by a learning algorithm. With given input-output data, ANFIS employs the back propagation gradient descent method to optimize these parameters.

As previously discussed, neural networks use back propagation algorithm to adjust the weights between neurons. In the ANFIS, premises and consequents parameters play the role of weights in neural networks. Back propagation algorithm adjusts the premise parameters which determine the shape of membership functions and the least square error (LSE) solve the consequent parameters. This process has forward and backward steps. In the forward pass, the consequent parameters are estimated by LSE method, while the premise parameters are fixed and in the backward pass the error signals are propagated in backwards to modify the premise parameters while the consequent parameters are fixed. Since it uses two very different algorithms to reduce the error, the training algorithm is also called hybrid algorithm.

3.3.5 INPUT SPACE PARTITIONING

For the large data sets, determining the fuzzy rules by an expert is not effective and relies on trial and error. There are several techniques for determining the numbers of rules and membership functions. In a fuzzy inference system, basically there are three types on input space partitioning to form the fuzzy rules: tree partitioning, grid partitioning and scatter partitioning (clustering). In this study the last two methods are applied for initializing fuzzy if-then rules.

3.3.5.1 Grid partitioning

Grid partitioning is an approach for initializing the design of a fuzzy inference system when the number of inputs are limited. Grid partitioning divides the data

space into rectangular sub-spaces using axis-parallel partition. This method generates rules for all possible combinations of membership functions of all inputs. Therefore, the number of rules is equal to m^k where k is the number of inputs and m is the number of membership functions. Figure 3.10 demonstrates a grid partition in a two dimensional input space.

When we have a relatively large number of inputs, using this method encounter problems as the number of fuzzy rules increases exponentially with the number of input variables. For instance a fuzzy model with ten input variables and two membership functions on each, leads to 1024 fuzzy if-then rules.

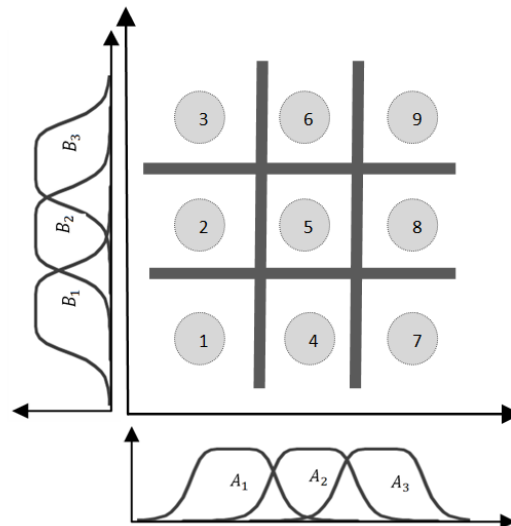


Figure 3.10 Grid partitioning of two inputs into 9 fuzzy rules.

3.3.5.2 Scatter partitioning (clustering)

The number of fuzzy rules increases with respect to the number of inputs. For minimizing the number of rules, the first step in generating the ANFIS is clustering the input-output data. Clustering partitions the data from a large data set into several groups by putting the most similar data in one group. To prevent grouping based on the ranges of the elements in the input vectors, input data must be normalized before clustering (Jang et al., 1997).

There are different kind of clustering techniques such as K-means, mountain, subtractive and C-means clustering. In this study subtractive and C-mean approaches are selected for clustering the inputs.

(a) Subtractive clustering

When there is no clear idea about the number of clusters for a large number of input data, subtractive clustering is fast, one-pass algorithm for finding the cluster centres in a data set. Subtractive clustering, which is first proposed by Chiu (1994), considers data points as the candidates for cluster centres and defines a density measure at data point x_i as follows:

$$D_i = \sum_{j=1}^n \exp \left(-\frac{\|x_i - x_j\|^2}{\left(\frac{r_a}{2}\right)^2} \right) \quad (3.34)$$

where r_a is a neighborhood radius which is determined by trial and error. A data point will have a high-density value if it has many neighboring data points. After calculating the density measure for each data point, the data point with the highest density measure (D_{c1}) is selected as the first cluster centre (x_{c1}) and the density measure of each data point x_i is:

$$D_i = D_i - D_{c1} \exp \left(-\frac{\|x_i - x_{c1}\|^2}{(r_b/2)^2} \right) \quad (3.35)$$

where r_b is a positive constant indicating neighborhood that has a measurable reduction in density measure and generally is equal to $1.5r_a$ (Chiu, 1994). The data points near the first cluster centre will have significantly reduced density measures. Thereby the next cluster centre is selected as the point has the greatest density value. This process continues until a sufficient number of cluster centres are achieved.

(b) Fuzzy C-mean clustering (FCM)

Fuzzy C-mean clustering algorithm (FCM) was proposed by Bezdek (1981). FCM algorithm is an extension of the C-mean clustering algorithm, which is based on a crisp clustering criterion (Abe, 2001). In this method, one piece of data could belong to more than one cluster with a degree of membership for each cluster. If x_j is a variable in the data set and m_{ij} is its membership degree for cluster i , the sum of the degrees of membership of x_j for all the clusters is equal to one:

$$\sum_{i=1}^c m_{ij} = 1, \quad (3.36)$$

where c is the number of clusters. In FCM partitioning the data is implemented by minimizing the sum of square error of each group using the following objective function:

$$\text{obj} = \sum_{i=1}^c \sum_{j=1}^n (m_{ij})^\alpha \|x_j - v_i\|^2 \quad (3.37)$$

where n is the number of variables in data set, v_i is the cluster centre of the cluster i and α is weighting exponent ($\alpha > 1$).

For solving the objective function, iterative algorithm is used. In each iteration, the values of v_i and m_{ij} are updated, using the equations given below:

$$v_i = \frac{\sum_{j=1}^n (m_{ij})^\alpha x_j}{\sum_{j=1}^n (m_{ij})^\alpha} \quad (3.38)$$

$$m_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_j - v_i\|^2}{\|x_j - v_k\|^2} \right)^{1/(\alpha-1)}} \quad (3.39)$$

3.4 WAVELET MULTI-RESOLUTION ANALYSIS

3.4.1 Introduction

Wavelet analysis is a signal processing technique which is able to determine the time-frequency-transformation characteristics of a signal. Before wavelet, the Fourier transform was the most widely used signal processing technique. Although Fourier transform is very suitable for stationary signal processing, it is not able to determine the time information of different events in a signal. In contrast to the Fourier transform, wavelet analysis can examine the signal simultaneously in both time and frequency domain by adjusting the window lengths automatically. This ability makes wavelet a very powerful tool for multi-resolution analysis (MRA) of complex non-stationary signals.

The concept of wavelet analysis was first proposed by Morlet et al. (1982). Later in 1980s, wavelet theory was improved to the fundamental level of wavelet application today. Mallat (1985) and Meyer (1987) discovered the multi-resolution theory of wavelet. Daubechies (1988) constructed a set of wavelet functions based on the Mallat's theory, which can be considered the most applicable wavelet function. The application of wavelet analysis in different fields of study commenced from early 1990s. In the last decade the number of publications in wavelet application in science and engineering has grown rapidly. Wavelet analysis has been utilized in many fields of engineering including water engineering.

Considering hydrological series as quasi-periodic signals contaminated by various noises, one of the recent developments in hydrological forecasting with data driven approach, is based on pre-processing the inputs with wavelet analysis. River flow time series, in particular, are complex and contain a wide range of frequency components. The multi-resolution analysis of wavelet transform can be utilised to decompose non-stationary signals of time series into their major sub-components. Then in the second stage, the decomposed signals, at different resolution levels, are taken as CI models inputs.

Application of wavelet-CI models in different field of hydrological forecasting have been recently investigated in a number of studies. Kim and Valdes (2003) applied

dyadic wavelet transforms in combination with neural networks to forecast droughts in the Conchos River Basin in Mexico, and showed that the hybrid model significantly improved the ability of neural networks to forecast regional drought. Wang and Ding (2003) also applied wavelet neural networks to predict shallow groundwater levels in Beijing and daily discharge of the Yangtze River in China, and concluded that the model could increase the lead time extension and accuracy of prediction. Partal and Kisi (2007) used a wavelet-neuro-fuzzy model to forecast the daily precipitation of three stations in Turkey. Kisi and Shiri (2012) developed wavelet and neuro-fuzzy model for water table depth fluctuations prediction. They found decomposing the inputs with wavelet significantly improve model efficiency provided that some of the wavelet components (the detail coefficients) are excluded from inputs and only approximation components are used.

The application of hybrid neural network models in river flow forecasting, in particular, has been also studied in the recent years. Cannas et al. (2006) applied hybrid wavelet neural networks for monthly runoff forecasting. They used both discrete and continuous wavelet for input and output preprocessing. They obtained better results using a hybrid model compared to a classical ANN model utilizing raw, noisy signals. They also concluded that preprocessing data with discrete wavelet lead to a better prediction compared to continuous wavelet. Kisi (2009) used neural networks wavelet model for forecasting intermittent river flow and the test results indicated that the discrete wavelet transform could significantly increase the accuracy of the ANN model in modeling intermittent river flows. Adamowski and Sun (2010) developed coupled wavelet transform neural networks model for 1 day and 3 days flow forecasting and reported improved performance over the ANN models. Wei et al. (2013) developed a wavelet neural network model for monthly river flow forecasting. They used discrete wavelet transformation (DWT) on a single river flow data and produced more accurate results for 48 months ahead river flow forecasting compared to ANN models.

However, the application of hybrid wavelet ANFIS approach in river flow forecasting, has been investigated in very limited studies. Nourani et al. (2011) preprocessed ANFIS rainfall-runoff models with DWT for both daily and monthly time scales. They concluded that combining discrete wavelet transform with ANFIS

model leads to promising result in runoff forecasting, especially for monthly forecasting. Ren et al. (2013) also developed an ANFIS model based on wavelet analysis for monthly runoff forecasting. Comparing the observed and predicted values, requirement for further improvements were noted.

3.4.2 Fourier Transform

The Fourier transform (FT) is one of the most popular signal processing tools. It is named after Joseph Fourier, who had the greatest contribution in developing the Fourier transform method (Fourier, 1808). He demonstrated that every periodic signal could be represented by a series of sinusoid functions. In fact, in Fourier transform a time-space domain signal is transferred to its finest frequency resolution for measuring its frequency components. Generally for the signal $s(t)$, FT pair is defined by the following equations:

$$FT[s(t)] = S(u) = \int_{-\infty}^{+\infty} s(t)e^{-jut} dt \quad (3.40)$$

$$FT^{-1}[s(t)] = s(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(u)e^{jut} du \quad (3.41)$$

where FT and FT^{-1} are Fourier transform (signal spectrum) and inverse Fourier transform, respectively.

FT is a very useful tool for analysing stationary signals and converting the complex function into the simple frequency domain. But FT is not able to analyse the non-stationary signals or providing the time frame of the signals (Labat, 2005).

3.4.3 Short-Time Fourier Transform

As mentioned before, the main problem of Fourier Transform is its inability to determine the time information of different events in a signal. Short-Time Fourier

Transform (STFT) is introduced to overcome this deficiency. In STFT, classical FT applies on a moving time window basis instead of whole signal. Narrow time intervals gives a good time resolution and weak frequency resolution, while wider time intervals provide a better frequency resolution and weak time information of the signal. In STFT, first a window function $W(t)$ with finite length is chosen (centred at $t=0$). Then the signal $s(t)$ is multiplied by this window function and then the FT of of the product is computed. The STFT equation can be described by the following equation:

$$STFT(\tau, u) = \int_{-\infty}^{+\infty} s(t)W^*(t - \tau) e^{-jut} dt \quad (3.42)$$

In STFT selecting the optimum window length is difficult as this cannot change while analysing. Moreover, STFT is still incapable to analyse non-stationary signals.

3.4.4 Continuous Wavelet Transform

Recently, wavelet analysis has become a common tool for multi-resolution analyzing of complex signals. The continuous wavelet transform (CWT) is first introduced by Morlet et al. (1982). The advantage of wavelet over FT is that it can examine the signal simultaneously in both time and frequency domains. Wavelet is also able to analyze highly non-stationary signals. In wavelet transform instead of multiplying the signal with a window function, signal is multiplied the wavelet functions. Wavelet functions are scaled (dilated or compressed) and shifted versions of the mother wavelet (ψ) as follows;

$$\psi_{\alpha,\beta} = \frac{1}{\sqrt{\alpha}} \psi\left(\frac{t - \beta}{\alpha}\right) \quad (3.43)$$

where α is the scale parameter and β is the translation parameter, $\alpha > 0$ and t is finite. Figure 3.11 is an illustration of shifted and scaled forms of a mother wavelet. The mother wavelet needs to meet the following conditions:

$$(a) \int_{-\infty}^{+\infty} \psi(t)dt = 0 \quad (3.44)$$

$$(b) \int_{-\infty}^{+\infty} \psi^2(t)dt = 1 \quad (3.45)$$

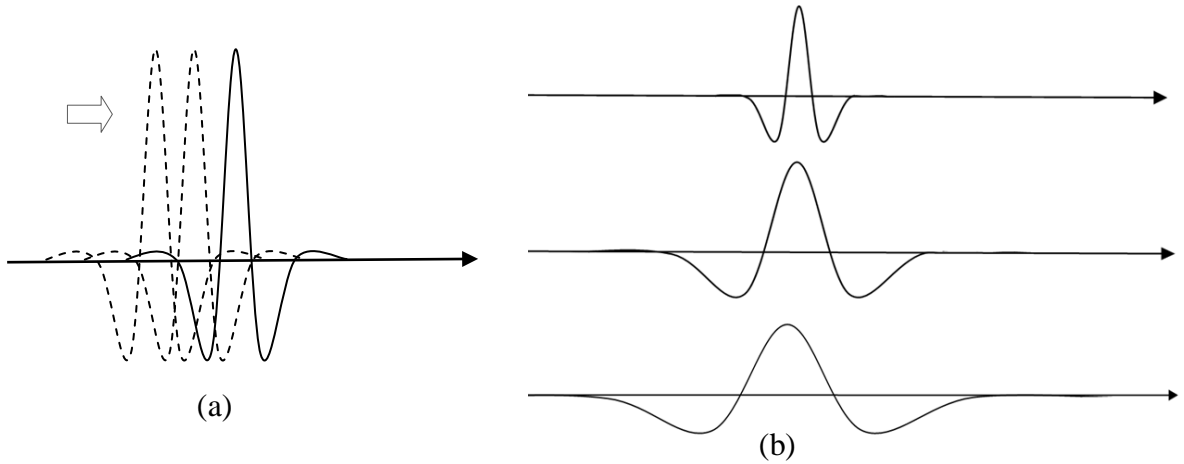


Figure 3. 11 (a) Shifted and (b) Scaled wavelet illustration.

If $s(t)$ is the continuous time series, the continuous wavelet transform (CWT) of $s(t)$ is defined as;

$$W(\alpha, \beta) = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{+\infty} s(t)\psi^*\left(\frac{t-\beta}{\alpha}\right)dt \quad (3.46)$$

where $W(\alpha, \beta)$ are the wavelet coefficients and ‘*’ corresponds the complex conjugate function of ψ which is the mother or original wavelet. The wavelet transform is a three dimensional space, including scale, time and wavelet spectrum ($[W(\alpha, \beta)]^2$).

3.4.5 Discrete Wavelet Transform

The CWT calculation requires a significant amount of computation time and resources. Considering the discrete nature of observed hydrological time series, discrete wavelet transformation (DWT) is preferred most hydrological forecasting. In

DWT, scale and translation parameters are usually based on powers of two (dyadic) instead of every possible scale and translation. Therefore, CWT equation can be written as follows:

$$W(\alpha, \beta) = \frac{1}{\sqrt{\alpha}} \sum_{t=1}^{N-1} s(t) \psi^* \left(\frac{t - \beta}{\alpha} \right) \quad (3.47)$$

where N is the length of discrete signals and $W(\alpha, \beta)$ is the wavelet coefficient for the discrete wavelet of scale $\alpha = 2^m$ and translation $\beta = 2^m n$.

The integers m and n control the wavelet dilation and translation respectively. Substituting for α and β , the equation becomes:

$$W(m, n) = 2^{-m/2} \sum_{t=1}^{N-1} s(t) \psi^*(2^{-m}t - n) \quad (3.48)$$

3.4.6 Mother wavelets

There are different types of mother wavelets. In water engineering, the most important characteristic of each mother wavelet is its shape. Appropriate selection of mother wavelet is the main concerns of hybrid wavelet modelling in hydrology. The best suggestion is trial and error procedure by applying different types of mother wavelets. Perhaps, the best selection could be based on similarity between the shape of river flow time series and its associated wavelet coefficients. However, this matter needs extensive investigation in future studies. Some of the most important mother wavelets are Haar, Daubechies (db), Mexican-hat, Symlets, Coiflets (Coifmann), Morlet, Mallat and Trous transform algorithm (Iyengar et al., 2002). Based on the nature of the signal and the purpose of analysis, the best mother wavelet can be applied.

Daubechies wavelet (db), the most popular mother wavelet, was first introduced by Ingrid Daubechies (Daubechies, 1992). Daubechies family is orthogonal and unsymmetrical. This mother wavelet is often presented by dbN (N is the applied order) . Figure 3.12 presents a few samples of Daubechies families.

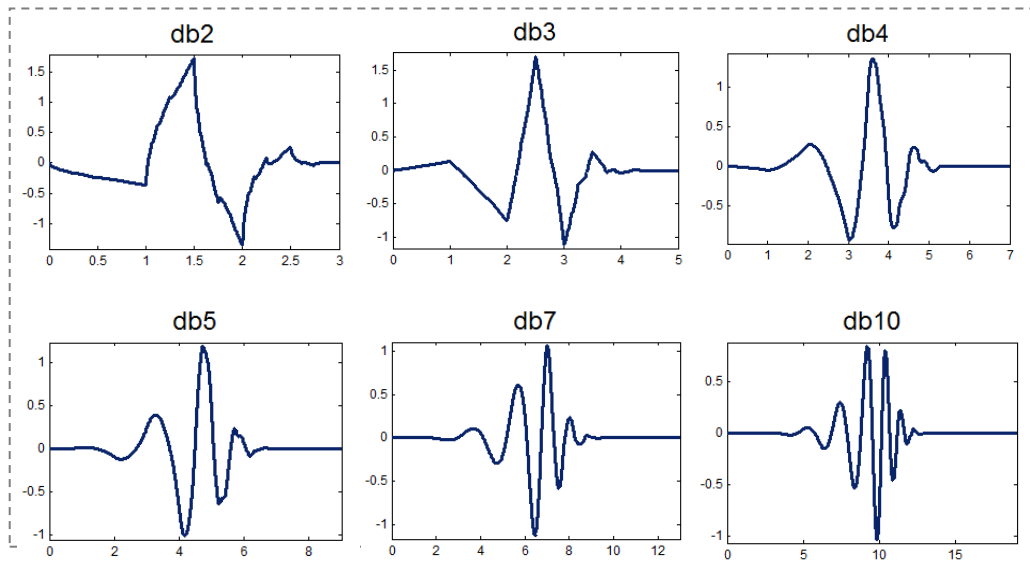


Figure 3.12 Daubechies family wavelet.

Practically the order of Daubechies wavelet is ranging from one to twenty. These orders represent the number of vanishing moments of the mother wavelet. However Daubechies order one (db1) is a distinction of other Daubechies order. It is called Haar mother wavelet, the first and simplest wavelet. The simple function of Haar wavelet is as follows;

$$\psi_t \begin{cases} 1 & \text{if } 0 \leq t < 0.5 \\ -1 & \text{if } 0.5 \leq t < 1 \\ 0 & \text{Other values of } t \end{cases} \quad (3.49)$$

This wavelet simplicity has the advantage of being fast and memory efficient. Figure 3.13 illustrates the Haar mother wavelet function.

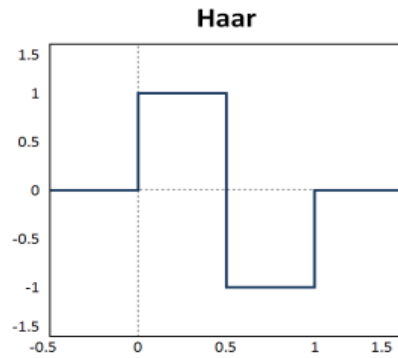


Figure 3. 13 Haar mother wavelet function.

For improving the unsymmetrical characteristic of Daubechies wavelet, Symmlets wavelet was proposed by Daubechies. The number of vanishing moments and the size of Symmlet wavelet are similar to the db wavelets, but it has more symmetry.

Another popular wavelet is the Coiflets, designed by Dubechies. This wavelet named after Ronald Coifman for his proposed wavelets with equal number of vanishing moments and scaling functions. Figure 3.14 presents the shape of five different mother wavelet.

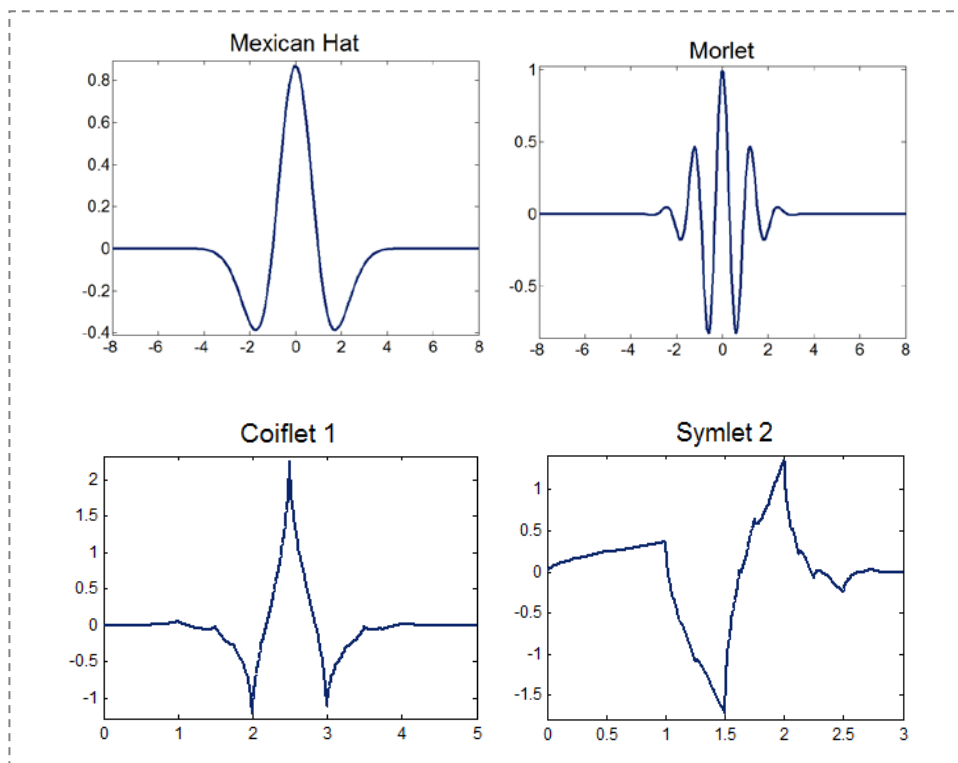


Figure 3. 14 Mexican Hat, Morlet, Coiflet1 and Symlet2 mother wavelets.

3.4.7 Time Series Decomposition by Wavelet

In DWT, the original signal, passes through two complementary filters and emerges as one approximation and one detail components. Approximation (A) is the high-scale, low frequency and details (D) are the low-scale, high frequency components of the signal. Normally approximation is the most important part of the signal that represents the background information of data. This decomposition process can be repeated to reach different resolution levels (n). Figure 3.15 shows the diagram of multi-resolution analysis of a signal to three levels of decomposition.

$$\text{Signal } (S) = \text{Approx}_n(t) + \sum_{i=1}^n \text{Detail}_i(t) \quad (3.50)$$

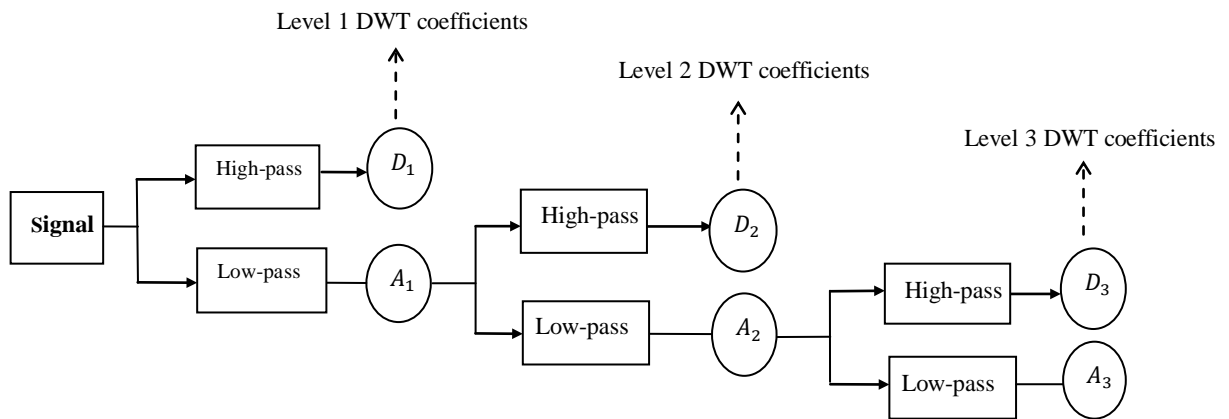


Figure 3. 15 Diagram of multi-resolution analysis of signal.

For more illustration, Figure 3.16 shows Daubechies order three wavelet coefficients of a noisy signal to four level of decomposition and Coiflet order one wavelet coefficients of another signal with three level of decomposition. It's clearly visible that the approximation presents the main pattern of the signal and details are the different frequencies that signal contains. By adding this wavelet coefficients the original signal would be reconstructed.

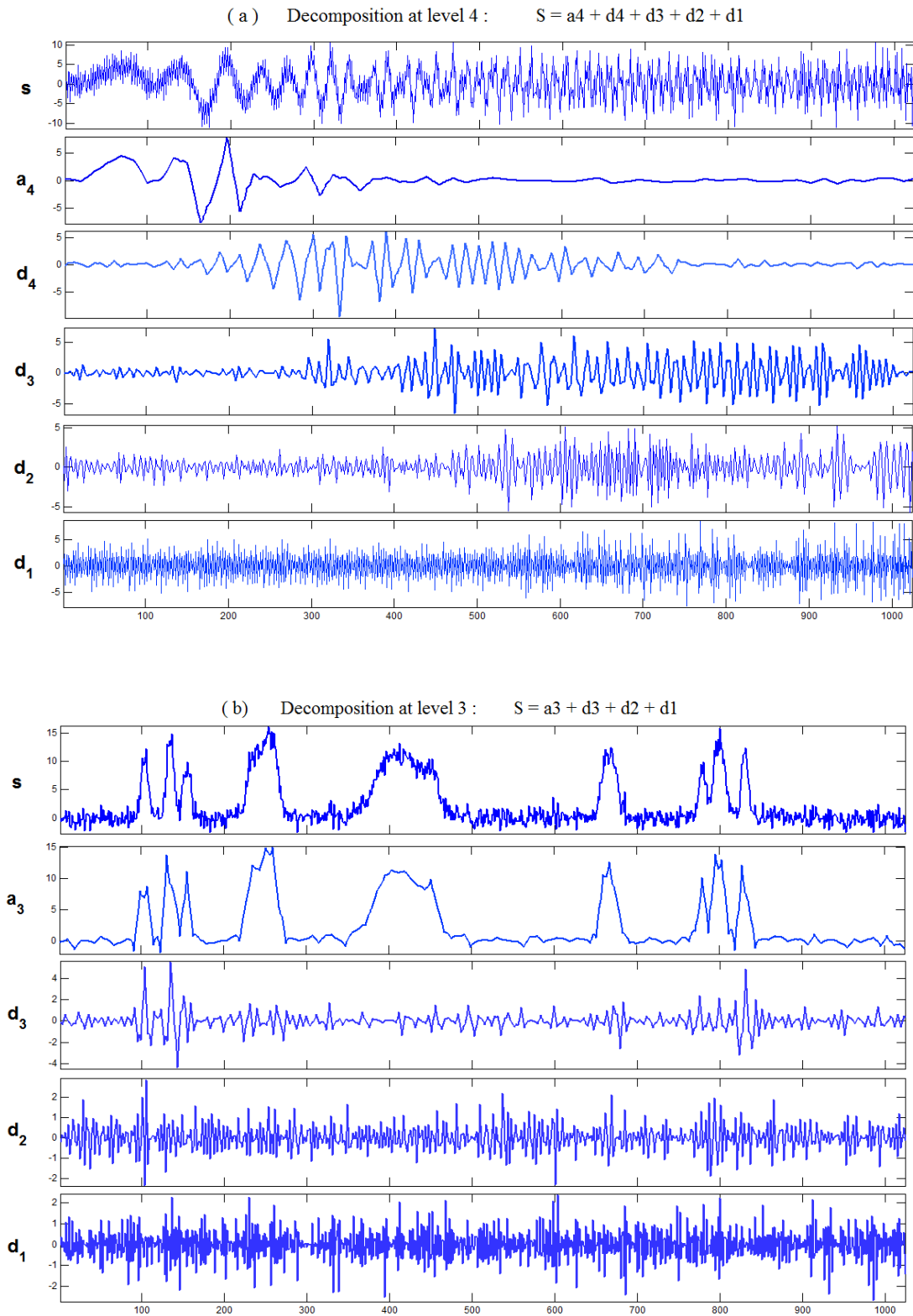


Figure 3. 16 Two noisy signals and their (a) Daubechies3 (b) Coiflet1 wavelet coefficients.

3.5 SUMMARY

In this chapter, the background theory of artificial neural networks, fuzzy modelling and hybrid wavelet models, with special references to their application in hydrology and river flow forecasting, were briefly reviewed. The structure and theory of selected CI methods, applied in developed models, were illustrated in more details. Therefore, detailed methodologies of three layered feed-forward, back propagation neural networks with Levenberg-Marquardt training algorithm were explained. The Structure of three different types of adaptive neuro-fuzzy inference system, namely, grid partitioning, subtractive and C-mean clustering, all based on Sugeno fuzzy, are given. Different type of mother wavelets and wavelet time series, including Haar, Daubechies (db), Mexican-hat, Symlets, Coiflets (Coifmann) and Morlet transform algorithm were illustrated. The continuous and discrete wavelet multi-resolution analysis and decomposition were also described.

Chapter 4

Structure of Proposed Hybrid Models

4.1 INTRODUCTION

As mentioned before in this study the performance of river flow forecasting is improved when computational intelligence models are applied in conjunction with wavelet as a signal processing tool. Four main hybrid models are developed in this study, namely, wavelet neural networks, wavelet neuro-fuzzy model with grid partitioning, wavelet neuro-fuzzy model with subtractive clustering and wavelet neuro-fuzzy model with C-mean clustering.

4.2 WAVELET NEURAL NETWORKS

The proposed wavelet neural network model (WNN) is an integrated model with the input pre-processed by the discrete wavelet transform (DWT). In other words, this hybrid model incorporates two main sub-models of neural networks and wavelet decomposition. The output of the wavelet sub-model is imposed as input to the artificial neural networks sub-model. The model output is un-decomposed, N step ahead river flow time series. Figure 4.1 illustrates the structure of hybrid wavelet neural networks model.

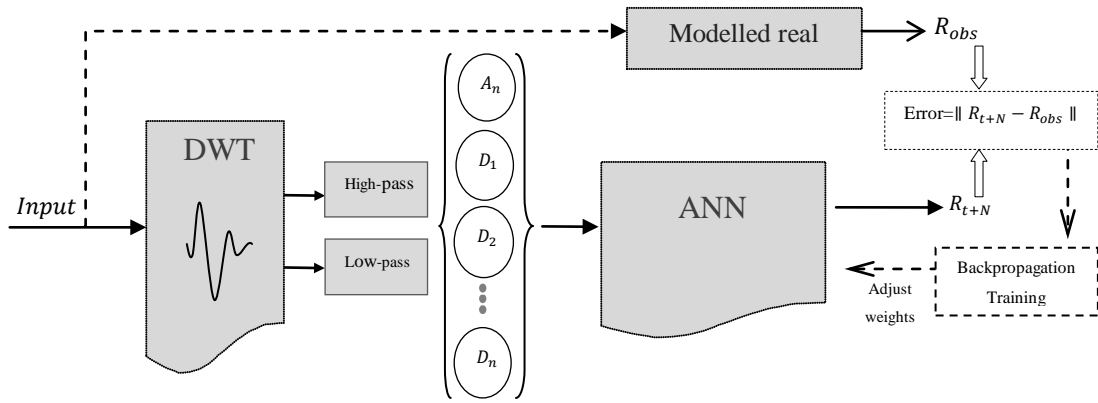


Figure 4. 1 Structure of the proposed hybrid WNN model for N step ahead forecasting.

4.2.1 Neural networks sub-model

The neural networks developed in this study is a three layered feed-forward backpropagation neural networks (BPNN), which is most commonly used in water resources engineering. To improve the network modelling performance and reduce the chance of training process trapped in a local minima, data normalization is applied in ANN input and output by using equation below:

$$x_n = 0.001 + 0.99 * \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (4.1)$$

where x_n is the normalized value of x_i , x_{max} and x_{min} are measured maximum and minimum values of the time series. The small number of 0.001 is also added to the time series to achieve feasible modelling in case of having zero values in the time series, especially for seasonal river flow modelling.

To avoid local minima problems, input data are also divided into two sets of training and validation. The ratio of training and validation data set can be defined as a model input. To verify the accuracy of model the validation set in this model is a completely independent data set, which has no role in training the networks, so by checking the performance criteria for validation data set, the reliability of the model can be confirmed.

To achieve optimal trainings of the ANNs, different training algorithms, namely Levenberg-Marquardt, gradient descent, gradient descent with the adaptive learning rule and Bayesian regularization are investigated. The transfer functions for the hidden and output layers are Tan-sigmoid and linear functions, respectively. As discussed in Chapter three, optimum number of the hidden neurons, need to be determined for different case studies. Having a greater number of hidden neurons, gives the network the required flexibility to solve more complex problems, while having too many neurons may cause over fitting (Tetko et al., 1995). In this study the number of hidden neurons is determined by trial and error. Considering the volume of input data, the minimum and maximum number of hidden neurons are defined for each study. The number of hidden neurons is increased from minimum to maximum number with a step size of one in each trial. The stopping criteria of each trial is set as the number of epochs. The flow chart of ANN modeling is given in the Figure 4.2.

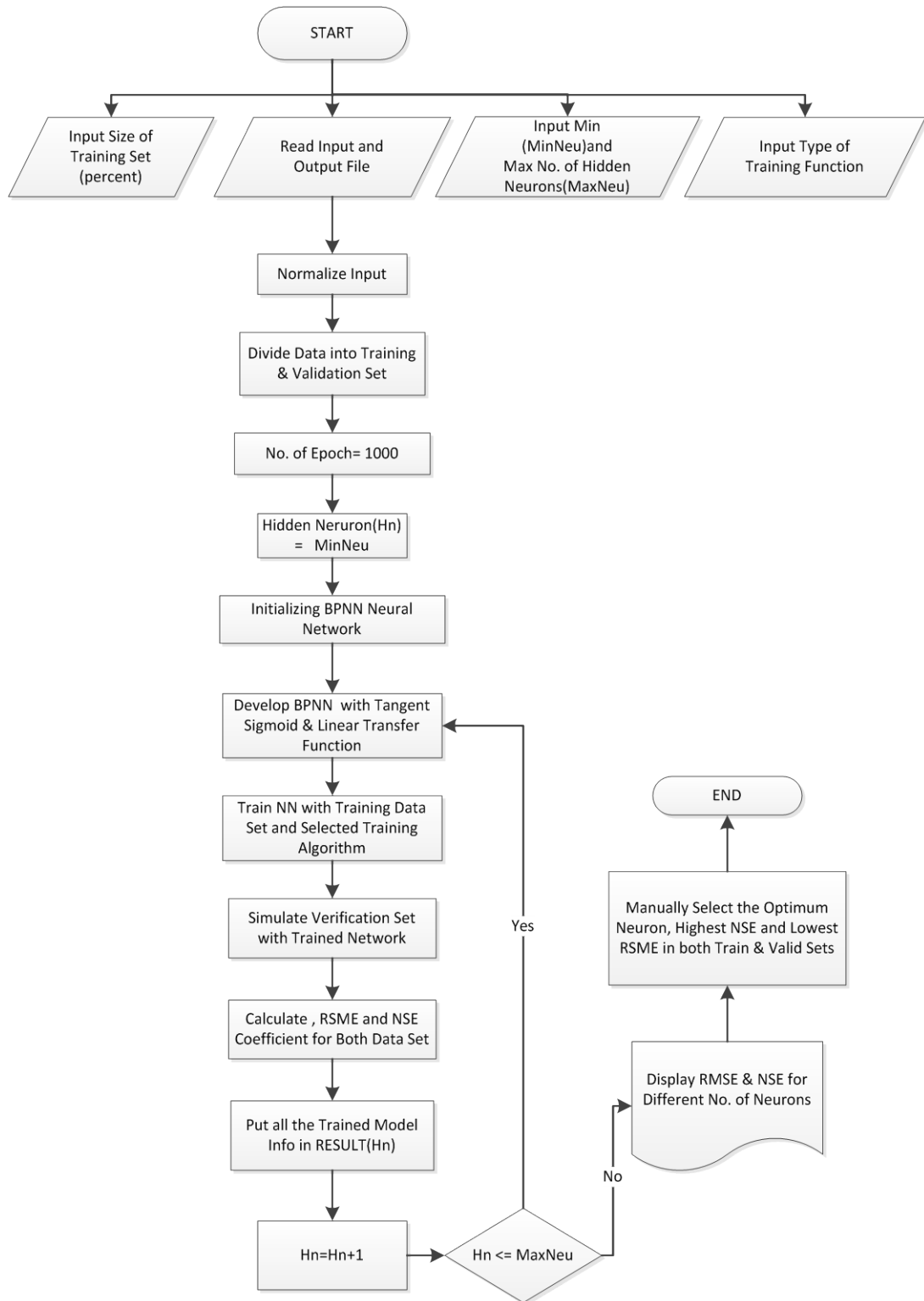


Figure 4. 2 Feed-forward neural network sub-model flow chart.

4.2.2 Wavelet sub-model

The wavelet multi-resolution analysis is implemented as a sub-model. The original time series are decomposed to their low and high frequency components (approximation and details). This sub-model has the ability to apply different mother wavelets and decompose the signal in different levels.

There is no direct way of determining the optimum mother wavelet or level of decomposition. The most important characteristic of each mother wavelet is its shape as wavelet coefficients are scaled and shifted version of mother wavelets. In this study, the appropriate type of mother wavelet is selected by comparing different hybrid models performance. Different mother wavelets of Haar, Daubechies, Symlet and Coiflet, with different shapes, are employed in hybrid modelling. In this study the optimum level of decomposition is determined by trial and error. However, some studies suggest $\text{int} [\log (n)]$ level of decomposition, where n is the length of time series (e. g. Wang and Ding, 2003). Therefore, for different river flow data, this equation is considered as the average optimum level of decomposition.

After decomposing the input signals to desired wavelet coefficients, these sub-series will be fed into the subsequent sub-model (ANN). Each of this sub-series has a different role in the time series and it is important to keep all of them as the neural networks' inputs. Figure 4.3 is the flowchart of hybrid wavelet neural networks model.

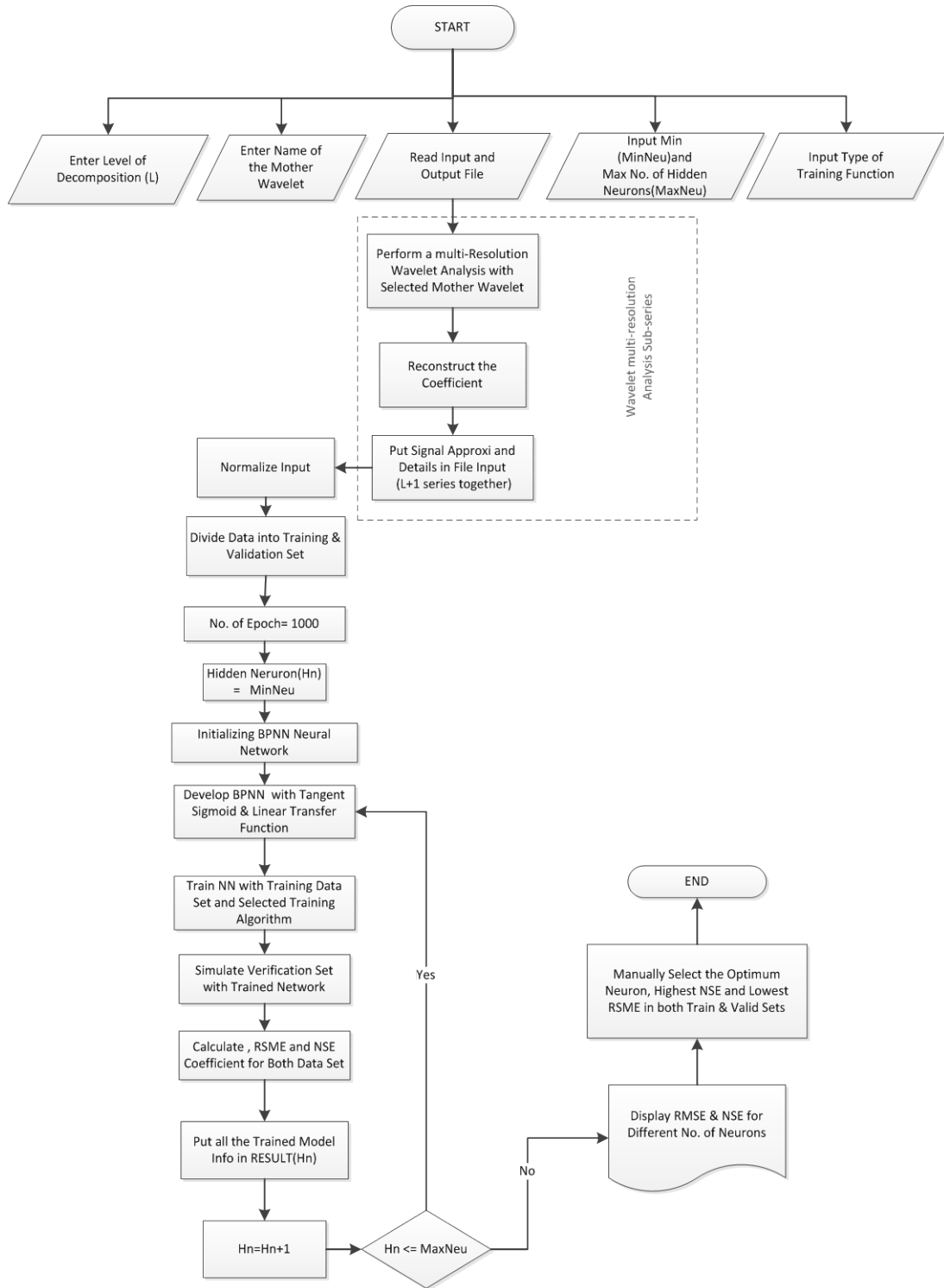


Figure 4. 3 hybrid wavelet neural networks model flow chart.

4.3 WAVELET NEURO-FUZZY WITH GRID PARTITIONING

The proposed wavelet neuro-fuzzy (WNFG) is an integrated one with the input pre-processed by the discrete wavelet transform (DWT). Similar to WNN this hybrid model has two main sub-models. The first sub-model is wavelet multi-resolution analysis and the second sub-model is adaptive neuro-fuzzy inference system. Figure 4.4 illustrates the structure of hybrid wavelet neuro-fuzzy model. The model output is the un-decomposed, N step ahead river flow time series

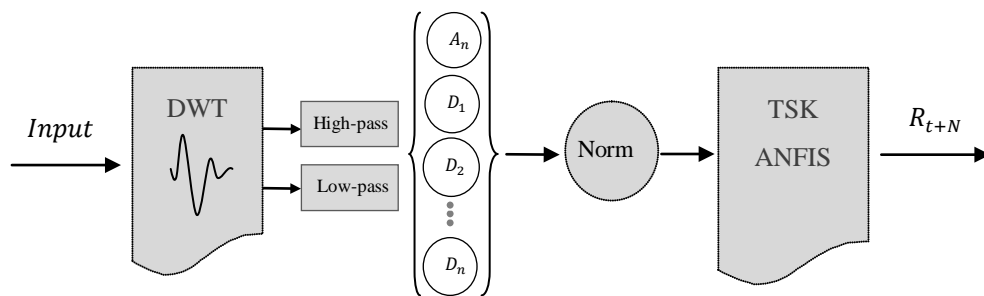


Figure 4. 4 Structure of Wavelet Neuro-Fuzzy hybrid model for N step ahead forecasting.

4.3.1 ANFIS sub-model with grid partitioning

The computational intelligence sub-model in this hybrid approach is ANFIS model with grid partitioning. As discussed in previous chapter, ANFIS is a feed-forward network that finds corresponding fuzzy if-then rules, for achieving optimum model. The structure of ANFIS is equivalent to the TSK first-order fuzzy model. In this hybrid model, Grid partitioning approach is applied for initializing the design of a fuzzy inference system. This sub-model is able to apply different types of membership functions such as generalized bell shaped, Gaussian and triangular. The optimum number of fuzzy rules is determined by trial and error. The number of fuzzy rules is increased from two to the maximum defined number of fuzzy rules.

At first step data normalization is applied to the input data as an essential part of the ANFIS training. The reliability of the model is established by evaluating performance criteria in the validation data set. The validation set in this model is completely independent data set and plays no role in the training process. Checking

the performance criteria in validation data set, the reliability of the model could be confirmed. The stopping criteria of each trial is set as the number of epochs. Figure 4.5 shows the flow chart of this sub-series.

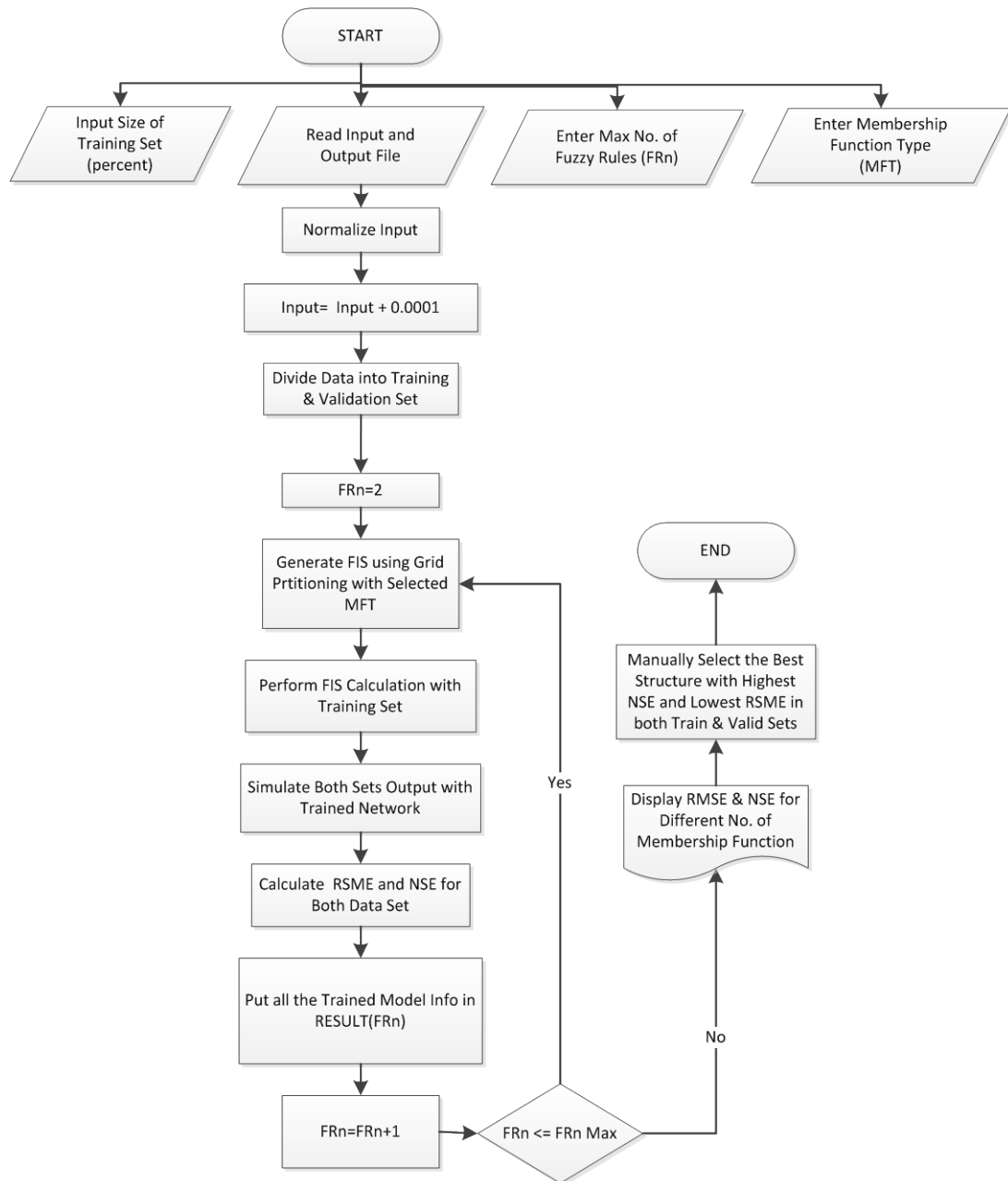


Figure 4.5 Adaptive neuro-fuzzy with grid partitioning sub-model flow chart.

4.3.2 Wavelet sub-model

The wavelet sub-model here is similar to WNN model. In wavelet sub-model the original time series are decomposed to their low and high frequency components (approximation and details). This model has the ability to apply different mother wavelet and is able to decompose the signal in different levels. The maximum feasible level of decomposition in a wavelet neuro-fuzzy model is however less than that in WNN model due to the fuzzy system constraints on the volume of input data. After decomposing the river flow signals to associated wavelet coefficients, these sub-series will be imposed to the ANFIS model. Figure 4.6 shows the flow chart of hybrid wavelet neuro-fuzzy model with grid partitioning.

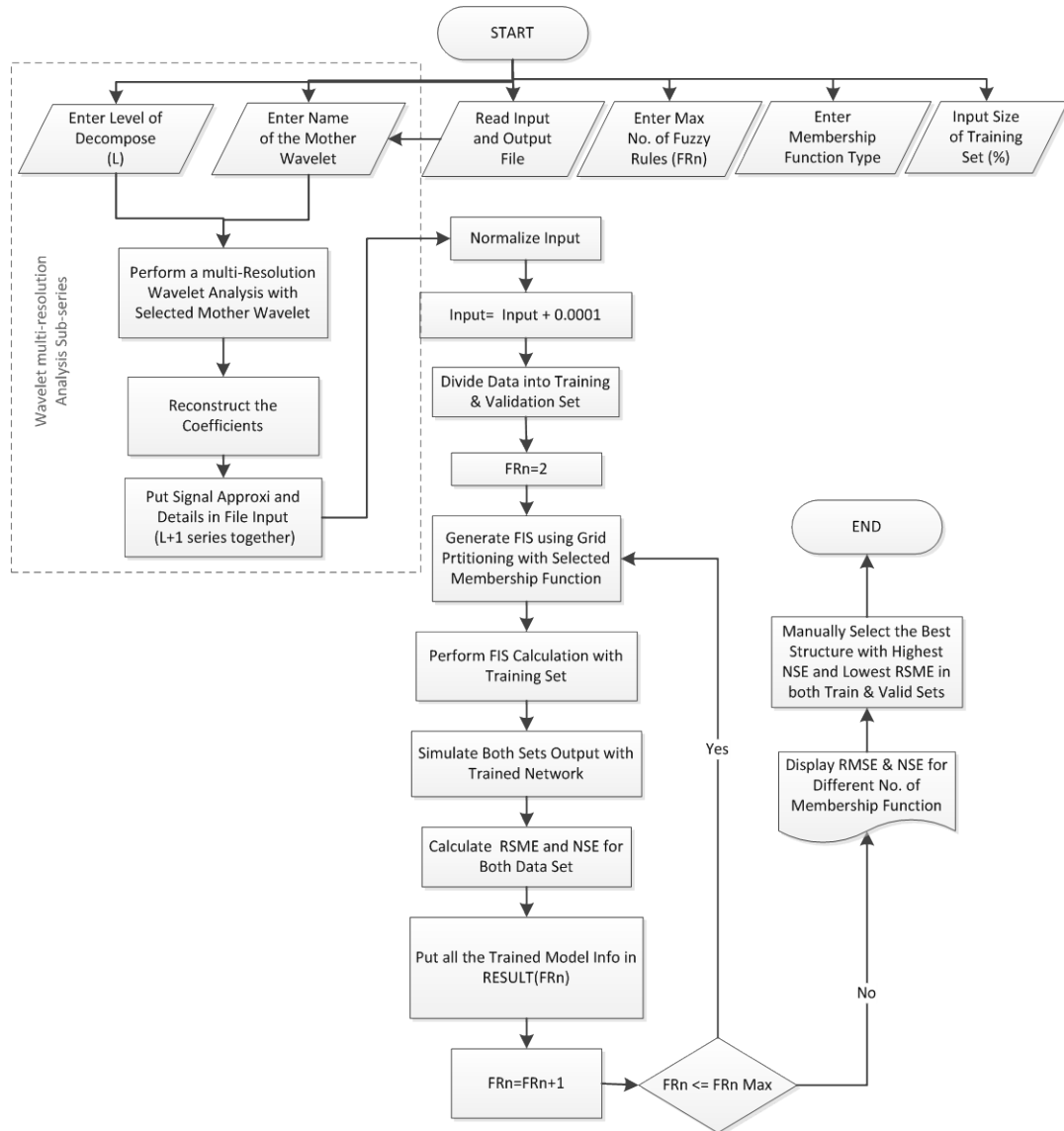


Figure 4. 6 Flow chart of hybrid wavelet neuro-fuzzy model with grid partitioning.

4.4 WAVELET NEURO-FUZZY WITH CLUSTERING

Two other hybrid WNF models are also developed with similar structures to the discussed hybrid WNFG model. In these models, subtractive clustering and C-mean clustering approach are applied for initializing the design of a fuzzy inference system in WNFS and WNFC hybrid models respectively.

4.4.1 Hybrid wavelet neuro-fuzzy model with subtractive clustering (WNFS)

The ANFIS sub-model in this hybrid model applies subtractive clustering method for initializing the FIS. Similar to WFNG, data normalization is the first and essential step of training the model, which improves the model performance. In this sub-model subtractive clustering method is used to generate a TSK fuzzy inference system. The subcluster function first determines the number of rules and antecedent membership functions and then uses linear least squares estimation to determine each rule's consequent equation. Each input and output has as many membership functions as the number of clusters that has been identified. Neighborhood radius (r_a), which specifies a cluster centre's range of influence in each input data dimension, are considered equal for all dimensions and determined by trial and error. In this sub-model, the neighborhood radius is increased from 0.1 to maximum selected amount (less than 1) with a defined step size (e.g. 0.05) in each trial to reach the best performance model. Figure 4.7 illustrates the flow chart of this sub-series.

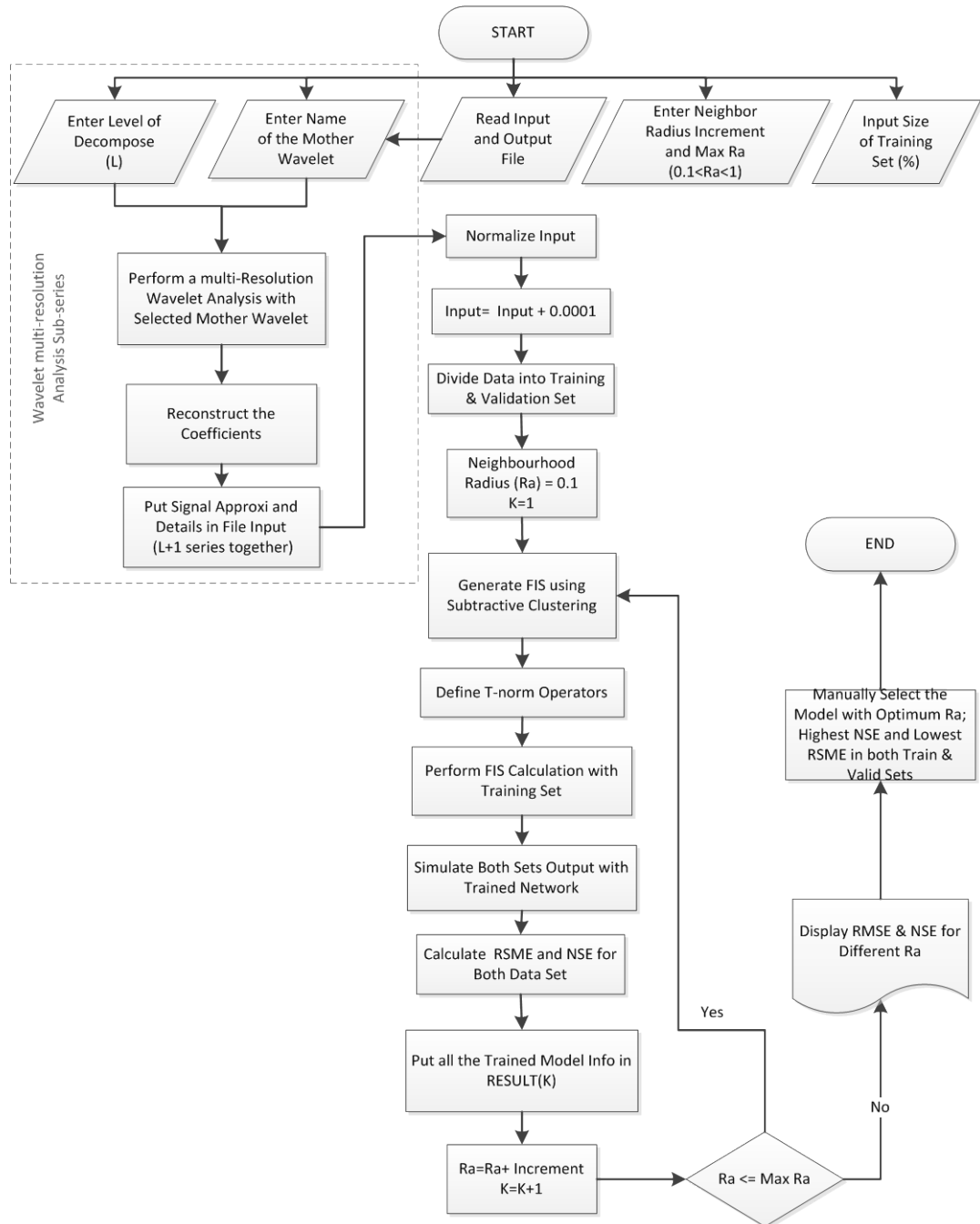


Figure 4. 7 Flow chart of hybrid wavelet neuro-fuzzy model with subtractive clustering.

4.4.2 Hybrid wavelet neuro-fuzzy model with fuzzy C-mean clustering (WNFC)

This model uses fuzzy C-mean clustering (FCM) for initialising Sugeno fuzzy inference system. The input of this hybrid model is also normalized wavelet coefficients. In FCM, each data belongs to more than one cluster with a different

degree of membership. Model extracts a set of rules by FCM algorithm. FCM determines the number of rules and membership functions based on the number of clusters. In the developed model, optimum number of clusters is determined by trial and error by increasing this number from selected minimum (e.g. Two) to the maximum allocated with a step size of one in each trial. The model with highest performance, offers the optimum number of clusters for the modeled river flow time series. The flow chart of this hybrid model is given in Figure 4.8.

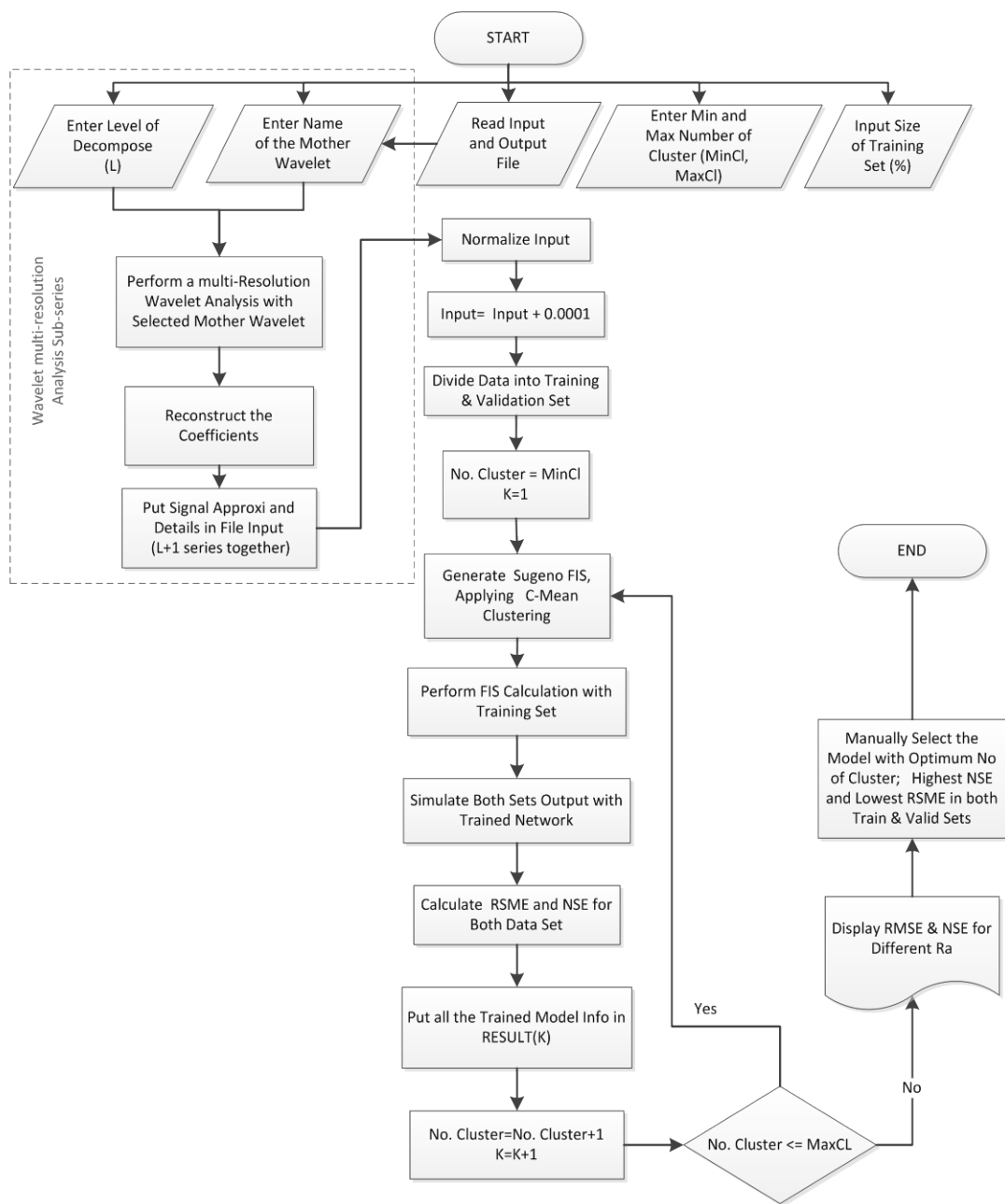


Figure 4. 8 Flow chart of hybrid wavelet neuro-fuzzy model with FCM clustering.

4.5 PERFORMANCE CRITERIA

The evaluation of model performance is based on the difference between simulated and observed values. There are a number of performance criteria such as the bias (B), percent bias (PB), mean absolute error (MAE), relative mean absolute error (RMAE), mean square error (MSE), root mean square error (RMSE), variance (Var), RMSE-Standard deviation ratio (RSR), Nash-Sutcliffe coefficient of efficiency (NSE), agreement index (d), coefficient of correlation (R) and coefficient of determination (R^2).

The bias of a model is the simplest performance criterion which measures the difference between mean observed and forecasted river flow as follows;

$$B = \bar{Q}_{obs} - \bar{Q}_{sim} \quad (4.2)$$

where \bar{Q}_{obs} and \bar{Q}_{sim} are the mean of the observed and forecasted river flow time series respectively.

Percent bias indicates how much forecasted river flow tends to be higher or lower than their associated observed river flow. The positive and negative value of PB detects model underestimation and overestimation, respectively. High-magnitude values indicates model inefficiency.

$$PB = \frac{[\sum_{i=1}^N (Q_{obs}(i) - Q_{sim}(i)) * 100]}{[\sum_{i=1}^N (Q_{obs}(i))]} = \frac{B}{\bar{Q}_{obs}} * 100 \quad (4.3)$$

where Q_{sim} is simulated river flow, Q_{obs} is observed river flow and N is the length of time series.

Mean absolute error, relative mean absolute error, mean square error and root mean square error are other common evaluation criteria which also indicate the error in forecasting. The value of these criteria varies from zero for a perfect simulation to infinity. The magnitude of this value depends on the value of the average and standard deviation of the observed river flow time series. Therefore, they are not suitable for comparing the efficiency of the model in different case studies. In

general MAE and RMAE are less sensitive to peak flows, but could have better representation of model efficiency when there are a very few errors with large gaps between observed and modelled values. These parameters are defined as follows;

$$MAE = \frac{1}{N} \sum_{i=1}^N (|Q_{sim}(i) - Q_{obs}(i)|) \quad (4.4)$$

$$RMAE = \frac{MAE}{\bar{Q}_{obs}} \quad (4.5)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (Q_{sim}(i) - Q_{obs}(i))^2 \quad (4.6)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_{sim}(i) - Q_{obs}(i))^2} = \sqrt{(MSE)} \quad (4.7)$$

RMSE-Standard deviation ratio (RSR) is another error index. RSR is the standardized version of RMSE when it divided by observed time series standard deviation.

$$RSR = \frac{RMSE}{STDV_{obs}} = \frac{\sqrt{\sum_{i=1}^N (Q_{sim}(i) - Q_{obs}(i))^2}}{\sqrt{\sum_{i=1}^N (\bar{Q}_{obs} - Q_{obs}(i))^2}} \quad (4.8)$$

Variance also shows the variability of forecast values and measures the random error.

$$Var = MSE - B^2 \quad (4.9)$$

The Nash-Sutcliffe coefficient of efficiency is one of the most popular criteria for evaluating the hydrological modelling. The value of NSE varies from minus infinity to one, which the model efficiency of one corresponds a perfect forecasting, NSE of zero means that the accuracy of forecasting is equal to the average observed data and the negative values of NSE indicate that using the average of the observed time series is more accurate than model forecasts. In general, the closer the NSE is to one,

more accurate the model becomes. This parameter is also very useful in depicting the scatter and residual plot between observed and modelled river flow time series (Han, 2011).

$$NSE = 1 - \frac{\sum_{i=1}^N (Q_{sim}(i) - Q_{obs}(i))^2}{\sum_{i=1}^N (Q_{obs}(i) - \bar{Q}_{obs})^2} = 1 - \frac{MSE}{STDV_{obs}^2} \quad (4.10)$$

Agreement index is the ratio of MSE and potential error (PE) which is defined by the following equation. The denominator of this equation is dependent on the forecasted and observed time series rang and is applied to standardized the MSE (Ji and Galoo, 2006). Agreement index varies from 0 for the model without correlation and 1 for the perfect model. However, this coefficient is not reliable and may have a near 1 value even for the poor model correlation.

$$d = 1 - \frac{\sum_{i=1}^N (Q_{sim}(i) - Q_{obs}(i))^2}{\sum_{i=1}^N (|Q_{sim}(i) - \bar{Q}_{sim}| + |Q_{obs}(i) - \bar{Q}_{obs}|)^2} \quad (4.11)$$

Correlation coefficient or Pearson correlation measures the strength of the linear relationship between forecasted and observed time series. R varies from -1 to 1. Zero value for R indicates that there is no linear relationship between modelled and observed time series. Coefficient of determination is the squared value of the Pearson correlation. Consequently, the range of coefficient of determination lies between 0 and 1. The efficiency of the model enhances as the value of R^2 increases and the optimal modelling occurs when R^2 reaches 1. In general a model with the R^2 greater than 0.5 is considered as an acceptable match to the real system (Moriasi et al., 2007). In the case of linear regression this coefficient is equivalent to NSE coefficient (Han, 2011).

$$R^2 = \left(\frac{\sum_{i=1}^N (Q_{obs}(i) - \bar{Q}_{obs})(Q_{sim}(i) - \bar{Q}_{sim})}{\sqrt{\sum_{i=1}^N (Q_{obs}(i) - \bar{Q}_{obs})^2} \sqrt{\sum_{i=1}^N (Q_{sim}(i) - \bar{Q}_{sim})^2}} \right)^2 \quad (4.12)$$

Selecting the ideal performance criteria depends on the application. Among all these performance criteria, the Nash-Sutcliffe coefficient of efficiency and coefficient of determination are very sensitive to the peak flows (Krause et al., 2005). Therefore, in this study root mean square error (RMSE) and Nash-Sutcliffe coefficient of efficiency (NSE) are considered as the two main performance criteria in developed models. Both correlations and errors, between observed and modelled variable, are clearly measurable using these performance criteria.

4.6 SUMMARY

In this chapter, the structures of developed CI models and different steps of training process were explained. The structure of ANN and ANFIS models along with proposed hybrid wavelet models were described. Different aspects of modelling such as data normalization, data partitioning and stopping criteria were defined. In ANN-based models, the method of determining optimum number of hidden neurons, selecting training algorithm and transfer function were discussed. In fuzzy-based models, the procedure of determining type and optimum number of membership functions, number of fuzzy rules, optimum size of neighbourhood radius and number of clusters were clarified. Selecting the mother wavelet and optimum number of decomposition levels, as the most important aspect of data pre-processing in hybrid wavelet models, were also discussed. Various types of performance criteria were addressed and compared. Root mean square error and Nash-Sutcliffe coefficient of efficiency are selected as two most efficient criteria to achieve research objectives. Schematics of different models' flow chart were also provided.

Chapter 5

Daily River Flow Forecasting Using Multivariate Inputs

Extended from:

Badrzadeh, H. and Sarukkalige, R., *River flow forecasting using an integrated approach of wavelet analysis and ANN*, Hydrology and water resources symposium, Sydney, 2012.

Badrzadeh, H. and Sarukkalige, R., *Combined wavelet-neural network model for intermittent stream flow prediction*, ASEA-SEC Conference, Perth, 2012.

5.1 INTRODUCTION

In this chapter the application of developed models for daily river flow forecasting is investigated. Back propagation feed-forward (BPFF) neural networks, ANFIS with FCM clustering, hybrid WNN and WNFC models are applied. Also the impact of having multivariate input on model performance is studied by adding the rainfall time series into the input as well as river flow time series (rainfall-runoff modelling). For this reason two different case studies are chosen to achieve a more reliable conclusion for different rivers with different flow regime and rainfall pattern.

5.2 CASE STUDIES

The data of two different rivers, Harvey and Avon River, in the South West of Western Australia (SWWA) are used for this study. SWWA is recognized as one of the top 25 biodiversity hotspots of the world and has the highest concentration of rare and endangered species on the Australian continent. Climate change predictions for the SWWA include decreased rainfall and runoff, increased temperatures, evaporation, seasonal variation and storm intensity which all will have a significant effect on water resources in SWWA (Wilke, 2006). In the selected study areas, like many other basins in Australia, the rural areas are facing surface water shortage that may cause reduction or dispersal of livestock, death of livestock through bogging in dam sediments at low water levels, increasing costs for farm businesses and the community and also reduction in crop area or crop failure particularly those irrigated from locally-sourced water. Furthermore, approximately 10% of agricultural land of Avon River Basin are endanger of moderate to high risk of flooding (Galloway, 2004).

The Harvey River is about 110 kilometers south of Perth city. The Harvey Basin is approximately 2000 km² and includes two irrigation districts. The area has a warm temperate Mediterranean climate. For this study data from the Dingo road station on Harvey River is used. Mean daily river flow and mean daily rainfall for 35 years, with an observation period from 1976 to 2011, are collected from Department of Water and the Bureau of Meteorology. For this reason first 25 years of data (9131 days, around 70% of the whole data set) are used for training and the rest 10 years (3742 days, around 30% of the whole data set) are used for validating the models.

The second river is Avon River, which is the largest inflow to Swan River and has the largest area of all Swan and Canning River subcatchments. Avon River length is 280 kilometers. The full Catchment area is approximately 120,000 km². Average of annual rainfall in Avon River is 850 mm per year and average annual flow of Avon River is 199 GL which is around 79% of total inflow to the Swan River. For this study data from the Northam weir station on Avon River upstream, is used. Mean daily river flow and mean daily rainfall for 32 years, with an observation period between 1978 and 2010, are collected from the Department of Water and Bureau of

Meteorology. Similarly, the first 70% of the whole data set is used for training and the rest is used for validation. In Table 5.1, the daily statistics of the stream flow and rainfall of the selected stations are presented. Historical river flow and rainfall time series of these two stations are also shown in Figure 5.1 to Figure 5.4. It can be seen that extreme values are placed in the training set rather than the verification set to improve model accuracy for extreme flow forecasting as trained model performs better within the training data range (Maier and Dandy, 2000).

Both selected rivers in this study are highly seasonal. The high seasonality trend of Harvey and Avon River is also depicted in Figure 5.5 which presents river flow hydrographs in a few selected years. It is evident that these rivers normally cease flowing during the summer months of the year.

Table 5. 1 Daily statistical parameters of stream flow and rainfall data sets of the Dingo road and Northam weir stations.

Station	Data Set	River flow (1000 m ³ /day)				Rainfall (mm/day)			
		I _{mean}	I _{min}	I _{max}	I _{stdv}	R _{mean}	R _{min}	R _{max}	R _{stdv}
Dingo Road	Training	88.09	0.01	1561	110.7	3.10	0	131.80	7.97
	Validation	55.57	0	703.8	78.76	2.88	0	65.32	7.53
	Total	78.8	0	1561	103.6	3.04	0	131.80	7.85
Northam Weir	Training	419.1	0	28957.2	1278.1	1.05	0	91.00	3.73
	Validation	159.5	0	9366.5	434.1	0.94	0	59.00	3.23
	Total	340.8	0	28957.2	1100.7	1.01	0	91.00	3.67



Figure 5. 1 Map of Northam weir and Dingo road station location in Western Australia (Bureau of meteorology, 2013)

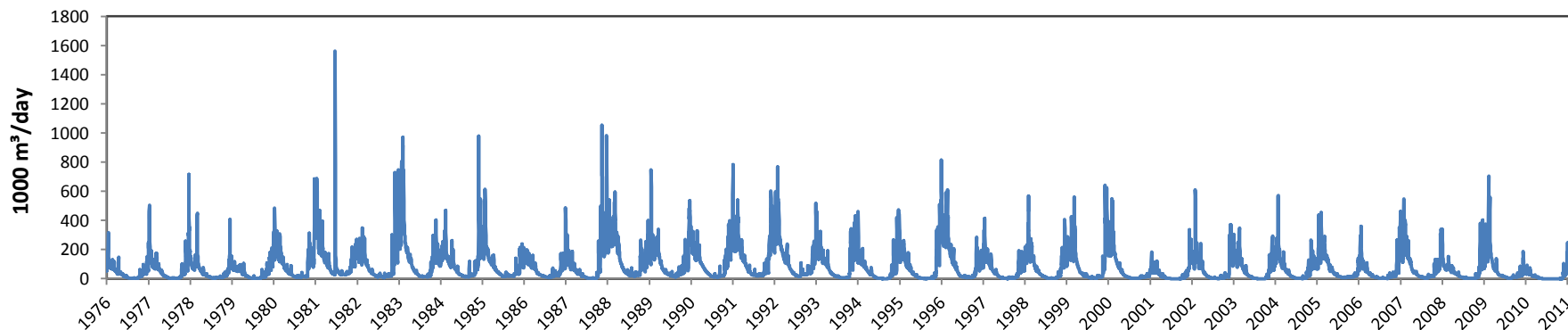


Figure 5. 2 Daily river flow time series at the Dingo road station in the Harvey River, Western Australia (1976-2011).

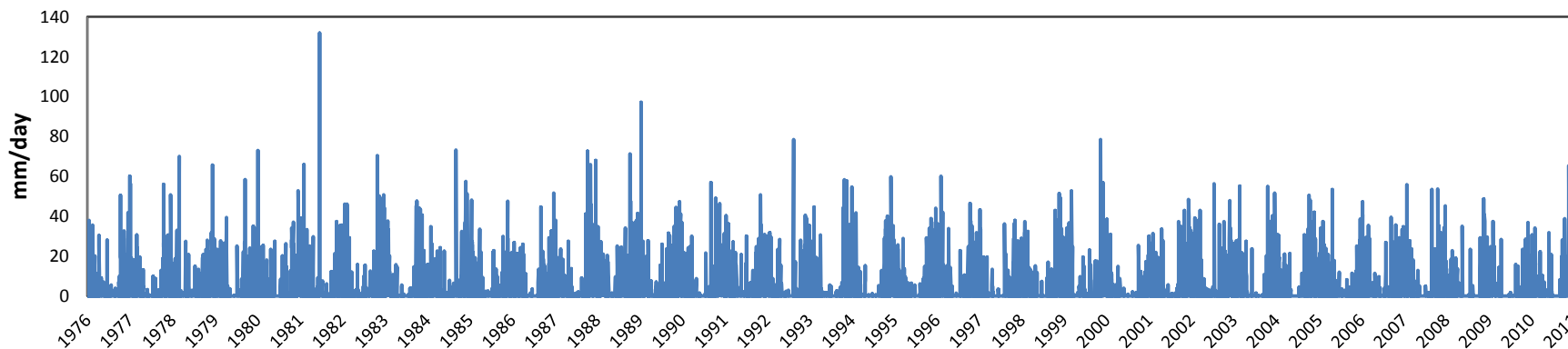


Figure 5. 3 Daily rainfall time series near the Dingo road station in the Harvey River, Western Australia (1976-2011).

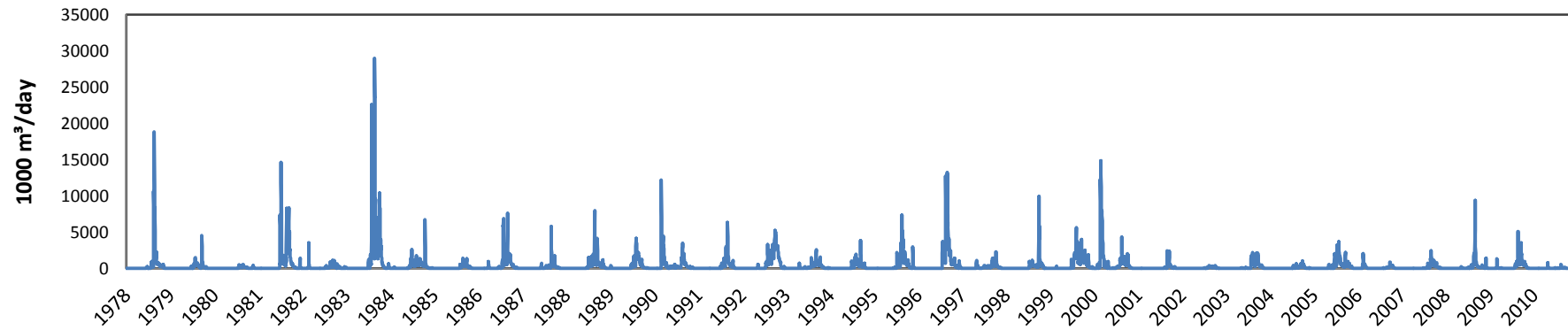


Figure 5. 4 Daily river flow time series at the Northam weir station in the Avon River, Western Australia (1978-2010).

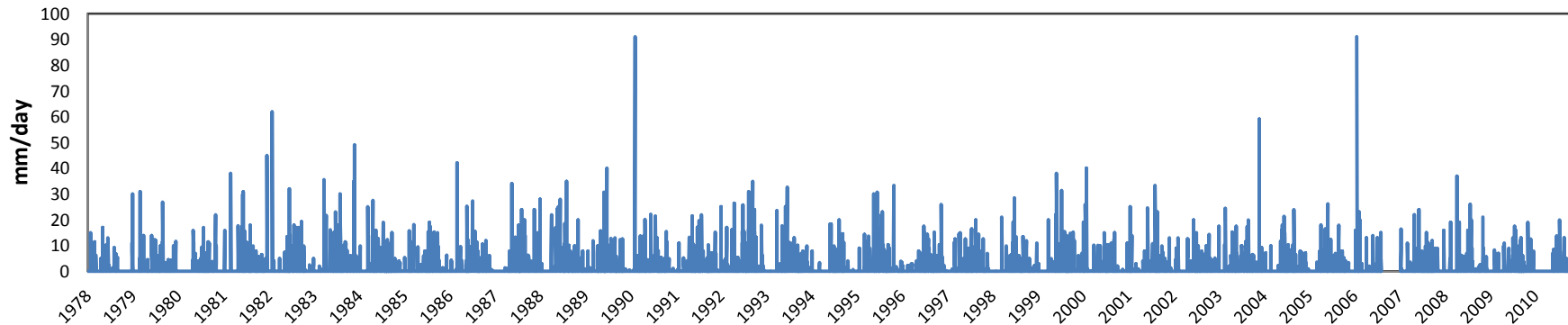


Figure 5. 5 Daily rainfall time series near the Northam weir station in the Avon River, Western Australia (1978-2010).

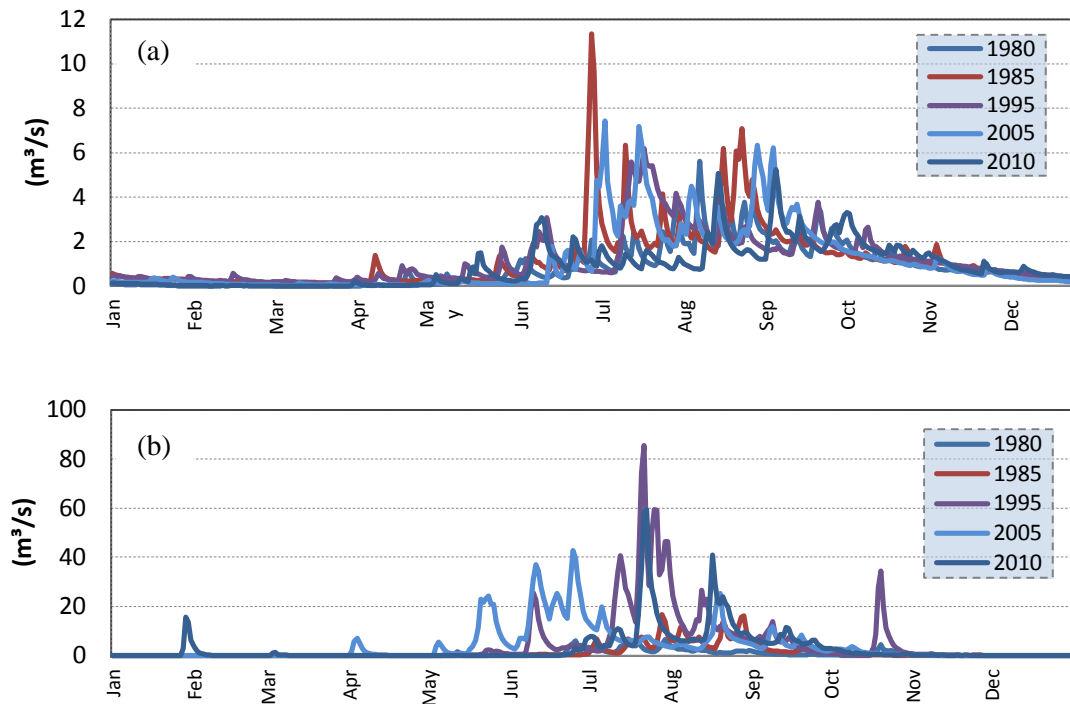


Figure 5. 6 (a) Harvey and (b) Avon River mean daily river flow hydrographs in selected years.

5.3 APPLICATION OF ANN

The three layer feed-forward back propagation neural networks are developed for forecasting. The structure of developed neural networks is described in Chapter four. The number of hidden neurons is determined by trial and error. Different training algorithm of gradient descent with the adaptive learning, Bayesian regularization and LM algorithm is applied for ANN training. Two different data sets are applied as ANN's input for each case study. The first model trained by using just river flow historical time series as its input. In the second model both river flow and rainfall time series are applied as the input. The structure and the performance of the models are presented in Table 5.2. Altering the training algorithm, as discussed in chapter three, forecasting accuracy remained nearly unchanged, while LM algorithm was the fastest among all.

Table 5.2 ANN models structure and performance.

Station	Model	Input Data	Model Structure	Calibration		Validation	
				<i>NSE</i>	RSME(m ³ /s)	<i>NSE</i>	RSME(m ³ /s)
Dingo Road	ANN1	Flow	1-14-1	0.899	0.409	0.893	0.464
	ANN2	Rain&Flow	2-24-1	0.965	0.233	0.951	0.182
Northam Weir	ANN1	Flow	1-13-1	0.849	5.73	0.845	1.98
	ANN2	Rain&Flow	2-16-1	0.851	5.70	0.825	2.11

The results indicate that having both rainfall and runoff as model input considerably improves the performance of the models for Dingo road station data set. The Nash-Sutcliffe coefficient of efficiency is improved from 0.89 to 0.96 and the root mean square error modified from 0.46 to 0.18 m³/s. However, having multivariate input did not have a considerable impact on Northam weir station forecasting model. Figure 5.6 and Figure 5.7 show the scatter plot between the observed and modeled values of daily river flow with these two ANN models.

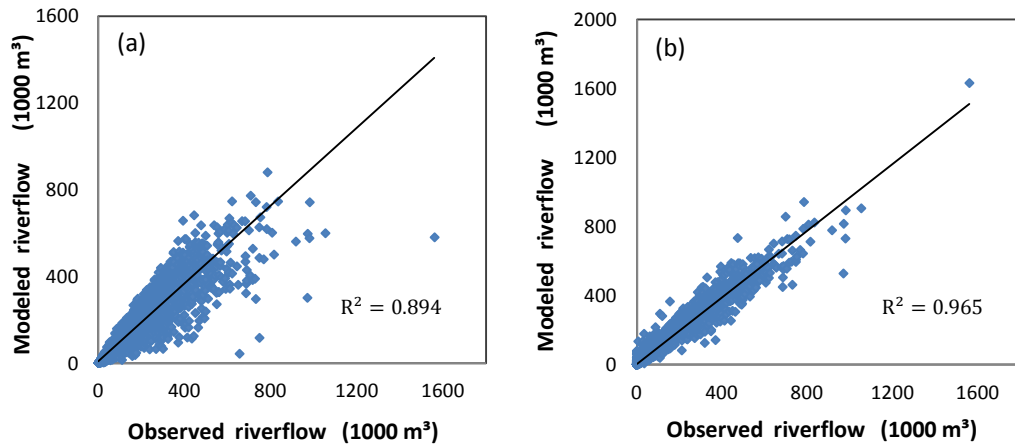


Figure 5.7 Scatter plots between Dingo road station observed and modelled daily river flow: (a) ANN single flow input; (b) ANN with multivariate input.

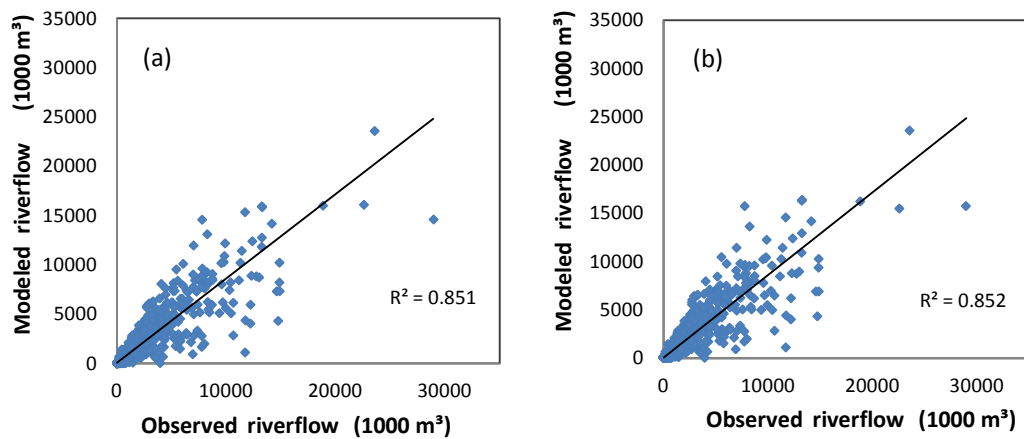


Figure 5.8 Scatter plots between Northam weir station observed and modelled daily river flow: (a) ANN with single input; (b) ANN with multivariate input.

5.4 APPLICATION OF ANFIS

Adaptive neuro-fuzzy inference system model was developed using the FCM clustering method to generate a TSK type fuzzy inference system. The structure of developed neuro-fuzzy model is described in Chapter four. Two different input data sets were used for daily river flow forecasting. Table 5.3 shows different models' performance.

Table 5.3 ANFIS models' performance.

Station	Model	Input Data	Calibration		Validation	
			<i>NSE</i>	RSME(m ³ /s)	<i>NSE</i>	RSME(m ³ /s)
Dingo Road	ANFIS1	Flow	0.916	0.265	0.892	0.438
	ANFIS2	Rain&Flow	0.969	0.162	0.963	0.256
Northam Weir	ANFIS1	Flow	0.829	5.73	0.837	2.257
	ANFIS2	Rain&Flow	0.842	4.173	0.849	2.201

The results indicate that using multivariate input leads to better performance compared to those with using single river flow time series for both case studies. In Dingo Road station, in particular, the Nash-Sutcliffe coefficient of efficiency is 8% improved and the root mean square error modified from 0.44 to 0.26 m³/s in validation set.

5.5 IMPROVING THE EFFICIENCY WITH HYBRID MODELS

As discussed in methodology, both ANN and ANFIS models are combined with wavelet multi-resolution analysis for the purpose of improving the forecasting accuracy. Hybrid wavelet neural network (WNN) and hybrid wavelet neuro-fuzzy model with C-mean clustering (WNFC) are developed. The structure of these hybrid models previously explained in chapter four. Two different mother wavelets are applied for each case study. Haar and db5 mother wavelets are applied on Harvey River and Coiflet1 and db4 mother wavelets are applied on Avon River flow time series. Due to fuzzy inference system input size restrictions, only wavelet coefficients of one level of decomposition lead to feasible training for the Avon River. Table 5.4 and Table 5.5 present the structure and performances of developed models for Harvey and Avon Rivers respectively. The best fitted hybrid WNN and WNFC models for each study area, are also highlighted in these tables.

Table 5. 4 Hybrid models' structure and performance for Dingo road station daily river flow forecast.

Model	Wavelet	Level ¹	Input ² data	Model structure	Calibration		Validation	
					<i>NSE</i>	RSME(m ³ /s)	<i>NSE</i>	RSME(m ³ /s)
WNN1	db5	2	I	3-22-1	0.954	0.276	0.943	0.214
WNN2	db5	3	I	4-24-1	0.961	0.256	0.943	0.214
WNN3	db5	4	I	5-20-1	0.960	0.257	0.938	0.222
WNN4	haar	2	I	3-25-1	0.949	0.290	0.946	0.207
WNN5	haar	3	I	4-19-1	0.959	0.263	0.948	0.205
WNN6	haar	4	I	5-23-1	0.961	0.254	0.936	0.226
WNN7	db5	2	I&R	6-19-1	0.988	0.139	0.983	0.117
WNN8	db5	3	I&R	8-23-1	0.992	0.117	0.984	0.113
WNN9	db5	4	I&R	10-22-1	0.993	0.107	0.982	0.121
WNN10	haar	2	I&R	6-20-1	0.990	0.128	0.982	0.119
WNN11	haar	3	I&R	8-25-1	0.993	0.110	0.982	0.121
WNN12	haar	4	I&R	10-24-1	0.994	0.106	0.983	0.116
WFNC1	db5	2	I	-	0.938	0.257	0.927	0.335
WFNC2	db5	3	I	-	0.943	0.156	0.928	0.550
WFNC3	db5	4	I	-	0.930	0.352	0.935	0.223
WFNC4	haar	2	I	-	0.978	0.129	0.957	0.188
WFNC5	haar	3	I	-	0.955	0.213	0.948	0.357
WFNC6	haar	4	I	-	0.957	0.229	0.944	0.368

Table 5. 5 Hybrid models' structure and performance for Northam weir station daily river flow forecast.

Model	Wavelet	level ¹	Input data ²	Model structure	Calibration		Validation	
					<i>NSE</i>	RSME(m ³ /s)	<i>NSE</i>	RSME(m ³ /s)
WNN1	coif1	2	I	3-16-1	0.94	3.61	0.922	1.41
WNN2	coif1	3	I	4-17-1	0.935	3.20	0.921	1.41
WNN3	coif1	4	I	5-18-1	0.963	2.82	0.921	1.41
WNN4	db4	2	I	3-15-1	0.936	3.73	0.879	1.75
WNN5	db4	3	I	4-14-1	0.957	3.06	0.894	1.63
WNN6	db4	4	I	5-19-1	0.962	2.87	0.915	1.46
WNN7	coif1	2	I&R	6-17-1	0.937	3.72	0.917	1.45
WNN8	coif1	3	I&R	8-20-1	0.957	3.06	0.894	1.63
WNN9	coif1	4	I&R	10-17-1	0.964	2.79	0.926	1.69
WNN10	db4	2	I&R	6-20-1	0.947	3.40	0.886	1.70
WNN11	db4	3	I&R	8-16-1	0.957	3.07	0.896	1.62
WNN12	db4	4	I&R	10-19-1	0.954	3.17	0.906	1.55
WFNC1	db4	1	I	-	0.937	3.43	0.893	1.72
WFNC2	coif1	1	I	-	0.929	3.74	0.891	1.85

These tables reveal that in both case studies, the performance criteria of hybrid models are much better than these of single ANN and ANFIS models. Figure 5.8 shows the scatter plot between the observed and modelled values of Dingo road daily river flow with four different models and illustrates how hybrid models outperform single ANN and ANFIS models.

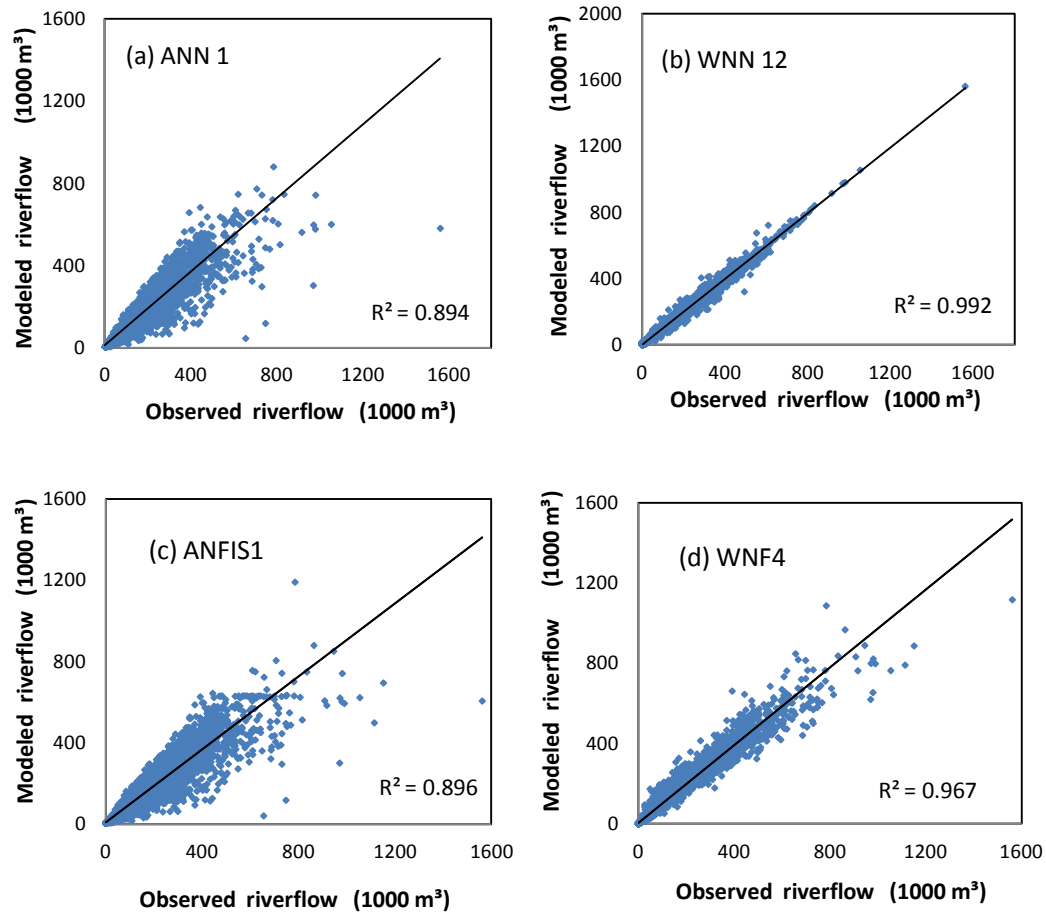


Figure 5.9 Scatter plots between Dingo road station observed and modelled daily river flow with: (a) ANN1; (b) Hybrid WNN12; (c) ANFIS1; (d) Hybrid WNF4.

It is also observed that although there is not considerable differences between performance of different wavelet models, for the first case study the model with Haar wavelet transform, four levels of decomposition and using both rainfall and river flow as ANN inputs, shows the lowest Root mean square error (0.106 to 0.116) and highest Nash-Sutcliffe coefficient of efficiency (0.994 to 0.983) during the calibration and validation period. Therefore, this hybrid wavelet neural networks model (bold in tables) is selected as the best fit model for the Dingo road station forecasting. Figure 5.9 presents the hydrograph of observed and modelled river flow with WNN12. Detailed hydrograph of the last 4 years of validation set from year 2007 to 2011 is also presented in Figure 5.10.

The best fit model for the Northam weir station is the WNN9 with Coiflet1 wavelet transform and 4 levels of decomposition when both rainfall and stream flow are used as ANN inputs. This model shows the lowest RMSE (1.69 to 2.79) and highest Nash-Sutcliffe coefficient of efficiency (0.964 to 0.926) during the calibration and validation stages. Comparing the results also demonstrate that unlike the first case study, adding rainfall time series had no considerable impact on model performance in Northan weir station. Figure 5.11 shows the observed and modelled river flow time series with WNN9. Figure 5.12 also presents the hydrograph of the last 4 years of validation set from year 2006 to 2010, for detailed comparison.

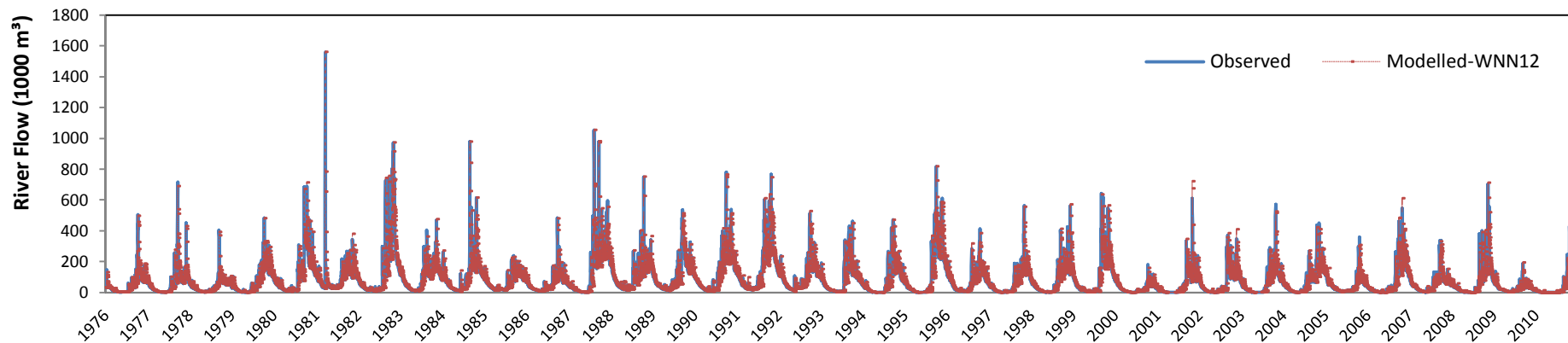


Figure 5. 10 Comparison of the Dingo road observed and predicted daily river flow with WNN12.

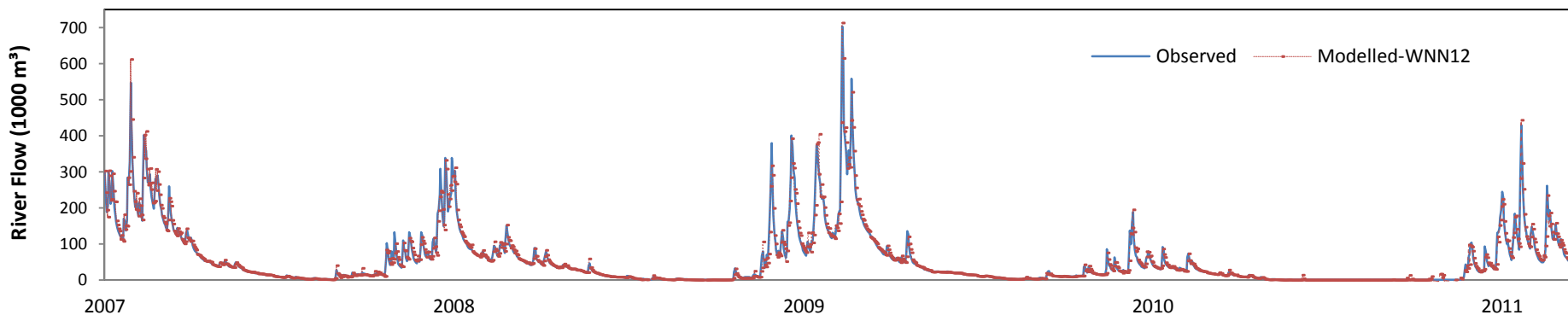


Figure 5. 11 Comparison of the Dingo road observed and predicted daily river flow with WNN12 in the validation set (2007-2011).

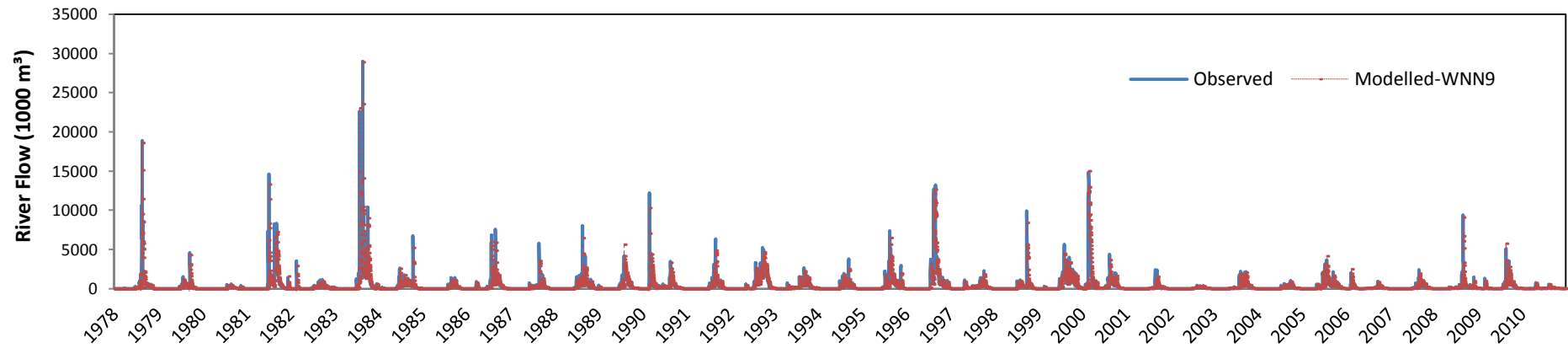


Figure 5.12 Comparison of the Northam weir observed and predicted daily river flow with WNN9.

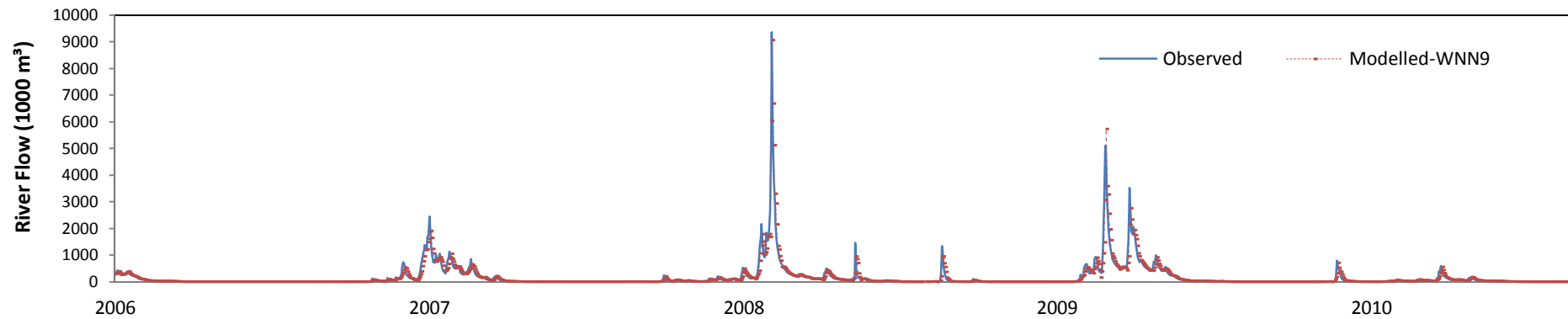


Figure 5.13 Comparison of the Northam weir observed and predicted daily river flow with WNN9 in the validation set (2006-2010).

5.6 CONCLUSION

In this study the application of different data-driven models for daily river flow forecasting of two rivers in Wester Australia was investigated. Having high quality, long historical data, as the most important element of data-driven modelling (Anctil, 2004), very satisfactory result was achieved. The results indicate that both ANN and ANFIS are promising approaches for daily river flow forecasting. Considering highly nonlinear and non-stationary characteristics of river flow time series, the accuracy of forecasting further improved to a very satisfactory level by applying DWT on input data time series.

Altering ANN training algorithm didn't have a notable impact on model accuracy while LM algorithm led to faster convergence.

ANFIS is restricted to the input size, therefore DWT with high level of decomposition could not be applied on hybrid WNFC model for some case studies.

Comparing hybrid models with different wavelet transform and level of decomposition also indicates that, altering the mother wavelet or decomposition level could only slightly improve the forecasting reliability. The most efficient wavelet transform and the optimum level of decomposition depends on the river flow time series characteristics and length and should be determined for each case study.

Adding another hydrological parameter to the model input could improve the model efficiency. This matter closely observed in this study by adding rainfall time series as another model input for two different case studies. However, the performance of the forecasting model for Avon River with higher river flow rate and very low rate of precipitation, didn't improve considerably. It can be concluded that it is important to find and add the effective parameter(s) to the model. The parameters could be upstream rainfall, temperature, evapotranspiration or any other effective hydrological parameter.

Chapter 6

Short Term and Long Term River Flow Forecasting

Extended from:

Badrzadeh, H., Sarukkalige, R. and Jayawardean, A. W., 2013. Improving ANN-based short term and long term seasonal river flow forecasting with signal processing techniques, *River Research and Applications Journal*, doi: 10.1002/rra.2865.

Badrzadeh, H., Sarukkalige, R. and Jayawardean, A. W., *Development of a Wavelet Neuro-Fuzzy Computational Model for Stream Flow Forecasting*, *Nonlinear processes in geophysics*, Under review.

6.1 INTRODUCTION

In this chapter, the application of developed models for both short and long term river flow forecasting is investigated. The performance of river flow forecasting is improved when different input combinations and signal processing techniques applied on multi-layer back propagation feed-forward neural networks and adaptive neuro-fuzzy inference system with grid partitioning. Haar, Coiflet and Daubechies wavelet analysis are coupled with BPNN and ANFIS model to develop hybrid WNN and WNFG models, respectively. Different models in terms of inputs and structure are developed for daily, weekly and monthly river flow forecasting for Ellen Brook River, Western Australia.

6.2 STUDY AREA AND DATA USED

The river flow data of the Railway parade station on Ellen Brook River is used as a case study. The Ellen Brook catchment is located in Western Australia. It is about 20 km from Perth city and 25 km from the coastline (Figure 6.1). The Ellen Brook catchment area is approximately 715 km² and three local governments including shire of Gingin, Chittering and Swan, administer the catchment. Ellen Brook is one of the Swan-Canning estuary sub-catchment which contributes 6% of the total stream flow of the estuary. The average catchment rainfall is 800 mm per year and its average annual river flow is 18.9 million m³. The climate of the catchment is warm temperate Mediterranean type. Climate change predictions for the Ellen Brook catchment include descending trend in rainfall and runoff, and ascending trend in the temperature, evaporation and storm intensity (Wilke, 2006).

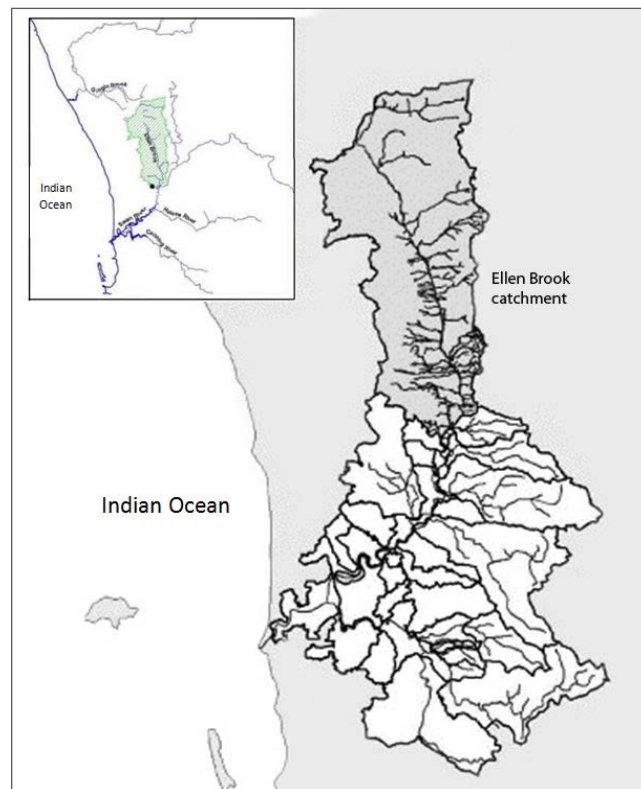


Figure 6. 1 Location of Ellen Brook catchment in the Western Australia.

For this study mean daily river flow discharge for 34 years, with an observation period from 1977 to 2010, are collected from the Department of Water. First 23 years

of data are used for training and the remaining 11 years are used for validation. It is checked that the extreme values are placed in the training set rather than validation set as CI models performance are higher in a forecasting within the data range being utilized during the training phase (Maier and Dandy, 2000). The average daily river flow of the Railway parade station is $0.88 \text{ m}^3/\text{s}$ with a maximum flow of $41.28 \text{ m}^3/\text{s}$ in July 1987 and a minimum flow of zero as expected for a seasonal river. For mid term and long term forecasting weekly and monthly time series are also prepared. Figure 6.2 shows the daily, weekly and monthly historical river flow time series. Statistical analyses of the Railway parade station daily, weekly and monthly river flow data set is given in Table 6.1, which contains the mean, minimum, maximum and standard deviation values. It is evident that extreme values are placed in the training set rather than the verification set to improve model accuracy for extreme flow forecasting as ANN perform better within the training date range (Maier and Dandy, 2000).

Table 6.1 Statistical parameters of Ellen Brook river flow data sets of the Railway parade station.

Railway parade Station	Data Set	Mean	Maximum	Minimum	Standard deviation
Daily River flow (m^3/s)	Training	1.04	41.29	0	2.77
	Validation	0.49	17.95	0	1.33
	Total	0.88	41.29	0	2.45
Weekly River flow (m^3/s)	Training	1.03	20.44	0	2.26
	Validation	0.38	7.66	0	1.01
	Total	0.84	20.44	0	1.99
Monthly River flow (m^3/s)	Training	1.01	8.96	0	1.77
	Validation	0.55	4.47	0	0.97
	Total	0.87	8.96	0	1.58

6.3 INPUT SELECTION FOR MODELS

The input selection for BPNN forecasting was chosen based on forward stepwise selection of inputs and considering the time series with high auto correlation function

(ACF) value. Considering decreasing ACF with increasing time lag (Figure 6.3), different input combination of time series up to 4 steps ($Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}$) for daily and weekly and up to 2 steps (Q_t, Q_{t-1}, Q_{t-2}) for monthly forecasting is applied (Table 6.2 and Table 6.3).

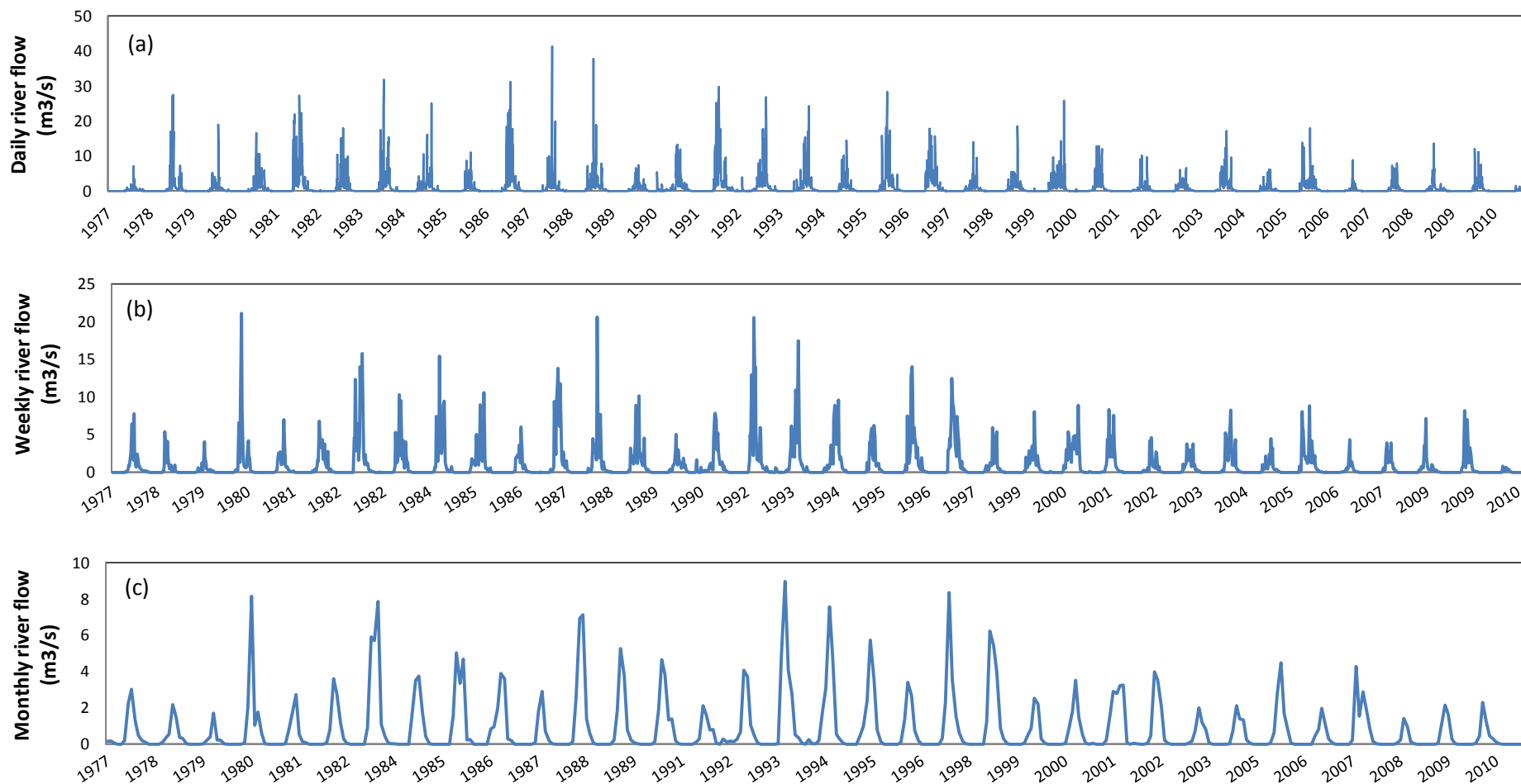


Figure 6. 2 (a) Daily; (b) Weekly and (c) Monthly river flow time series at the Railway Parade station on the Ellen Brook River, Western Australia (1977-2010).

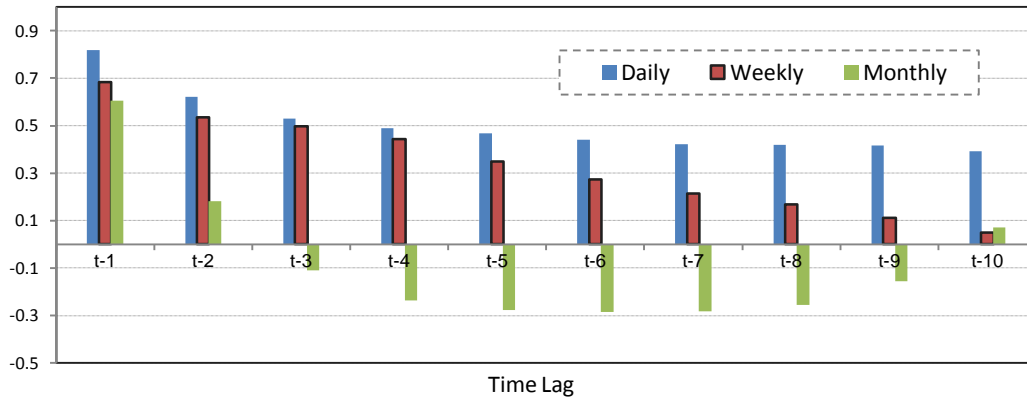


Figure 6.3 ACF of Ellen Brook River daily, weekly and monthly flow time series (1977-2010).

Table 6.2 Input selection for different BPNN models.

Time series	Model Name	Input Structure
Daily	BPNN-D1	Q_{dt}
	BPNN-D2	Q_{dt}, Q_{dt-1}
	BPNN-D3	$Q_{dt}, Q_{dt-1}, Q_{dt-2}$
	BPNN-D4	$Q_{dt}, Q_{dt-1}, Q_{dt-2}, Q_{dt-3}$
	BPNN-D5	$Q_{dt}, Q_{dt-1}, Q_{dt-2}, Q_{dt-3}, Q_{dt-4}$
Weekly	BPNN-W1	Q_{wt}
	BPNN-W2	Q_{wt}, Q_{wt-1}
	BPNN-W3	$Q_{wt}, Q_{wt-1}, Q_{wt-2}$
	BPNN-W4	$Q_{wt}, Q_{wt-1}, Q_{wt-2}, Q_{wt-3}$
	BPNN-W5	$Q_{wt}, Q_{wt-1}, Q_{wt-2}, Q_{wt-3}, Q_{wt-4}$
Monthly	BPNN-M1	Q_{mt}
	BPNN-M2	Q_{mt}, Q_{mt-1}
	BPNN-M3	$Q_{mt}, Q_{mt-1}, Q_{mt-2}$

Table 6.3 Input selection for different ANFIS models.

Time series	Model Name	Input Structure
Daily	ANFIS-D1	Q_{dt}
	ANFIS -D2	Q_{dt}, Q_{dt-1}
	ANFIS -D3	$Q_{dt}, Q_{dt-1}, Q_{dt-2}$
	ANFIS -D4	$Q_{dt}, Q_{dt-1}, Q_{dt-2}, Q_{dt-3}$
	ANFIS-D5	$Q_{dt}, Q_{dt-1}, Q_{dt-2}, Q_{dt-3}, Q_{dt-4}$
Weekly	ANFIS-W1	Q_{wt}
	ANFIS -W2	Q_{wt}, Q_{wt-1}
	ANFIS -W3	$Q_{wt}, Q_{wt-1}, Q_{wt-2}$
	ANFIS -W4	$Q_{wt}, Q_{wt-1}, Q_{wt-2}, Q_{wt-3}$
	ANFIS-W5	$Q_{wt}, Q_{wt-1}, Q_{wt-2}, Q_{wt-3}, Q_{wt-4}$
Monthly	ANFIS-M1	Q_{mt}
	ANFIS -M2	Q_{mt}, Q_{mt-1}
	ANFIS -M3	$Q_{mt}, Q_{mt-1}, Q_{mt-2}$

Three different mother wavelets were chosen for developing the hybrid wavelet neural network model. As concluded in chapter 4, the optimum level of decomposition for different case studies should be reached by trial and error procedure. However, some studies suggest that the level of decomposition is based on the time series length as follows;

$$\text{Number of level of decomposition} = \text{int} [\log(n)] \quad (6.1)$$

Considering the length of daily, weekly and monthly time series (12410, 1773 and 408), the suggested level of decomposition for each time series would be 4, 3 and 3, respectively. Therefore, various levels of decomposition based on the data length and Equation 6.1, were employed for ANN daily, weekly and monthly modelling (Table 6.4). However, for fuzzy modelling fewer level of decomposition is considered for feasible modelling (Table 6.5). The input of the hybrid models would be the wavelet coefficients which are the wavelet decomposition outputs. The number of input time series is $N+1$ for N levels of decomposition as the wavelet coefficients are one approximation and N details.

Table 6. 4 Input pre-processing type for hybrid WNN models.

Time series	Model Name	Wavelet	Level of decomposition
Daily	WNN-D1,2,3,4	db5	3,4,5,6
	WNN-D5,6,7,8	coiflets1	3,4,5,6
	WNN-D9,10,11,12	haar	3,4,5,6
Weekly	WNN-W1,2,3,4	db5	2,3,4,5
	WNN-W5,6,7,8	coiflets1	2,3,4,5
	WNN-W9,10,11,12	haar	2,3,4,5
Monthly	WNN-M1,2,3,4	db5	2,3,4,5
	WNN-M5,6,7,8	coiflets1	2,3,4,5
	WNN-M9,10,11,12	haar	2,3,4,5

Table 6. 5 Input pre-processing type for hybrid WNFG models.

Time series	Model Name	Wavelet	Level of decomposition
Daily	WNFG-D1,2,3,4	db5	2,3,4,5
	WNFG-D5,6,7,8	coiflets1	2,3,4,5
	WNFG-D9,10,11,12	haar	2,3,4,5
Weekly	WNFG-W1,2,3	db5	2,3,4
	WNFG-W4,5,6	coiflets1	2,3,4
	WNFG-W7,8,9	haar	2,3,4
Monthly	WNFG-M1,2,3	db5	2,3,4
	WNFG-M4,5,6	coiflets1	2,3,4
	WNFG-M7,8,9	haar	2,3,4

6.4 RESULTS AND DISCUSSION

6.4.1 Performance of ANN-based models in river flow forecasting

After developing BPNN and WNN frameworks and determining the input selection, models were applied to forecast Ellen Brook River flow. Overall 13 different BPNN and 36 different WNN models with different input combination (Table 6.2 and Table 6.4) and structure were developed for forecasting daily, weekly and monthly river

flow of the study area. The hybrid model results were compared with classical BPNN results for both short term and long term forecasting.

The best structure of models with different input combinations were achieved by increasing the number of hidden neurons from 10 to 25 for daily and weekly and from 1 to 15 for monthly time series. For evaluating the performance of the models, root mean square error (RMSE) and Nash-Sutcliffe coefficient of efficiency (NSE) are considered as the two main criteria in both WNN and BPNN models.

Tables 6.6 – 6.8 present developed BPNN and WNN model structures and forecasting performance for daily, weekly and monthly forecasting. Comparing the performance of all models, best performed models with highest NSE and lowest RSME, in validation set, were chosen (shown bold in the tables). Overall WNN model efficiency is high which made them quite reliable for forecasting. Compared to BPNN, hybrid model make significantly less error and have higher regression with observed flow.

Table 6. 6 BPNN and WNN models structure and performance for daily river flow forecasting.

Model	Best Neuron Structure	Training			Validation		
		NSE	RMSE		NSE	RMSE	
			(m3/s)	(MCM)		(m3/s)	(MCM)
BPNN-D1	1-19-1	0.70	0.78	0.07	0.69	1.56	0.13
BPNN-D2	2-23-1	0.73	0.74	0.06	0.73	1.46	0.13
BPNN-D3	3-22-1	0.75	0.73	0.06	0.73	1.40	0.12
BPNN-D4	4-24-1	0.76	0.71	0.06	0.75	1.37	0.12
BPNN-D5	5-23-1	0.77	0.73	0.06	0.73	1.35	0.12
WNN-D1	4-13-1	0.85	0.64	0.06	0.79	1.10	0.10
WNN-D2	5-22-1	0.88	0.50	0.04	0.82	0.99	0.09
WNN-D3	6-25-1	0.87	0.60	0.05	0.82	1.00	0.09
WNN-D4	7-22-1	0.88	0.60	0.05	0.82	0.98	0.08
WNN-D5	4-17-1	0.86	0.61	0.05	0.80	0.99	0.09
WNN-D6	5-21-1	0.87	0.62	0.05	0.81	1.03	0.09
WNN-D7	6-21-1	0.85	0.63	0.05	0.80	1.10	0.09
WNN-D8	7-20-1	0.86	0.63	0.05	0.80	1.05	0.09
WNN-D9	4-12-1	0.86	0.68	0.06	0.79	1.06	0.09
WNN-D10	5-14-1	0.87	0.64	0.06	0.80	1.02	0.09
WNN-D11	6-15-1	0.87	0.63	0.05	0.80	1.03	0.09
WNN-D12	7-19-1	0.88	0.64	0.06	0.80	0.98	0.08

Table 6. 7 BPNN and WNN models structure and performance for weekly flow forecasting.

Model	Best Neuron Structure	Training			Validation		
		NSE	RMSE		NSE	RMSE	
			(m3/s)	(MCM)		(m3/s)	(MCM)
BPNN-W1	1-14-1	0.60	0.76	0.46	0.43	1.48	0.89
BPNN-W2	2-5-1	0.54	0.77	0.47	0.42	1.53	0.93
BPNN-W3	3-17-1	0.68	0.79	0.48	0.39	1.57	0.95
BPNN-W4	4-19-1	0.68	0.79	0.48	0.39	1.43	0.86
BPNN-W5	5-16-1	0.65	0.78	0.47	0.40	1.63	0.98
WNN-W1	3-19-1	0.85	0.64	0.38	0.65	0.86	0.52
WNN-W2	4-17-1	0.87	0.62	0.37	0.66	0.85	0.51
WNN-W3	5-20-1	0.85	0.64	0.39	0.67	0.86	0.52
WNN-W4	6-25-1	0.87	0.60	0.36	0.64	0.80	0.49
WNN-W5	3-21-1	0.88	0.62	0.38	0.69	0.74	0.45
WNN-W6	4-18-1	0.89	0.59	0.36	0.72	0.75	0.45
WNN-W7	5-24-1	0.84	0.59	0.36	0.74	0.95	0.57
WNN-W8	6-16-1	0.85	0.60	0.36	0.76	0.91	0.55
WNN-W9	3-21-1	0.82	0.57	0.34	0.72	0.89	0.54
WNN-W10	4-19-1	0.80	0.67	0.41	0.60	1.05	0.64
WNN-W11	5-22-1	0.84	0.61	0.37	0.61	1.00	0.61
WNN-W12	6-24-1	0.83	0.59	0.36	0.71	0.90	0.54

Table 6. 8 BPNN and WNN models structure and performance for monthly flow forecasting.

Model	Best Neuron Structure	Training			Validation		
		NSE	RMSE		NSE	RMSE	
			(m ³ /s)	(MCM)		(m ³ /s)	(MCM)
BPNN-M1	1-1-1	0.41	0.83	2.15	0.23	1.41	3.65
BPNN-M2	2-8-1	0.71	0.76	1.99	0.34	0.95	2.46
BPNN-M3	3-3-1	0.66	0.77	2.00	0.34	1.15	2.98
WNN-M1	3-10-1	0.90	0.48	1.24	0.74	0.59	1.53
WNN-M2	4-11-1	0.90	0.61	1.57	0.76	0.58	1.50
WNN-M3	5-9-1	0.90	0.50	1.30	0.75	0.60	1.54
WNN-M4	6-4-1	0.85	0.50	1.29	0.72	0.72	1.87
WNN-M5	3-6-1	0.89	0.62	1.62	0.54	0.72	1.87
WNN-M6	4-9-1	0.85	0.63	1.64	0.56	0.71	1.84
WNN-M7	5-9-1	0.77	0.63	1.63	0.55	0.61	1.58
WNN-M8	6-10-1	0.84	0.63	1.64	0.55	0.71	1.85
WNN-M9	3-9-1	0.78	0.59	1.53	0.59	0.86	2.22
WNN-M10	4-11-1	0.81	0.60	1.56	0.62	0.88	2.28
WNN-M11	5-10-1	0.78	0.58	1.50	0.59	0.74	1.92
WNN-M12	6-4-1	0.84	0.61	1.58	0.58	0.87	2.26

Figure 6.4 compares the Nash-Sutcliffe coefficient of efficiency of models in both training and validation sets which illustrates how WNN model always outperform BPNN models. In the training set the largest gap between original and hybrid model efficiency is for weekly forecasting. The Nash-Sutcliffe coefficient of efficiency of best fitted BPNN model for weekly forecasting, improved from 0.43 to 0.76 and the RMSE of the model decreased from 1.48 to 0.91 m^3 / s .

As mentioned before, in the developed models, validation sets have no role in the training process of the model. It means that investigation the performance of the validation set lead us to the most reliable evaluation. Figure 6.4b illustrates the greater difference between original and hybrid model performance in the validation set. This improvement is more substantial for the long term forecasting as the BPNN model efficiency dramatically decreases with increasing length of forecasting. The NSE of validation set of best performed BPNN model drops from 0.73 to 0.34 from daily to monthly forecasting. 0.34 demonstrate a very weak correlation and indicates that the monthly river flow cannot be predicted with BPNN method. Applying the

hybrid model, this value improved to 0.76 for monthly forecasting, which correspond to a strong correlation between observed and simulated river flow.

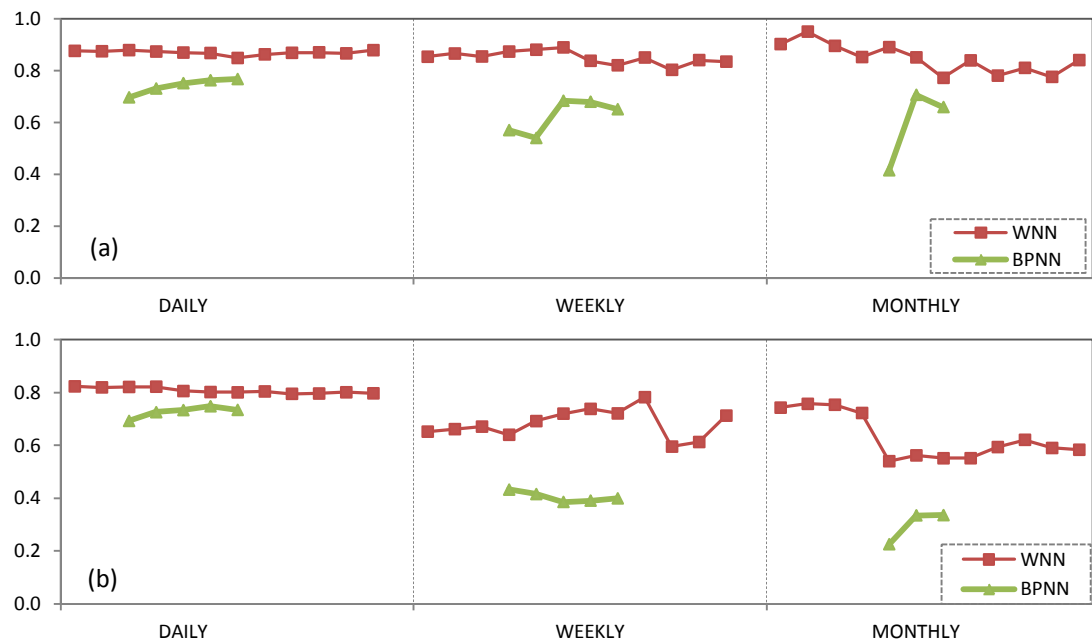


Figure 6. 4 Nash-Sutcliffe coefficient of efficiency of (a) training and (b) validation set, for different BPNN and WNN models.

The results show that applying different time-lagged river flow time series with high ACF as model inputs, improves daily and monthly forecasting. For weekly forecasting, using only current river flow time series leads to the best fitted simulation in the study area. Results also indicate that the type of mother wavelet and the level of decomposition do not have a significant impact on model efficiency. However, applying Daubechies wavelet leads to the best daily and monthly forecasting and Coiflet is the best choice for weekly forecasting of Ellen Brook River flow. Figure 6.5 illustrates the wavelet coefficients of Ellen Brook weekly time series, with Coiflet1 wavelet to 5 level of decompositions which are the best fitted weekly model (WNN-W8) inputs. It can be seen that with five levels of decomposition one time series divided to six sub series, one approximation and five details.

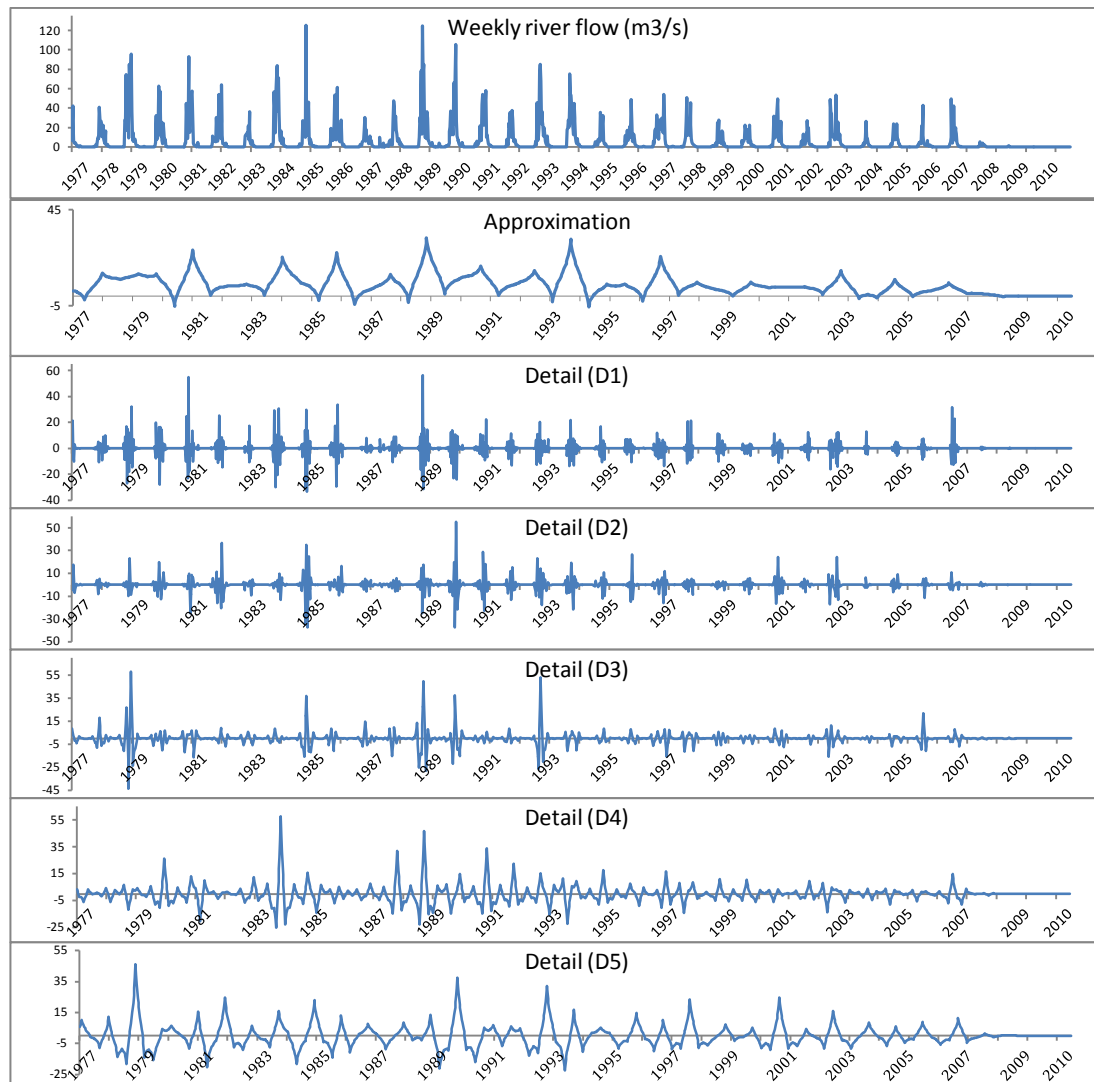


Figure 6.5 Ellen Brook weekly river flow time series and its wavelet coefficients with Coif1 wavelet.

Figure 6.6 shows the scatter plots of observed and forecasted river flow with the best fitted BPNN and WNN models for different lead times. These scatter plots clearly illustrate the performance of different models. It can be seen that unlike WNN model, accuracy of BPNN models decreases in the longer term modeling. The correlation between observed and forecasted river flow with hybrid model is always higher than with BPNN model. This figure also displays that BPNN models frequently fail to simulate extreme events.

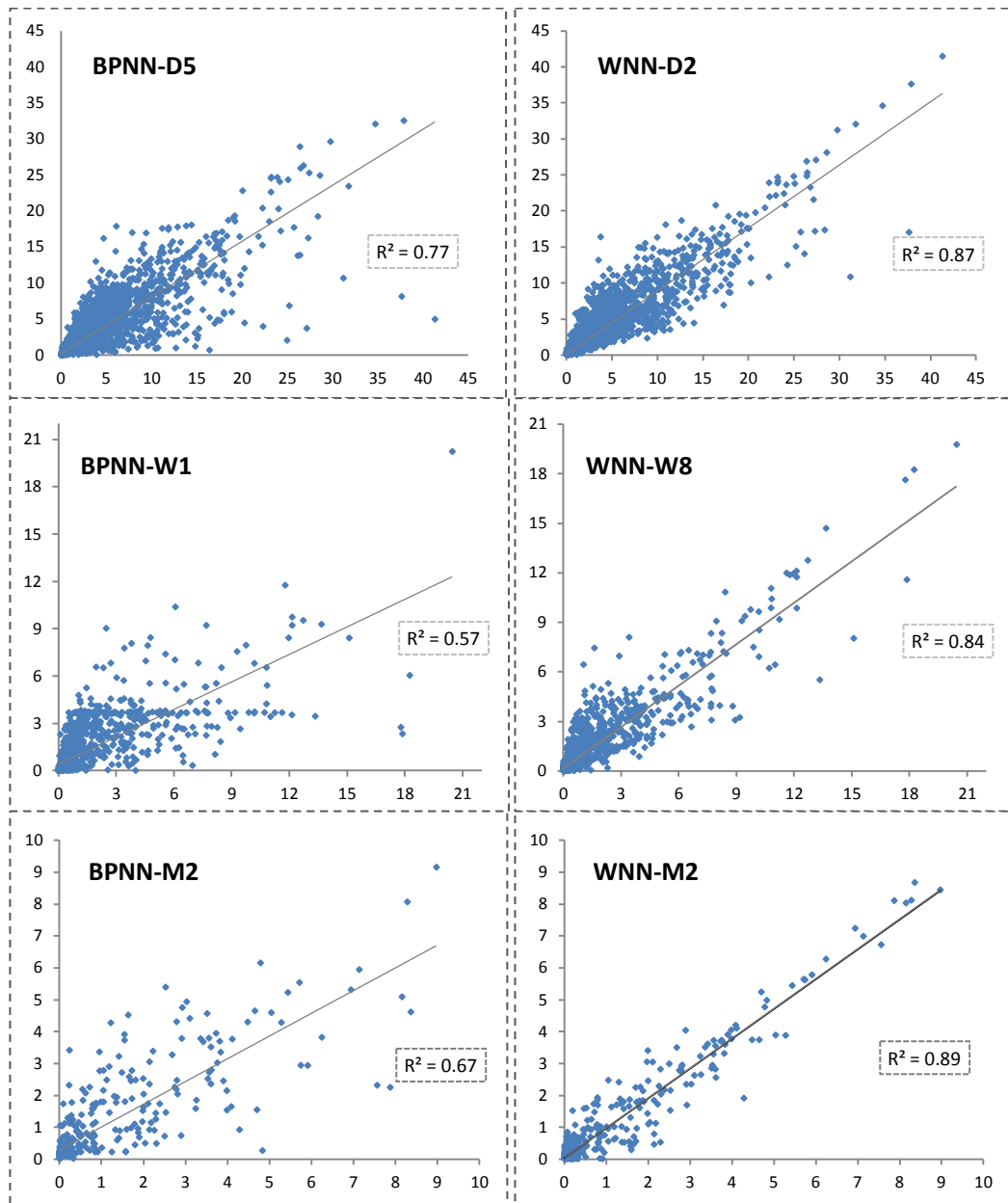


Figure 6. 6 Scatter plots of observed and forecasted river flow with the best fitted BPNN and WNN models for daily, weekly and monthly forecasting.

Having a visual comparison of the model performance, Figure 6.7 and Figure 6.8 compare the original neural networks (BPNN) and best fitted hybrid (WNN) outputs with the observed monthly river flow. Figure 6.9 and Figure 6.10 compares the best fitted hybrid (WNN) and original neural networks (BPNN) outputs with the observed weekly river flow. It can be seen that WNN provides a better match with the

observed time series. These figures also demonstrate that WNN forecasted time series closely meet the extreme values while in many cases BPNN forecasted time series fail to simulate these conditions.

In order to investigate the ability of models in simulating the peak values, the first twenty highest observed river flow in the 34 years of observation (flows greater than $0.50 Q_{\max}$ for each time series) is compared with their simulated values. Table 6.9 shows the relative error between observed daily, weekly and monthly river flow and their best fitted simulated values, which clearly illustrate the reliability of WNN models over the BPNN. The relative error between WNN modelled and observed river flow is always considerably less than the relative error with BPNN modelled and observed ones. For instance, when BPNN model totally fails to simulate the $41.3 \text{ m}^3/\text{s}$ observed river flow (with 88% error), WNN forecast this peak value with a high accuracy of $41 \text{ m}^3/\text{s}$. The weakness of BPNN model in forecasting the extreme values, make this approach ineffective for flood and drought analysis. The outcomes of this study confirm the reliability and accuracy of the proposed wavelet neural networks model. Considering the growing interest in applying data-driven methods, WNN would be a desirable approach for water resources management studies.

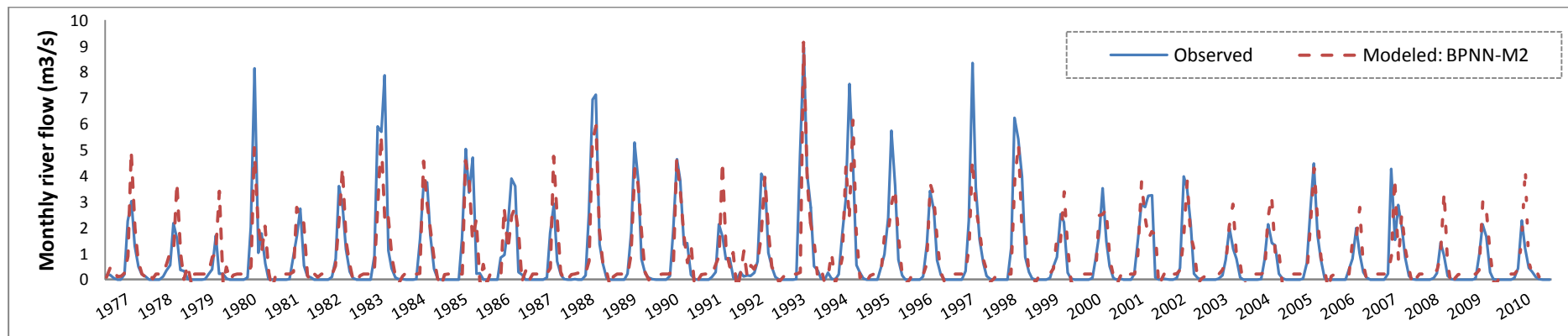


Figure 6. 7 Comparing observed versus modeled monthly river flow with best fitted BPNN model.

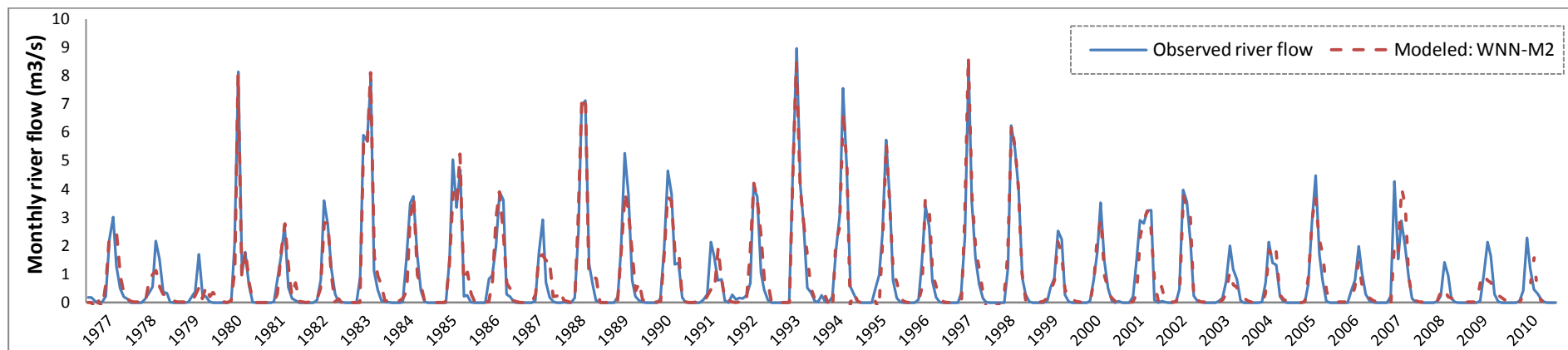


Figure 6. 8 Comparing observed versus modeled monthly river flow with best fitted WNN model.

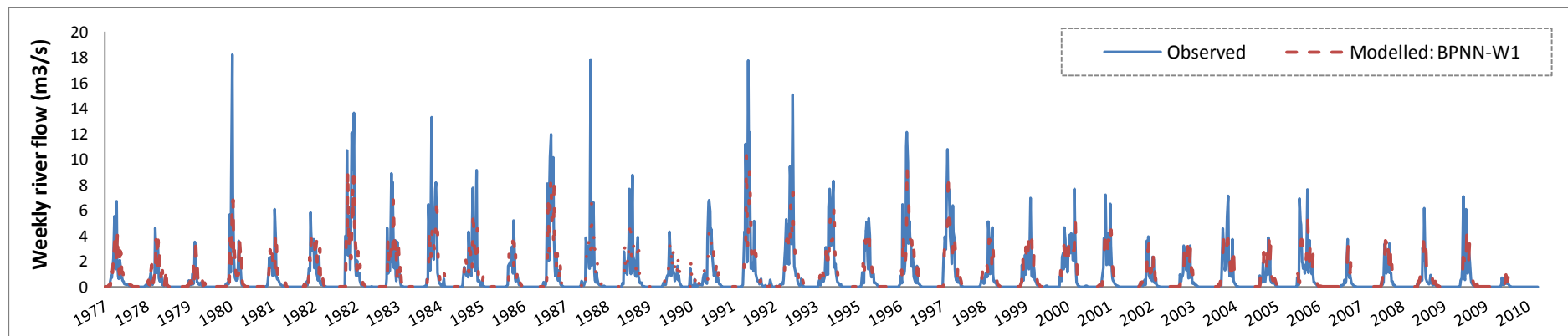


Figure 6. 9 Comparing observed versus modeled weekly river flow with best fitted BPNN model.

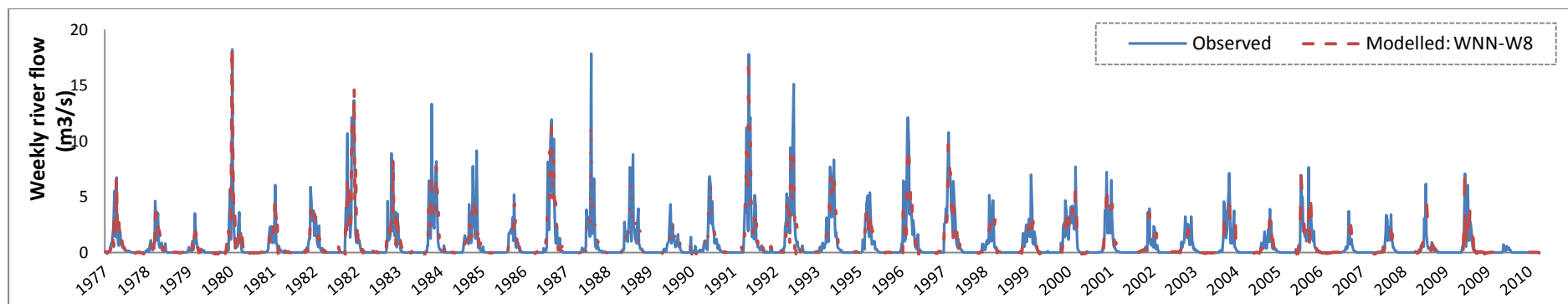


Figure 6. 10 Comparing observed versus modeled weekly river flow with best fitted WNN model.

Table 6.9 Accuracy of developed ANN-based models in simulating daily, weekly and monthly extreme flow values.

No.	Daily stream flow modeling					Weekly stream flow modeling					Monthly stream flow modeling				
	Observed	BPNN		WNN		Observed	BPNN		WNN		Observed	BPNN		WNN	
		Modeled	Error (%)	Modeled	Error (%)		Modeled	Error (%)	Modeled	Error (%)		Modeled	Error (%)	Modeled	Error (%)
1	41.3	4.9	88	41.0	1	20.4	19.8	3	20.2	1	9	8.4	6	9.2	2
2	37.8	32.5	14	38.2	1	18.2	6.0	67	18.2	0	8.4	4.6	45	8.7	4
3	37.6	8.1	78	34.2	9	17.9	2.3	87	11.6	35	8.3	8.1	3	8.1	2
4	34.7	32.0	8	32.7	6	17.8	2.8	84	17.6	1	8.1	5.1	38	8.0	1
5	31.8	23.4	26	29.6	7	15.1	8.0	47	8.4	44	7.9	2.2	71	8.1	3
6	31.2	10.6	66	27.0	13	13.6	9.3	32	14.7	8	7.6	2.3	69	7.4	3
7	29.7	26.5	11	29.6	1	13.3	3.5	74	5.5	59	7.1	5.9	17	7.2	2
8	28.6	24.9	13	26.2	8	12.7	9.5	25	12.7	0	6.9	5.3	23	7.2	4
9	28.4	19.2	32	25.9	9	12.1	9.7	20	9.8	19	6.2	3.8	39	6.3	1
10	27.4	25.2	8	25.6	7	12.1	9.2	24	11.7	3	5.9	2.9	50	5.8	2
11	27.3	16.2	41	25.0	8	12.1	3.5	71	12.1	0	5.7	2.9	49	5.6	2
12	27.1	3.7	87	24.6	9	11.9	8.4	29	11.9	0	5.7	5.5	3	5.6	1
13	26.8	24.3	9	26.3	2	11.8	11.8	0	11.9	1	5.4	5.2	4	5.4	0
14	26.4	23.6	10	25.9	2	11.6	3.7	68	12.0	3	5.3	4.3	19	4.7	11
15	26.4	13.8	48	23.8	10	11.2	3.7	67	9.2	18	5.0	4.6	9	4.6	13
16	26.4	28.9	10	23.8	10	11.0	3.4	69	6.4	42	4.8	0.3	95	5.0	3
17	26.1	13.7	47	23.2	11	10.8	5.4	50	10.4	4	4.8	6.2	29	4.8	0
18	25.7	17.7	31	22.9	11	10.8	6.5	39	11.0	2	4.7	1.5	67	5.2	12
19	25.2	6.8	73	22.5	11	10.8	4.2	61	9.9	9	4.6	4.1	12	4.6	0
20	25.1	22.4	10	24.3	3	10.7	3.7	65	6.2	42	4.5	4.0	10	4.3	4

6.4.2 Performance of Fuzzy-based models in river flow forecasting

In this study ANFIS models developed based on Takagi-Sugeno-Kang (TSK) fuzzy rule based system. Generalized bell membership function and grid partitioning are applied for initializing the fuzzy rule-based structure. In order to develop the hybrid WNFG models, wavelet multi-resolution analysis is coupled with ANFIS model. The Ellen Brook River time series is decomposed into multi-frequency time series by using Haar, Coiflet order 1 and Daubechies order 5 mother wavelets. Then the wavelet coefficients are imposed as input data to the neuro-fuzzy model. Overall 13 different ANFIS and 30 different WNFG models with different input combination (Table 6.3 and Table 6.5) and structure were developed for forecasting daily, weekly and monthly river flow of the study area. The hybrid model results were compared with original ANFIS results for both short term and long term forecasting. Figure 6.11 shows the structure of hybrid WNFG-M2 model, generated by grid partitioning FIS, as an example. Inputs of this model, the best fitted hybrid neuro-fuzzy model for monthly forecasting in the study area, are db5 wavelet coefficients including one approximation and three details for three levels of decomposition, with generalized bell membership function and 16 rules.

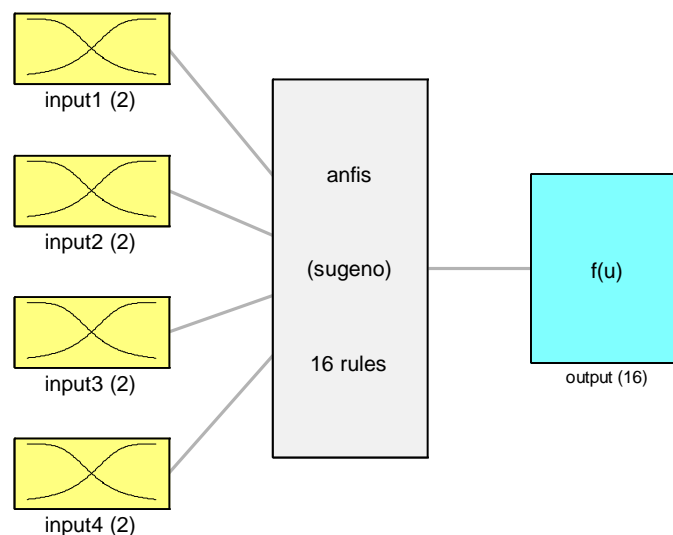


Figure 6. 11 Best fitted hybrid neuro-fuzzy model (WNFG-M2) structure for monthly forecasting.

For the ANFIS models, the optimum number of membership functions is reached by increasing this number from 2 to 5. Considering the large size of the hybrid models input and fuzzy system restrictions, only 2 membership functions are defined for each model. Root mean square error and Nash-Sutcliffe coefficient of efficiency are also considered as the main performance criteria for both ANFIS and WNFG models.

Tables 6.10 – 6.12 show developed ANFIS and WNFG models structure and forecasting performance for daily, weekly and monthly forecasting. Comparing the performance of all models, best fitted models with highest NSE and lowest RSME, were chosen (Bold in tables).

Table 6. 10 ANFIS and WNFG models structure and performance for daily river flow forecasting.

Model	No. MFs	No. Nods	Linear PAR	Non Linear PAR	Fuzzy Rules	Training			Validation		
						NSE	RMSE		NSE	RMSE	
							(m3/s)	(MCM)		(m3/s)	(MCM)
ANFIS-D1	7	32	14	21	7	0.71	0.25	0.021	0.69	0.51	0.044
ANFIS-D2	5	75	75	30	25	0.73	0.24	0.021	0.72	0.48	0.041
ANFIS-D3	4	158	256	36	64	0.76	0.23	0.020	0.74	0.45	0.039
ANFIS-D4	3	193	405	36	81	0.77	0.24	0.020	0.74	0.44	0.038
ANFIS-D5	2	92	192	30	25	0.76	0.23	0.020	0.74	0.45	0.039
WNFG-D1	3	78	108	27	27	0.85	0.22	0.019	0.83	0.36	0.031
WNFG-D2	2	55	80	24	16	0.85	0.19	0.016	0.83	0.36	0.031
WNFG-D3	2	92	192	30	32	0.89	0.19	0.017	0.83	0.23	0.020
WNFG-D4	2	161	448	36	64	0.91	0.22	0.019	0.76	0.34	0.030
WNFG-D5	3	78	108	27	27	0.86	0.23	0.019	0.82	0.35	0.030
WNFG-D6	2	55	8	24	16	0.86	0.20	0.017	0.82	0.34	0.029
WNFG-D7	2	92	192	30	32	0.87	0.21	0.018	0.79	0.33	0.029
WNFG-D8	2	161	448	36	64	0.87	0.23	0.020	0.79	0.29	0.025
WNFG-D9	3	78	108	27	27	0.86	0.25	0.021	0.79	0.38	0.033
WNFG-D10	2	55	80	24	16	0.86	0.21	0.018	0.80	0.34	0.029
WNFG-D11	2	92	192	30	32	0.87	0.22	0.019	0.78	0.33	0.028
WNFG-D12	2	161	448	36	64	0.91	0.23	0.020	0.74	0.31	0.027

Table 6. 11 ANFIS and WNFG models structure and performance for weekly river flow forecasting.

Model	No. of MFs	No. of Nods	Linear PAR	Non Linear PAR	Fuzzy Rules	Training			Validation		
						NSE	RMSE		NSE	RMSE	
							(m3/s)	(MCM)		(m3/s)	(MCM)
ANFIS-W1	5	24	10	15	5	0.54	1.54	0.93	0.36	1.22	0.74
ANFIS-W2	4	53	48	24	16	0.61	0.81	0.49	0.35	1.42	0.86
ANFIS-W3	3	78	108	27	27	0.71	1.22	0.74	0.41	0.88	0.53
ANFIS-W4	2	55	80	24	16	0.66	0.82	0.49	0.34	1.49	0.90
ANFIS-W5	2	92	192	30	25	-	-	-	-	-	-
WNFG-W1	2	34	18	18	8	0.77	0.57	0.34	0.57	1.07	0.65
WNFG-W2	2	55	80	24	16	0.82	0.57	0.35	0.68	0.95	0.58
WNFG-W3	2	92	192	30	32	0.87	0.66	0.40	0.68	0.76	0.46
WNFG-W4	2	34	32	18	8	0.76	0.60	0.36	0.63	0.92	0.56
WNFG-W5	2	55	80	24	16	0.79	0.63	0.38	0.61	0.88	0.53
WNFG-W6	2	92	192	30	32	-	-	-	-	-	-
WNFG-W7	2	34	32	18	8	0.75	0.60	0.36	0.65	1.13	0.69
WNFG-W8	2	55	80	24	16	0.77	0.71	0.43	0.51	1.09	0.66
WNFG-W9	2	92	192	30	32	0.86	0.77	0.46	0.42	0.85	0.51

Table 6. 12 ANFIS and WNFG models structure and performance for monthly river flow forecasting.

Model	No. of MFs	No. of Nods	Linear PAR	Non Linear PAR	Fuzzy Rules	Training			Validation		
						NSE	RMSE		NSE	RMSE	
							(m3/s)	(MCM)		(m3/s)	(MCM)
ANFIS-M1	2	12	4	6	2	0.41	1.41	3.65	0.23	0.83	2.14
ANFIS-M2	5	75	75	30	25	0.64	0.82	2.12	0.15	0.89	2.31
ANFIS-M3	2	34	18	50	8	0.69	0.79	2.05	0.30	0.87	2.25
WNFG-M1	2	34	32	18	8	0.79	0.56	1.45	0.61	0.84	2.16
WNFG-M2	2	55	80	24	16	0.84	0.60	1.55	0.59	0.67	1.74
WNFG-M3	2	92	192	30	32	-	-	-	-	-	-
WNFG-M4	2	34	32	18	8	0.77	0.79	2.05	0.30	0.89	2.31
WNFG-M5	2	55	80	24	16	-	-	-	-	-	-
WNFG-M6	2	92	192	30	32	-	-	-	-	-	-
WNFG-M7	2	34	32	18	8	0.78	0.56	1.46	0.56	0.98	2.55
WNFG-M8	2	55	80	24	16	0.83	0.38	0.98	0.38	0.77	1.98
WNFG-M9	2	92	192	30	32	-	-	-	-	-	-

The Nash-Sutcliffe coefficient of efficiency (NSE) of the models illustrated in Figure 6.12. This figure clearly shows that hybrid WNFG model outperforms ANFIS models.

It can be observed that the ANFIS models' efficiency dramatically decreases with increasing the length of forecasting. Considering the validation set as the most reliable set for evaluating the performance, ANFIS models almost fail to forecast the weekly and monthly river flows. However, by adding the DWT to the ANFIS models the performance of models significantly improves. Overall, in the whole data set, the Nash-Sutcliffe coefficient of efficiency of ANFIS model increased from 0.69 to 0.82 and 0.86 for weekly and monthly forecasting respectively.

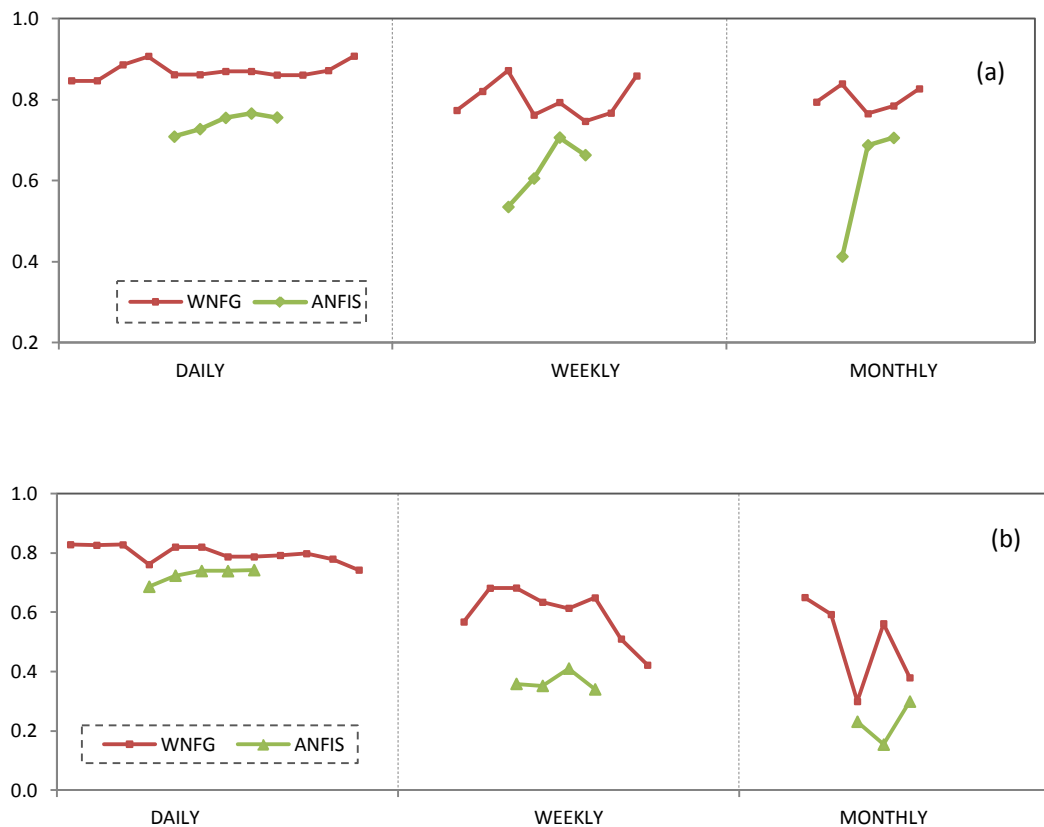


Figure 6. 12 Nash-Sutcliffe coefficient of efficiency of (a) training and (b) validation set, for different ANFIS and WNFG models.

The results show that applying different time-lagged river flow time series with high ACF as model inputs, improves accuracy of forecasting in the study area.

Results also indicate that the type of mother wavelet and the level of decomposition could have a significant impact on weekly and monthly model efficiency. Applying db5 DWT with 3 or 4 level of decomposition leads to the best forecasting of Ellen Brook River flow. Whereas, decomposing the monthly river flow time series with Coiflet1 wavelet, leads to a very poor simulation with NSE of 0.30.

As it is shown in Table 6.11 and Table 6.12, due to restricted structure of fuzzy modelling, some models failed to simulate weekly and monthly river flow.

Figure 6.13 illustrates the wavelet coefficients of Ellen Brook daily time series, with db5 wavelet to 4 level of decompositions which are the best fitted daily model (WNFG-D3) inputs.

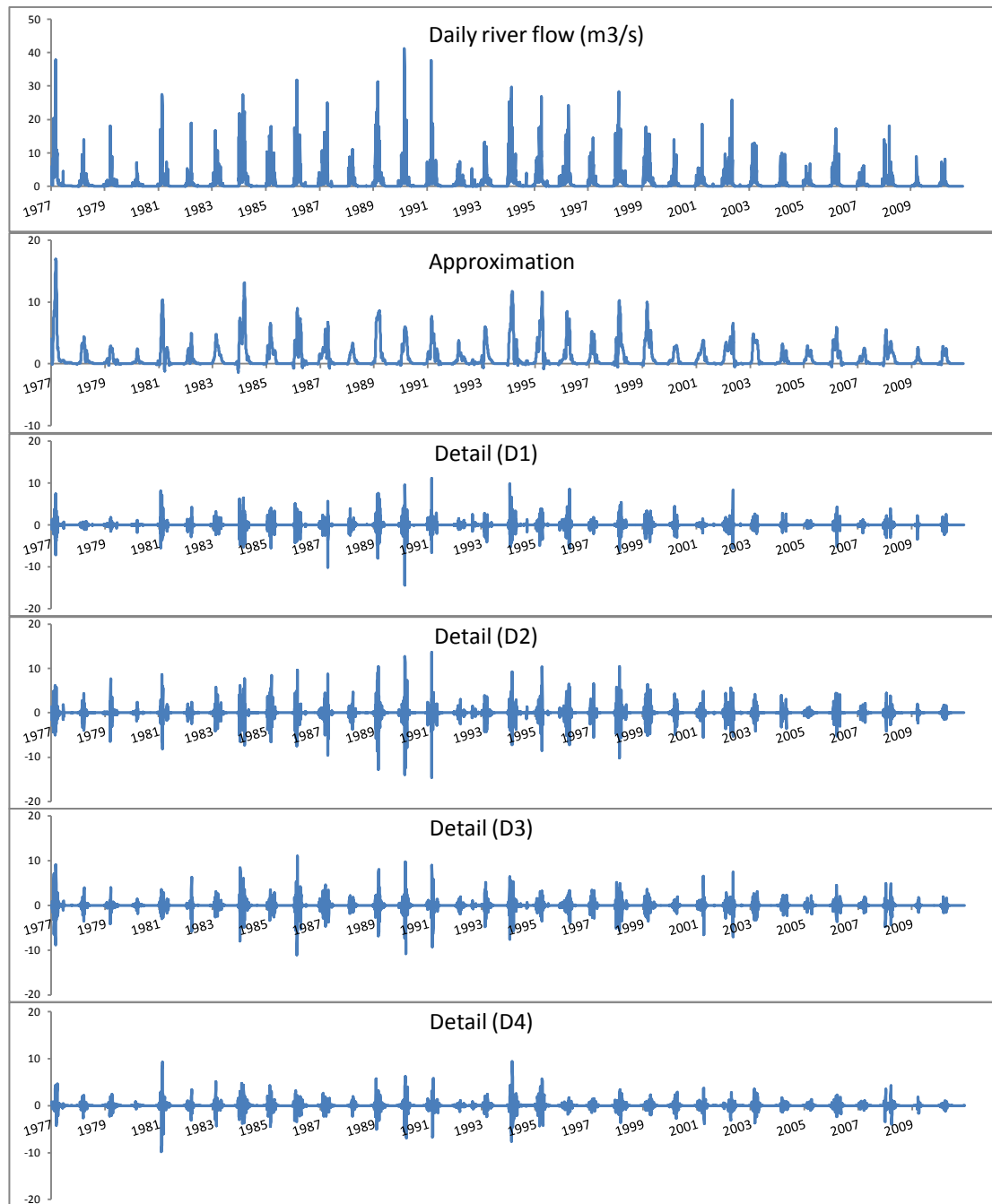


Figure 6.13 Ellen Brook daily river flow signal and its wavelet coefficients with db5 wavelet.

Figure 6.14 shows the scatter plots of observed and forecasted river flow with the best fitted ANFIS and WNFG models for different lead times. These scatter plots clearly illustrate the performance of different models. It can be seen that unlike WNFG model, the accuracy of ANFIS models decreases in the weekly and monthly forecasting.

This figure also demonstrates that in spite of relatively high correlation between observed and ANFIS modelled river flow, these models frequently fail to simulate the extreme events.

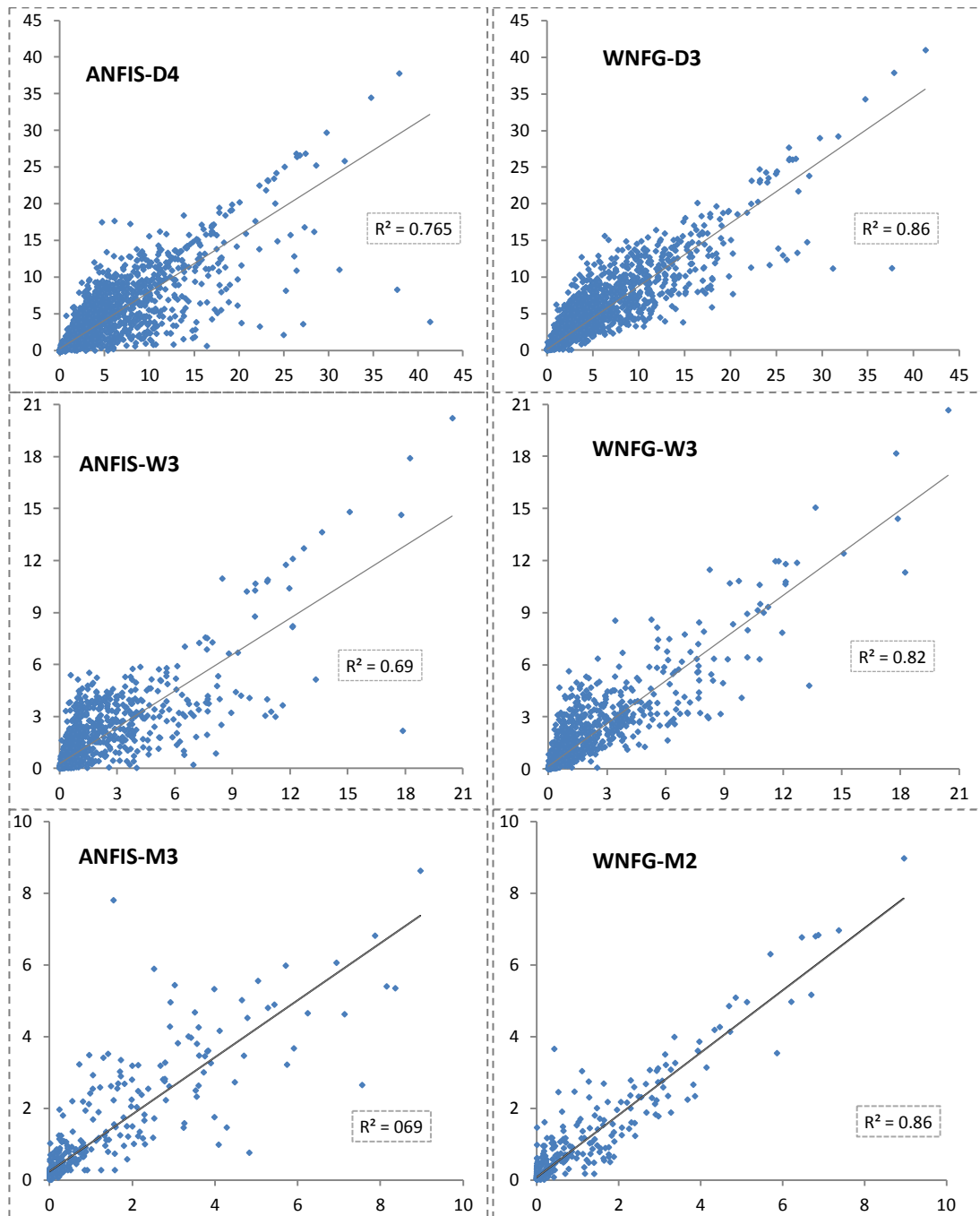


Figure 6. 14 Scatter plots of observed and forecasted river flow with the best fitted ANFIS and WNFG models for daily, weekly and monthly forecasting.

Figures 6.15 – 6.16 compare the best fitted ANFIS and hybrid WNFG model outputs with the observed weekly river flows. Figure 6.17 and 6.18 also compare the best fitted hybrid WNFG and ANFIS outputs with the observed monthly river flows. It can be seen that hybrid models provide a better match with the observed time series.

To investigate the ability of models in forecasting the extreme values, the first twenty highest observed river flows in the 34 years of observation are compared with their simulated values. Table 6.13 shows the relative error between observed daily, weekly and monthly river flows and their best fitted simulated values, which illustrates that hybrid models have relatively smaller errors. The errors between WNF modelled and observed river flow are often considerably lower than those of ANFIS model. For instance, when ANFIS model totally fails to simulate the $41.3 \text{ m}^3/\text{s}$ observed daily river flow (with 91 % relative error), WNF forecast this peak value with a high accuracy.

Table 6. 13 Accuracy of developed Fuzzy-based models in simulating daily, weekly and monthly extreme flow values.

No.	Daily stream flow modeling					Weekly stream flow modeling					Monthly stream flow modeling				
	Observed	ANFIS		WNFG		Observed	ANFIS		WNFG		Observed	ANFIS		WNFG	
		Modeled	Error (%)	Modeled	Error (%)		Modeled	Error (%)	Modeled	Error (%)		Modeled	Error (%)	Modeled	Error (%)
1	41.3	3.9	91	41.0	1	20.4	20.2	1	20.7	1	9.0	8.6	4	9.0	0
2	37.9	37.7	0	37.9	0	18.2	11.3	38	17.9	2	8.4	5.3	36	7.9	6
3	37.6	8.3	78	21.2	43	17.9	2.2	88	14.4	19	8.1	5.4	34	8.1	1
4	34.7	34.5	1	34.2	1	17.8	14.6	18	18.2	2	7.9	6.8	13	7.9	0
5	31.8	25.8	19	29.2	8	15.1	12.4	18	14.8	2	7.5	2.6	65	5.8	23
6	31.2	11.0	65	17.5	42	13.6	13.6	0	13.1	4	7.1	4.6	35	7.5	5
7	29.7	29.7	0	28.9	3	13.3	5.1	61	5.8	56	6.9	6.1	13	6.5	6
8	28.6	22.8	17	25.2	12	12.7	12.7	0	11.9	6	6.2	4.6	25	5.0	20
9	28.4	16.2	43	19.7	30	12.1	8.2	32	10.8	11	5.9	3.7	38	6.5	10
10	27.4	26.8	2	26.2	4	12.1	12.1	0	11.8	3	5.7	3.2	44	5.5	3
11	27.3	16.8	38	17.3	37	12.1	8.2	33	10.7	12	5.7	6.0	5	5.9	4
12	27.1	3.6	87	26.1	4	11.9	10.4	13	11.9	1	5.4	4.9	10	4.9	9
13	26.8	26.7	0	26.0	3	11.8	11.8	0	12.0	2	5.3	4.8	9	5.4	3
14	26.4	24.4	8	26.1	1	11.6	3.6	69	12.0	3	5.0	5.6	10	4.8	5
15	26.4	10.9	59	25.9	2	11.2	3.0	73	9.3	17	4.8	0.8	84	4.7	4
16	26.4	26.9	2	26.6	1	11.0	3.3	70	9.0	18	4.8	3.6	24	4.5	5
17	26.1	12.9	51	13.3	49	10.8	10.9	1	10.8	0	4.7	3.5	26	4.6	3
18	25.7	15.8	38	17.9	30	10.8	10.8	0	10.6	2	4.6	5.0	8	4.3	8
19	25.2	8.2	68	13.9	45	10.8	4.0	63	6.3	41	4.5	2.7	40	2.7	30
20	25.1	25.1	0	24.4	3	10.7	3.0	71	9.1	14	4.3	1.5	66	3.0	30

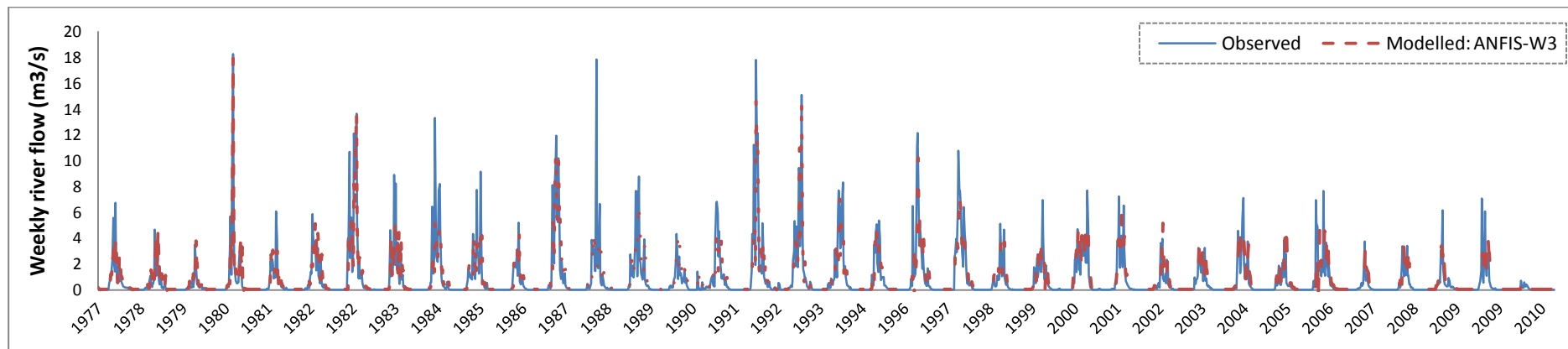


Figure 6. 15 Comparing observed versus modeled weekly river flow with best fitted ANFIS model.

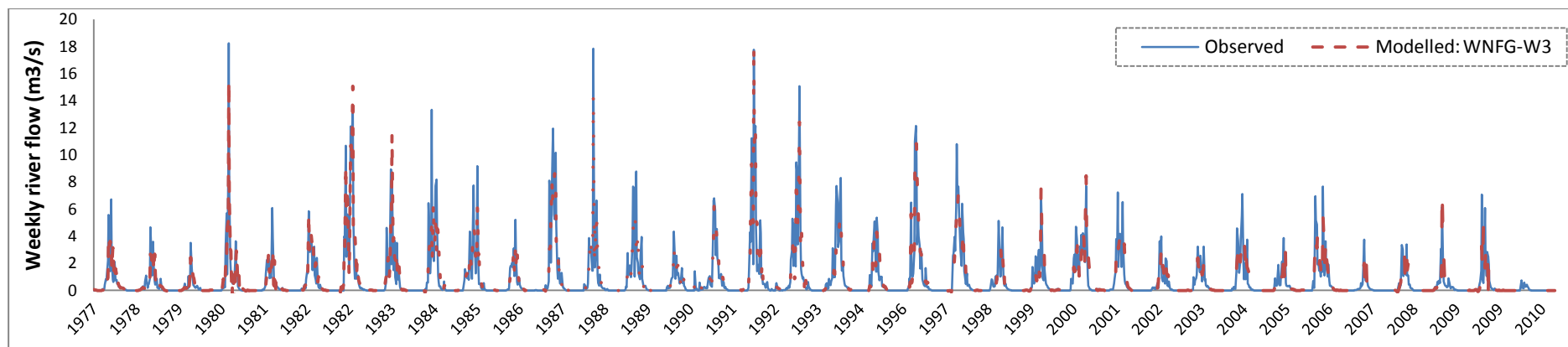


Figure 6. 16 Comparing observed versus modeled weekly river flow with best fitted WNFG model.

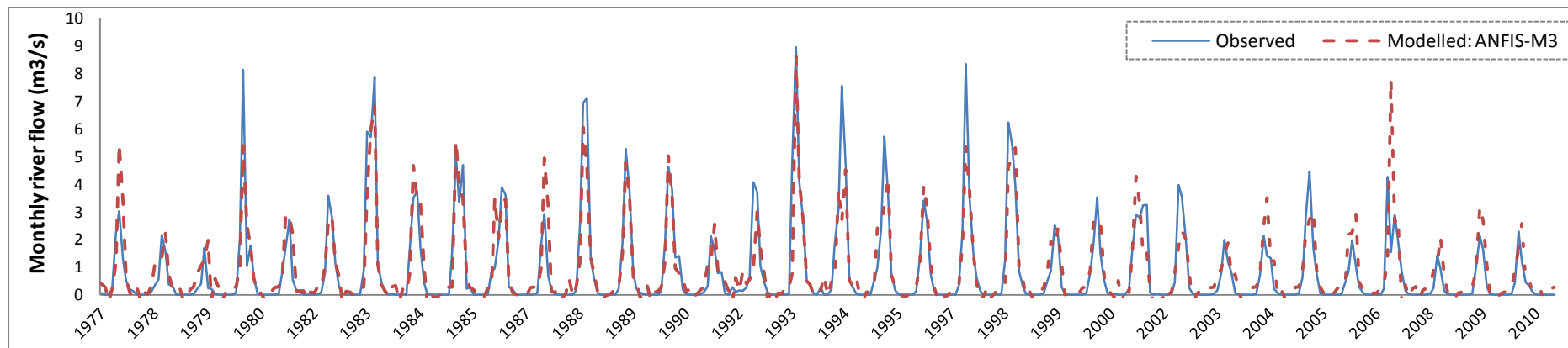


Figure 6. 17 Comparing observed versus modeled monthly river flow with best fitted ANFIS model.

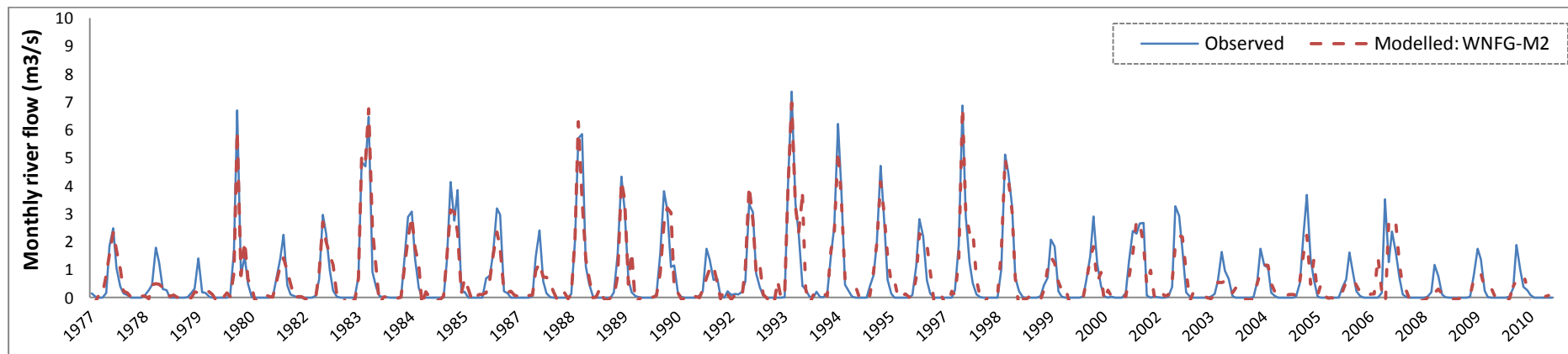


Figure 6. 18 Comparing observed versus modeled monthly river flow with best fitted WNFG model.

6.5 CONCLUSION

In this section different computational intelligent models for short and long term river flow forecasting have proposed. Application of multi-resolution analysis of the input data on BPNN and ANFIS model performance for forecasting one step ahead of daily, weekly and monthly river flow has been investigated. Haar, Daubechies order five and Coiflet order one wavelets were applied on Ellen Brook River flow time series to decompose the time series in different levels of resolution. Different wavelet coefficients were imposed to BPNN and ANFIS models as their inputs.

The overall results show that pre-processing the raw data with wavelet has significantly improved the accuracy of forecasting. The results also indicated that the performance improvement was more substantial in longer lengths of forecasting. Where BPNN and ANFIS models almost fail to forecast monthly river flow, hybrid models simulate the time series with quite high accuracy.

Although using the right selection of the different time series with different time-lag and high autocorrelation function (ACF) improves the BPNN and ANFIS models efficiency, the improvement is considerably less than pre-processing the data with discrete wavelet transform.

Among all models, hybrid neural networks (WNN-D5) is selected as the best fitted model for daily river forecasting and the hybrid neuro-fuzzy model (WNFG-M2) achieved the best performance for monthly river flow forecasting. However, there is no significant difference between best fitted ANN-based and fuzzy-based hybrid models' performance for river flow forecasting.

Furthermore the results verified that unlike WNFG models, altering mother wavelet or the level of decomposition does not have a considerable impact on WNN models' performance.

These results are based on the unique characteristics of Ellen Brook River flow time series and different DWT might be more compatible for modelling different case studies. However, considering the similar characteristics of Western Australia rivers,

with high seasonal trend, the same method would eventuate the best prediction result in this region.

Usually computational approaches fail to simulate sudden extreme conditions as they use current and few previous data as their inputs. Considering the transient nature of hydrological signals, applying DWT on input data and extracting different frequencies from historical data, helps more accurate prediction of extreme values. This matter is well observed in this research, where WNN forecasted time series are highly matched with the observed time series at the extreme values. Since BPNN and ANFIS failed to simulate the peak conditions most of the time, these models are not recommended for flood and drought studies.

Chapter 7

Multi-Step Ahead River Flow Forecasting

Extended from:

Badrzadeh, H., Sarukkalige, R. and Jayawardena, A. W., 2013. *Impact of multi-resolution analysis of artificial intelligence models Inputs on river flow forecasting*, Journal of hydrology, (507) 75-85.

7.1 INTRODUCTION

In this chapter, an attempt is made to show that the performance of longer lead-time forecasting is improved when data-preprocessing techniques are used in conjunction with computational intelligence methods. One of the inherent problems in all forecasting methods is that the forecasting reliability decreases with increasing the lead-time. The developed ANN and ANFIS model performance are compared against hybrid wavelet neural networks and hybrid wavelet neuro-fuzzy with subtractive clustering methods. Different models with a combination of the different input data sets are developed for 1, 2, 3, 4 and 5 days ahead forecasting in Harvey River, Western Australia. Daubechies and Symlet wavelets are used to decompose river flow time series to different levels. Comparing the results with those of the original ANN and ANFIS models indicates that the hybrid models produce significantly better forecasts, especially for the peak values and longer lead-times.

7.2 STUDY AREA AND DATA USED

In this study, the river flow data of the Dingo road station on Harvey River are used as a case study. The information of Harvey River and statistical analyses of flow and rainfall time series at the Dingo road station is available in Chapter five.

For this study mean daily river flow discharge and mean daily rainfall for 39 years, with an observation period from 1972 to 2011, are collected from the Department of Water and Bureau of Meteorology. First 27 years of data (9971 days, around 70% of the whole data set) are used for training and the remaining 12 years (4273 days, around 30% of the whole data set) are used for validation.

The input combinations for forecasting the river flow time series was chosen based on forward stepwise selection of inputs and considering the auto correlation function (ACF) of time series to optimize the volume of input data (Table 7.1 and Figure 7.1). Considering the huge amount of input data, especially after decomposing each data set to sub-series, time series with a maximum three-day lag of river flow was designated in input combinations.

Table 7.1 ACF of Harvey River daily flow and rainfall time series (1972-2011).

ACF	$t-1$	$t-2$	$t-3$	$t-4$	$t-5$	$t-6$	$t-7$	$t-8$	$t-9$	$t-10$
Q_t	0.944	0.869	0.825	0.795	0.757	0.746	0.738	0.724	0.714	0.697
R_t	0.283	0.127	0.109	0.094	0.101	0.111	0.093	0.088	0.106	0.099

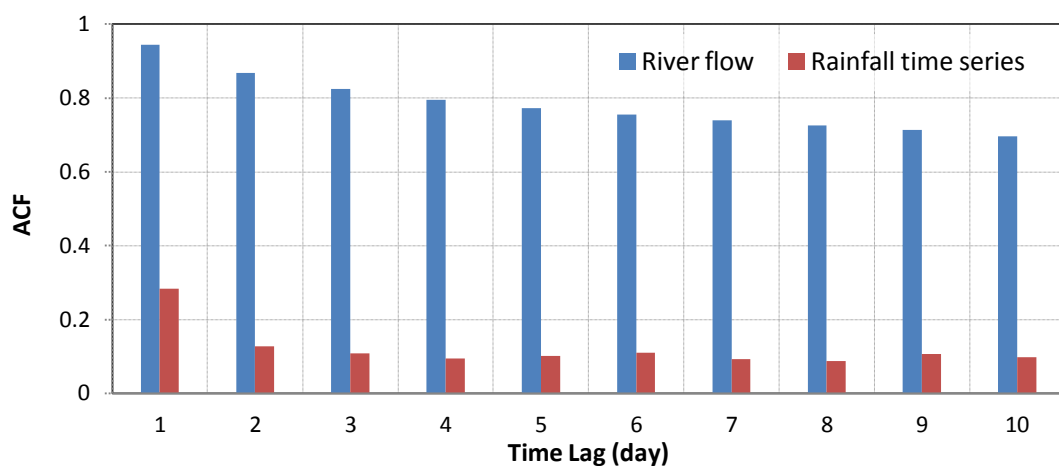


Figure 7.1 ACF of Harvey River daily flow and rainfall time series.

Table 7.2 shows the different input combination that applied to the models for 1 to 5 days ahead (Q_{t+1} , Q_{t+2} , Q_{t+3} , Q_{t+4} , Q_{t+5}) forecasting. The Q and R are the river flow and rainfall time series and QDWT and RDWT are their wavelet coefficient time series, including details and approximation sub-series for n level of decomposition (D_1, D_2, \dots, D_n and A_n).

Table 7. 2 Different input combinations.

No.	Input data combination
i	Q_t
ii	Q_t, Q_{t-1}
iii	Q_t, Q_{t-1}, Q_{t-2}
iv	$Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}$
v	Q_t, R_t
vi	$QDWT_t$
vii	$QDWT_t, QDWT_{t-1}$
viii	$QDWT_t, QDWT_{t-1}, QDWT_{t-2}$
ix	$QDWT_t, QDWT_{t-1}, QDWT_{t-2}, QDWT_{t-3}$
x	$QDWT_t, RDWT_t$

7.3 RESULTS AND DISCUSSION

After developing ANN, ANFIS, WNN and WNF frameworks, they were applied to forecast Harvey River flow. Overall 215 different models for various lead-times of 1 to 5 days ahead, with different input combination were developed for forecasting daily river flow of the study area. The hybrid model results were compared with classical ANN and ANFIS model results.

7.3.1 Application of ANN

A three layered feed-forward backpropagation ANN model was used without preprocessing the data. Different ANN models with different input combinations have been developed. Considering the volume of the input data, each ANN was

trained by increasing the number of hidden neurons from 15 to 30 to reach the best ANN structure. The result indicates that the model efficiency varies with the input selection. Table 7.3 shows the best ANN model's structure for each of the five lead-time forecasting based on the input selection.

Table 7.3 ANN structure and performance for different lead time.

Lead time	Model	Input Dataset	Neuron structure	Training		Validation	
				<i>NSE</i>	RSME (m ³ /s)	<i>NSE</i>	RSME (m ³ /s)
t+1	ANN1-1	i	1-16-1	0.913	0.283	0.897	0.437
	ANN1-2	ii	2-23-1	0.927	0.259	0.912	0.404
	ANN1-3	iii	3-21-1	0.934	0.245	0.923	0.379
	ANN1-4	iv	4-26-1	0.934	0.246	0.928	0.367
	ANN1-5	v	2-24-1	0.975	0.216	0.969	0.169
t+2	ANN2-1	i	1-30-1	0.795	0.538	0.770	0.653
	ANN2-2	ii	2-24-1	0.801	0.427	0.781	0.637
	ANN2-3	iii	3-20-1	0.805	0.423	0.790	0.624
	ANN2-4	iv	4-28-1	0.800	0.609	0.796	0.433
	ANN2-5	v	2-29-1	0.850	0.434	0.841	0.382
t+3	ANN3-1	i	1-25-1	0.714	0.749	0.707	0.750
	ANN3-2	ii	2-25-1	0.718	0.727	0.716	0.513
	ANN3-3	iii	3-23-1	0.731	0.707	0.719	0.508
	ANN3-4	iv	4-29-1	0.739	0.696	0.715	0.512
	ANN3-5	v	2-29-1	0.736	0.512	0.736	0.738
t+4	ANN4-1	i	1-30-1	0.664	0.789	0.651	0.566
	ANN4-2	ii	2-28-1	0.671	0.781	0.652	0.566
	ANN4-3	iii	3-25-1	0.690	0.758	0.651	0.566
	ANN4-4	iv	4-24-1	0.702	0.743	0.680	0.542
	ANN4-5	v	2-28-1	0.685	0.765	0.682	0.568
t+5	ANN5-1	i	1-25-1	0.634	0.805	0.612	0.597
	ANN5-2	ii	2-21-1	0.639	0.819	0.615	0.595
	ANN5-3	iii	3-21-1	0.662	0.795	0.616	0.597
	ANN5-4	iv	4-18-1	0.670	0.784	0.613	0.601
	ANN5-5	v	2-29-1	0.651	0.805	0.629	0.584

It can be seen from the results that model performance changes with respect to lead-time forecast. For one day ahead forecasting the result is quite satisfactory, but as the lead-time increases the model efficiency is decreasing dramatically (Nash-Sutcliffe coefficient of efficiency decreases from 0.97 to 0.65). Also, using only current river flow time series (input combination (i) of Table 7.2) gives the worst result, while using a combination of both current river flow and rainfall (input combination (v) of Table 7.2) often gives the best result. Figure 7.2 shows the scatter plots between observed and forecasted river flow with the best fitted ANN models, for different lead time. This figure illustrates how the performance of the forecasting decreases with increasing the lead time. Also, it can be seen that for lead times greater than 3 days, the ANN model totally fails in simulating the extreme conditions.

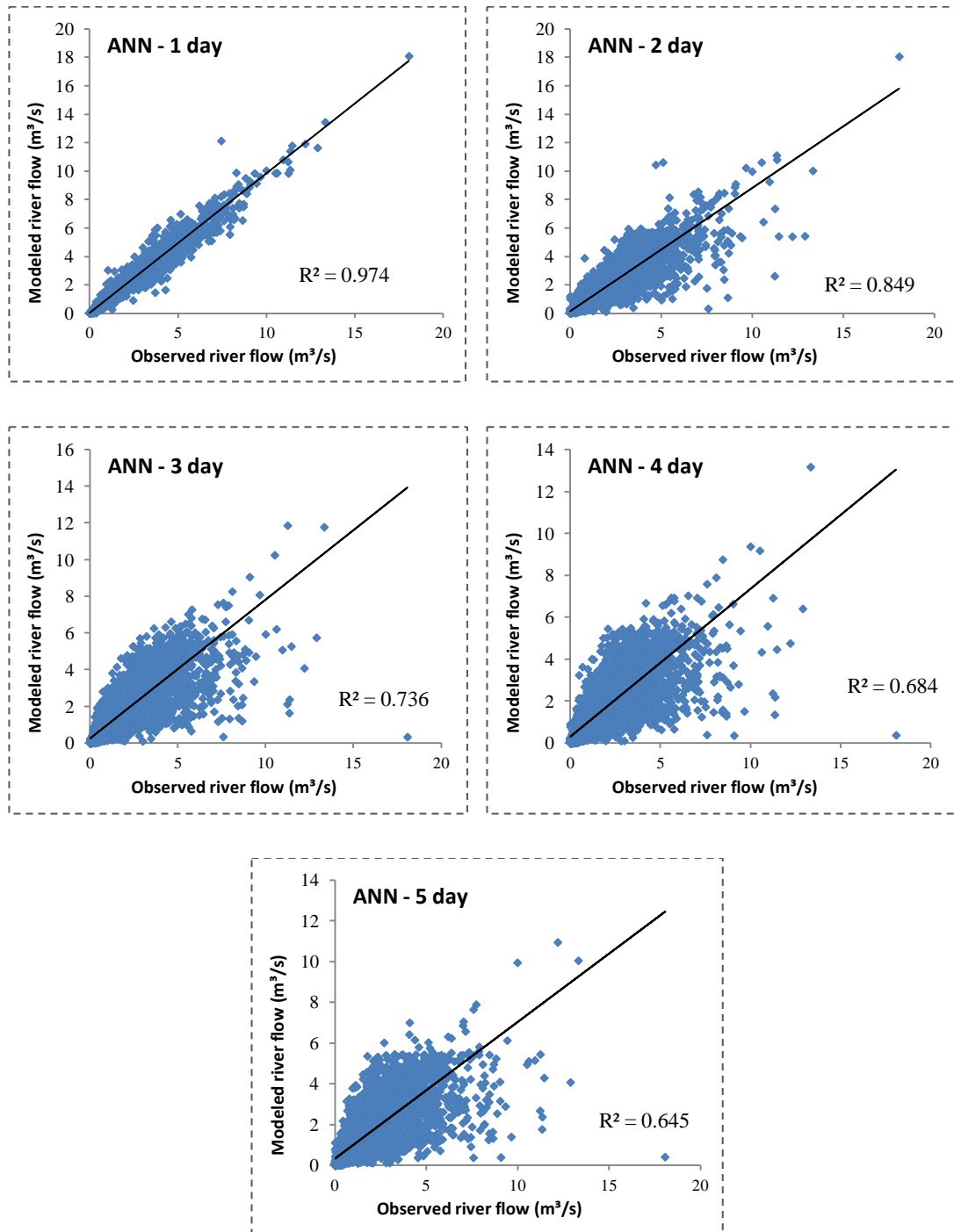


Figure 7.2 Scatter plots of observed and ANN forecasted flow for different lead time.

7.3.2 Improving the efficiency of ANN with WNN

To improve the model efficiency, hybrid wavelet neural network model was applied. Both Daubechies and Symlet mother wavelets were used for decomposing input time

series into 3,4 and 5 levels of decomposition. Figure 7.3 shows river flow time series and its sub-series, which are 'db5' wavelet coefficient with 4 levels of resolution.

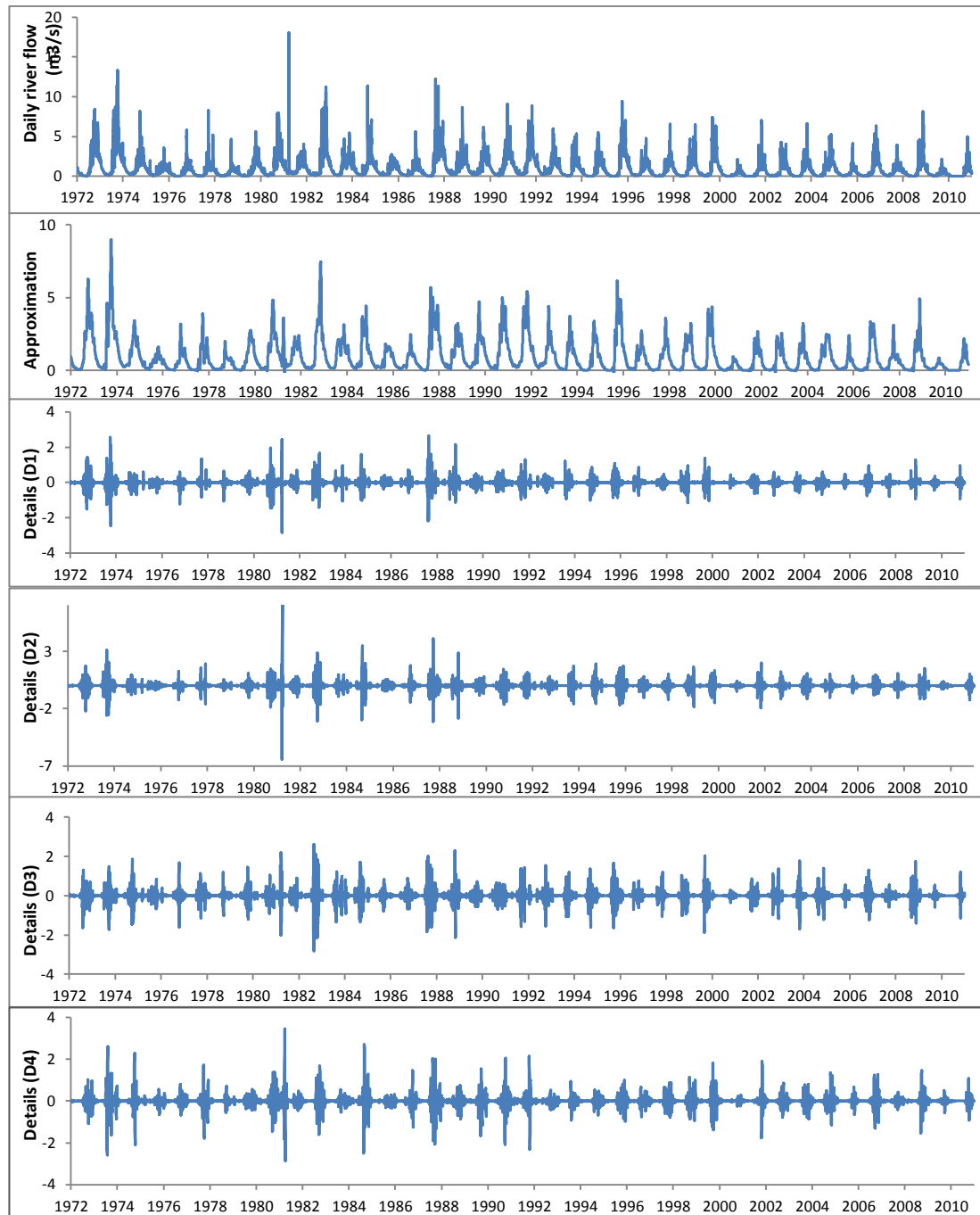


Figure 7.3 Daily river flow time series and its db5 wavelet coefficients with four level of resolution.

Similar to ANN models, various WNN models with different input combinations and structure were developed. Table 7.4 to Table 7.8 show the best WNN model structure based on the input selection, for one to five day lead-time forecasting respectively.

Table 7. 4 Different WNN model’s structure and performance for 1 day ahead lead-time.

Model	Input Dataset	Wavelet	Level of Decompos	Neuron Structure	Training		Validation	
					<i>NSE</i>	RSME (m ³ /s)	<i>NSE</i>	RSME (m ³ /s)
WNND1-1	vi	db5	3	4-26-1	0.952	0.306	0.947	0.221
WNND1-2	vii	db5	3	8-25-1	0.987	0.128	0.991	0.111
WNND1-3	viii	db5	3	12-26-1	0.996	0.086	0.992	0.083
WNND1-4	ix	db5	3	16-28-1	0.989	0.076	0.998	0.068
WNND1-5	x	db5	3	8-27-1	0.987	0.126	0.990	0.117
WNND1-6	vi	db5	4	5-29-1	0.960	0.271	0.944	0.228
WNND1-7	vii	db5	4	10-25-1	0.993	0.114	0.988	0.106
WNND1-8	viii	db5	4	15-20-1	0.995	0.095	0.992	0.084
WNND1-9	ix	db5	4	20-25-1	0.998	0.068	0.995	0.071
WNND1-10	x	db5	4	10-24-1	0.993	0.113	0.986	0.112
WNND1-11	vi	db5	5	6-27-1	0.961	0.269	0.947	0.221
WNND1-12	vii	db5	5	12-29-1	0.993	0.113	0.987	0.110
WNND1-13	viii	db5	5	18-19-1	0.996	0.094	0.994	0.084
WNND1-14	ix	db5	5	24-27-1	0.998	0.067	0.994	0.073
WNND1-15	x	db5	5	12-26-1	0.994	0.112	0.988	0.112
WNNS1-1	vi	sym2	3	4-16-1	0.944	0.337	0.943	0.228
WNNS1-2	vii	sym2	3	8-19-1	0.980	0.180	0.984	0.281
WNNS1-3	viii	sym2	3	12-29-1	0.994	0.103	0.986	0.113
WNNS1-4	ix	sym2	3	16-22-1	0.984	0.136	0.990	0.110
WNNS1-5	x	sym2	3	8-17-1	0.984	0.148	0.988	0.126
WNNS1-6	vi	sym2	4	5-17-1	0.952	0.299	0.940	0.235
WNNS1-7	vii	sym2	4	10-19-1	0.986	0.160	0.980	0.270
WNNS1-8	viii	sym2	4	15-28-1	0.993	0.114	0.986	0.114
WNNS1-9	ix	sym2	4	20-17-1	0.992	0.121	0.986	0.114
WNNS1-10	x	sym2	4	10-16-1	0.991	0.132	0.984	0.121
WNNS1-11	vi	sym2	5	6-18-1	0.952	0.296	0.943	0.229
WNNS1-12	vii	sym2	5	12-22-1	0.986	0.159	0.980	0.279
WNNS1-13	viii	sym2	5	18-27-1	0.993	0.113	0.987	0.114
WNNS1-14	ix	sym2	5	24-19-1	0.992	0.120	0.985	0.117
WNNS1-15	x	sym2	5	12-19-1	0.991	0.131	0.985	0.121

Table 7.5 Different WNN model's structure and performance for 2 day ahead lead-time.

Model	Input Dataset	Wavelet	Level of Decompos	Neuron Structure	Training		Validation	
					<i>NSE</i>	RSME (m ³ /s)	<i>NSE</i>	RSME (m ³ /s)
WNND2-1	vi	db5	3	4-24-1	0.930	0.361	0.904	0.297
WNND2-2	vii	db5	3	8-29-1	0.977	0.206	0.960	0.193
WNND2-3	viii	db5	3	12-19-1	0.982	0.181	0.971	0.163
WNND2-4	ix	db5	3	16-27-1	0.980	0.141	0.981	0.143
WNND2-5	x	db5	3	8-22-1	0.981	0.166	0.956	0.190
WNND2-6	vi	db5	4	5-25-1	0.940	0.332	0.900	0.303
WNND2-7	vii	db5	4	10-25-1	0.981	0.213	0.965	0.191
WNND2-8	viii	db5	4	15-26-1	0.986	0.160	0.966	0.177
WNND2-9	ix	db5	4	20-30-1	0.991	0.129	0.977	0.146
WNND2-10	x	db5	4	10-28-1	0.984	0.172	0.961	0.189
WNND2-11	vi	db5	5	6-27-1	0.946	0.315	0.901	0.301
WNND2-12	vii	db5	5	12-29-1	0.982	0.182	0.963	0.202
WNND2-13	viii	db5	5	18-27-1	0.990	0.144	0.965	0.182
WNND2-14	ix	db5	5	24-20-1	0.992	0.111	0.975	0.154
WNND2-15	x	db5	5	12-25-1	0.988	0.155	0.961	0.194
WNNS2-1	vi	sym2	3	4-26-1	0.913	0.409	0.884	0.325
WNNS2-2	vii	sym2	3	8-24-1	0.957	0.261	0.937	0.232
WNNS2-3	viii	sym2	3	12-17-1	0.968	0.259	0.953	0.202
WNNS2-4	ix	sym2	3	16-25-1	0.970	0.200	0.946	0.226
WNNS2-5	x	sym2	3	8-22-1	0.976	0.188	0.947	0.211
WNNS2-6	vi	sym2	4	5-27-1	0.924	0.376	0.881	0.331
WNNS2-7	vii	sym2	4	10-20-1	0.961	0.270	0.942	0.230
WNNS2-8	viii	sym2	4	15-18-1	0.972	0.228	0.948	0.219
WNNS2-9	ix	sym2	4	20-27-1	0.982	0.184	0.942	0.230
WNNS2-10	x	sym2	4	10-28-1	0.980	0.195	0.952	0.210
WNNS2-11	vi	sym2	5	6-30-1	0.930	0.357	0.881	0.330
WNNS2-12	vii	sym2	5	12-24-1	0.962	0.232	0.941	0.243
WNNS2-13	viii	sym2	5	18-19-1	0.976	0.206	0.947	0.225
WNNS2-14	ix	sym2	5	24-19-1	0.983	0.158	0.941	0.243
WNNS2-15	x	sym2	5	12-25-1	0.983	0.176	0.952	0.215

Table 7. 6 Different WNN model's structure and performance for 3 day ahead lead-time.

Model	Input Dataset	Wavelet	Level of Decompos	Neuron Structure	Training		Validation	
					<i>NSE</i>	RSME (m ³ /s)	<i>NSE</i>	RSME (m ³ /s)
WNND3-1	vi	db5	3	4-28-1	0.913	0.401	0.861	0.357
WNND3-2	vii	db5	3	8-28-1	0.961	0.269	0.939	0.237
WNND3-3	viii	db5	3	12-21-1	0.967	0.246	0.951	0.213
WNND3-4	ix	db5	3	16-29-1	0.968	0.195	0.947	0.216
WNND3-5	x	db5	3	8-28-1	0.959	0.278	0.930	0.256
WNND3-6	vi	db5	4	5-26-1	0.927	0.367	0.865	0.352
WNND3-7	vii	db5	4	10-27-1	0.969	0.238	0.945	0.224
WNND3-8	viii	db5	4	15-29-1	0.979	0.196	0.948	0.219
WNND3-9	ix	db5	4	20-30-1	0.983	0.179	0.951	0.213
WNND3-10	x	db5	4	10-24-1	0.967	0.246	0.936	0.242
WNND3-11	vi	db5	5	6-26-1	0.930	0.360	0.856	0.353
WNND3-12	vii	db5	5	12-24-1	0.972	0.234	0.935	0.225
WNND3-13	viii	db5	5	18-28-1	0.982	0.192	0.937	0.220
WNND3-14	ix	db5	5	24-24-1	0.986	0.175	0.940	0.213
WNND3-15	x	db5	5	12-24-1	0.970	0.241	0.926	0.243
WNNS3-1	vi	sym2	3	4-23-1	0.884	0.475	0.848	0.375
WNNS3-2	vii	sym2	3	8-30-1	0.933	0.381	0.896	0.336
WNNS3-3	viii	sym2	3	12-16-1	0.941	0.370	0.917	0.273
WNNS3-4	ix	sym2	3	16-24-1	0.940	0.316	0.907	0.290
WNNS3-5	x	sym2	3	8-28-1	0.945	0.332	0.906	0.301
WNNS3-6	vi	sym2	4	5-21-1	0.898	0.435	0.852	0.369
WNNS3-7	vii	sym2	4	10-26-1	0.942	0.338	0.903	0.318
WNNS3-8	viii	sym2	4	15-20-1	0.953	0.295	0.914	0.281
WNNS3-9	ix	sym2	4	20-24-1	0.955	0.289	0.911	0.286
WNNS3-10	x	sym2	4	10-24-1	0.954	0.294	0.912	0.284
WNNS3-11	vi	sym2	5	6-21-1	0.901	0.426	0.842	0.370
WNNS3-12	vii	sym2	5	12-24-1	0.944	0.331	0.893	0.319
WNNS3-13	viii	sym2	5	18-20-1	0.956	0.289	0.904	0.282
WNNS3-14	ix	sym2	5	24-21-1	0.958	0.283	0.901	0.287
WNNS3-15	x	sym2	5	12-24-1	0.956	0.287	0.902	0.285

Table 7. 7 Different WNN model's structure and performance for 4 day ahead lead-time.

Model	Input Dataset	Wavelet	Level of Decompos	Neuron Structure	Training		Validation	
					<i>NSE</i>	RSME (m ³ /s)	<i>NSE</i>	RSME (m ³ /s)
WNND4-1	vi	db5	3	4-30-1	0.876	0.480	0.809	0.419
WNND4-2	vii	db5	3	8-27-1	0.934	0.350	0.903	0.299
WNND4-3	viii	db5	3	12-28-1	0.953	0.296	0.912	0.284
WNND4-4	ix	db5	3	16-30-1	0.935	0.291	0.885	0.297
WNND4-5	x	db5	3	8-26-1	0.935	0.345	0.886	0.329
WNND4-6	vi	db5	4	5-29-1	0.900	0.441	0.838	0.389
WNND4-7	vii	db5	4	10-30-1	0.952	0.300	0.919	0.273
WNND4-8	viii	db5	4	15-28-1	0.962	0.266	0.918	0.274
WNND4-9	ix	db5	4	20-27-1	0.961	0.268	0.917	0.276
WNND4-10	x	db5	4	10-29-1	0.953	0.295	0.902	0.301
WNND4-11	vi	db5	5	6-30-1	0.910	0.408	0.838	0.386
WNND4-12	vii	db5	5	12-28-1	0.952	0.297	0.919	0.271
WNND4-13	viii	db5	5	18-26-1	0.964	0.132	0.915	0.136
WNND4-14	ix	db5	5	24-29-1	0.963	0.263	0.923	0.265
WNND4-15	x	db5	5	12-23-1	0.955	0.288	0.912	0.284
WNNS4-1	vi	sym2	3	4-22-1	0.843	0.542	0.794	0.434
WNNS4-2	vii	sym2	3	8-19-1	0.894	0.475	0.858	0.373
WNNS4-3	viii	sym2	3	12-26-1	0.922	0.390	0.847	0.381
WNNS4-4	ix	sym2	3	16-26-1	0.906	0.382	0.834	0.381
WNNS4-5	x	sym2	3	8-24-1	0.916	0.412	0.866	0.361
WNNS4-6	vi	sym2	4	5-21-1	0.866	0.499	0.824	0.403
WNNS4-7	vii	sym2	4	10-20-1	0.910	0.407	0.873	0.341
WNNS4-8	viii	sym2	4	15-25-1	0.931	0.351	0.852	0.368
WNNS4-9	ix	sym2	4	20-21-1	0.931	0.351	0.865	0.354
WNNS4-10	x	sym2	4	10-27-1	0.933	0.353	0.881	0.330
WNNS4-11	vi	sym2	5	6-22-1	0.876	0.461	0.823	0.399
WNNS4-12	vii	sym2	5	12-19-1	0.911	0.404	0.873	0.338
WNNS4-13	viii	sym2	5	18-24-1	0.933	0.174	0.850	0.182
WNNS4-14	ix	sym2	5	24-25-1	0.932	0.345	0.871	0.341
WNNS4-15	x	sym2	5	12-22-1	0.935	0.344	0.891	0.312

Table 7. 8 Different WNN model's structure and performance for 5 day ahead lead-time.

Model	Input Dataset	Wavelet	Level of Decompos	Neuron Structure	Training		Validation	
					<i>NSE</i>	RSME (m ³ /s)	<i>NSE</i>	RSME (m ³ /s)
WNND5-1	vi	db5	3	4-23-1	0.840	0.544	0.768	0.461
WNND5-2	vii	db5	3	8-30-1	0.914	0.400	0.861	0.357
WNND5-3	viii	db5	3	12-16-1	0.928	0.369	0.891	0.332
WNND5-4	ix	db5	3	16-29-1	0.920	0.328	0.848	0.329
WNND5-5	x	db5	3	8-21-1	0.888	0.486	0.838	0.397
WNND5-6	vi	db5	4	5-30-1	0.871	0.489	0.826	0.400
WNND5-7	vii	db5	4	10-26-1	0.944	0.321	0.898	0.307
WNND5-8	viii	db5	4	15-26-1	0.951	0.302	0.901	0.301
WNND5-9	ix	db5	4	20-20-1	0.953	0.295	0.912	0.285
WNND5-10	x	db5	4	10-19-1	0.918	0.391	0.874	0.341
WNND5-11	vi	db5	5	6-26-1	0.887	0.457	0.818	0.409
WNND5-12	vii	db5	5	12-27-1	0.911	0.326	0.943	0.285
WNND5-13	viii	db5	5	18-21-1	0.951	0.303	0.909	0.290
WNND5-14	ix	db5	5	24-23-1	0.961	0.270	0.905	0.296
WNND5-15	x	db5	5	12-23-1	0.934	0.349	0.886	0.323
WNNS5-1	vi	sym2	3	4-21-1	0.815	0.598	0.746	0.492
WNNS5-2	vii	sym2	3	8-30-1	0.868	0.543	0.799	0.457
WNNS5-3	viii	sym2	3	12-19-1	0.897	0.474	0.829	0.425
WNNS5-4	ix	sym2	3	16-27-1	0.821	0.432	0.855	0.427
WNNS5-5	x	sym2	3	8-28-1	0.881	0.506	0.816	0.432
WNNS5-6	vi	sym2	4	5-26-1	0.844	0.537	0.802	0.427
WNNS5-7	vii	sym2	4	10-29-1	0.897	0.436	0.833	0.392
WNNS5-8	viii	sym2	4	15-30-1	0.919	0.388	0.838	0.385
WNNS5-9	ix	sym2	4	20-18-1	0.919	0.389	0.851	0.370
WNNS5-10	x	sym2	4	10-30-1	0.911	0.406	0.851	0.370
WNNS5-11	vi	sym2	5	6-24-1	0.860	0.502	0.794	0.436
WNNS5-12	vii	sym2	5	12-20-1	0.866	0.443	0.874	0.365
WNNS5-13	viii	sym2	5	18-25-1	0.918	0.390	0.845	0.371
WNNS5-14	ix	sym2	5	24-21-1	0.857	0.356	0.912	0.384
WNNS5-15	x	sym2	5	12-28-1	0.927	0.363	0.863	0.351

When wavelet coefficients are used as inputs, the number of input neurons as well as the number of weights increase. Like ANN, WNN model efficiency varies by input selection. Figure 7.4 depicts the variation of different develop hybrid WNN models efficiency with different input selection and for different lead-time (L). Having only current river flow wavelet coefficient (input combination vi of Table 7.2) gives the worst result, while a combination of four river flow wavelet coefficient with different time lag (input combination ix of Table 7.2) gives the best result. Although there is not considerable difference between different WNN models, in most cases WNN models with db5 have better performance while those with sym2 and 3 levels of decomposition have the worst performance.

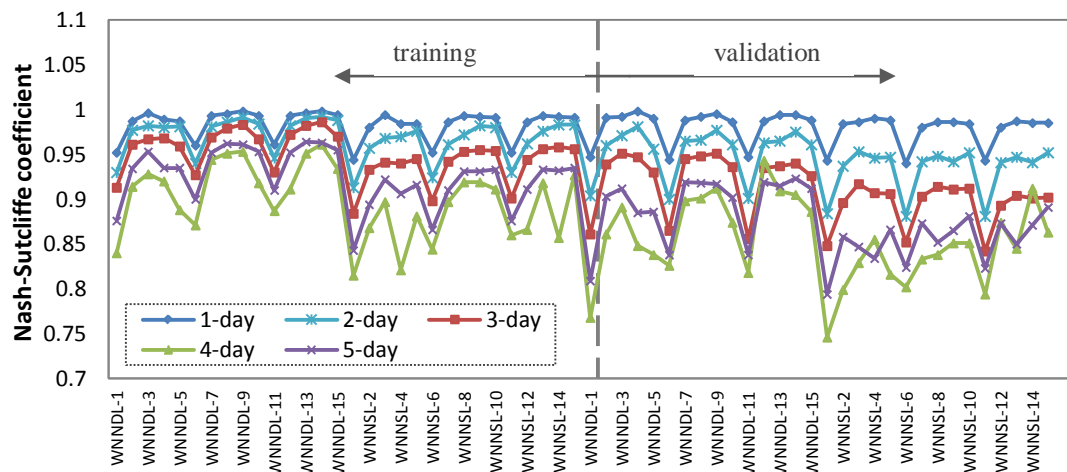


Figure 7. 4 Different hybrid WNN model efficiency for different lead-time (L) in training and validation set.

The results clearly revealed that the performance of the hybrid WNN models in both low and high lead time is better than ANN models. In particular, the higher lead time WNN performance is very satisfactory compared to ANN. This is also illustrated in Figure 7.5 which compares the RMSE of best fitted ANN and WNN. It can be seen that by increasing the lead-time the RMSE of ANN model dramatically increases compared to the WNN model. The superiority of hybrid WNN is also depicted in Figure 7.6 where 5-day ahead modelled river flow with best fitted classic ANN and hybrid WNN models are compared with observed values.

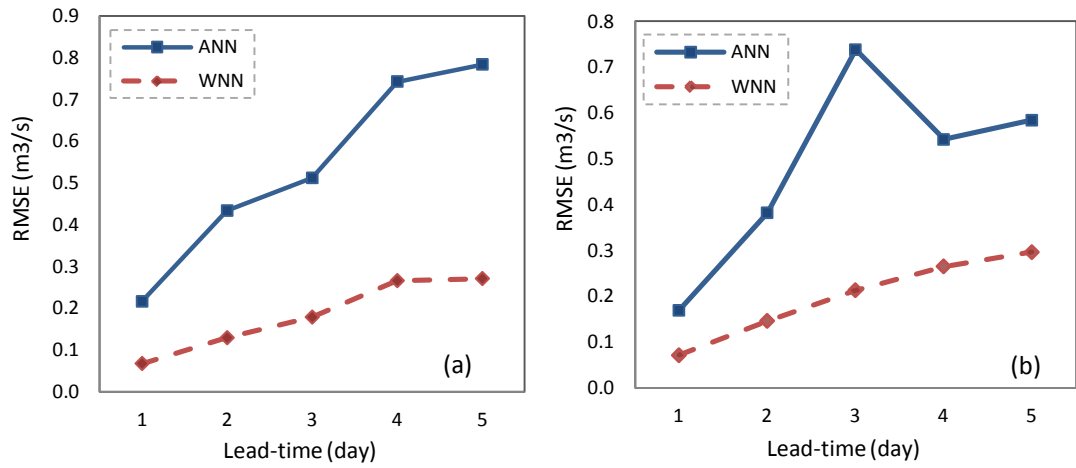


Figure 7.5 Best fitted ANN and WNN model efficiency (RMSE) variation over the lead time in (a) training; (b) verification set.

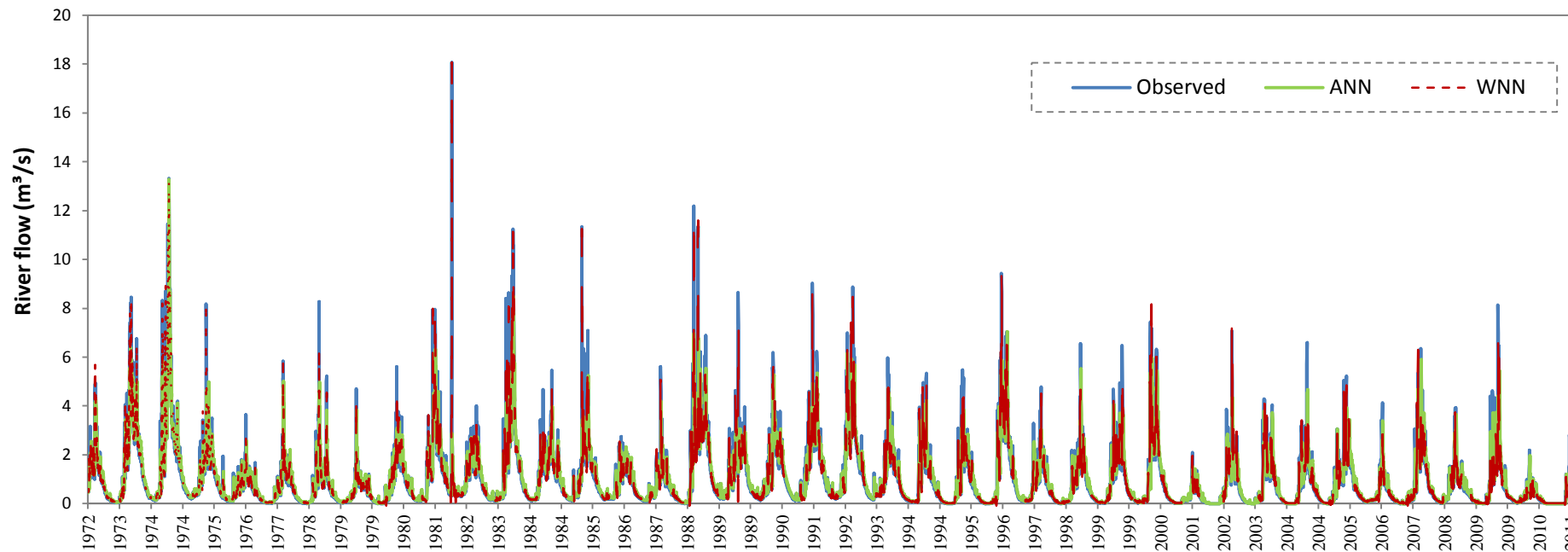


Figure 7. 6 Comparison of the observed and modeled river flow for 5-day ahead with ANN5-5 and WNND5-14 models (1972-2011).

7.3.3 Application of ANFIS

Adaptive neuro-fuzzy inference system model was developed with subtractive clustering method to generate a TSK type FIS structure. Two different input data sets were used for different lead-time forecasting. Table 7.9 shows different models' performance. The results indicate that the models with both river flow and rainfall time series have better performance compared to those with using only river flow time series.

Table 7.9 ANFIS performance for different lead time.

Lead time	Model	Input Dataset	Training		Validation	
			<i>NSE</i>	RSME (m ³ /s)	<i>NSE</i>	RSME (m ³ /s)
t+1	ANFIS1-1	i	0.916	0.265	0.892	0.438
	ANFIS1-2	v	0.969	0.162	0.963	0.256
t+2	ANFIS2-1	i	0.800	0.462	0.763	0.649
	ANFIS2-2	v	0.856	0.408	0.823	0.561
t+3	ANFIS3-1	i	0.718	0.499	0.698	0.733
	ANFIS3-2	v	0.734	0.485	0.720	0.705
t+4	ANFIS4-1	i	0.659	0.673	0.656	0.782
	ANFIS4-2	v	0.683	0.633	0.668	0.768
t+5	ANFIS5-1	i	0.628	0.569	0.612	0.813
	ANFIS5-2	v	0.634	0.553	0.635	0.806

Also, it can be seen from the results that model performance declines drastically with increasing lead-time. The Nash-Sutcliffe coefficient of efficiency decreases from 0.97 for one step ahead to 0.63 for five steps ahead forecasting. Figure 7.7 shows the scatter plots between observed and forecasted river flow with the best fitted ANFIS models, for different lead time. This figure illustrates that like ANN models, ANFIS models fail in simulating the extreme river flow for the lead times greater than 3 days.

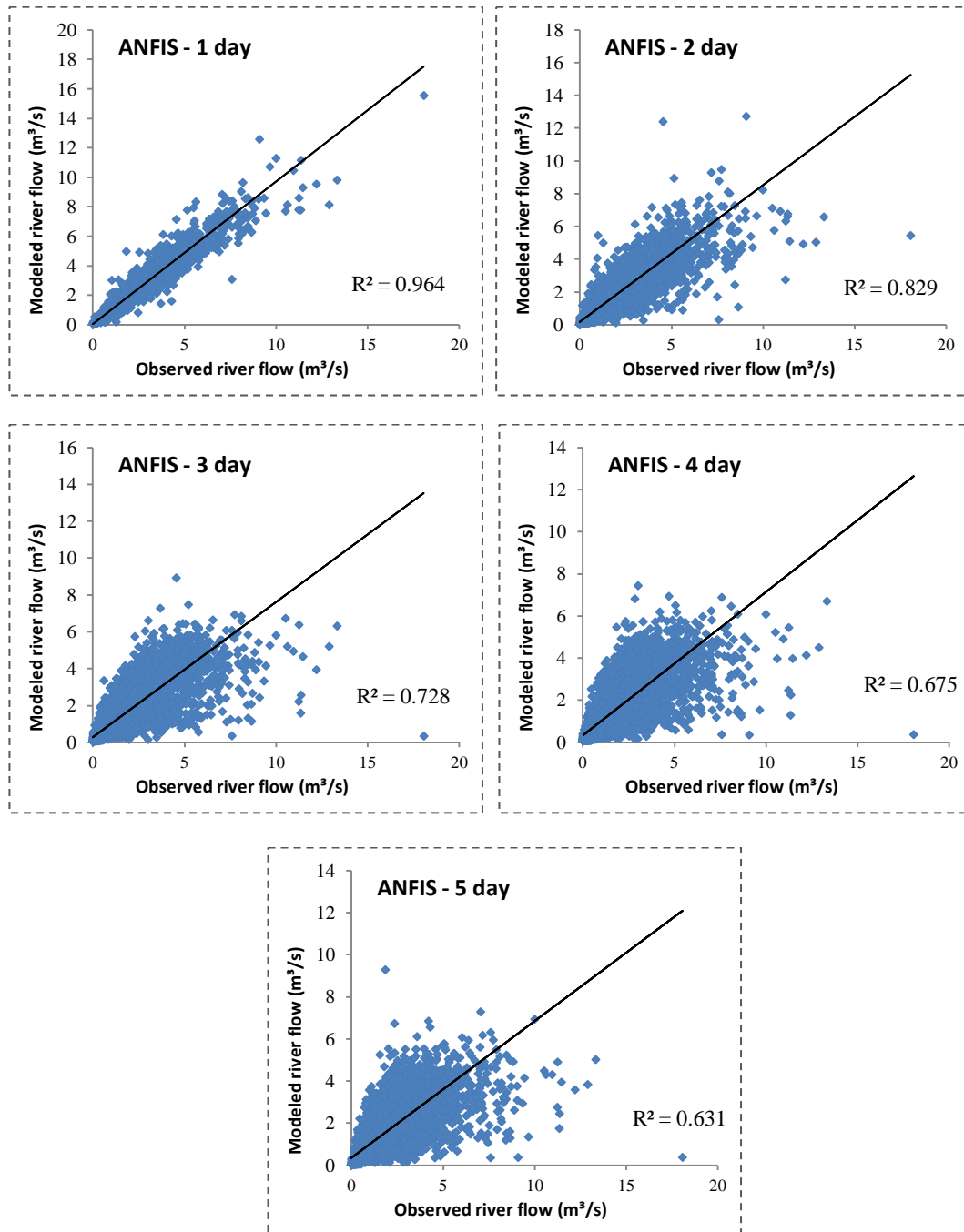


Figure 7.7 Scatter plots of observed and ANFIS forecasted flow for different lead time.

7.3.4 Improving the efficiency of ANFIS with WNF

Wavelet neuro-fuzzy hybrid model was developed by using the wavelet sub-series as ANFIS input. In the WNF model, same db5 and sym2 wavelet coefficient with 3,4 and 5 levels of resolution was used to make a comparison just based on the artificial

intelligence approach (ANN and ANFIS) performances. Figure 7.8 shows the structure of hybrid WNFD1-1 model, generated by subtractive clustering as an example. Inputs are db5 wavelet coefficients including one approximation and three details for three levels of decomposition, with generalized bell membership function and two rules.

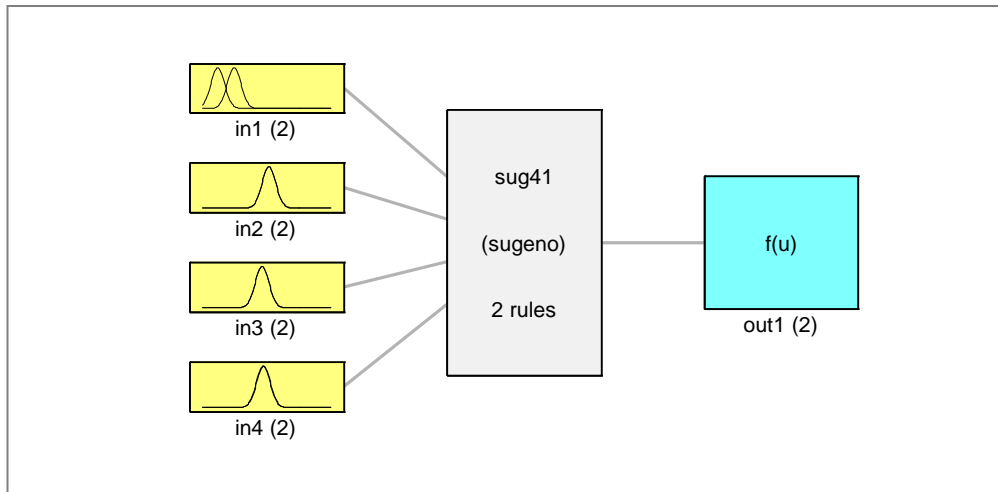


Figure 7. 8 Hybrid WNFD1-1 model structure, generated with subtractive clustering approach.

The performances of different WNF models are very similar for each lead time, but usually models using sym2 wavelet with 3 level of resolution have the worst performance among all. Table 7.10 shows WNF models' performance for each of the five lead-times forecasting based on the mother wavelet and level of decomposition.

Table 7. 10 WNF models' structure and performance for different lead time.

Lead time	Model	Input Dataset	Wavelet	Level of Decompos	Training		Validation	
					<i>NSE</i>	RSME (m ³ /s)	<i>NSE</i>	RSME (m ³ /s)
t+1	WNFD1-1	vi	db5	3	0.943	0.156	0.928	0.550
	WNFD1-2	vi	db5	4	0.944	0.135	0.928	0.357
	WNFD1-3	vi	db5	5	0.935	0.352	0.930	0.233
	WNFS1-1	vi	sym2	3	0.982	0.125	0.942	0.336
	WNFS1-2	vi	sym2	4	0.937	0.368	0.924	0.229
	WNFS1-3	vi	sym2	5	0.982	0.489	0.938	0.122
t+2	WNFD2-1	vi	db5	3	0.871	0.452	0.879	0.317
	WNFD2-2	vi	db5	4	0.881	0.341	0.882	0.320
	WNFD2-3	vi	db5	5	0.880	0.461	0.878	0.319
	WNFS2-1	vi	sym2	3	0.875	0.459	0.879	0.342
	WNFS2-2	vi	sym2	4	0.880	0.486	0.867	0.316
	WNFS2-3	vi	sym2	5	Failed	-	-	-
t+3	WNFD3-1	vi	db5	3	0.862	0.496	0.829	0.378
	WNFD3-2	vi	db5	4	0.868	0.499	0.836	0.370
	WNFD3-3	vi	db5	5	0.872	0.378	0.845	0.362
	WNFS3-1	vi	sym2	3	0.870	0.487	0.817	0.537
	WNFS3-2	vi	sym2	4	0.840	0.538	0.837	0.365
	WNFS3-3	vi	sym2	5	Failed	-	-	-
t+4	WNFD4-1	vi	db5	3	0.826	0.556	0.780	0.418
	WNFD4-2	vi	db5	4	0.838	0.421	0.810	0.406
	WNFD4-3	vi	db5	5	0.837	0.539	0.806	0.402
	WNFS4-1	vi	sym2	3	0.801	0.630	0.797	0.421
	WNFS4-2	vi	sym2	4	0.811	0.588	0.806	0.397
	WNFS4-3	vi	sym2	5	Failed	-	-	-
t+5	WNFD5-1	vi	db5	3	0.798	0.609	0.739	0.466
	WNFD5-2	vi	db5	4	0.822	0.478	0.792	0.465
	WNFD5-3	vi	db5	5	0.798	0.612	0.763	0.448
	WNFS5-1	vi	sym2	3	0.796	0.573	0.714	0.663
	WNFS5-2	vi	sym2	4	0.787	0.619	0.784	0.422
	WNFS5-3	vi	sym2	5	Failed	-	-	-

It can be seen from the results that the coefficient of efficiency values change with respect to the lead times. This change is more drastic in ANFIS model (0.97 to 0.634 for *NSE*) compare to WNF (0.98 to 0.82 for *NSE*).

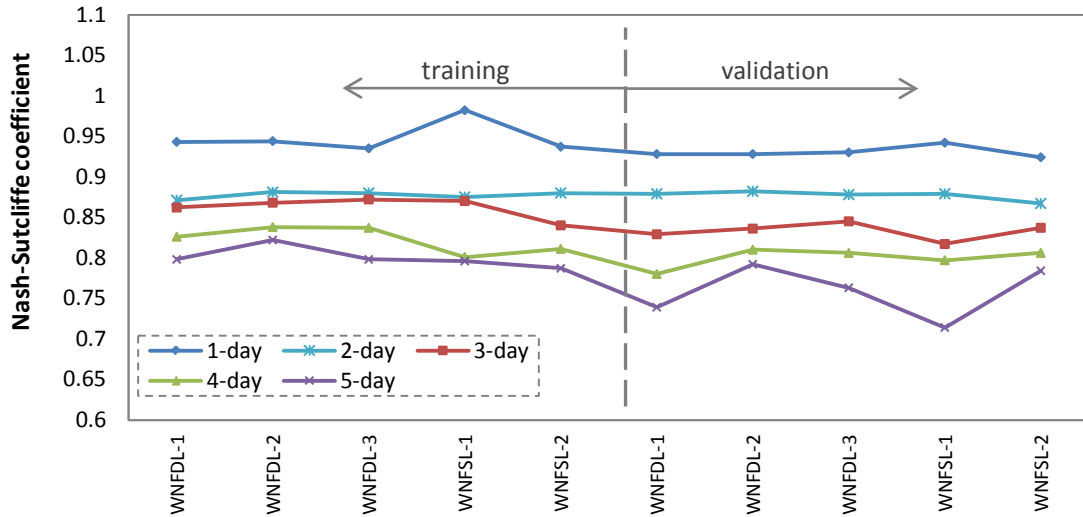


Figure 7.9 Different hybrid WNF model efficiency for different lead-time (L) in training and validation set.

Figure 7.9 demonstrates how hybrid neuro-fuzzy model efficiency varies by altering the wavelet transform and level of decomposition for different lead times. Figure 7.10 also compares the RMSE of best fitted ANFIS and WNF models. It can be seen that the increasing rate of RMSE by the lead-time is higher for ANFIS model compared to the WNN model. The results revealed that the WNF model performance is much better than ANFIS especially for the longer lead time forecasting. Figure 7.11 compares the hydrograph of observed and modelled river flow for 5-day ahead with best fitted ANFIS and WNF models.

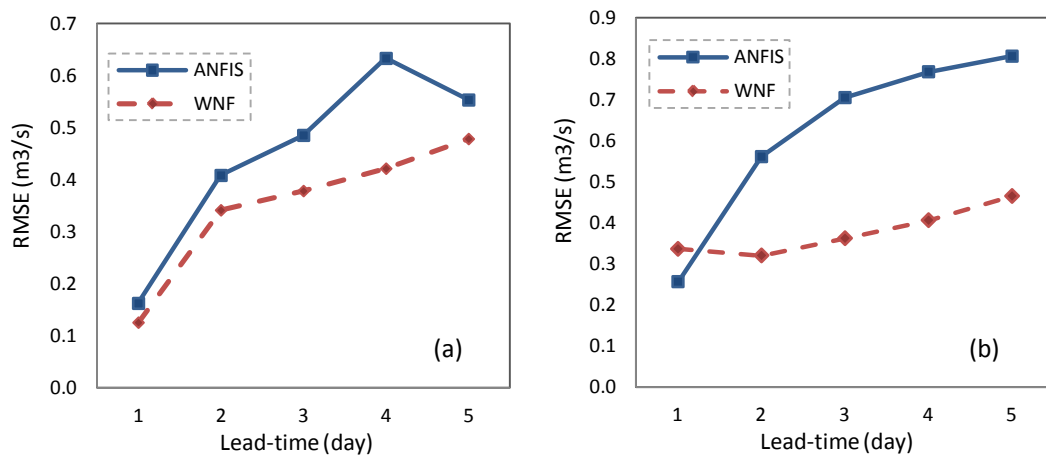


Figure 7.10 ANFIS and WNF models' efficiency (RMSE) variation over the lead time in (a) training; (b) verification set.

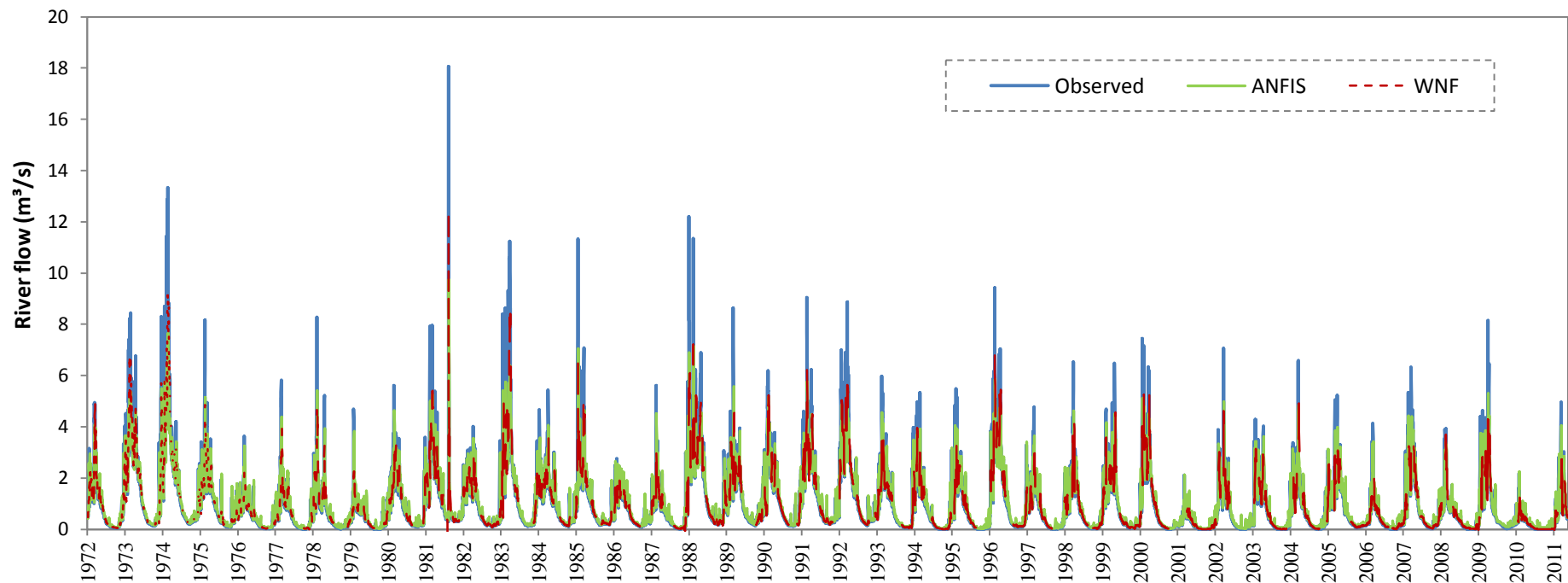


Figure 7. 11 Comparison of the observed and modeled river flow for 5-day ahead with ANFIS5-2 and WNF5-2 models (1972-2011).

7.3.5 Model Comparison

The results confirm that pre-processing the data improves artificial intelligence models' efficiency. It is also observed that neural network models' performance is slightly better than fuzzy approach performance (Figure 7.12). It could be because of constrained structure of FIS as mentioned before. That is especially relevant in this study with the very large size of input data set that makes fuzzy approach more restricted. Table 7.11 summarized the best fitted ANN, ANFIS, WNN and WNF models performance for different lead time.

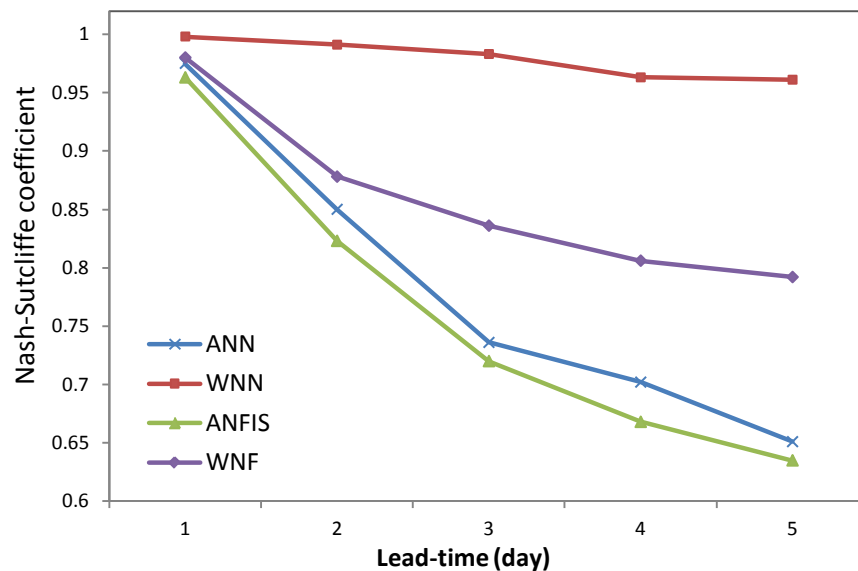


Figure 7. 12 Variation of different models' performance (R^2) over the lead time.

Best fitted WNN and WNF models' estimations and the observed river flow for the last 5 years of data (from the validation data set) are shown in Figure 7.13. It can be seen that the modeled time series match with the observed time series, while WNN provides better performance compare to the WNF.

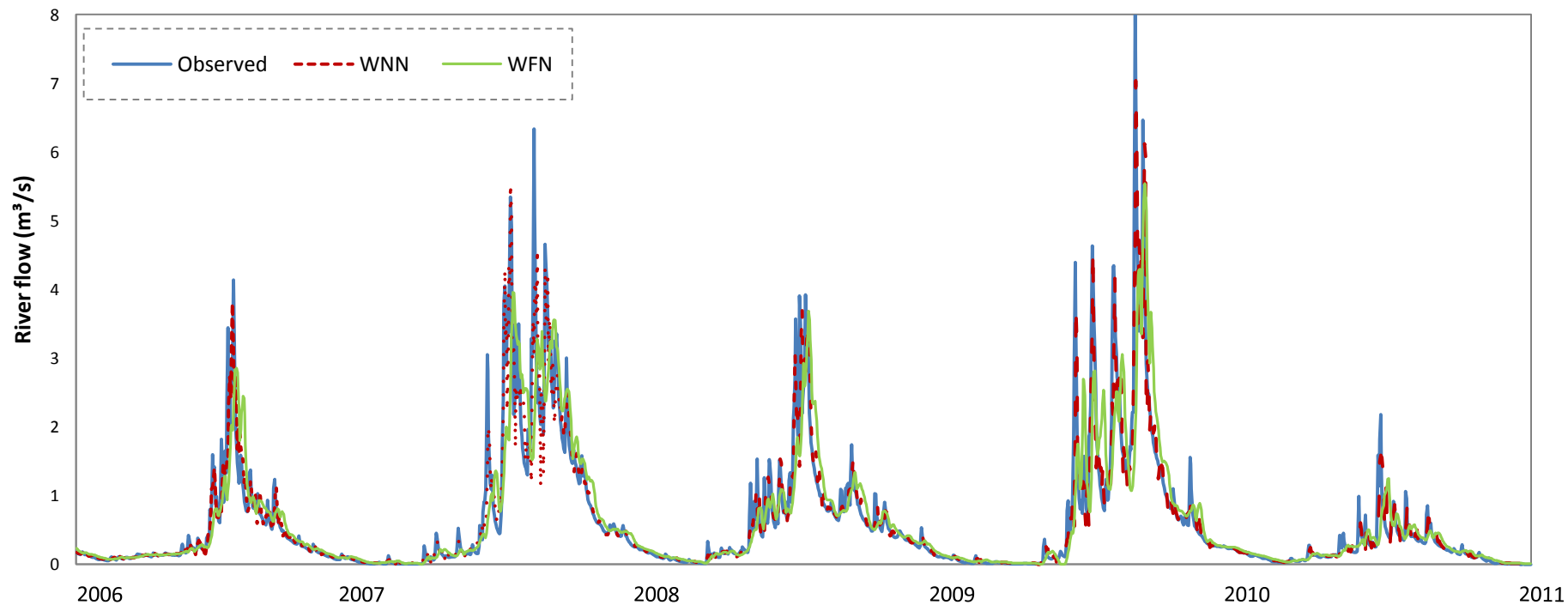


Figure 7. 13 Comparison of the observed and modeled river flow for 5-day ahead with best fitted WNN and WFN models (2006-2011).

Table 7. 11 Best fitted models performances for different lead time.

Lead time	Model	Training		Validation	
		<i>NSE</i>	RSME(m ³ /s)	<i>NSE</i>	RSME(m ³ /s)
t+1	ANN1-5	0.975	0.216	0.969	0.169
	ANFIS1-2	0.969	0.162	0.963	0.256
	WNND1-9	0.998	0.068	0.995	0.071
	WNFS1-1	0.982	0.125	0.942	0.336
t+2	ANN2-5	0.850	0.434	0.841	0.382
	ANFIS2-2	0.856	0.408	0.823	0.561
	WNND2-9	0.991	0.129	0.977	0.146
	WNFD2-2	0.881	0.341	0.882	0.320
t+3	ANN3-5	0.736	0.512	0.736	0.738
	ANFIS3-2	0.734	0.485	0.720	0.705
	WNND3-9	0.983	0.179	0.951	0.213
	WNFD3-3	0.872	0.378	0.845	0.362
t+4	ANN4-4	0.702	0.743	0.680	0.542
	ANFIS4-2	0.683	0.633	0.668	0.768
	WNND4-14	0.963	0.263	0.923	0.265
	WNFD4-2	0.838	0.421	0.810	0.406
t+5	ANN5-5	0.651	0.805	0.629	0.584
	ANFIS5-2	0.634	0.553	0.635	0.806
	WNND5-14	0.961	0.270	0.905	0.296
	WNFD5-2	0.822	0.478	0.792	0.465

Figure 7.14 shows the scatter plots of observed and best fit modelled river flow for 5-day ahead forecasting, which clearly illustrates the different model performance and the accuracy of WNN model. Also, it shows that unlike WNN model, other models are failing to simulate extreme conditions, especially for flows greater than 5 m³/s. This matter has been investigated more closely by comparing the first ten highest observed river flow in the 39 years of historical data with their modelled value (Table 7.12). The maximum relative error between simulated and observed peak flow is 3% with WNN, whereas this error is up to 72% for WNF and 98% for ANN and ANFIS.

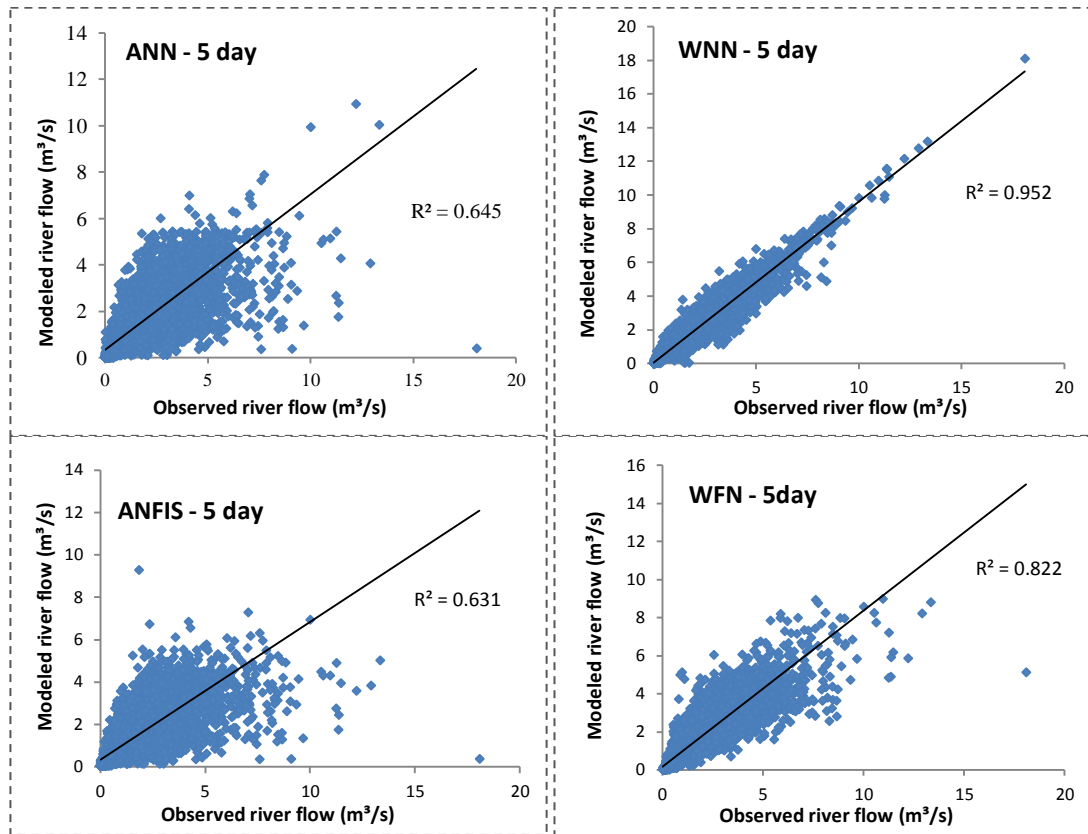


Figure 7. 14 Scatter plots of five-day ahead forecasting of the best fit ANN, WNN, ANFIS and WNF model.

Table 7. 12 Accuracy of different model in simulating five-day ahead extreme flow values.

Observed peak (m³/s)	ANN		WNN		ANFIS		WNF	
	Modeled (m³/s)	Relative Error	Modeled (m³/s)	Relative Error	Modeled (m³/s)	Relative Error	Modeled (m³/s)	Relative Error
18.07	0.41	%98	18.10	%0	0.37	%98	5.12	%72
13.32	10.04	%25	13.18	%1	5.02	%62	8.80	%34
12.89	4.07	%68	12.77	%1	3.83	%70	8.21	%36
12.19	10.94	%10	12.15	%0	3.58	%71	5.86	%52
11.45	4.29	%63	11.07	%3	3.94	%66	6.17	%46
11.35	2.37	%79	11.51	%1	2.44	%78	5.91	%48
11.33	1.76	%84	11.56	%2	1.74	%85	4.89	%57
11.24	5.44	%52	10.90	%3	4.89	%56	7.20	%36
10.94	5.14	%53	10.85	%1	4.30	%61	8.98	%18

Figure 7.15 shows how developed models estimated the highest river flow in the historical time series, which happened on January 1982. This figure again confirms the reliability of WNN model for forecasting the sudden extreme events.

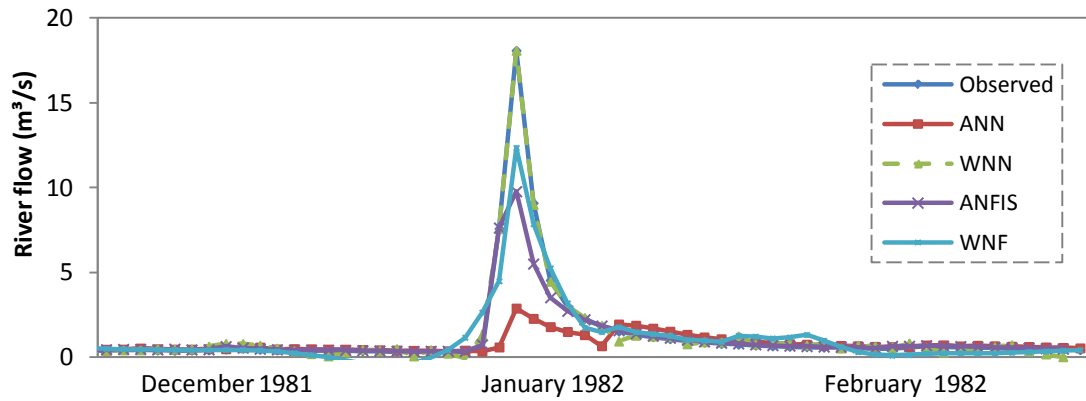


Figure 7.15 Highest observed river flow in historical time series and its estimation with developed models.

7.4 CONCLUSION

This section introduces a novel approach of applying wavelet neuro-fuzzy with a subtractive clustering method for river flow simulation and forecasting. Application of multi-resolution analysis of the input data on ANN and ANFIS model performance for forecasting multi-step ahead daily river flow has been investigated. Daubechies5 and Symlet2 wavelets were applied on Harvey River flow and rainfall time series to decompose the time series in 3, 4 and 5 levels of resolution. Different time lag combinations of wavelet coefficients were fed into ANN and ANFIS models as their inputs. The overall results show that pre-processing the raw data with wavelet has significantly improved the accuracy of forecasting.

The results also indicated that the performance improvement is more substantial for longer lead-times. Although combining the wavelet with fuzzy neural network improves ANFIS model efficiency, WNN is the best fit model among all. In this study having both river flow and rainfall time series as inputs (input combination v),

lead to the best fit ANN and ANFIS models in terms of performance criteria. Whereas, best input selection for hybrid WNN models is the combination of up to four different time lag of river flow wavelet coefficients (input combination ix). Reaching very satisfactory result, even for higher lead-times could be possible because of the strong correlation structure of Harvey River flow time series, as its autocorrelation coefficients decrease very gradually by increasing the time lag.

The results further verified that altering mother wavelet or the level of decomposition does not have a considerable impact on models' performance for each lead time, though using symlet2 with three levels of decomposition has the lowest efficiency among other hybrid models.

Applying wavelets multi-resolution analysis, which extract different frequencies from historical data, helps predicting extreme values more accurately. In this case study, WNN model was the only model which had a very satisfactory performance in the extreme conditions simulation, even for five-day ahead forecasting. The quality of the historical time series plays an important role in this situation. Since in this case study, we had 39 years of daily data, the reliability of the WNN method in simulating the peak flows needs to be verified by studying more cases with different quality of available datasets.

The outcome of this study will be useful for hydrologists, hydrological designers and decision makers in forecasting river flows and developing sustainable water distribution plan.

Chapter 8

Real Time Runoff Flow Forecasting for Flood Risk Management

8.1 INTRODUCTION

Undoubtedly, having appropriate flood-warning systems could save lives and reduce damaging effects of floods (Penning-Rowsell et al., 2014). Improving flood protection plan has high priority in many countries' political agenda (Cloke and Pappenberger, 2009). However, this goal cannot be achieved unless through accurate and timely flood forecasting system.

In this chapter an hourly rainfall-runoff model for the purpose of timely flood warning is developed. The application of different data-driven approaches for real time flood forecasting of the Richmond River, NSW, Australia, is investigated. Richmond River is highly prone to flooding and its characteristics are totally different with intermittent Western Australian rivers. The application of classic feed-forward artificial neural networks and adaptive neuro-fuzzy model with grid partitioning are first investigated. Model performance further improved by applying discrete wavelet transform (DWT). Daubechies mother wavelet is selected for multi resolution analysis of the hybrid wavelet neural network (WNN) and wavelet neuro-fuzzy model (WNFG) input. Hourly river flow and rainfall data of the Casino gauging station of the Richmond River, including historical flood data, is collected. Developed models are applied for forecasting 1, 6, 12, 24, 36 and 48 hour ahead of

river flow at this station. The performance of models is also examined when an upstream river flow data (Wiangaree station), is employed as additional input.

8.2 STUDY AREA AND DATA USED

In this study, the Casino observation station of Richmond River is considered as a case study. The Richmond River is one of the largest rivers, located in the Northern part of New South Wales, Australia. The reason for selecting this study area is that Richmond River is highly prone to flooding and has experienced flooding a number of times. The Richmond River catchment area is approximately $6,900 \text{ km}^2$, discharging into the Pacific Ocean. The catchment has the steep mountainous topography in the upper boundary. The major flow is forming by integration of mountain streams reaching the floodplain at the Casino town and then passes the region in a large flow path until it reaches to the Coraki (Caddis, 2010).

For this study mean hourly rainfall and river flow discharge for 5 years, with an observation period from 2009 to 2014 (time series with a length of 43800 data), are collected from NSW water information website. The first 70 percent of data are used for training and the remaining are used for validation. The average hourly river flow of the Casino station is $23.36 \text{ m}^3/\text{s}$. The maximum flow during the study period is $1276.86 \text{ m}^3/\text{s}$ which caused flood on 22nd May 2009. For improving the accuracy of the forecasting, river flow data from an upstream station, Wiangaree station, is also collected. Figure 8.1 illustrates the location of these two stations in the Richmond River catchment and highlights the flood plain map of the catchment.

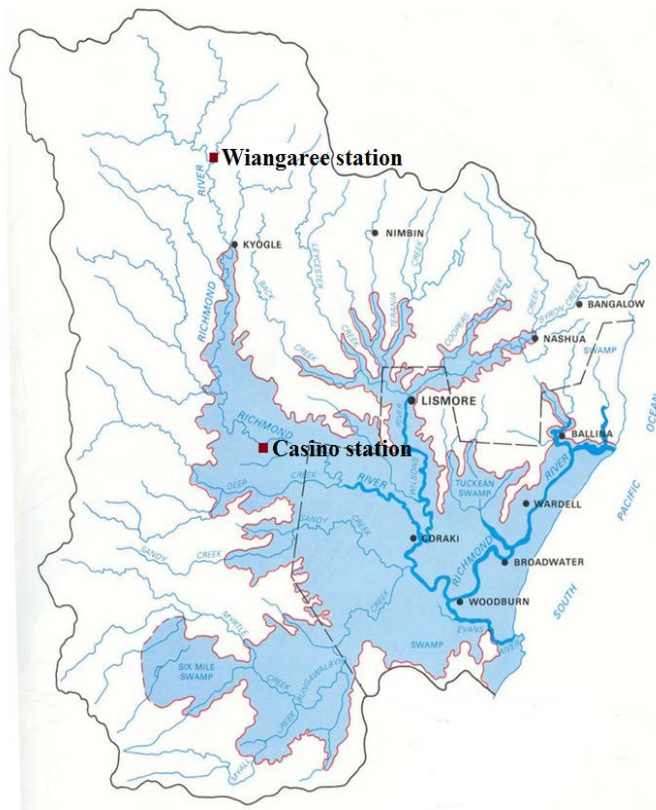


Figure 8. 1 Richmond River catchment and its flood plain
 (<http://australiasevereweather.com/floods>, Bath, 2014)

Statistics of river flow time series of both Casino and Wiangaree stations is shown in Table 8.1, including mean, minimum, maximum and standard deviation. Historical hourly river flow and rainfall time series for both training and validation sets of these two stations are also shown in Figure 8.2 and Figure 8.3.

Table 8. 1 Statistical parameters of Richmond River flow data sets of the Casino and Wiangaree stations.

Station	Data Set	River flow (1000 m ³ /day)			
		Q _{mean}	Q _{min}	Q _{max}	Q _{stdv}
Casino	Training	24.613	0.276	1276.86	72.836
	Validation	20.427	0.823	758.63	67.69
	Total	23.359	0.276	1276.861	71.357
Wiangaree	Training	10.511	0.181	1437.495	37.66
	Validation	9.211	0.516	728.26	37.58
	Total	10.121	0.181	1437.495	37.64

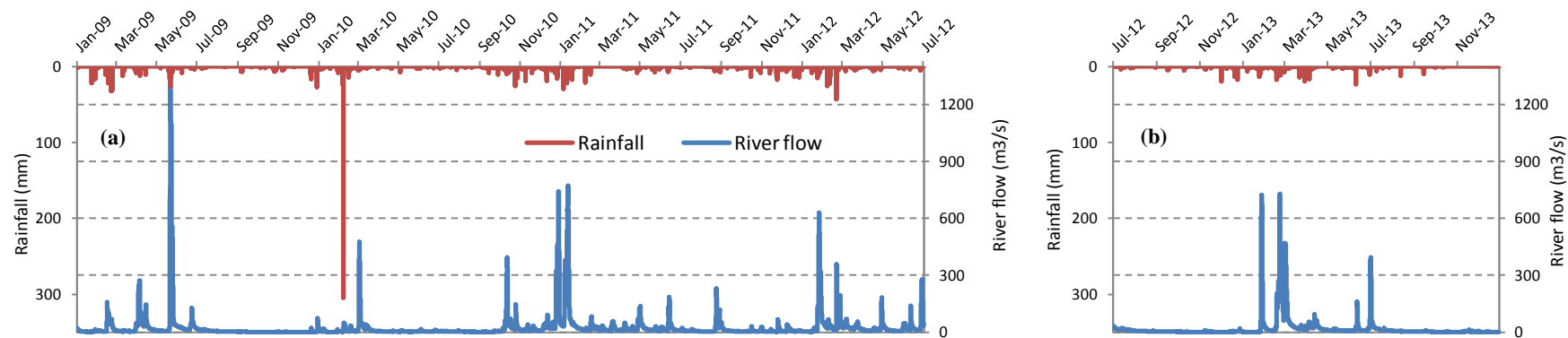


Figure 8.2 Hourly river flow and rainfall time series of Casino station for (a) training and (b) validation set.

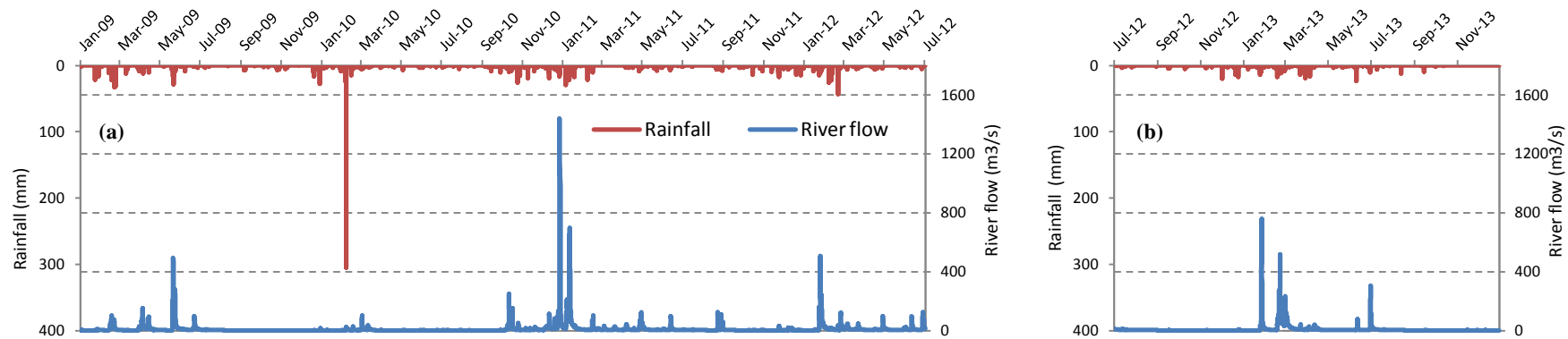


Figure 8.3 Hourly river flow and rainfall time series of Wiangaree station for (a) training and (b) validation set.

8.3 ANALYSIS AND RESULTS

8.3.1 ANN-based models

For hourly river flow forecasting, first a three layered ANN with back propagation algorithm is developed. As explained in Chapter three, the optimum structure of a network is obtained by trial and error. The number of hidden neurons is increased from one to twenty in each trial to achieve the optimum number. The model is also applied for various lead times of 1, 6, 12, 24, 36 and 48 hour ahead, for runoff forecasting. In the first step, only rainfall and river flow of Casino station time series is used as the model input. The best structure of ANN for each lead time and their performances are given in Table 8.2. It is evident that the efficiency of the model decreases by increasing the lead time where the model is unreliable for 36 and 48 hour ahead forecasting, as the Nash Sutcliffe efficiency is less than 0.5.

Table 8. 2 ANN models structure and performance using only Casino station data.

Model	Lead time	Model Neuron Structure	Calibration		Validation	
			NSE	RSME (m ³ /s)	NSE	RSME (m ³ /s)
ANN1S-1	1	2-3-1	0.998	244.8	0.999	212.6
ANN1S-2	6	2-15-1	0.972	1055.4	0.974	950.0
ANN1S-3	12	2-15-1	0.881	2170.5	0.884	1990.1
ANN1S-4	24	2-12-1	0.634	3808.8	0.627	3570.4
ANN1S-5	36	2-20-1	0.441	4706.7	0.409	4496.0
ANN1S-6	48	2-18-1	0.321	5188.7	0.255	5048.2

For improving the rainfall-runoff modelling, data of an upstream station is also integrated to the model input. Hourly river flow time series of Wiangaree station is used for this reason. The best structure and performances of the neural networks models, with the input of rainfall and river flow data of both stations are given in Table 8.3. The results show that the model efficiency is significantly improved by adding the upstream data, especially for 36 and 48 hours lead time.

Table 8. 3 ANN models structure and performance using Casino and Wiangaree stations data.

Model	Lead time	Model Neuron Structure	Calibration		Validation	
			NSE	RSME (m ³ /s)	NSE	RSME (m ³ /s)
ANN2S-1	1	3-19-1	0.999	155.4	0.999	162.1
ANN2S-2	6	3-7-1	0.988	685.9	0.986	695.6
ANN2S-3	12	3-9-1	0.957	1308.3	0.910	1751.25
ANN2S-4	24	3-6-1	0.875	2230.3	0.740	2983.49
ANN2S-5	36	3-12-1	0.699	3453.6	0.630	3274.6
ANN2S-6	48	3-2-1	0.506	4426.4	0.395	4548.0

In the next step, hybrid WNN model is applied for forecasting. Considering previous studies in the literature, this research selected db3 mother wavelet as the multi-resolution analysis of input time series. Both rainfall and river flow time series are decomposed by db3 mother wavelet into five level of decomposition. All of this wavelet coefficient (12 altogether) is imposed into the neural network model for different lead time forecasting. The best fitted model performance and structure with optimum number of hidden neurons is given in Table 8.4.

Table 8. 4 Hybrid WNN models structure and performance using only Casino station data.

Model	Lead time	Wavelet Level	Model Neuron Structure	Calibration		Validation	
				NSE	RSME (m ³ /s)	NSE	RSME (m ³ /s)
WNN1S-1	1	db3-5	12-10-1	0.999	136.605	0.999	178.33
WNN1S-2	6	db3-5	12-15-1	0.989	659.419	0.980	833.44
WNN1S-3	12	db3-5	12-7-1	0.962	1221.297	0.946	1352.8
WNN1S-4	24	db3-5	12-19-1	0.888	2107.750	0.849	2268.4
WNN1S-5	36	db3-5	12-12-1	0.724	3305.438	0.716	3106.45
WNN1S-6	48	db3-5	12-16-1	0.562	4169.415	0.493	4309.12

Comparing the result with original ANN and same data set, significant improve in model efficiency is achieved. However, the performance of a hybrid model with single station data is almost as good as the ANN model with added upstream data for shorter lead-time and better for longer lead-time. Therefore, for further improvement of forecasting accuracy, data of both stations are used as input for the hybrid model.

Again, rainfall and two river flow time series are decomposed by db3 wavelet into five level of decomposition (18 wavelet coefficients), and imposed into ANN model. The result of hybrid WNN model with added upstream station data is given in the Table 8.5.

Table 8.5 Hybrid models structure and performance using Casino and Wiangaree stations data.

Model	Lead time	Wavelet Level	Model Neuron Structure	Calibration		Validation	
				NSE	RSME (m ³ /s)	NSE	RSME (m ³ /s)
WNN2S-1	1	db3-5	18-8-1	0.999	90.96	0.999	141.180
WNN2S-2	6	db3-5	18-12-1	0.995	462.82	0.993	476.199
WNN2S-3	12	db3-5	18-11-1	0.988	676.71	0.977	891.067
WNN2S-4	24	db3-5	18-12-1	0.958	1287.19	0.934	1502.815
WNN2S-5	36	db3-5	18-6-1	0.760	2926.75	0.754	2897.018
WNN2S-6	48	db3-5	18-3-1	0.593	3888.53	0.516	4152.275

Adding another river flow time series as extra hybrid WNN model input, shows the best results in terms of performance criteria. Figure 8.4 and Figure 8.5 compare NSE and RMSE of developed models for different lead time, respectively. These figures depict superiority of hybrid models and the positive impact of adding extra effective parameter as the model input.

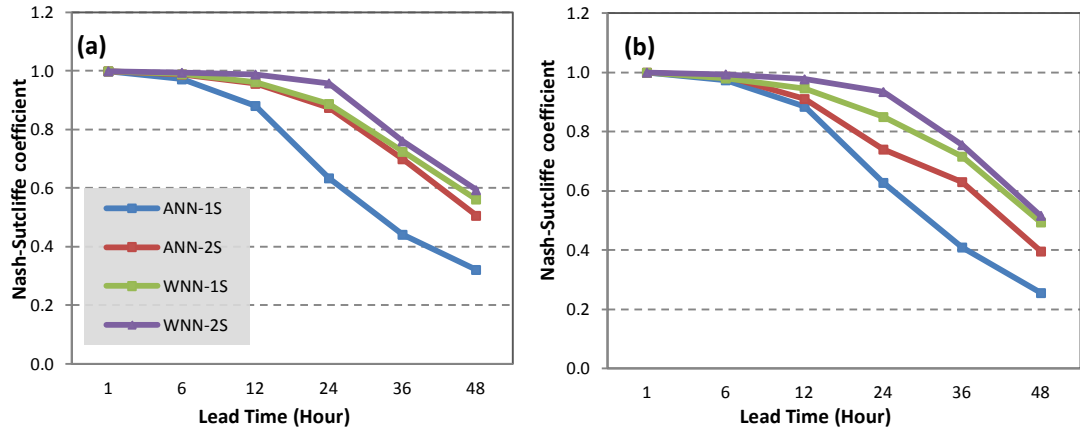


Figure 8. 4 Nash-Sutcliffe coefficient of (a) training and (b) validation set of ANN-based models.

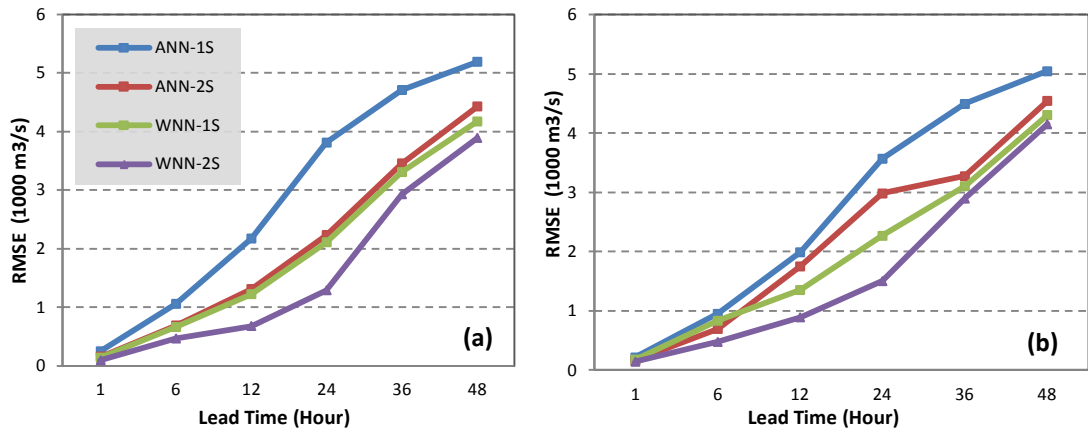


Figure 8. 5 Root mean square error of (a) training and (b) validation set of ANN-based models.

Figure 8.6 shows the scatter plot between the observed and modelled values of Casino station hourly river flow of the 24 hour lead time. This scatter plots show that adding upstream data, improves prediction reliability . It is also clearly evident that hybrid WNN model significantly outperform ANN model, particularly in extreme events forecasting. Therefore, WNN-2S model with NSE of 0.96, is a powerful tool for real time river flow forecasting, which provides reliable warning at least 24 hours before flood events.

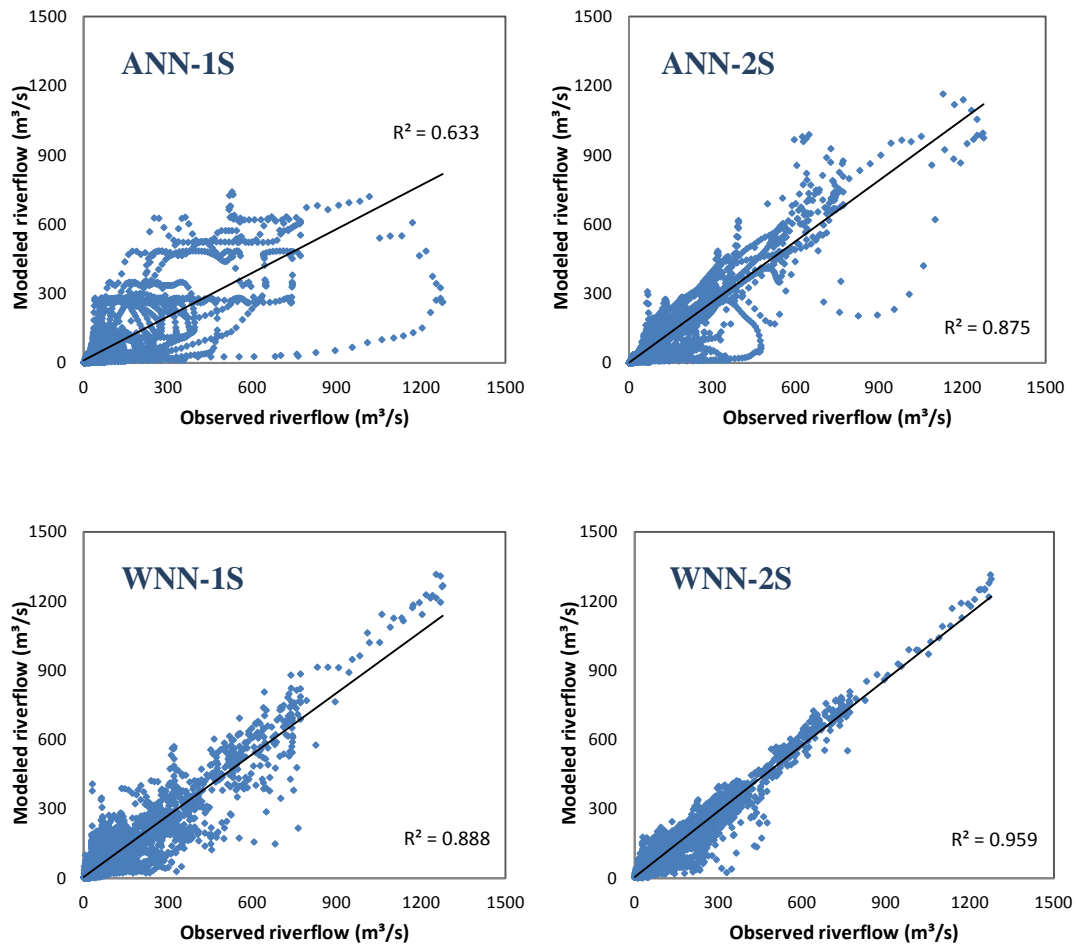


Figure 8. 6 Scatter plots of observed and simulated river flow for 24 hour lead time with different ANN-based models.

As extreme events forecasting is one of the most important and critical application of hydrological forecasting, the ability of hybrid models in simulating extreme events is examined using the recent flood events. As mentioned earlier, Richmond River has experienced flooding in Casino town in May 2009. Figure 8.7 compares observed hourly river flow and simulated flow (with 24 hour lag) with different models in the period of 20th to 24th of May 2009. This graph clearly illustrates the ability of hybrid models in simulating the sudden flood event. Also Figure 8.8 demonstrates the ability of different models in simulating peak flow by comparing observed versus simulated river flow values for the four highest flow during the study period. The large gap between observed flow and single input ANN model outcomes, again clarifies incapability of ANN in flood forecasting and flood risk mitigation applications.

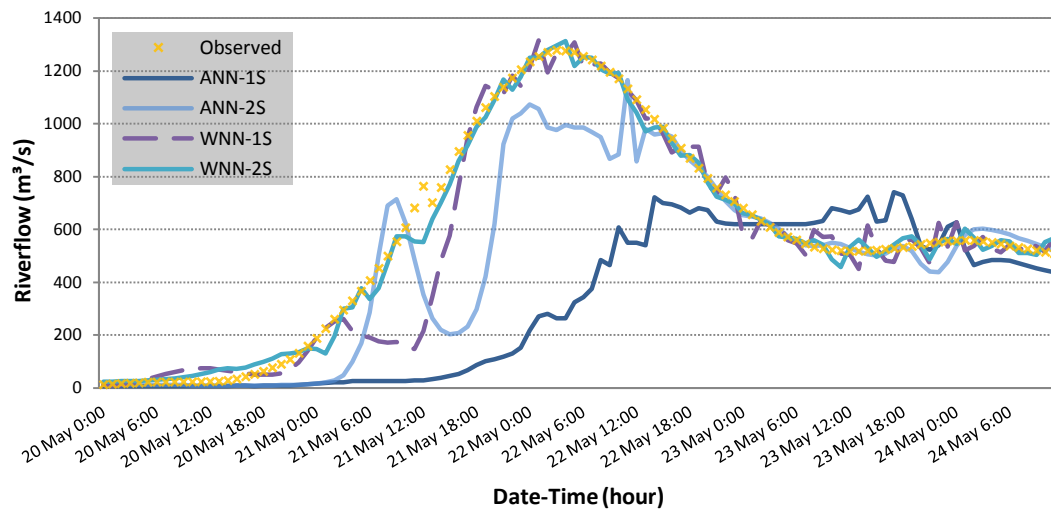


Figure 8.7 Comparing observed flood versus ANN-based modeled hourly river flow (24 hour ahead forecasts).

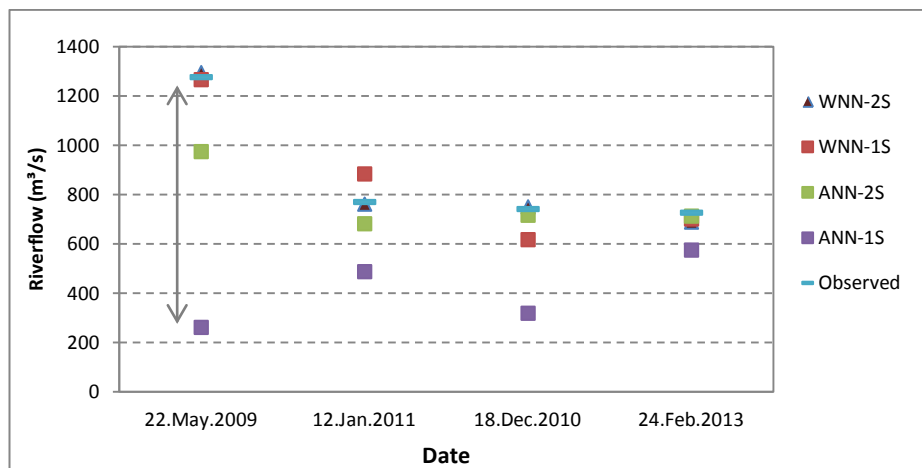


Figure 8.8 Comparing four highest observed peak flow versus ANN-based modeled values (24 hour ahead forecasts).

8.3.2 Fuzzy-based models

Adaptive neuro-fuzzy inference system (ANFIS) is used as another data-driven approach for hourly rainfall-runoff modelling of the Richmond River. Grid partitioning is applied to initialize the FIS if-then rule system. Generalized bell-

shaped function is selected as the membership function. The optimum number of membership functions with the highest model performance is achieved by trial and error and increasing the number from 2 to maximum feasible membership functions. Both rainfall and river flow data of Casino station are used as the model input for different lead time river flow forecasting. Table 8.6 shows the performances of ANFIS model and the optimum number of membership functions and correspondent fuzzy rules for two input datasets.

Table 8. 6 ANFIS models structure and performance using only Casino station data.

Model	Lead time	No. MFs	Fuzzy Rules	Calibration		Validation	
				NSE	RSME (m ³ /s)	NSE	RSME (m ³ /s)
ANF1S-1	1	6	36	0.998	230.8	0.998	192.7
ANF1S-2	6	7	49	0.970	1085.0	0.974	947.0
ANF1S-3	12	7	49	0.876	2213.7	0.885	1982.8
ANF1S-4	24	8	64	0.630	3829.7	0.628	3564.1
ANF1S-5	36	5	25	0.435	4734.8	0.410	4489.6
ANF1S-6	48	8	16	0.318	5201.0	0.255	5047.2

Results show that model efficiency is very high up to 12 hours ahead and acceptable for 24 hours ahead runoff forecasting. However, it fails to forecast lead-time of 36 hours or more.

In the next step, upstream river flow time series of Wiangaree station is also added to the model. Table 8.7 shows the results of ANFIS models using data from both Casino and Wiangaree stations . By increasing the number of inputs, the maximum feasible number of membership functions is reduced up to four. Therefore, adding another input time series just slightly improved the model efficiency.

Table 8.7 ANFIS models structure and performance using Casino and Wiangaree stations data.

Model	Lead time	No. MFs	Fuzzy Rules	Calibration		Validation	
				NSE	RSME (m ³ /s)	NSE	RSME (m ³ /s)
ANF2S-1	1	3	27	0.999	127.618	0.998	225.166
ANF2S-2	6	4	64	0.981	869.411	0.975	921.277
ANF2S-3	12	4	64	0.957	1818.534	0.891	1926.832
ANF2S-4	24	3	27	0.676	3447.144	0.649	3464.284
ANF2S-5	36	3	27	0.421	4793.193	0.408	4496.919
ANF2S-6	48	3	27	0.326	5171.172	0.296	4904.844

Application of hybrid wavelet neuro-fuzzy model is also investigated by applying db3 mother wavelet on input data. As a result of fuzzy modelling restrictions, two level of decomposition is obtained as the maximum feasible level for rainfall-runoff modelling with two input data set (rainfall and one river flow time series). Adding upstream data, this number reduced to only one level of decomposition. The reason as explained in Chapter three, is that the number of fuzzy rules increases exponentially with the number of input variables. For example, if a rainfall and two river flow time series are decomposed to two level of decomposition, the number of input data increases to nine time series, which require 512 fuzzy rules for the minimum of two membership functions. Table 8.8 and Table 8.9 show hybrid WNF models performance of different time-lead forecasting with single and double stations input data, respectively. The maximum level of decomposition and the number of membership functions for feasible modelling is considered for WNF modelling.

Table 8. 8 WNF models structure and performance using only Casino station data.

Model	Lead time	Wavelet Level	No. MFs	Fuzzy Rules	Calibration		Validation	
					NSE	RSME (m ³ /s)	NSE	RSME (m ³ /s)
WNF1S-1	1	db3-2	2	64	0.999	177.2	0.999	131.7
WNF1S-2	6	db3-2	2	64	0.986	741.0	0.979	838.4
WNF1S-3	12	db3-2	2	64	0.952	1378.5	0.930	1547.1
WNF1S-4	24	db3-2	2	64	0.851	2426.9	0.917	1680.4
WNF1S-5	36	db3-2	2	64	0.667	3631.0	0.751	2919.5
WNF1S-6	48	db3-2	2	64	0.459	4631.3	0.532	4001.9

Table 8. 9 WNF models structure and performance using Casino and Wiangaree stations data.

Model	Lead time	Wavelet Level	No. MFs	Fuzzy Rules	Calibration		Validation	
					NSE	RSME (m ³ /s)	NSE	RSME (m ³ /s)
WNF2S-1	1	db3-1	2	64	0.999	150.584	0.999	160.507
WNF2S-2	6	db3-1	2	64	0.974	792.837	0.980	833.270
WNF2S-3	12	db3-1	2	64	0.949	1336.500	0.931	1538.896
WNF2S-4	24	db3-1	2	64	0.827	2616.283	0.790	2681.384
WNF2S-5	36	db3-1	2	64	0.624	3756.337	0.500	3735.770
WNF2S-6	48	db3-1	2	64	0.417	4957.059	0.395	4595.445

Comparing the result with the original ANFIS models, the forecasting accuracy has significantly improved, especially for 24 and 36 hour ahead forecasts. However, hybrid models are also failing to forecast 48 hour ahead of runoff. Figures 8.9 and 8.10 compare NSE and RMSE of developed models for different lead time, respectively.

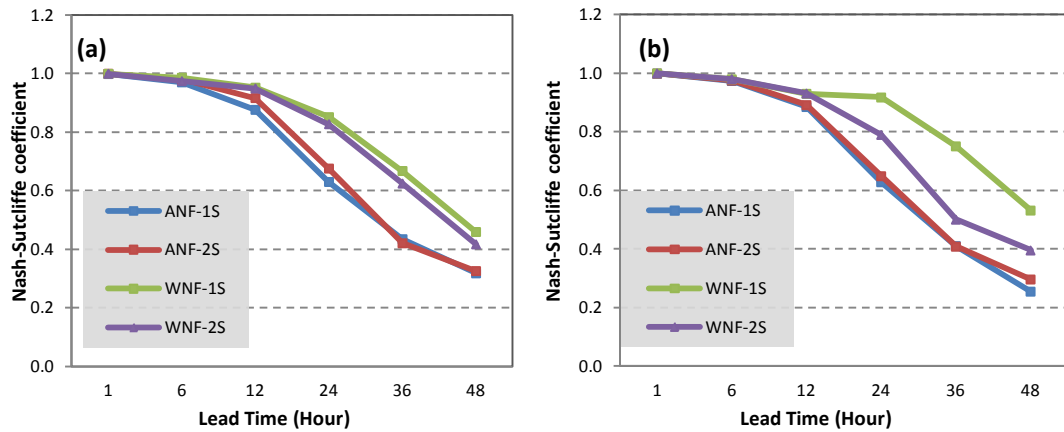


Figure 8.9 Nash-Sutcliffe coefficient of (a) training (b) validation set of fuzzy-based models.

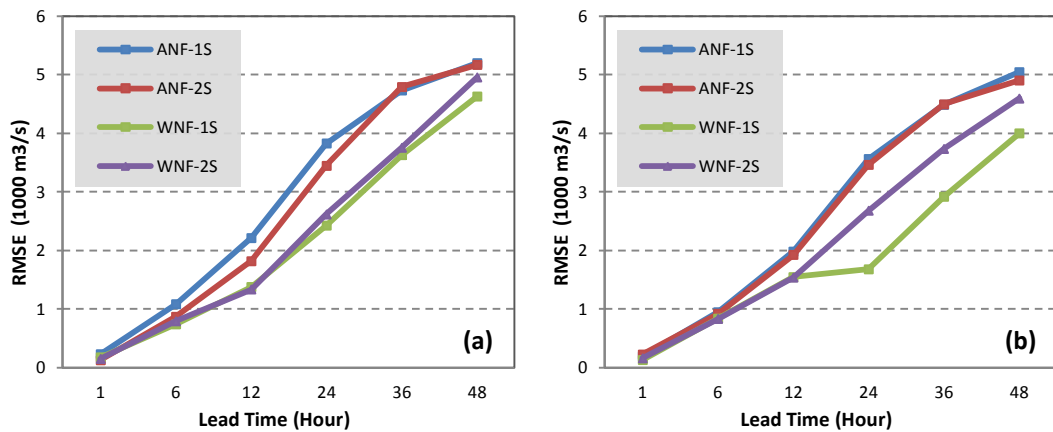


Figure 8.10 Root mean square error of (a) training and (b) validation set of fuzzy-based models.

These figures indicate that hybrid model, with only Casino station data, performs better than hybrid model with added upstream data. This pattern again confirms FIS limitation when the size of input-output variables is extremely huge and the fact that only one level of decomposition is not efficient enough for enhancing the forecasting. Figure 8.11 shows the scatter plot between the observed and modelled values of Casino station hourly river flow for 24-hour lead time. It is evident that hybrid models outperform single ANFIS model. However, they are not reliable for one day ahead flood warning.

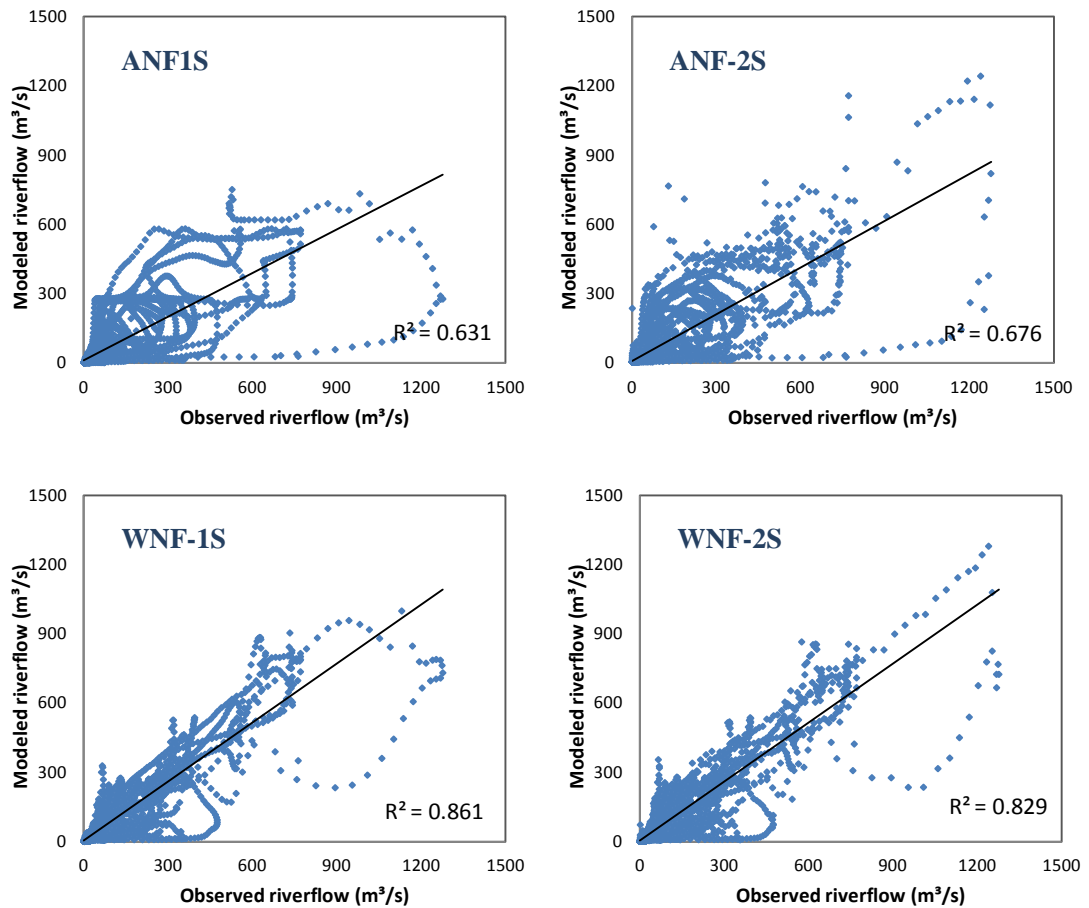


Figure 8.11 Scatter plots of observed and simulated river flow for 24 hour lead time with different fuzzy-based models.

For investigation of model performance in simulating May 2009 flood, observed and modelled river flow (with 24 hour lag) during the flood period is plotted in Figure 8.12. This graph shows applying even one or two level of decomposition and limited structure of FIS, WNF model still able to forecast 24 hour ahead flood with an acceptable accuracy. Figure 8.13 also demonstrates the ability of different fuzzy-based models in simulating peak flow by comparing observed versus simulated river flow values (24 hour lag), for the four highest flow during the study period. Generally, hybrid model outputs are closer to observed peak flow and the gap between observed flow and ANFIS models outcome is larger.

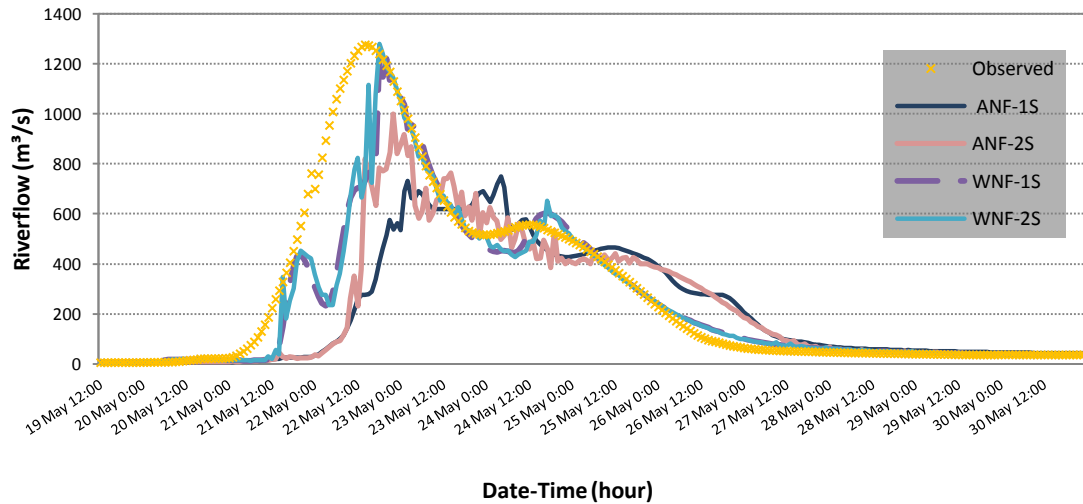


Figure 8. 12 Comparing observed flood versus fuzzy-based modeled hourly river flow (24 hour ahead forecasts).

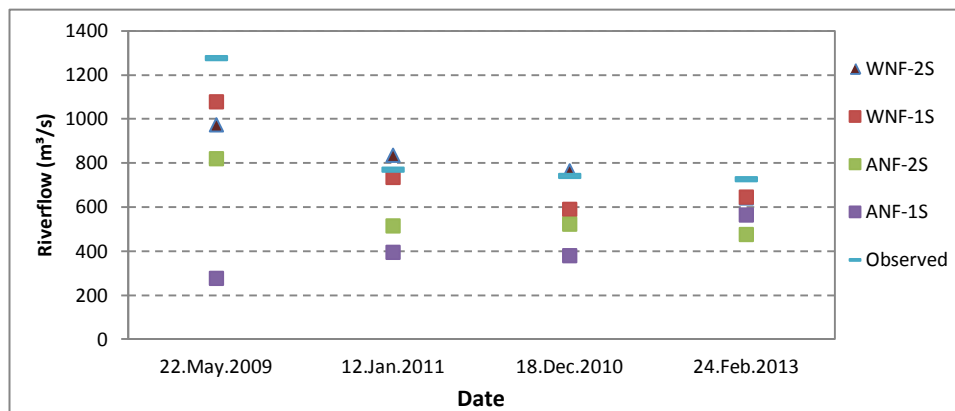


Figure 8. 13 Comparing four highest observed peak flow versus fuzzy-based modeled values (24 hour ahead forecasts).

8.3.3 Models comparison

Comparing all models, hybrid WNN model is the most efficient among all, especially for longer than 24 hour ahead forecasting. ANN model efficiency is generally higher than ANFIS model for different lead times in this case study. Whilst NSE of the best fitted ANN model for 36 hour ahead flow forecasting is 0.70, where ANFIS model fails to do this forecast (NSE=0.42). Figure 8.14 to Figure 8.17 compares scatter plot between observed and the best fitted ANN, ANFIS, WNN and

WNF modelled runoff for 12, 24, 36 and 48 hour lead times, respectively. The superiority of ANN-based models is clearly evident in these figures. However, comparing flood hydrographs (Figure 8.7 and Figure 8.12) indicates that neither ANN nor ANFIS models are reliable for flood forecasting. Applying wavelet decomposition on ANN inputs (rainfall and river flow time series) leads to a highly accurate flood forecasting of 24 hours prior to the event.

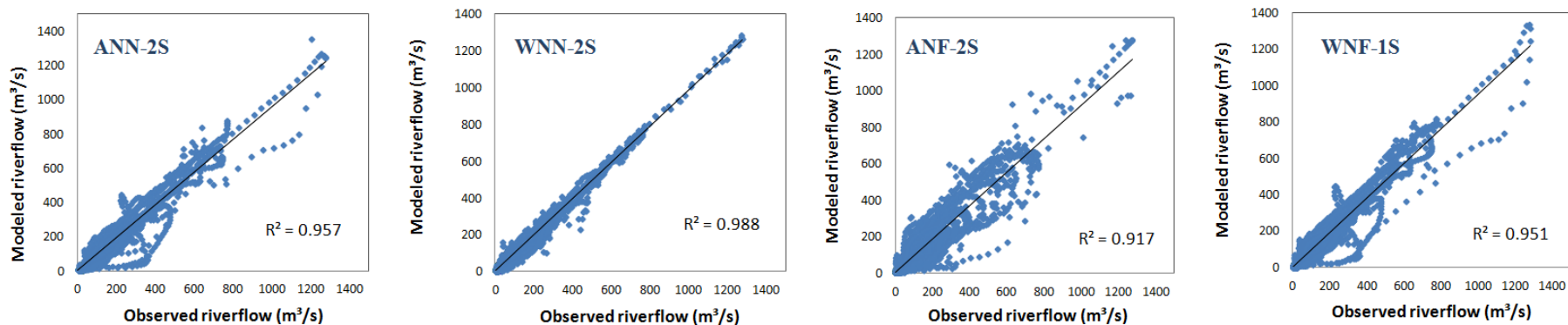


Figure 8.14 Scatter plots of observed and simulated river flow for 12 hour lead time with best fitted ANN, WNN, ANFIS and WNF models.

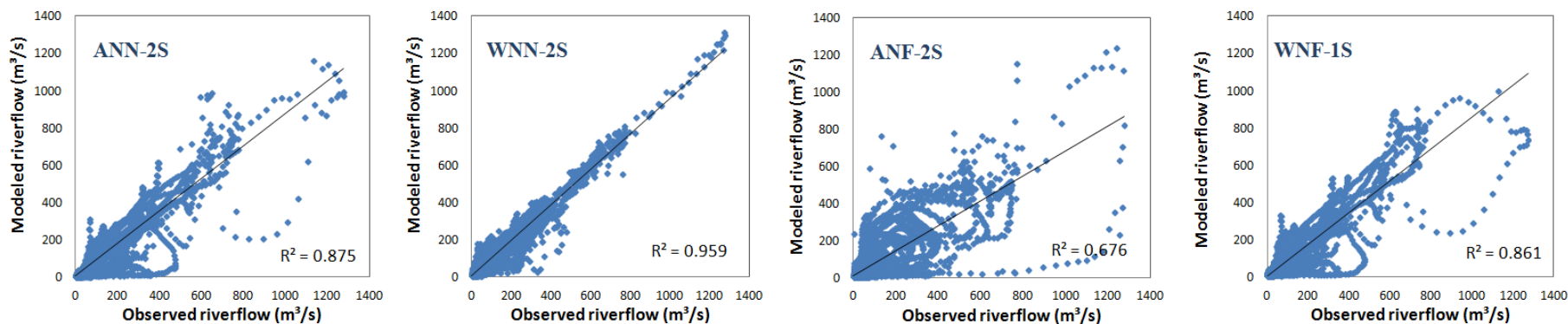


Figure 8.15 Scatter plots of observed and simulated river flow for 24 hour lead time with best fitted ANN, WNN, ANFIS and WNF models.

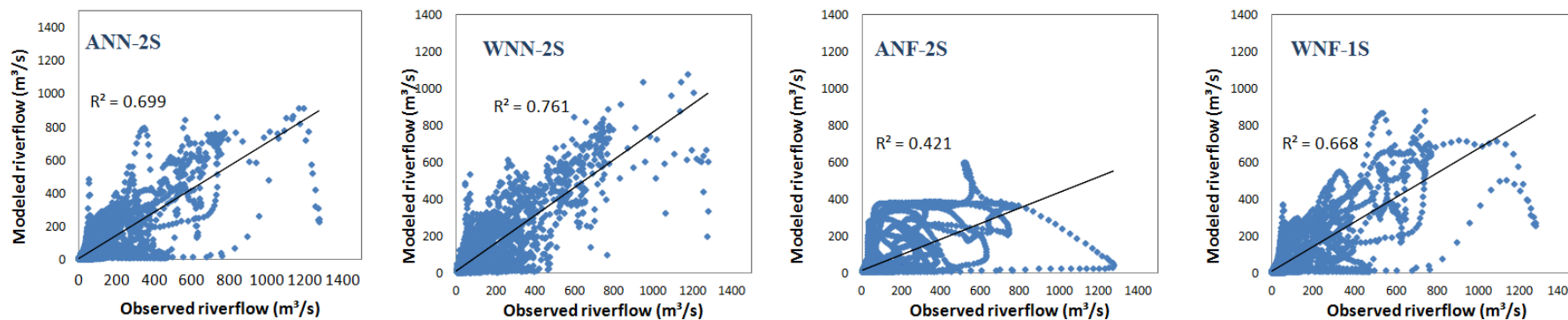


Figure 8.16 Scatter plots of observed and simulated river flow for 36 hour lead time with best fitted ANN, WNN, ANFIS and WNF models.

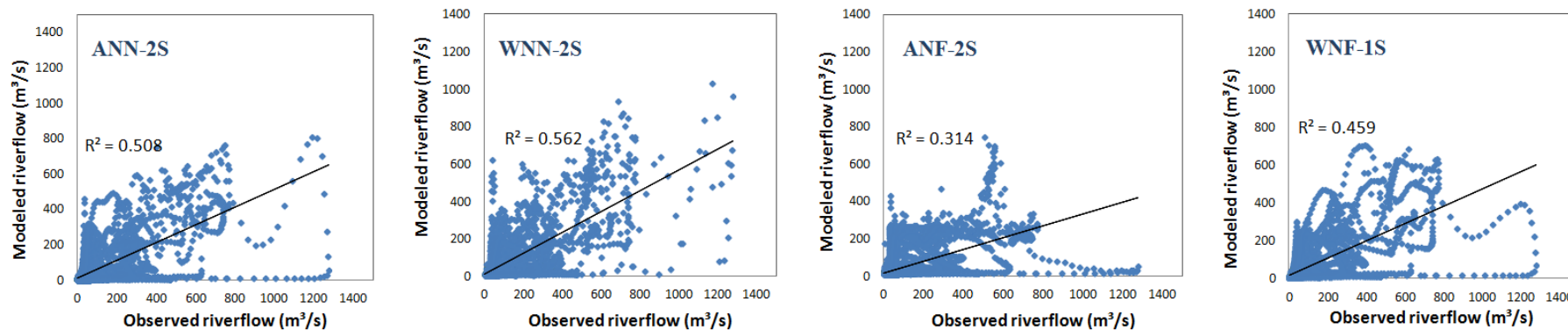


Figure 8.17 Scatter plots of observed and simulated river flow for 48 hour lead time with best fitted ANN, WNN, ANFIS and WNF models.

8.4 CONCLUSION

In this section, different data-driven approaches are developed for real time runoff forecasting. Artificial neural networks, adaptive neuro-fuzzy inference system and their associated hybrid models, in conjunction with wavelet multi-resolution analysis, are applied for hourly rainfall-runoff modeling. Application of the developed models in forecasting the different lead time is investigated. All models were highly efficient in forecasting 1, 6 and 12 hour ahead forecasting. By increasing the lead time from 12 to 48 hour ahead, the accuracy of forecasting decreases for all models with different trends.

Hybrid wavelet models significantly outperform the classic ANN and ANFIS models in forecasting the longer lead times. Applying wavelet multi resolution analysis and adding upstream river flow time series, the performance of 36 hour ahead forecasting is 80% improved (from 0.41 to 0.74 of the Nash-Sutcliffe efficiency in the testing set). Hybrid wavelet neural network model (WNN) has the best performance among all models. This model is able to forecast hourly river flow, 24 hour ahead with the accuracy of 0.95 for NSE. This ability makes WNN model a reliable tool for flood warning.

This study further concludes that the performance of ANN-based models is generally better than fuzzy-based models. The reason is mainly restricted structure of FIS with too many rules for higher number of inputs. Using very long historical time series (with length of almost 44000 for each), the choice of a number of membership function and level of decomposition was limited to the minimum. It can be inferred that applying one or two level of decomposition in multi-resolution analysis of input data could not improve the model efficiency significantly. The application of ANFIS and WNF models in different case studies with different size of input variables need to be investigated for a credible comparison between ANN and fuzzy-based models. Since adding one upstream river flow data had a significant impact on the accuracy of the forecasting, adding more effective variable should be considered in future studies. Overall, the results of this study confirm the robustness of the proposed structure of the hybrid models, WNN in particular, for the real time rainfall-runoff forecasting in the study area.

Chapter 9

Conclusions and Future Work

9.1 CONCLUSION

In this study an attempt is made to develop highly accurate river flow forecasting models using innovative computational intelligence based methods. Unlike conceptual or physically-based models, CI models do not need various number of variables for modelling complex rainfall-runoff process. Computational intelligence models are able to extract information from only river flow time series to achieve accurate future values. Application of various computational intelligence methods, including various types and structure of artificial neural networks, adaptive neuro-fuzzy inference system and hybrid wavelet models, are investigated. Four different rivers with different characteristics are selected as case studies; Harvey River, Avon River, Ellen Brook River in WA and Richmond River in NSW, Australia. Firstly, the impact of multivariate input selection on daily river flow forecasting is investigated in two different study areas. Secondly, different CI models are applied for short, mid and long term river flow forecasting. Then the application of different CI models for forecasting multi-step ahead of daily river flow is studied and improved. Finally, developed models are applied for hourly river flow forecasting of Richmond River with high potential of flooding to investigate the application of proposed models for timely flood warning.

In modelling and forecasting of short term, long term, multi-step ahead, seasonal and extreme river flow, CI models are found to be very promising alternative to traditional river flow forecasting models. Following is the summarized conclusions of this study;

- Computational intelligence models performance highly depends on the quality of data. Having long historical data is essential for reliable forecasting.
- Input selection has a very important role in data-driven modelling. This study confirms that forward stepwise selection of input data with high auto correlation function, improves river flow time series modelling.
- Adding other effective variables such as rainfall or upstream flow time series, could significantly enhance forecasting accuracy.
- Normalization and preprocessing the data is essential in fuzzy modelling. Application of neural networks modelling to the selected study areas also shows that data normalization has a very positive impact on neural networks training.
- There is no precise method for defining the optimum number of neurons in the hidden layer of neural networks. For each study area, this number needs to be determined by trial and error procedure. However, the optimum number increases by model input size.
- Considering highly nonlinear and non-stationary characteristics of river flow time series, pre-processing the input data with discrete wavelet analysis, significantly improves the forecasting reliability. Investigating the application of different CI models with different structures, hybrid wavelet neural networks and wavelet neuro-fuzzy models considerably outperformed classical ANN and ANFIS models in river flow forecasting.

- As one of the inherent problems in all forecasting methods, the forecasting reliability decreases with increasing the lead time. Integration of wavelet multi-resolution analysis into proposed hybrid models, the accuracy of long term and higher step ahead forecasting substantially improved.
- Despite high correlation between modelled and observed river flow, ANN and ANFIS failed to predict sudden extreme conditions. These models are not reliable for extreme event forecasting and sudden flood warning.
- ANN-based models often outperformed fuzzy-based models. This is attributed to the restricted structure of fuzzy inference systems. The number of fuzzy if-then rules increases exponentially with the number of inputs. The maximum input size for a fuzzy model to achieve feasible training is less than that for an ANN model. Therefore, adding other effective parameters or different time lag of time series with high ACF, is very restricted in ANFIS modelling. Decomposing inputs to a number of wavelet coefficients the gap between hybrid neural networks and hybrid neuro-fuzzy model performance is even higher.
- Adopting an appropriate mother wavelet in hybrid models improves the forecasting performance. The most effective type of mother wavelet depends on the river flow time series characteristics and could be defined with trial and error procedure. In this study, several mother wavelets are applied for multi-resolution analysis of input time series, including Haar, Daubechies, Coiflet and Symlet. In most case studies, Daubechies wavelet resulted most efficient modelling.
- The optimum level of wavelet decomposition in hybrid models also needs to be determined by trial and error. Increasing the resolution level, up to the optimum number, will improve forecasting accuracy. Whereas, having a large level of decomposition, the model might fail to reach feasible convergence or become inaccurate due to oversized network.

- Hybrid wavelet neural networks approach gives the highest accuracy among all developed models for extreme flow forecasting and timely flood warning.

9.2 RECOMMENDATION FOR FUTURE WORKS

The outcome of this study confirms the robustness and reliability of proposed hybrid CI approaches for river flow forecasting. However, there are still many issues to be investigated in future research. Following is a brief recommendation for further research;

In this study the type of mother wavelet and optimum level of decomposition for reaching best fitted hybrid model are determined by trial and error. Investigating the relation between river flow characteristics and suitable shape of mother wavelet or the length of data and the optimum level of decomposition, would be helpful in future forecasting.

It is observed that adding effective time series such as rainfall or upstream river flow data, have a significant impact on model performance. Investigating the role of other effective climatic or hydrological parameter (e.g. air temperature, soil moisture, evaporation, solar radiation,...) in enhancing model efficiency could be considered in future work.

Current studies often use observed data for future forecasting. The application of forecasted rainfall (by weather forecast methods) as CI models' input in improving the accuracy of river flow forecasts also needs to be investigated. This method could be especially effective in improving the reliability of multi step ahead hourly and daily river flow forecasts.

Since application of hybrid wavelet neuro-fuzzy approach with subtractive clustering in river flow forecasting introduced in this research, the limitations and capabilities of this approach need further investigation by applying this method in different regions.

Considering the emerging development in optimization approach, one suggestion for future work could be investigating the application of evolutionary optimization approaches (e.g. Swarm intelligence optimization or genetic programming) in training the CI models.

For more convenient utilization of CI methods in different regions, selecting the best fitted model structure, based on river characteristics could be studied. For this reason data clustering should initially apply on different rivers. Then the best fitted model with a fixed structure for rivers in same cluster could be selected.

Considering the proven ability of hybrid wavelet models in simulating highly non-stationary and nonlinear river flow time series, the application of this approach in other hydro-environmental fields such as groundwater or environmental modelling and forecasting could be studied.

REFERENCES:

- Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986). An introduction to the European hydrological system, SHE. Part2: Structure of a physically-based distributed modelling system. *Journal of Hydrology*, 87, 61-77.
- Abe, S. (2001) *Pattern classification : neuro-fuzzy methods and their comparison*. Springer-Verlag New York.
- Abrahart, R., & See, L. (2000). Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. *Hydrological Processes*, 14(11-12), 2157-2172.
- Adamowski, J., & Sun, K. (2010). Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. *Journal of Hydrology*, 390 (1-2), 85-91.
doi: 10.1016/j.jhydrol.2010.06.033
- Anctil, F., Lauzon N., & Filion, M. (2008). Added gains of soil moisture content observations for streamflow predictions using neural networks. *Journal of Hydrology*, 359(3-4):225-234.
- Anctil, F., Perrin, C., & Andréassian, V. (2004). Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. *Environmental Modelling and Software*, 19(4), 357-368.
- Aqil, M., Kita, I., Yano, A., & Nishiyama, S. (2007). A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff. *Journal of Hydrology*, 337(1-2), 22-34.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2007.01.013>
- ASCE Task Committee on application of artificial neural networks in hydrology. (2000a). *Artificial neural networks in hydrology I: Preliminary concepts*. *Journal of Hydrologic Engineering* 5(2): 115-123.
- ASCE Task Committee on application of artificial neural networks in hydrology. (2000b). *Artificial neural networks in hydrology II: Hydrologic applications*. *Journal of Hydrologic Engineering* 5(2): 124-137.

- Bates, B., Kundzewicz, Z. W., Wu, S., & Palutikof, J. (2008). Climate change and water. Technical paper of the intergovernmental panel on climate change. Geneva(VI):210.
- Beale, M. H., Hagan, M.T., & Demuth, H.B. (2012). MATLAB neural network toolbox user's guide. MathWorks, Inc.
- Berthet, L., Andréassian, V., Perrin, C., & Javelle, P. (2009). How crucial is it to account for the Antecedent Moisture Conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. *Hydrol. Earth Syst. Sci. Discuss.*, 6(2), 1707-1736.
doi: 10.5194/hessd-6-1707-2009
- Besaw, L. E., Rizzo, D. M., Bierman, P. R., & Hackett, W. R. (2010). Advances in ungauged streamflow prediction using artificial neural networks. *Journal of hydrology*, 386(1-4), 27-37.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2010.02.037>
- Beven, K. (1985). *Distributed models*: John Wiley and Sons, Chichester, Engl.
- Beven, K. J., & Kirkby, M. J. (1979). A physically-based variable contributing area model of basin hydrology. *Hydrology and Earth System Science. Bull.* 24:43-69.
- Beven, K., Warren, R., & Zaoui, J. (1980). SHE: Towards a methodology for physically-based distributed forecasting in hydrology. *Hydrological Forecasting: IAHS-AISH Publ. No. 129.*
- Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York.
- Birikundavyi, S., Labib, R., Trung, H., & Rousselle, J. (2002). Performance of Neural Networks in Daily Streamflow Forecasting. *Journal of Hydrologic Engineering*, 7(5):392-398.
- Boughton, W. C. (2004). The Australian water balance model. *Environmental Modelling and Software*, 19, 943-956.
- Box, G., & Jenkins, G. (1970). *Time series analysis: forecasting and control*. Holden-Day: San Francisco.

- Broomhead, D., & Lowe, D. (1988). Multivariable functional interpolation and adaptive networks. *Complex Systems*, 2, 321-355.
- Burnash, R. J. C., Ferral, R. L. & McGuire, R. A. (1973). A Generalised streamflow simulation system—conceptual modelling for digital computers. Joint Federal and State River Forecast Centre, Sacramento, Technical Report, 204.
- Butts, M. B., Payne, J. T., Kristensen, M. & Madsen, H. (2004). An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow prediction. *Journal of Hydrology*, 298, 242-266.
- Caddis, B. (2010). Richmond River flood mapping study. BMT WBM Pty Ltd.
- Cannas, B., Fanni, A., See, L., & Sias, G. (2006). Data preprocessing for river flow forecasting using neural networks: Wavelet transforms and data partitioning. *Physics and Chemistry of the Earth, Parts A/B/C*, 31(18), 1164-1171.
doi: 10.1016/j.pce.2006.03.020
- Carpenter, T. M., Sperflage, J. A., Georgakakos, K. P., Sweeney, T., & Fread, D. L. (1999). National threshold runoff estimation utilizing GIS in support of operational flash flood warning systems. *Journal of hydrology*, 224(1–2), 21-44.
doi: [http://dx.doi.org/10.1016/S0022-1694\(99\)00115-8](http://dx.doi.org/10.1016/S0022-1694(99)00115-8)
- Chang, F. J., & Chang, Y. T. (2006). Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Advances in Water Resources*, 29(1):1-10.
- Chartres C., & Williams, J. (2006). Can Australia overcome its water scarcity problems. *Journal of Developments in Sustainable Agriculture*, 1:17-24.
- Chau, K., Wu, C., & Li, Y. (2005). Comparison of Several Flood Forecasting Models in Yangtze River. *Journal of Hydrologic Engineering*, 10(6), 485-491.
doi:10.1061/(ASCE)1084-0699(2005)10:6(485)
- Chiang, Y. M., Chang, L. C., & Chang, F.J. (2004). Comparison of static-feedforward and dynamic-feedback neural networks for rainfall–runoff modeling. *Journal of Hydrology*, 290(3–4):297-311.
- Chiew, F. H. S., Peel, M. C., & Western, A. W. (2002). Application and testing of the simple rainfall-runoff model SIMHYD. *Mathematical models of small watershed hydrology and applications*, 335-367.

- Chiu, S. (1994). Fuzzy Model Identification Based on Cluster Estimation. *Journal of Intelligent & Fuzzy Systems*, 2(3).
- Cigizoglu, H. K. (2003a). Estimation, forecasting and extrapolation of river flows by artificial neural networks. *Hydrological Sciences Journal*, 48(3), 349-361.
- Cigizoglu, H. K. (2003b). Incorporation of ARMA models into flow forecasting by artificial neural networks. *Environmetrics*, 14(4), 417-427.
doi: 10.1002/env.596
- Cigizoglu, H. K., & Kisi, O. (2005). Methods to improve the neural network performance in suspended sediment estimation. *Journal of Hydrology*, 317(3-4):221-238.
- Clarke, R. T. (1973). A review of some mathematical models used in hydrology, with observations on their calibration and use. *Journal of hydrology*, 19(1), 1-20.
doi: [http://dx.doi.org/10.1016/0022-1694\(73\)90089-9](http://dx.doi.org/10.1016/0022-1694(73)90089-9)
- Cloke, H. L., & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of hydrology*, 375(3-4), 613-626.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2009.06.005>
- Crawford, N. H., & Linsley, R.K. (1966). Digital simulation in hydrology: Stanford watershed model IV. Technical report No. 39, Department of Civil engineering. Stanford University:210.
- Croke, B., Smith, A. B., & Jakeman, A. J. (2002). A one parameter groundwater discharge model linked to the IHACRES rainfall-runoff model, Proceedings of the 1st Biennial meeting of the international environmental modelling and software society, University of Lugano, Switzerland, 1, 428-433.
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematical Control, Signals and Systems*, 2, 303-314.
- Daniell, T.M. (1991). Neural networks - applications in hydrology and water resources engineering. *International Conference on Hydrology and Water Resources Symposium*, 3(3):797-802.
- Daubechies, I. (1988). Time-frequency localization operators: a geometric phase space approach. *IEEE Ttransactions of information theory*, 34(605-12).

- Daubechies, I. (1992). Ten lectures on wavelets. Society for industrial and applied mathematics, 357.
- Dawson, C. W., & Wilby, R. L. (2001). Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, 25(1), 80-108.
- Dawson, C. W., Abrahart, R. J., Shamseldin, A. Y., & Wilby, R. L. (2006). Flood estimation at ungauged sites using artificial neural networks. *Journal of Hydrology*, 319(1-4), 391-409.
- Demicco, R., & Klir, G. (2004). *Fuzzy logic in Geology*: Elsevier Inc.
doi: 10.1016/B978-012415146-8/50000-7
- Dibike, Y. B., & Solomatine, D. P. (2001). River flow forecasting using artificial neural networks. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 26(1), 1-7.
doi: 10.1016/s1464-1909(01)85005-x
- Donn, M. J., Barron, O. V., & Barr, A. D. (2012). Identification of phosphorus export from low-runoff yielding areas using combined application of high frequency water quality data and MODHMS modelling. *Science of The Total Environment*, 426(0), 264-271.
doi: <http://dx.doi.org/10.1016/j.scitotenv.2012.03.021>
- Downer, C., & Ogden, F. (2004). GSSHA: model To simulate diverse stream flow producing processes. *Journal of Hydrologic Engineering*, 9(3), 161-174.
doi:10.1061/(ASCE)1084-0699(2004)9:3(161)
- Fernando, D. A. K., & Jayawardena, A. W. (1998). Runoff forecasting using RBF networks with OLS algorithm. *Journal of Hydrologic Engineering*, 3(3), 203-209.
- Firat, M. (2008). Comparison of artificial intelligence techniques for river flow forecasting. *Hydrology and Earth System Sciences*, 12(1), 123-139.
- Firat, M., Turan, M.E., Yurdusev, M.A.,. (2009). Comparative analysis of fuzzy inference systems for water consumption time series prediction. *Journal of Hydrology*, 374(3-4), 235-241.
- Fortin, J., Turcotte, R., Massicotte, S., Moussa, R., Fitzback, J., & Villeneuve, J. (2001a). Distributed watershed model compatible with remote sensing and

- GIS data. I: Description of model. *Journal of Hydrologic Engineering*, 6(2), 91-99.
doi:10.1061/(ASCE)1084-0699(2001)6:2(91)
- Fortin, J., Turcotte, R., Massicotte, S., Moussa, R., Fitzback, J., & Villeneuve, J. (2001b). Distributed watershed model compatible with remote sensing and GIS data. II: Application to Chaudière watershed. *Journal of Hydrologic Engineering*, 6(2), 100-108.
- Fourier, J. (1808). Memoire sur la propagation de la chaleur dans les corps solides. *Nouveau Bulletin des sciences par la Societe philoatque de Paris.* , I, 112-116.
- Galloway, P. (2004). Natural resource management issues in the Avon River basin. Resource management technical report, Department of agriculture, Perth, 288.
- Ghanbarpour, M. R., Abbaspour, K. C., & Hipel, K. W. (2009). A comparative study in long-term river flow forecasting models. *International Journal of River Basin Management*, 7(4), 403-413.
doi: 10.1080/15715124.2009.9635398
- Googhari, S. H. K., & Lee, T. S. (2011). Applicability of adaptive neuro-fuzzy inference systems in daily reservoir inflow forecasting. *International Journal of Soft Computing*, 6(3), 75-84.
- Gosain, A. K., Mani, A. & Dwivedi, C. (2009). Hydrological modeling review. *Climawater*, Report No. 1.
- Graham, D. N., & Butts, M. B. (2005). Flexible intergrated watershed modeling with MIKE SHE. *Watershed Models*. Ed. V.P.Singh and D.K. Frevert, CRC Press, P245-271.
- Han, D. (2011). Flood risk assessment and management. Bentham Science Publishers.
- Huang, M., & Liang, X. (2006). On the assessment of the impact of reducing parameters and identification of parameter uncertainties for a hydrologic model with applications to ungauged basins. *Journal of hydrology*, 320(1–2), 37-61.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2005.07.010>

- Hundecha, Y., Bardossy, A., & Werner, H. W. (2001). Development of a fuzzy logic-based rainfall-runoff model. *Hydrological Sciences Journal*, 46(3), 363-376.
doi: 10.1080/02626660109492832
- Imrie, C. E., Durucan, S., & Korre, A. (2000). River flow prediction using artificial neural networks: generalisation beyond the calibration range. *Journal of Hydrology*, 233(1-4), 138-153.
doi: 10.1016/s0022-1694(00)00228-6
- Iyengar, S. S., Cho, E. C. & Phoha, V. V. (2002). *Foundations of Wavelet Networks and Applications*. Chapman & Hall/CRC Press.
- Jacquin, A. P., & Shamseldin, A. Y. (2009). Review of the application of fuzzy inference systems in river flow forecasting. *Journal of hydroinformatics*, 11(3-4), 202-210.
- Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics*, 23(3), 665.
- Jang, J. S., & Gulley, N. (1995). *The Fuzzy Logic Toolbox for Use with MATLAB*. MathWorks, Inc.
- Jang, J. S., Sun, C.T., & Mizutani, E. (1997). Neuro-Fuzzy and Soft Computing-A Computational Approach to Learning and Machine Intelligence [Book Review]. *IEEE Transactions on Automatic Control* 42, 1482-1484.
- Jayawardena, A. W. (2014). *Environmental and hydrological systems modelling*, CRC Press.
- Jayawardena, A. W., Perera, E. D. P., Zhu, B., Amarasekara, J. D., & Vereivalu, V. (2014). A comparative study of fuzzy logic systems approach for river discharge prediction. *Journal of hydrology*, 514(0), 85-101.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2014.03.064>
- Jayawardena, A. W., & Zhou, M. C. (2000). A modified spatial soil moisture storage capacity distribution curve for the Xinanjiang model. *Journal of Hydrology*, 227(93-113).
- Ji, L., Gallo, K. (2006). An agreement coefficient for image comparison. *Photogrammetric Engineering and Remote Sensing*, 72(7), 823-833.

- Jiang, S., Ren, L., Hong, Y., Yong, B., Yang, X., Yuan, F., & Ma, M. (2012). Comprehensive evaluation of multi-satellite precipitation products with a dense rain gauge network and optimally merging their simulated hydrological flows using the Bayesian model averaging method. *Journal of Hydrology*, 452–453(0), 213-225.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2012.05.055>"
- Johanson, R. C., Imhoff, J. C., Davis, H. H. (1980). User's manual for the hydrologic simulation program FORTRAN (HSPF). EPA-600/9-80-105, U.S. EPA environmental research laboratory, Athens, GA.
- Julien, P.Y. & Saghaffian B., (1991). A two-dimensional watershed rainfall-runoff model. Civil Eng. Rep. CER90-91PYJ-BS-12Colorado State University, Fort Collins.
- Kachroo, R. K. (1992a). River flow forecasting part5: Applications of a conceptual model. *Journal of Hydrology*, 133, 141-178.
- Kachroo, R. K. (1992b). River flow forecasting. Part 1. A discussion of the principles. *Journal of Hydrology*, 133(1–2), 1-15.
doi: [http://dx.doi.org/10.1016/0022-1694\(92\)90146-M](http://dx.doi.org/10.1016/0022-1694(92)90146-M)"
- Kalman, B. L., Kwasny, S.c. (1992). Why Tanh? Choosing a sigmoidal function. *Proceedings of the International Joint Conference on Neural Networks*. Baltimore, MD IEEE, New York.
- Karunanithi, N., Grenney, W. J., Whitley, D., & Bovee, K. (1994). Neural networks for river flow prediction. (Special Issue: Neural Networks). *Journal of Computing in Civil Engineering*, 8, 201.
- Kermani, M. Z., & Teshnehlab, M. (2008). Using adaptive neuro-fuzzy inference system for hydrological time series prediction. *Applied Soft Computing*., 8(2), 928-936.
doi: 10.1016/j.asoc.2007.07.011
- Keskin, M. E., Taylan, D., & Terzi, Ö. (2006). Adaptive neural-based fuzzy inference system (ANFIS) approach for modelling hydrological time series. *Hydrological Sciences Journal*, 51(4), 588-598.
doi: 10.1623/hysj.51.4.588

- Khan, M. S., & Coulbaly, P. (2006). Bayesian neural network for rainfall-runoff modeling. *Water Resources Research*, 42(7).
- Kim, G., & Barros, A. P. (2001). Quantitative flood forecasting using multisensor data and neural networks. *Journal of Hydrology*, 246(1-4), 45-62.
- Kim, T.W., & Valdés, J. B. (2003). Nonlinear Model for Drought Forecasting Based on a Conjunction of Wavelet Transforms and Neural Networks. *Journal of Hydrologic Engineering*, 8(6), 319-328.
doi:10.1061/(ASCE)1084-0699(2003)8:6(319)"
- Kingston, G. B., Lambert, M. F. & Maier, H. R. (2005). Bayesian training of artificial neural networks used for water resources modeling. *Water Resources Research*, 41(12):W12409.
- Kisi, O. (2006). Daily pan evaporation modelling using a neuro-fuzzy computing technique. *Journal of Hydrology*, 329(3-4), 636-646.
- Kisi, O. (2009). Neural networks and wavelet conjunction model for intermittent streamflow forecasting. *Journal of Hydrologic Engineering*, 14(8), 773-782.
- Kisi, O., Shiri, J., 2012. Wavelet and neuro-fuzzy conjunction model for predicting water table depth fluctuations. *Nordic Hydrology*, 43, 286-300.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps, *Biological cybernetics*, 43, 59-69.
- Krause P., B., Boyle, D.P., & Base, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5, 89-97.
- Krycha, K. A., & Wagner, U. (1999). Applications of artificial neural networks in management science: a survey. *Journal of Retailing and Consumer Services*, 6(4), 185-203.
doi: [http://dx.doi.org/10.1016/S0969-6989\(98\)00006-X](http://dx.doi.org/10.1016/S0969-6989(98)00006-X)
- Kumar, A. R. S., Sudheer, S. K. P., Jain, K., & Agarwal, P. K. (2005). Rainfall-runoff modelling using artificial neural networks: comparison of network types. *Hydrological Processes*, 19(6), 1277-1291.
- Labat, D. (2005). Recent advances in wavelet analyses: Part 1. A review of concepts. *Journal of Hydrology*, 314(1-4), 275-288.

doi: 10.1016/j.jhydrol.2005.04.003

Lam, H. K., Ling, S., & Nguyen, H. T. (2012). Computational intelligence and its applications: Evolutionary computation, fuzzy logic, neural network and support vector machine techniques. World Scientific Publishing Company.

Lin, G. F., & Wu, M. C. (2011). An RBF network with a two-step learning algorithm for developing a reservoir inflow forecasting model. *Journal of Hydrology*, 405(3–4), 439-450.

doi: 10.1016/j.jhydrol.2011.05.042

Linsley, R. K., Crawford, N. H. (1960). Computation of a Synthetic Streamflow Record on a Digital Computer. In surface water-Proceedings of the Helsinki symposium, IAHS Publication, 51, 526-538.

Liong, S. Y., Lim, W. H., Kojiri, T., & Hori, T. (2000). Advance flood forecasting for flood stricken Bangladesh with a fuzzy reasoning method. *Hydrological Processes*, 14(3), 431-448.

doi: 10.1002/(sici)1099-1085(20000228)14:3<431:aid-hyp947>3.0.co;2-0

Lippmann, R. P. (1987). An Introduction to Computing with Neural Nets. *IEEE ASSP Magazine*, 4-22(April).

Liu, Y., Brown, J., Demargne, J., & Seo, D.-J. (2011). A wavelet-based approach to assessing timing errors in hydrologic predictions. *Journal of hydrology*, 397(3–4), 210-224.

doi: <http://dx.doi.org/10.1016/j.jhydrol.2010.11.040>

Machado, F., Mine, M., Kaviski, E., & Fill, H. (2011). Monthly rainfall–runoff modelling using artificial neural networks. *Hydrological Sciences Journal*, 56(3), 349-361.

doi: 10.1080/02626667.2011.559949

Maier, H. R., & Dandy, G. C. (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software*, 15(1), 101-124.

doi: [http://dx.doi.org/10.1016/S1364-8152\(99\)00007-9](http://dx.doi.org/10.1016/S1364-8152(99)00007-9)

Maier, H. R., Jain, A., Dandy, G. C., & Sudheer, K. P. (2010). Methods used for the development of neural networks for the prediction of water resource variables

in river systems: Current status and future directions. *Environmental Modelling & Software*, 25(8), 891-909.
doi: <http://dx.doi.org/10.1016/j.envsoft.2010.02.003>

- Mallat, S. (1985). A compact multiresolution representation: the wavelet model. *Proceedings of IEEE workshop computer Vision*, Miami, FL.
- Mamdani, E.H., Assilian S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of Man-Machine Studies*, 7(1), 1-13.
- Martins, O., Sadeeq M. & Ahaneku, I. (2011). ARMA Modelling of Benue River Flow Dynamics: Comparative Study of PAR Model. *Journal of Modern Hydrology*, 1(1) 1-9.
doi: 10.4236/ojmh.2011.11001.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 5: 115-133.
- Mehrotra, K., Mohan, C. K., & Ranka, S. (1997). *Elements of artificial neural networks*. MIT Press.
- Meyer, Y. (1987). *Wavelet and operators*. Rapport CEREMADE, 8704. Paris-Dauphine University, Paris.
- Moradkhani, H., Hsu, K. I., Gupta, H. V., & Sorooshian, S. (2004). Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. *Journal of Hydrology*, 295(1-4), 246-262.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2004.03.027>
- Moriasi, D. N., Arnold, J.G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., 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.
- Morlet, J., Arens, G., Fourgeau, E., & Glard, D. (1982). Wave propagation and sampling theory—Part I: Complex signal and scattering in multilayered media. *Geophysics*, 47(2), 203-221.
doi:10.1190/1.1441328

- Motovilov, Y. G., Gottschalk, L., Engeland, K., & Rodhe, A. (1999). Validation of a distributed hydrological model against spatial observations. *Agricultural and Forest Meteorology*, 98–99(0), 257-277.
doi: [http://dx.doi.org/10.1016/S0168-1923\(99\)00102-1](http://dx.doi.org/10.1016/S0168-1923(99)00102-1)
- Mulvany, W. T. (1845). Observation on regulating weirs. *Transactions of the institution of civil engineers, Ireland*, I, 83-93.
- Nash, J. E., Sutcliffe, J. V. (1970). River flow forecasting through conceptual models. Part 1- A discussion of principles. *Journal of Hydrology*, 10, 282-290.
- Nash, J. E. (1959). Synthetic determination of unit hydrograph parameters. *Journal of Geophysics research*, 64(1), 111-115.
- Nawi, N., Khan, A., & Rehman, M. (2013). A New Back-Propagation Neural Network Optimized with Cuckoo Search Algorithm. In *Computational science and its applications– ICCSA 2013*, 7971, 413-426, Springer Berlin Heidelberg.
- Nayak, P. C., Sudheer, K. P., Rangan, D. M., & Ramasastri, K. S. (2005). Short-term flood forecasting with a neuro-fuzzy model. *Water Resources Research*, 41(4), W04004.
doi: 10.1029/2004wr003562
- Nayebi, M., Anmin, S., & Parsa, S. Z. (2006). Daily stream flow prediction capability of artificial neural networks as influenced by minimum air temperature data. *Biosystems Engineering*, 95(4), 557-567.
- Nor, N., Harun, S., & Kassim, A. (2007). Radial basis function modeling of hourly streamflow hydrograph. *Journal of Hydrologic Engineering*, 12(1), 113-123.
doi:10.1061/(ASCE)1084-0699(2007)12:1(113)
- Nourani, V., Baghanam, A. H., Adamowski, J., & Gebremichael, M. (2013). Using self-organizing maps and wavelet transforms for space–time pre-processing of satellite precipitation and runoff data in neural network based rainfall–runoff modeling. *Journal of Hydrology*, 476(0), 228-243.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2012.10.054>

- Nourani, V., Kisi, O., & Komasi, M. (2011). Two hybrid artificial Intelligence approaches for modeling rainfall–runoff process. *Journal of Hydrology*, 402(1-2), 41-59.
doi: 10.1016/j.jhydrol.2011.03.002
- O'Connell, P. E. (1970). River flow forecasting through conceptual models part II - The Brosna catchment at Ferbane. *Journal of hydrology*, 10(4), 317-329.
- O'Connor K. M., 2006. River flow forecasting. In *River basin modelling for flood risk mitigation*: Taylor & Francis.
- Panday, S., & Huyakorn, P. S. (2004). A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. *Advances in Water Resources*, 27(4), 361-382.
doi: <http://dx.doi.org/10.1016/j.advwatres.2004.02.016>
- Partal, T. (2009). River flow forecasting using different artificial neural network algorithms and wavelet transform. *NRC Research Press* 36, 26-39.
- Partal, T., & Kisi, O. (2007). Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. *Journal of Hydrology*, 342(1-2), 199-212.
doi: 10.1016/j.jhydrol.2007.05.026
- Penning-Rowsell, E. C., Tunstall, S. M., Tapsell, S. M., & Parker, D. J. (2000). The Benefits of Flood Warnings: Real But Elusive, and Politically Significant. *Water and Environment Journal*, 14(1), 7-14.
doi: 10.1111/j.1747-6593.2000.tb00219.x
- Perrin, C., Michel, C. & Andreassian, V. (2003). Improvement of a parsimonious model for streamflow simulations. *Journal of Hydrology*, 279, 275-289.
- Piotrowski, A. P., & Napiorkowski, J. J. (2011). Optimizing neural networks for river flow forecasting- Evolutionary Computation methods versus the Levenberg–Marquardt approach. *Journal of Hydrology*, 407(1–4), 12-27.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2011.06.019>
- Podger, G. (2004). *Rainfall runoff library*. CRC for catchment hydrology. Australia.
- Poff, N. L., Tokar, S., & Johnson, P. (1996). Stream Hydrological and ecological responses to climate change assessed with an artificial neural network. *Limnology and Oceanography*, 41(5), 857-863.

- Pramanid, N., Rabindra, K., Singh, A., (2011). Daily river flow forecasting using wavelet ANN hybrid models. *Journal of Hydroinformatics*, 13(1), pp49.
- Refsgaard, J. C. (1997). Parameterisation, calibration and validation of distributed hydrological models. *Journal of hydrology*, 198(1–4), 69-97.
doi: [http://dx.doi.org/10.1016/S0022-1694\(96\)03329-X](http://dx.doi.org/10.1016/S0022-1694(96)03329-X)
- Refsgaard, J. C., & Storm, B. (1995). MIKE SHE. In V. P. Singh, *Computer models in hydrology* (pp. 809-846), Water Resources Publications. Colorado, USA.
- Ren, L., Xiang, X., & Ni, J. (2013). Forecast Modeling of Monthly Runoff with Adaptive Neural Fuzzy Inference System and Wavelet Analysis. *Journal of Hydrologic Engineering*, 18(9), 1133-1139.
doi:10.1061/(ASCE)HE.1943-5584.0000514
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386-408.
- Rumelhart, D. E., Hintont, G. E., & Williams, R. J. (1986). Learning representation by back-propagating errors. *Nature*, 323 (9), 533-536.
- Sen, Z., & Altunkaynak, A. (2006). A comparative fuzzy logic approach to runoff coefficient and runoff estimation. *Hydrological Processes*, 20(9), 1993-2009.
doi: 10.1002/hyp.5992
- Sene, K. (2010). *Hydrometeorology: Forecasting and applicaions*: Springer.
doi:10.1007/978-90-481-3403-8
- Shamseldin, A. Y. (1997). Application of a neural network technique to rainfall-runoff modelling. *Journal of Hydrology*, 199(3–4), 272-294.
doi: [http://dx.doi.org/10.1016/S0022-1694\(96\)03330-6](http://dx.doi.org/10.1016/S0022-1694(96)03330-6)
- Sivakumar, B., & Berndtsson, R. (2010). *Advances in Data-based Approaches for Hydrologic Modeling and Forecasting*: World Scientific. GB656.2.H9.
- Sivakumar, B., Jayawardena, A. W., & Fernando, T. M. K. G. (2002). River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches. *Journal of Hydrology*, 265(1-4), 225-245.
- Srinivasulu, S., & Jain, A. (2006). A comparative analysis of training methods for artificial neural network rainfall–runoff models. *Applied Soft Computing*, 6(3), 295-306.

doi: <http://dx.doi.org/10.1016/j.asoc.2005.02.002>

Sudheer, C., Maheswaran, R., Panigrahi, B. K., & Mathur, S. (2014). A hybrid SVM-PSO model for forecasting monthly streamflow. *Neural Computing and Applications*, 24(6), 1381-1389.

doi: 10.1007/s00521-013-1341-y

Sugawara, M. (1961). On the analysis of runoff structure about several Japanese rivers. *Japanese Journal of Geophysics*, 2(4), 210-216.

Sugeno, M., & Kang, G.T. (1988). Structure identification of fuzzy model. *Fuzzy Set and System*, 28, 15-33.

Takagi, T., Sugeno, M., (1985). Fuzzy Identification of Systems and its Applications to Modeling and Control. *IEEE Transactions on Systems, Man and Cybernetics*, 15, 116-132.

Talei, A., Chua, L. H. C., & Wong, T. S. W. (2010). Evaluation of rainfall and discharge inputs used by Adaptive Network-based Fuzzy Inference Systems (ANFIS) in rainfall-runoff modeling. *Journal of Hydrology*, 391(3-4), 248-262.

doi: <http://dx.doi.org/10.1016/j.jhydrol.2010.07.023>"

Tawfik, M., Ibrahim, A., & Fahmy, H. (1997). Hysteresis sensitive neural network for modeling rating curves. *Journal of Computing in Civil Engineering*, 11(3), 206-211.

doi: 10.1061/(asce)0887-3801(1997)11:3(206)

Tetko, I. V., Livingstone, D. J., & Luik, A. I. (1995). Neural network studies. 1. Comparison of overfitting and overtraining. *Journal of Chemical Information and Computer Sciences*, 35(5), 826-833.

doi: 10.1021/ci00027a006

Todini, E. (1988). Rainfall-runoff modeling - Past, present and future. *Journal of hydrology*, 100(1-3), 341-352.

doi: [http://dx.doi.org/10.1016/0022-1694\(88\)90191-6](http://dx.doi.org/10.1016/0022-1694(88)90191-6)

Turan, M. E., & Yurdusev, M. A. (2009). River flow estimation from upstream flow records by artificial intelligence methods. *Journal of Hydrology*, 369(1-2), 71-77.

doi: <http://dx.doi.org/10.1016/j.jhydrol.2009.02.004>

- Tuteja, N. K., & Cunnane, C. (1999). A quasi physical snowmelt runoff modelling system for small catchments. *Hydrological Processes*, 13(12-13), 1961-1975. doi: 10.1002/(sici)1099-1085(199909)13:12/13<1961:aid-hyp887>3.0.co
- Vaze, J., Jorda, P., Beecham, R., Frost, A. Summerell, G. (2012). Guidelines for rainfall-runoff modelling: Towards best practice model application. eWater cooperative research centre. ISSN- 978-1-921543-51-7
- Valenca, M., Ludermir, T., & Valenca, A. (2005). River flow forecasting for reservoir management through neural networks. HIS '05. Fifth International Conference on Hybrid Intelligent Systems. doi: 10.1109/ICHIS.2005.95
- Wang, K., & Altunkaynak, A. (2012). Comparative case study of rainfall-runoff modeling between SWMM and fuzzy logic approach. *Journal of Hydrologic Engineering*, 17(2):283-291.
- Wang, W. (2006). Stochasticity, nonlinearity and forecasting of streamflow processes. Scitech Book News, 31.
- Wang, W., & Ding, J. (2003). Wavelet network model and its application to the prediction of hydrology. *Nature and Science*, 1(1), 67-71.
- Wei, S., Song, J., & 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. doi: 10.1002/hyp.8227
- Wei, S., Yang, H., Song, J., Abbaspour, K., & Xu, Z. (2013). A wavelet-neural network hybrid modelling approach for estimating and predicting river monthly flows. *Hydrological Sciences Journal*, 58(2), 374-389. doi: 10.1080/02626667.2012.754102
- Weilin, L., Lina, L., & Zengchuan, D. (2011). Neural network model for hydrological forecasting based on multivariate phase space reconstruction. In *proceedings of Seventh International Conference on Natural Computation (ICNC) 2*: 663-667. doi: 10.1109/icnc.2011.6022232.

- Werner, M., Reggiani, P., De Roo, A., Bates, P., & Sprokkereef, E. (2005). Flood forecasting and warning at the river basin and at the european scale. *Natural Hazards*, 36(1-2), 25-42.
- Wilke, S. J. (2006). The Harvey River Restoration Taskforce: a Novel Community-based Management Scheme. 9th international river symposium, Brisbane, Australia.
- Wu, C. L., Chau, K. W., & Li, Y. S. (2009). Methods to improve neural network performance in daily flows prediction. *Journal of Hydrology*, 372(1-4), 80-93.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2009.03.038>
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338-353.
- Zeng, X., Kiviat, K. L., Sakaguchi, K., & Mahmoud, A. M. A. (2012). A toy model for monthly river flow forecasting. *Journal of Hydrology*, 452-453(0), 226-231.
doi: <http://dx.doi.org/10.1016/j.jhydrol.2012.05.053>
- Zhang, W. J., & Barrion, A. T. (2006). Function approximation and documentation of sampling data using artificial neural networks. *Environmental Monitoring and Assessment*, 122, 185-201.
- Zhao, R. J. (1977). Flood forecasting method for humid regions of China. East China college of hydraulic engineering Nanjing, 19-51.
- Zhou, H. Ch., Peng, Y., Liang, G.H. (2008). The research of monthly discharge predictor-corrector model based on wavelet decomposition. *Water Resource Mange*, 22, 217-227.

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.