Seasonal Rainfall Forecasting Using Large Scale Climate Drivers: An Artificial Intelligence Approach

A Thesis Submitted in Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Fatemeh Mekanik

Faculty of Science, Engineering and Technology

Swinburne University of Technology

Hawthorn, VIC, 3122

Australia

July

2015

Abstract

As Australia is exposed to severe droughts and floods, seasonal rainfall forecasting is crucial for water resources management, food production and mitigating flood risks. However, rainfall prediction models have not been very satisfactory in terms of accuracy. The main focus of this study is the development of a non-linear rainfall forecast model for Victoria, Australia using antecedent large-scale climate predictors. Artificial Neural Network (ANN) was chosen as the primary modelling technique due to its capability to extract complex relationships from the data. ANN has been rarely used in rainfall forecasting in Australia in conjunction with the use of large-scale climate modes. In order to compare the forecast results of ANN models with those of linear and non-linear models, Multiple Linear Regression (MLR) models and Adaptive Network-based Fuzzy Inference System (ANFIS) were developed respectively. To the best of the author's knowledge, this research is the first study using ANFIS technique in conjunction with large-scale climate modes to forecast rainfall in Australia.

Australian climate is highly influenced by the large scale-climate modes. The large-scale climate modes taking place in the Pacific and Indian Oceans were considered in this study as potential rainfall predictors; El Nino Southern Oscillation (ENSO) and Interdecadal Pacific Oscillation (IPO) which occur in the Pacific Ocean and Indian Ocean Dipole (IOD) which occurs in the Indian Ocean were examined as rainfall predictors when developing the models. Furthermore, this study investigated the concurrent and antecedent relationship between seasonal rainfall and large-scale climate modes; classification analysis and Pearson correlation analysis were used in this regard. Three distinct regions in Victoria, Australia were considered as case studies; from each region three rainfall stations were selected. The classification analysis revealed that the dry phases of each climate modes consistently result in dry conditions, while the wet phases of these phenomena are highly variable. With the use of Pearson correlation analysis the magnitude of the strength of the effective phases of climate modes on Victoria's seasonal rainfall was shown. Pearson correlation analysis was further used in order to find the relationship between spring seasonal rainfall and lagged climate modes prior to spring. It was revealed that only three months of June, July and August of climate modes have statistical relationship with spring rainfall. The statistical significant

lagged climate modes were used as model inputs when developing ANN, MLR and ANFIS models.

Two different scenarios where considered in choosing the ANN model inputs. In order to evaluate the extent in time where the antecedent climate modes can affect rainfall predictability, the first scenario was developed based on two different input sets: a) three months' single climate indices and b) nine months' single climate indices as rainfall predictors. The results revealed that nine months single IOD models outperform the ANN models with three antecedent months' climate modes (ENSO/IOD) as model inputs. In order to evaluate the potential of combined ENSO-IOD and ENSO-IPO in rainfall forecasting the second scenario was defined based on combined climate modes. The results of the second scenario revealed that antecedent combined ENSO-IOD are better predictors for seasonal rainfall forecasting than combined ENSO-IPO. Comparing the results of the first and second scenario revealed that in east Victoria, ANN models based on combined ENSO-IOD are superior to single climate mode models. As east Victoria is closer to the Pacific Ocean where ENSO occurs, the effect of both ENSO and IOD can therefore be felt in this region. However, towards central and west Victoria and closer to the Indian Ocean, the models based on nine antecedent single IOD were showing better forecasts compared to the combined ENSO-IOD models.

The results of the MLR models revealed that MLR models were able to forecast spring rainfall with acceptable errors for some stations, while showing less accurate results for the others. It was shown that MLR modelling approach has many statistical limitation and care needs to be taken when developing reliable models. By comparing ANN and ANFIS models, it was found that ANN models have lower errors in east Victoria, and the results of ANN and ANFIS models are comparable for central and west Victoria.

The results of the nonlinear models (ANN and ANFIS) were compared with the Predictive Ocean Atmosphere Model for Australia (POAMA) which is the official model used by the Australian Bureau of Meteorology (BoM) to produce daily to seasonal forecasts. By comparing the forecasts of the ANN and POAMA models, it was shown that the ANN models are comparable with the POAMA in regards to model errors, but in most stations the ANN models are superior to the POAMA model in

regards to correlation coefficient of the models. It was shown that the ANFIS models outperform the POAMA model in west and part of central Victoria. In summary, this research has revealed the potential of artificial intelligence techniques in seasonal rainfall forecasting in Australia.

Acknowledgements

First and foremost, I wish to express my greatest gratitude to God Almighty, for all his blessings on me during the course of life and for bestowing me with enough courage and strength to pursue my studies.

I would like to express my sincere appreciation and gratitude to my principal supervisor, Associate Professor Monzur Imteaz for his continuous support and invaluable guidance during my study. I would also like to thank my internal supervisory committee member, Dr. Shirley. Gato-Trinidad as well as my external supervisory committee member Dr. Amgad Elmahdi from Climate and Water division, Bureau of Meteorology, Melbourne, VIC, Australia for their guidance and support throughout the duration of this research. I would like to extend my appreciation to Dr. Amin Talei from Monash university, Malaysia campus for his invaluable guidance and constructive comments for the ANFIS analysis in this research. I would like to thank Dr. Eun-Pa Lim and Mr. Griffith Young from the Australian Bureau of Meteorology for their valuable comments and advice on obtaining and analyzing POAMA forecasts. I would also like to express my deep gratitude to Professor Gary Meyers for his invaluable comments.

I also acknowledge Swinburne University of Technology for supporting and funding this research project through a Swinburne University Postgraduate Research Award (SUPRA). I wish to thank all my friends who made my stay at Swinburne University of Technology an enjoyable experience.

Last but not least, I wish to express my sincere gratitude to my family whose love, patience and encouragement has given me power, enthusiasm, and motivation not only during my Ph.D. study but throughout my life.

Declaration

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at any other institutions. I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

Fatemeh Mekanik

July 2015

Dedication

I would like to dedicate this thesis to my beloved parents.

List of Publications

Book Chapters

1) Mekanik, F., and Imteaz, M. (2014). "Evaluating the Effect of Single and Combined Climate Modes on Rainfall Predictability". Chapter 40, IAENG Transactions on Engineering Technologies (Vol. 247, pp. 571-581): Springer publisher. DOI: 10.1007/978-94-007-6818-5_40

Journal Papers

1) Mekanik, F., Imteaz, M. A. and Talei A. accepted. "Seasonal Rainfall Forecasting by Adaptive Network-based Fuzzy Inference System (ANFIS) using Large Scale Climate Signals". Climate Dynamics.

2) Mekanik, F., Imteaz, M. A., Gato-Trinidad, S., and Elmahdi, A. (2013). "Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes". Journal of Hydrology, 503(0), 11-21.

Peer-Reviewed Conference Papers:

1) Mekanik, F. and Imteaz, M. "Analysing lagged ENSO and IOD as potential predictors for long-term rainfall forecasting using multiple regression modelling". 20th International Congress on Modelling and Simulation, Adelaide, Australia, December 2013. Indexed in Scopus

2) Mekanik, F. and Imteaz, M. "Capability of Artificial Neural Networks for predicting long-term seasonal rainfalls in east Australia". 20th International Congress on Modelling and Simulation, Adelaide, Australia, December 2013. Indexed in Scopus

3) Mekanik, F. and Imteaz, M. "Artificial Neural Networks (ANN) modelling of spring rainfall using dual-climate indices for Victoria, Australia". Hydrology and Water Resources Symposium, Sydney, Australia, November 2012.

4) Mekanik, F. and Imteaz, M. "A Multivariate Artificial Neural Network Approach for Rainfall Forecasting: Case Study of Victoria, Australia". World Congress on Engineering and Computer Science, San Francisco, USA, October, 2012. Indexed in Springer

5) Mekanik, F. and Imteaz, M. "Forecasting Victorian spring rainfall using ENSO and IOD: A comparison of linear multiple regression and nonlinear ANN". International Conference on Uncertainty Reasoning and Knowledge Engineering, Jakarta, Indonesia, August 2012. Indexed in Scop

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ALI	Atmosphere And Land Initialization
ANN	Artificial Neural Network
ANFIS	Adaptive Network-Base Fuzzy Inference System
AOI	Antarctic Oscillation Index
BMA	Bayesian Model Averaging
BoM	Australian Bureau Of Meteorology
BP	Back-Propagation
CGCM	Coupled General Circulation Model
CLLJ	Caribbean Low-Level Jet
CSIRO	Commonwealth Scientific And Industrial Research Organization
DMI	Dipole Mode Index
DW	Durbin-Watson
ENSO	El Nino Southern Oscillation
FIS	Fuzzy Inference System
FNN	Feedforward Neural Networks
GCM	Global Climate Model
IOD	Indian Ocean Dipole
IPO	Interdecadal Pacific Oscillation
ISMR	Indian Summer Monsoon Rainfall
ISMRI	Indian Summer Monsoon Rainfall Index
KNMI	Royal Netherlands Meteorological Institute
LEPS	Linear Error In Probability Space
LMBP	Levenberg-Marquardt Back-Progation
LMS	Least Mean Square
MAE	Mean Absolute Error
MDB	Murray Darling Basin
MF	Membership Function
MLFN	Multi Layer Feed-Forward Neural Network
MLR	Multiple Linear Regression
MSE	Mean Square Error
MSLP	Mean Sea Level Pressure
NFS	Neuro-Fuzzy System
NRM	Natural Resources Management

NWP	Numerical Weather Prediction
PCA	Principal Component Analysis
PDO	Pacific Decadal Oscillation
PNA	Pacific North America
POAMA	Predictive Ocean Atmosphere Model For Australia
QB	Quasi-Biennial
RBF	Radial Basis Function
RMSE	Root Mean Square Error
SAM	Southern Annular Mode
SEA	South East Australia
SLP	Sea Level Pressure
SO	Southern Oscillation
SOI	Southern Oscillation Index
SOM	Self Organizing Map
SST	Sea Surface Tempurature
SWWA	South West Western Australia
TDNN	Time Delay Neural Network
TIO	Tropical Indian Ocean
YRV	Yangtze River Valley

List of Notations

SOI	Standardized difference between Darwin and Tahiti barometric pressures
P_{diff}	The difference between average Tahiti and Darwin mean SLP
$P_{diff(av)}$	Long-term average of P_{diff}
$SD_{(Pdiff)}$	Long-term standard deviation of P_{diff}
$\hat{x}_i$	Normalized observation
x_i	Observed data
x_{max}	Maximum observation
x_{min}	Minimum observation
O_i	Simulated (predicted) output
n	Number of observations
$x_{i(anom)}$	Observation anomaly
$\bar{x}$	Average of observations
X_1 ,X_2	First and second independent variable for MLR
b_1 and b_2	Coefficients of first and second independent variable
a	MLR model constant
c	MLR model error
ssr	Regression sum of squares
sst	Total sum of squares
R^2	Coefficient of multiple determination
y	ANN target output
$\hat{y}$	ANN predicted output
E	Sum of nonlinear least squares between the observed and the Predicted Output
net_{pj}	ANN weighted input into the jth hidden unit
W_{ji}	ANN weight from input unit *i* to the hidden unit *j*
y_{pi}	The value of the ith input for pattern p in ANN

W_{jo}	The threshold or bias for neuron *j*
$f(.)$	Tansigmoid activation function
M	The number of hidden units in ANN
W_{kj}	Weight connecting the hidden node j to the output k
W_{ko}	The treshold value for neuron *k*
$\hat{y}_{pk}$	The kth predicted output
h (.)	The purelin activation function
J	Jacobian matrix
e	Vector of network errors
ETR	Error trend ratio
$\mu_A(x)$	Membership function of x in A
$\bar{w}_i$	The normalized firing strength in ANFIS
p_i, q_i	Linear parameters in the consequent part of the first-order Sugeno fuzzy model
μ_{Ai} (or μ_{Bi-2})	Fuzzy set associated with node i
O_i^j	The output of node *i* in layer *j* in ANFIS

Chapter 1

Introduction

1.1 Background

The Australian climate is highly affected by changes of sea surface temperature (SST) and sea level pressure (SLP) in the surrounding oceans, particularly the Pacific and Indian Oceans. The El Nino Southern Oscillation (ENSO) occurring in the Pacific Ocean and the Indian Ocean Dipole (IOD) which takes place in the Indian Ocean are among the most important climate drivers of Australia. Australia is also affected by the Southern Annular Mode (SAM), which is the principal mode of atmospheric variability in the mid and high latitudes (Risbey et al. 2009). It is also believed that the Interdecadal Pacific Oscillation (IPO), a low frequency (15–35 years) form of variability of the tropical and extra-tropical Pacific Ocean (Verdon-Kidd &Kiem 2009a) modulates Australian rainfall.

The relationship between large-scale climate modes (ENSO, IOD, SAM and IPO) which take place in the surrounding oceans and the Australian climate has been studied for many years. Although some insight into the effect of large-scale climate modes on Australia has been gained after years of research, due to the complexity of these relationships, many aspects of these phenomena remain unknown. One main component of the atmosphere which is affected by the SSTs and SLPs of oceans around Australia is Australian rainfall. Australia is a vast continent where different regions experience different climate, and the climate can vary significantly from one year to another (BoM, 2014). Australian rainfall is highly variable both in space and time. Previous studies have examined the simultaneous relationship between Australian rainfall and large-scale climate mode(s) in different parts of Australia. However, less consideration has been given to understanding the effect that these phenomena have on rainfall prediction. Forecasting and monitoring of rainfall values are highly important for numerous aspects of human life. Unforeseen flash floods produced by severe rainfall result in natural hazards threatening human lives and properties. Of the significant natural causes of national disasters and famines around the world, large-scale floods and droughts account for a large percentage of the calamities.

Furthermore, Australia's agriculture and population are highly concentrated in the south-east of the continent (Murphy & Timbal 2008). With the effect of rainfall on water resources as a foregone conclusion, more accurate prediction of rainfall would enable more efficient utilization of water resources and power generation. The Australian Bureau of Meteorology (BoM), together with the Commonwealth Scientific and Industrial Research Organization (CSIRO), has developed a dynamic prediction model by the name of Predictive Ocean Atmosphere Model for Australia (POAMA). POAMA is a complex dynamic model with the initial focus of forecasting ENSO (Kiem & Verdon-Kidd 2009). POAMA is a state-of-the-art dynamic model which produces daily to seasonal rainfall and temperature forecasts.

Generally, rainfall prediction models are categorized into either dynamic prediction models or statistical models**.** Dynamic models are deterministic and do not require information about a specific situation beyond the initial and boundary conditions. Although significant improvement has been made in the dynamic modelling of rainfall patterns, dynamic models are still much below their desired level of accuracy. In contrast, with statistical methods one does not know the dynamic relationship between the cause and effect of the system, nor one retains certain conceptions about the related roles of different processes that govern a phenomenon (Chakraverty & Gupta 2008). Statistical forecasts are an active area of research and with the use of non-linear techniques new developments are promised. While statistical prediction systems rely on the relationships between the variables, dynamic prediction systems are based on numerical simulation of the physical processes. However, in spite of considerable research effort and technological advances, sophisticated dynamic prediction systems are not able to out-perform simple statistical prediction systems (Schepen et al. 2012).

The Artificial Neural Network (ANN) approach is a non-linear statistical technique that has become popular among scientists as an alternative technique for predicting and modelling complicated time series, weather phenomena and climate variables. The performance of non-linear mapping between inputs and outputs has made ANN a suitable candidate for the prediction of rainfall which its formation involves rather complex physics. ANN is an adaptive system which changes its structure based on external or internal information that flows through the network during the learning

phase and can deal with large amounts of dynamic and non-linear noisy data (Nourani et al. 2009).

1.2 Statement of Problem

As Australia is exposed to severe droughts and floods, seasonal rainfall forecasting is crucial for water resources management, food production and mitigating flood risks. In recent years, Australia has experienced dramatic flood events due to inclement climate conditions. From December 2010 until early 2011, major floods due to heavy rainfall occurred in several states, including Victoria, New South Wales and Queensland (AbdulRauf &Zeephongsekul 2014). According to the Australian Bureau of Meteorology (BoM) flooding is Australia's most costly natural disaster. Although floods do not have as wide an effect on Australia as droughts, they still cause damage to dwellings, transportation networks, and other infrastructure. In Victoria, flooding occurs mostly in winter and spring. Seasonal rainfall forecasting for Victoria in south-east Australia has not been as successful as in other parts of Australia, particularly Queensland (Verdon-Kidd & Kiem 2009b). The relationship between the major large-scale climate modes affecting the Australian climate and simultaneous daily to monthly rainfall has been studied for many years. However, a strong relationship between simultaneous climate mode(s) and rainfall does not essentially mean that there is also lagged relationship. For the purpose of rainfall forecasting, knowledge of the significant lagged relationships is essential, but very few studies have examined the lagged relationship between large-scale climate mode(s) and Australian rainfall. On the other hand, using antecedent climate modes as potential predictors of future rainfall has not received enough attention. Therefore, further research in this area is necessary for the following reasons: firstly, the findings are not yet conclusive about the time extent in which the climate modes influence rainfall prediction in different seasons and different regions. Secondly, the seasonal forecast results for Victoria are poor compared to other parts of Australia. According to Verdon-Kidd and Kiem (2009b), in comparison to eastern Australia and particularly Queensland, the performance of past studies on Victorian seasonal rainfall predictability have been low and a maximum predictability of only 30% was achieved. Finally, non-linear methods have been rarely used in the area of seasonal forecasting using large-scale climate mode(s) in Australia; most studies have used linear regression analysis or probabilistic/categorical analysis between

rainfall and simultaneous large-scale climate modes. Of the limited studies focused on the relationship between rainfall and past values of climate modes, only a few have used non-linear techniques (Schepen et al. 2011; Abbot and Marohasy 2012; Mekanik et al. 2013).

Furthermore, seasonal rainfall prediction models have not been very satisfactory in terms of accuracy when compared with daily or monthly rainfall prediction models. The probable reasons that make conducting seasonal rainfall prediction difficult are the complexity of the atmospheric processes and the uncertainty of the relationships between rainfall and hydro-meteorological variables. In addition, seasonal rainfall prediction with the use of numerical models has not demonstrated useful performance since rainfall prediction from such models is an average over grid point values, and therefore is a function of the model's spatial resolution only; in this case neglecting the temporal variation will lead to consistent inaccuracies since rainfall is highly variable both in space and time. Hence, there is a need to develop seasonal rainfall prediction models using new data-driven tools like Artificial Neural Networks to represent the non-linear dynamic relationships among the data. Thus, the development of a sophisticated non-linear forecast model for Victoria using the potential climate modes as predictors is essential for accurate rainfall forecasting for this region.

1.3 Aims and Objectives

The main objective of this study is the development of a reliable seasonal rainfall prediction model for Victoria with the use of large-scale climate modes as potential predictors. The Artificial Neural Network (ANN) technique is chosen as the main modelling approach, followed by the Adaptive Network-based Fuzzy Inference System (ANFIS) and Multiple Linear Regression (MLR) modelling as bench marks for comparison with the developed ANN models. To the best of the author knowledge the present study is the first study in Australia that applies ANFIS approach for rainfall forecasting in conjunction with the use of large-scale climate modes. The detailed objectives of the study are as follows:

- To find the relative contributions of concurrent and antecedent large-scale climate modes (i.e. ENSO, IOD, and IPO) on Victoria's seasonal rainfall

- To classify and investigate the effects of different climate mode phases on Victoria's seasonal rainfall
- To investigate how these contributions vary by location across Victoria
- To develop reliable nonlinear ANN models for seasonal rainfall forecast in Victoria
- To develop MLR models for comparison with ANN results
- To develop ANFIS model for the first time for Victoria in order to examine its potential in forecasting seasonal rainfall and also to have a nonlinear benchmark for comparison with the ANN models
- To investigate the extent in time (e.g. one month, three months or further antecedent (lagged) climate modes) for which the antecedent months' climate modes are able to provide more accurate rainfall forecasts
- To compare the results of the developed model with the official forecast model of the Australian Bureau of Meteorology (BoM) in order to explore the shortcomings, effectiveness and advantages of each model.

1.4 Research Scope

The research scope of this study is as follows:

Three distinct regions in Victoria are selected as case studies, and monthly rainfall data for three rainfall stations in each region are obtained. Monthly climate modes data are also gathered and pre-processing of the data is done, based on the requirements of each model. Rainfall anomalies are constructed based on different phases of climate modes in order to investigate the effect of different climate phases on seasonal rainfall in Victoria and classification analysis is carried out. Further, the simultaneous and antecedent relationships between large-scale climate modes (ENSO, IOD, etc.) are investigated with the use of Pearson correlation analysis.

MLR, ANN and ANFIS models were developed. As a pioneer study, the model development is initiated by selecting the appropriate inputs for each type of model based on linear correlation analysis. The models are then calibrated and tested and the best models for each region is selected. The author acknowledges that with the use of a more sophisticated input selection techniques (e.g. genetic algorithm, partial mutual

information, etc.) better forecast results might be achieved. The results of the non-linear models (ANN and ANFIS) are then compared with the results of the official forecast model of the BoM and the advantages and limitations of the developed models are discussed. Due to software limitation, the rainfall forecasts in this study are achieved based on rainfall stations as opposed to grid rainfall forecasts. With the use of grid data set and more sophisticated software this study can be expanded in order to produce grid rainfall forecasts.

1.5 Outline of the Thesis

The thesis outline is as follows:

A thorough literature review is conducted and discussed in Chapter 2. Chapter 3 introduces the study area and the model verification criteria. The analysis of the effect of phases of climate modes on rainfall is discussed in Chapter 4. Chapter 5 discuses the MLR modelling approach and results. The ANN and ANFIS modelling methodology and results are discussed in Chapter 6 and 7 respectively. Chapter 8 discusses the comparison between ANN and ANFIS models with the POAMA forecasts followed by summary, conclusion and recommendation in Chapter 9.

Chapter 2

Literature Review

2.1 Introduction

One of the most important challenges for sustainable water resources management in many parts of the world is managing a greatly variable climate in conjunction with increasing demand for natural resources; the rainfall and stream flow regimes of Australia rank among the most variable and therefore Australia is no exception (Nicholls et al. 1997). This variability happens during different time scales, from annual to multidecadal and possibly longer (Verdon-Kidd & Kiem 2009b). The variation of climate around the globe is related to the variation of sea surface temperatures (SSTs) and sea level pressures (SLPs) of the oceans around the world. The fluctuation between SSTs and SLPs create the so called large scale climate modes. Among the most influential large scale climate modes are the El Nino southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD) which occur in Pacific and Indian Ocean respectively. In addition to these two major modes, the Interdecadal Pacific Oscillation (IPO) and Southern Annular Mode (SAM) also affect the world climate. IPO is a form of variability similar to ENSO which takes place in Pacific Ocean on a multidecadal time scale. SAM is another important climate mode affecting mostly the southern hemisphere. Many studies have tried to investigate the existent relationship between these modes and the climate (particularly rainfall) around the world (Lau et al. 2001; Yufu et al. 2002; Barsugli & Sardeshmukh 2002; Hartmann et al. 2008; Chattopadhyay et al. 2010; Shukla et al. 2011). In the following sections a brief description of each mode is given, followed by a review of the studies which have examined the relative and independence role of these phenomena around the world. Finally, the research on the impact of large scale climate mode in Australia and particularly southeast Australia is reported.

2.1.1 El Nino Southern Oscillation

El Nino Southern Oscillation (ENSO) which results in climatic changes across the tropics and subtropics, refers to the influences of a band of sea surface temperatures that

are unexpectedly warm or cold for long periods of time which develops off the western coast of South America. Deviations in the temperature of the surface of the tropical eastern Pacific Ocean and in air surface pressure in the tropical western Pacific Ocean is referred to as the Southern Oscillation (Bamston et al. 1997). Warming and cooling of the tropical eastern Pacific Ocean are known as El Nino and La Nina, respectively. The variations in surface temperature and air surface pressure are joined together; the El Nino (the warm oceanic phase) is accompanied by the high air surface pressure in the western Pacific and the La Nina (the cold oceanic phase) is accompanies by the low air surface pressure in the western Pacific (Verdon et al. 2004;) It is this ocean-atmosphere fluctuation that is referred to as the El Nino Southern Oscillation. It should be noted that mechanisms which result in the oscillation still remain under study. Extreme weather such as floods and droughts which happens in many regions of the world is caused by the extremes of this climate pattern's oscillations.

The initiating reasons of an ENSO warm or cool event are not exactly recognized. However, sea surface temperature and atmospheric pressure as the two components of ENSO are intensely correlated. ENSO is generally represented by two types of indicators, the SLP indicator and the SST indicator. ENSO conditions are monitored in 3 geographic regions of the equatorial Pacific using SST anomalies defined as Nino3 ($5^oS – 5^oN$, $150^o – 90^oW$), Nino3.4 (5^o S – 5^oN, $170^o – 120^oW$) and Nino4 ($5^oS – 5^oN$, $160^o – 150^oW$) (Risbey et al. 2009).

The SLP indicator is defined as a bimodal variation in sea level barometric pressure between observation stations at Darwin, Australia and Tahiti and is referred to as the Southern Oscillation Index (SOI). SOI is a standardized difference between the two barometric pressures. According to Australian Bureau of Meteorology (BoM) SOI is calculated as follow:

$$SOI = 10 \times \frac{P_{diff} - P_{diff(av)}}{SD(P_{diff})} \quad (1\text{-}1)$$

where

P_{diff} = (average Tahiti mean SLP for the month) - (average Darwin mean SLP for the month)

$P_{diff(av)}$ = long term average of P_{diff} for the month in question
$SD_{(Pdiff)}$ = long term standard deviation of P_{diff} for the month in question.

The complicated interactions between the ocean, atmosphere and neighbouring regions across the Pacific indicate that ENSO events have influences on weather in areas outside the tropical Pacific region. The different climatic conditions around the Pacific are related to El Nino and La Nina events (Hoerling & Kumar 2000). ENSO's warm phase (El Nino) conditions refers to SST anomalies equal to or greater than 0.5°C in the Nino 3.4 region including portions of Nino regions 3 and 4. However, cool phase (La Nina) conditions are related to anomalies less than or equal to –0.5°C. Generally, lower pressure over Darwin and higher pressure over Tahiti causes a circulation of air from east to west, drawing warm surface water westward and bringing precipitation to Australia and the western Pacific. El Nino conditions are intensely coincidental with reduction in the pressure difference which causes severe drought in parts of the western Pacific, such as Australia. However, the west coast of equatorial South America can experience flooding due to the heavy precipitation across the ocean (NOAA-2014). The ocean near Australia is cooler than usual during El Nino events which causes lower than average winter–spring rainfall over eastern and northern regions. Most significant Australian droughts have been linked with El Nino events. However, presence of an El Nino does not definitely results in a prevalent drought (Verdon-Kidd & Kiem 2009a; Gallant et al. 2012)

The most important driver of ENSO is the temperature gradients both at the surface and below the surface across the Pacific and particularly at the thermocline. Thermocline is a Greek term meaning the heat slope; thermocline refers to the region separating warm and well-mixed surface water from cool and deep ocean water. In general water temperatures above the thermocline are more than 25°C while those below the thermocline are 15°C or less. Greater convection over the warmer ocean to north of Australia are connected with La Nina events which usually result in higher than average rainfall and occasionally causing floods across much of Australia, specifically inland eastern and northern regions. In the case of having neither El Nino nor La Nina (a neutral state) trade winds blow east to west across the surface of the tropical Pacific Ocean. In the neutral state the western Pacific experiences warm moist air and warmer surface waters, while the central Pacific Ocean remains reasonably cool. The trade

winds are the east to southeasterly winds in the Southern Hemisphere which affect the northern areas of Australia.

In other words, warm sea surface temperatures in the western Pacific bring heat and moisture into the atmosphere above. Atmospheric convection refers to the process of rising this warm air into the atmosphere. When the air is sufficiently humid, it results in cumulonimbus clouds and rain. The air which is now drier moves to east before falling over the cooler eastern tropical Pacific. The Walker Circulation is referred to the pattern of air rising in the west and descending in the east with westward moving air at the surface. Trade winds become less strong or may even reverse during an El Nino event which allows the movement of the area of warmer than normal water into the central and eastern tropical Pacific Ocean. Deepening of the thermocline in the central to eastern Pacific are linked with these warmer than normal ocean temperatures. Warmer sea surface temperatures are related to a weaker upwelling of cooler ocean waters from below. SST around northern Australia are cooler than normal; the convection drifts away from Australia eastward towards the central tropical Pacific Ocean and thereby, causing more rainfall for regions such as Kiribati and Peru, but less rainfall over Australia. Inside eastern Australia usually has the greatest influences; however, areas such as southwest Western Australia and coastal New South Wales can experience different impacts from event to event. In western Tasmania effects are usually less felt (BoM-2014).

The Walker Circulation strengthens during a La Nina event, with stronger trade winds and more convection over the western Pacific. The pool of warmer water is restrained to the far western tropical Pacific when the trade winds intensify which leads to warmer than usual SST in north of Australia. SST over the central and eastern tropical Pacific Ocean become cooler than usual and the thermocline migrates closer to the surface. As upwelling intensifies, cool waters from the deep ocean are drawn to the surface. Stronger winds deliver more moisture to the overlying atmosphere and the Walker Circulation increases, resulting in increase in convection and also cloudiness over the north of Australia. This situation intensifies the Australian monsoon. Further, if the conditions are right, it increases humidity and rainfall over Australia. Increased rainfall over much of northern and eastern Australia are linked with La Nina events. The effects

of La Nina is superior to that of El Nino in parts of northern and central Australia.(Cai et. 2011; Murphy & Timbal 2008; Nicholls 1989)

2.1.2 Indian Ocean Dipole

Indian Ocean Dipole (IOD), similar to ENSO, is a coupled oceanic atmospheric event in the equatorial Indian Ocean (Saji et al. 1999). The IOD is represented by the difference in SST between two poles (i.e. a dipole) in which a western pole is in the western Indian Ocean and an eastern pole is in the eastern Indian Ocean south of Indonesia. The IOD has significant effects on climate and rainfall variability in Australia and other countries surrounded by the Indian Ocean Basin. Similar to an ENSO event, the variation in temperature gradients across the Indian Ocean changes the preferred regions of rising and descending moisture and air.

The IOD and ENSO events are interrelated through an extension of the Walker Circulation to the west and associated Indonesian warm tropical ocean water flowing from the Pacific into the Indian Ocean. According to Meyers et al. (2007), the IOD and ENSO can sometimes occur together in such a way that strengthens each other. Therefore, positive IOD events are frequently linked with El Nino, while negative events are associated with La Nina. When the IOD and ENSO are out of phase the influences of El Nino and La Nina events can be weakened. However, if they are in phase the effects of El Nino and La Nina events are commonly extreme over Australia (BoM-2014).

According to Risbey et al. (2009) the IOD usually peaks in Australian spring (September-November), nevertheless it can occur from May to November. The dipole mode index (DMI) is a measure of the IOD. Saji et al. (1999) defined DMI as the difference in SST anomaly between the tropical western Indian Ocean (10^{o}S–10^{o}N, 50^{o}–70^{o}E) and the tropical south-eastern Indian Ocean (10^{o}S–equator, 90^{o}–110^{o}E). There are arguments over the independence extent of IOD from ENSO (Saji et al. 1999; Ashok et al. 2003; Meyers et al. 2007). An index of IOD was developed by Meyers et al. (2007) using a lagged empirical orthogonal function (EOF) approach in which variation in ENSO in defining the IOD is taken into account.

The understanding of the dynamics of the IOD has improved rapidly; this has been comprehensively summarized by Yamagata et al. (2004). Similar to ENSO, the growth and maintenance of positive and negative patterns are significantly influenced by the change of subsurface temperature in the depth range of the thermocline. According to Yamagata et al. (2004), based on evidences from coupled, numerical models the IOD can grow by ocean–atmosphere interaction involving the thermocline. Although there are limited number of observational studies on the depth of the thermocline; however, it has been demonstrated through these observations that the thermocline experiences large vertical displacements beneath both poles of the dipole. Further, the displacements are correlated to the local SST anomalies (Meyers 1996; Rao et al. 2002; Feng and Meyers 2003). Similar to ENSO, it is shown in the previous studies that the depth of the thermocline is mainly forced by remote winds, from both the Indian and the Pacific Oceans (Wijffels & Meyers 2004). According to Feng and Meyers (2003), the remote forcing and the local wind are elements in the generation of the SST of the eastern pole, in such a way that cool SST anomalies (i.e., positive IOD) develop when the thermocline is shallow due to remote forcing and the easterly wind is helping the upwelling along the coast of Java. Therefore, similar to ENSO, upwelling in the Java–Sumatra region appears to be a vital controlling process in the growth of IOD anomalies.

Positive IOD event results in warmer SST in the western Indian Ocean relative to the east and easterly wind anomalies across the Indian Ocean. It also leads to less cloudiness to northwest of Australia as well as less rainfall over southern Australia. On the other hand, negative event are associated with cooler sea surface temperatures in the western Indian Ocean relative to the east, bringing more westerly winds, increasing cloudiness to Australia's northwest, and more rainfall in the southern Australia (Murphy & Timbal 2008; Gallant et al. 2012).

2.1.3 Southern Annular Mode

The Southern Annular Mode (SAM), also known as the Antarctic Oscillation (AAO), refers to the north–south movement of the westerly wind belt that circles Antarctica. According to Thompson and Solomon (2002), the SAM is the governing mode of atmospheric variability in the mid-and high latitudes of the Southern Hemisphere. The

SAM represents the north–south shifts in mass between the pole and mid-latitudes and zonal wind anomalies between about 30^{o} and 60^{o} latitude. Thompson (2014) defines the positive (i.e. high index) SAM phase as where pressures are lower than normal in the polar region with improved westerly winds along 55^{o} and 60^{o} latitude. A simple index of the SAM was defined as the difference between normalized monthly zonal mean sea level pressure at 40^{o} and 65^{o}S which has been calculated from station pressures by Marshall (2003) for the period from 1957 to the present (Gong & Wang 1999).

It is believed that variations in the SAM are associated with rainfall variability in each of the Southern Hemisphere regions (Hendon et al. 2007). An analysis of SAM contributions to rainfall variability in Australia has been carried out by Hendon et al. (2007). A daily SAM index was used and it was found that up to about 15% of weekly rainfall variance in parts of south-western and south-eastern Australia are explained by the SAM. Hendon et al. (2007) noted that although ENSO has substantial rainfall relationships for broader sections of the continent; however, SAM contribution in rainfall variance is comparable to the amount of rainfall variance associated with ENSO for these regions.

The belt of strong westerly winds contracts towards Antarctica during positive SAM events, which leads to weaker than normal westerly winds and higher pressures over southern Australia. Further, it limits the penetration of cold fronts inland. On the contrary, an expansion of the belt of strong westerly winds towards the equator indicates a negative SAM event. Stronger storms and low pressure systems over southern Australia occur as a result of this shift in the westerly winds. A positive SAM value during autumn and winter translates to elimination of rainfall in southern Australia. On the other hand, a strong positive SAM during spring and summer indicates that southern Australia is affected by the northern half of high pressure systems; thereby, more easterly winds carry moist air from the Tasman Sea which could bring more rain as the winds hit the coast. An important contributor to the “big dry” observed in southern Australia from 1997 to 2010 has been a high positive SAM dominating during autumn–winter in the recent years (Hendon et al. 2007; Risbey et al. 2009).

2.1.4 Interdecadal Pacific Oscillation

Similar to ENSO, a sustained pattern of Pacific climate variability is referred to as the Interdecadal Pacific Ocean (IPO) (Power et al. 1999) and Pacific Decadal Oscillation (PDO) (Zhang et al. 1997). According to Mantua et al. (1997), PDO (IPO) is the phenomenon responsible for these multi-decadal step changes in climate. Variable periods of warming (i.e. positive phase) and cooling (i.e. negative phase) in both hemispheres of the Pacific Ocean are associated with the PDO and IPO (Folland et al. 2002). However, the persistence of PDO/IPO periods (15–30 yr) and the fact that the climatic fingerprint of the PDO is most prevailing in the north Pacific sector with a secondary signature in the tropics are the two characteristics differentiate the PDO (and IPO) from ENSO (Mantua & Hare 2002). Decadal and annual-scale fluctuations in maximum temperature, rainfall, water volume transport and wheat crop yield and the general climate variability in Australia are linked to the IPO/PDO phenomena (Power et al. 1999; Kiem et al. 2003; Verdon et al. 2004). According to Folland et al. (2002), the IPO/PDO mainly affects the eastern Australian climate during the austral spring, summer and autumn by inducing variations in the South Pacific Convergence Zone that tends to be active during these months. During the mid-1940 through to the mid-1970 a period of higher rainfall and stream flow has occurred across much of eastern Australia (Verdon-Kidd & Kiem 2009b) which is related to IPO. Climate patterns around the world are also influenced by the PDO and IPO (e.g. Kiem et al. 2003; Verdon et al. 2004). From at least the 15th Century the IPO/PDO has been known as a dominant climate mode in the Pacific sector. Therefore, climate in the future is likely to continue to be influenced by the IPO/PDO (Verdon-Kidd & Kiem 2009a).

The IPO/PDO indirectly controls the eastern Australian climate by modulating the magnitude and frequency of ENSO effects (Power et al. 1999; Kiem et al. 2003; Verdon et al. 2004). The relationship between ENSO and Australian rainfall is declining when the IPO/PDO is in a warm phase. However, it is reinforced during the cool phase (Power et al. 1999).

An increased response of rainfall and streamflow to La Nina events during a cool IPO/PDO phase is the utmost impact of this modulation. Wet events are possibly to be wetter and more frequent during the negative (i.e. cool) IPO/PDO phase compared to a

neutral or warm IPO/PDO phase; which increases the flood risk in the Murray darling Basin (MDB) located at southeast Australia (Kiem et al. 2003; Verdon et al. 2004). On the contrary, wet events are less frequent and not as wet as they are during the IPO/PDO cool phase during the positive (i.e. warm) IPO/PDO phase. This leads to an elevated risk of drought across the MDB and other parts of eastern Australian (Verdon-Kidd & Kiem 2009a). Examining paleoclimate reconstructions of the two climate modes confirmed that the relationships between IPO/PDO phase and the frequency of ENSO events is consistent over the past 450 year. According to Lough (2007), the relationship between ENSO, IPO and rainfall/streamflow in northeast Queensland is consistent for at least the last 400 year.

2.2 Effect of Large Scale Climate Modes on Global Rainfall

Rainfall, evapotranspiration, temperature and humidity are considered as hydro-climate variables. Rainfall is the most significant and most investigated hydro-climate variable. Several research have been carried out to evaluate rainfall characteristics, mechanisms, their spatiotemporal changes, pattern and variability around the world (Ventura et al. 2002; Cheng et al. 2004; Rio et al. 2005; Kim et al. 2008; Grimm 2011; Niu 2013). The variability of rainfall around the world has been linked to large scale climate modes (Ashok et al. 2001; Mason & Goddard 2001; Manatsa et al. 2012). Researchers have used different methods and modelling techniques in order to extract and model the relationships between these large-scale climate modes and rainfalls in different parts of the world (Lau et al. 2001; Yufu et al. 2002; Barsugli & Sardeshmukh 2002; Hartmann et al. 2008; Chattopadhyay et al. 2010; Shukla et al. 2011). In this section some of these studies are reviewed.

The most significant source of variability for South American summer rainfall is known to be ENSO (Grimm 2011). According to Niu (2013), the main variabilities of rainfall were shown to be related to the IOD phenomenon in the Pearl River basin in China, particularly in the central and eastern part of the basin. The effect of large-scale climate oscillations such as ENSO and PDO on rainfall of the Colorado River basin is recognized by Kim et al. (2008) . Barsugli and Sardashmukh (2002) evaluated the global atmospheric response to SST anomalies' sensitivity via the general circulation model (GCM) through the tropical Indian and Pacific Ocean basins. For a uniform array

of 42 localized SST anomaly patches the responses of the model in January over the domain were calculated. To produce sensitivity maps, a statistically based smoothing procedure was used to combine the results from the individual forcing experiments. Target quantities of interest comprise the geopotential height response over the Pacific–North American (PNA) region and regional precipitation responses over North America, South America, Africa, Australia, and Indonesia. The analysis results revealed that that many significant targets for seasonal forecasting, comprising the PNA response, are most sensitive to SST anomalies in the Nino4 region of the central tropical Pacific. However, they have lesser and sometimes opposite sensitivities to SST anomalies in the Nino3 region of the eastern tropical Pacific. On the other hand, certain important targets including Indonesian rainfall are most sensitive to SST anomalies outside both the Nino4 and Nino3 regions. These results were also pertinent in evaluating atmospheric sensitivity to variations in tropical SSTs on decadal to centennial scales related to natural and anthropogenic forcing. The authors also revealed that warm SST anomalies in one-third of the Indo-Pacific domain result in a decrease of global mean precipitation.

Lau and Weng (2001) recognized three coherent modes of summertime rainfall variability over China and global SST during 1955–98 using Singular Value Decomposition. The influences of the El Nino in 1997-98 on main drought and flood incidences over China were evaluated according to these modes. The first mode, determined with the growing phase of El Nino superimposed on a warming trend since the mid-1950s, significantly affects rainfall over northern China. The second mode included a quasi-biennial (QB) variability manifested in alternate wet and dry years over the Yangtze River Valley (YRV) of central China. The third mode which had an opposite trend in southern China, was dominated by a quasi-decadal oscillation in eastern China between the Yangtze River and the Yellow River. The effects of these three modes on the 1997 and 1998 observed rainfall anomalies were assessed based on a mode-by-mode reconstruction. The results revealed that the influence of anomalous SST forcing during the growing phase of the 1997–98 El Nino possibly resulted in severe drought in northern China and the flood in southern China in 1997. It was also found that the severe flood over YRV in 1998 is related to the biennial tendency of basin-scale SST anomaly throughout the transition from El Nino to La Nina in 1997–98.

Moreover, it was revealed that the extended dry pattern over northern China and wet trend over YRV since the 1970s could be due to a long-term warming pattern in the tropical Indian Ocean and western Pacific. The long-term dry background intensified the drought situation over northern China in 1997. The wet background worsened the flood situation over YRV in 1998 due to the influences of the 1997–98 El Nino. On the other hand, neither El Nino nor QB signals showed clear dominance on the rainfall variability in southern China. Lau and Weng (2001) also discussed the significance, reliability and stability of the aforementioned results.

Guo et al. (2004) also studied the relationships between the floods in the Yangtze River valley and SST anomalies in the Pacific and Indian Oceans in 1998. According to their findings, their model was able to produce the heavy precipitation in the summer of 1998 over the valley of Yangtze River affected by global observational SST. It was revealed that the main characteristics of the observed subtropical high anomalies over the western Pacific was also possible to be produced by the model. For the different areas of the ocean and different periods the experiments with the observed SST were produced. Comparison of the influence of SST anomalies of different ocean areas on the floods revealed that they are significantly influenced by SST anomalies in the Indian Ocean. Based on their findings a much closer relationship exists between the SST anomalies in the Indian Ocean and the western Pacific and the strong anomalies of the subtropical high over the western Pacific compared to the SST anomalies in other concerned regions. Moreover, it was found out that the floods and subtropical high anomalies in the summer of 1998 were more influenced by the concurrent summertime SST anomalies compared to SST anomalies in the preceding winter and spring seasons.

Murphy et al. (2014) examined the relationship between monthly rainfall in the Panama Canal and SST anomalies. According to Murphy et al. (2014) in regards to monthly accumulated rainfall, the tropical Panama Canal Watershed has the largest inter-annual variability in December with the recorded wettest month being December 2010. They found that the December accumulated rainfall is associated with the SST anomalies in both the tropical North Atlantic and equatorial Pacific oceans. Nevertheless, a significant relationship with different SST anomalies were found. The flux of low-level moisture over the Caribbean Sea, southern Central America, and the eastern Pacific Ocean was found to be significantly influenced by the configuration of SST anomalies

in these two ocean basins during December. This effect is through the modification of the Caribbean Low-Level Jet (CLLJ) and the Chorro del Occidente Colombiano (CHOCO) jet. Murphy et al. (2014) found out that cool SST anomalies in the tropical Pacific, a fading (reinforcing) of the CLLJ (CHOCO jet), and increased moisture convergence over and around Panama are related to wet Decembers in the watershed. On the other hand, the opposite conditions are associated with dry Decembers. In other words, dry Decembers in the watershed are associated with warm SST anomalies in the tropical Pacific (North Atlantic) and increased moisture convergence over and around Panama. During December the distribution of daily rainfall is differently influenced by the SST anomalies in these two ocean basins. In other words, the SST anomalies in the Pacific (Atlantic) are mainly associated with variations in the frequency of heavier (lighter) rainfall.

In South America the variation of modern and past spatiotemporal precipitation is intensely influenced by ENSO cycles. The exact spatiotemporal rainfall pattern is complicated. However, in parts of the west coast of South America rainfall is usually increased during positive ENSO anomalies (El Nino). On the other hand, increased rainfall in parts of central South America and in the Amazon drainage basin is associated with negative ENSO anomalies (La Nina). Since precipitation affect erosion processes and discharge; thereby, in the Andes and in the neighbouring regions the global-scale ENSO phenomenon significantly influences sediment flux and aggradations/deposition cycles. On the other hand, there are only a limited number of rain- and river gauge networks in the remote areas of the Andes and Amazon drainage basin. Therefore, the variation in magnitude of the spatiotemporal rainfall and discharge remains weakly restrained between contrasting ENSO cycles (Bookhagen & Strecker, 2010). Mariotti et al. (2002) demonstrated that effect of ENSO on rainfall is significant in the Euro-Mediterranean areas in which their characteristics change seasonally. Similar to the ENSO-Europe connection in the spring, Mariotti et al. (2002) found a considerable correlation in the autumn. The absolute anomalies are small compared to tropical regions. However, the influence is relevant particularly for the regions around the Mediterranean with rare rainfall occurrence.

Several attempts have been made in the past to recognize the influence of Indian monsoon and large scale climate modes over the Indian continent. It has been shown

that there is a dominantly inverse relationship between ENSO and Indian rainfall using different methods and analysis of datasets comprising the last 140 years. However, in recent decades the dominant coupling between the ENSO and Indian rainfall is weakened. Sarkar et al. (2004) analysed pattern of rainfall over the Indian subcontinent and its relation with ENSO. Their results revealed a stronger circulation trend over the Indian region in the last two decades. It was also found that in recent years the effect of ENSO has increased; however, it has failed to influence the Indian rainfall due to the stronger circulation pattern dominant over India during this time. Sarkar et al. (2004) suggested that the effect of ENSO on Indian precipitation has increased in recent times. Further, it is proposed that the effect of ENSO is secondary to the local dynamics in the area. According to Sarkar et al. (2004), the strength of ENSO is only recognized during low bipolarity over the Indian Ocean which is when the local dynamics declines and the Walker circulation over this area becomes weak. The Walker circulation is further weakened with a stronger ENSO with the moving of the circulation cell towards the central Pacific which results in an overall low rainfall year. A strong local dynamics and circulation pattern has offset any impact of ENSO in the last two decades as a result of typically strong and persistent development of bipolarity in the Indian Ocean. The increased land-sea thermal contrast observed in recent years and a stronger dipole mode activity in the Indian Ocean (Sarkar et al. 2004) have possibly worked together; thereby, a good rainfall during the last few ENSO events is confirmed (Sarkar et al. 2004).

Ashok et al. (2001) examined the effect of the IOD on the inter-annual variability of the Indian summer monsoon rainfall (ISMR) from 1958 to 1997. The ISMR has been affected by the ENSO and the IOD during the last four decades. Whenever, the ENSO-ISMR correlation is low, the IOD-ISMR correlation is high. On the other hand, if the ENSO-ISMR correlation is high, the IOD-ISMR correlation is low. The IOD, as a modulator of the Indian monsoon rainfall, has significant impact on the correlation between the ISMR and ENSO. According to Ashok et al. (2001), the Indian summer monsoon is affected by the IOD events on their own. Thereby, it seems that IOD declines or reinforces the effect of the ENSO on the ISMR. The effect on ISMR relies on the phase and amplitude of the IOD and ENSO due to presence of positive and negative events in the two major tropical climate phenomena. It is probable that some IOD events could be associated with some ENSO events. However, Ashok et al. (2001)

found out that the approach in which considers the IOD as one of the main coupled modes in the tropics appears to be effective in assessing the effect of IOD on ISMR. Further studies on ISMR was conducted by Ashok et al. (2004). They examined the comparative effects of ENSO and the IOD events on the ISMR through analysing observations and experimental results. During the pure IOD years, pure ENSO years, and co-occurring years, composite analysis of the ISMR anomalies revealed that the influence of the El Nino on the Indian monsoon is significantly decreased by positive IOD. However, negative IOD considerably reduces the impact of the La Nina on the Indian monsoon. This endorses the hypothesis proposed by Ashok et al. (2001). In order to assess the effect of the El Nino, the IOD, and their combined influence on the ISMR and related circulation, numerous multi-ensemble sensitivity experiments were conducted using an AGCM with different types of SST fields as lower boundary forcing. It was found that both poles of the IOD contribute to the excess rainfall over India during the positive IOD event; thereby it decreases the effect of ENSO. Using AGCM experiments, it was also revealed that positive IOD events intensify the ENSO-induced subsidence and rainfall deficiency over the Indonesian region. As noted in Ashok et al. (2001), the net combined effect of these ocean processes relies on the relative phases and it strengths. The IODMI is a potentially suitable predictor for the Indian summer monsoon rainfall due to the fact that the IOD seems to weaken the relationship between ENSO and monsoon.

It is well recognized that SST anomalies in the central-eastern Pacific which are related to the ENSO, act as the main forcing of the rainfall variability. Nevertheless, it is difficult to simulate the aforementioned fact. Chattopadhay et al. (2010) evaluated the relationship between rainfall and SST anomalies for the period of the winter monsoon over India by means of scatter plot matrices and autocorrelation functions. It was revealed that the coefficient of determination for the linear trend was very low even when a six degree polynomial trend was adopted. In order to forecast the average winter monsoon rainfall of a given year an exponential regression equation and an artificial neural network (ANN) were produced. Substantial variables were selected and the rainfall amounts and the SST anomalies in the winter monsoon months of the previous year were chosen as predictors. Levenberg- Marquardt algorithm was used to produce the regression coefficients for the multiple exponential regression equation. The ANN

in the form of a multilayer perceptron with sigmoid non-linearity and genetic-algorithm based variable selection were made. The Willmott's index, percentage error of prediction, and prediction yields were used to evaluate both predictive models statistically and the potential of ANN over exponential regression was demonstrated.

2.3 Effect of Large Scale Climate Modes on Australian Rainfall

The significance of rainfall for sustainable water, agriculture and ecological management is well recognized In Australia; thereby, rainfall is the most significant and most investigated hydro-climate variables. Several research have been carried out to evaluate rainfall mechanism in Australia and the simultaneous relationships between large scale climate modes and Australian rainfall (Nazemosadat & Cordery 1997; Chambers 2003; Murphy & Timbal 2008; Chowdhury & Beecham 2010; Beecham & Chowdhury 2010; Evans et al. 2009). In general, rainfall in Australia has a high degree of spatiotemporal variability (Chowdhury & Beecham 2010) and demonstrates persistence characteristics (Simmonds & Hope 1997) which are affected by several natural climate phenomena originating from the Pacific, Indian and Southern Oceans. SST variability in the Pacific and Indian Ocean, the Southern Annular Mode (SAM) and the Interdecadal Pacific Oscillation are considered as key climate drivers in Australia (Cai et al. 2011; Chowdhury & Beecham 2010; McBride & Nicholls 1983; Drosdowsky 1993; Drosdowsky & Chambers 2001; Hendon et al. 2007; Meneghini et al. 2007; Power et al. 1999). According to Cai et al. (2011), eastern and southern Australian rainfall are considerably influenced by the ENSO and IOD phenomena, respectively. The current research interest in Australia is focused on the relative effects of these climate phenomena on rainfall and their teleconnection pathways which vary spatially within Australia. According to Cai et al. (2011), ENSO affects the Australian climate; nevertheless, it is still unclear whether SST variations of Pacific Ocean in the tropics is accountable for climate variability or that of Indian Ocean. According to Saji et al. (1999), the following two mechanisms are well identified: The development of the Southern Oscillation in the Pacific Ocean influences the lower latitudes of eastern Australia; however, the Indian Ocean SST anomalies affect higher latitudes. The SST in the tropical Indian Ocean and in the tropical Pacific Ocean change relatively during ENSO phenomena (Cai et al. 2011). Chowdhury and Beecham (2013) found that during the El Nino phenomenon winter rainfall is reduced in the western and southern parts.

However, upon removal of covariance with the IOD the effect of El Nino on SA rainfall disappears. (Chowdhury & Beecham 2013). Some of these studies will be discussed in more details.

Operational seasonal prediction schemes in Australia, forecast regional rainfall and temperature probabilities in Australia either based on SOI phases (Stone et al. 1996) or SST variability trend in the Pacific and Indian Ocean (Drosdowsky & Chambers, 2001). Nicholls (1989) found that ENSO affect rainfall over most of Australia. They found that based on a rotated principal component analysis of Australian winter (June-August) rainfall two large-scale variation patterns are responsible for more than half of the total rainfall variance. A broadband which stretches from the northwest to the southeast corners of the country was introduced as the first pattern, while the second pattern was centered in the eastern third of the continent. These two patterns were associated with the SST in the Indian and Pacific oceans. The difference in SST between Indonesian region and the central Indian Ocean is related to the first pattern, while the equatorial Pacific SST is associated to the second rainfall pattern. This relationship indicates the effect of the Southern Oscillation on both SSTs and Australian rainfall. However, the relationship between the difference between Indonesian and central Indian Ocean SSTs and the first rainfall pattern is mainly independent of the Southern Oscillation. The SSTs variation which is relatively distinct from the well-known effect of the Southern Oscillation could be another factor which affects rainfall in Australia (Nicholls 1989).

Equatorial Pacific SSTs directly influences one of the two principal modes of Australian inter-annual rainfall variability which is regionally centered on central eastern Australia. According to McBride and Nicholls (1983), in the winter and spring months rainfall is significantly influenced by ENSO. For the period December 1932 to November 1974 McBride and Nicholls (1983) calculated the correlations between indices of the Southern Oscillation (SO) and areal average rainfall for 107 Australian rainfall regions. According to simultaneous correlations between the SO and rainfall, there was a clear annual cycle in which the best relationship happened in spring (September-November). Summer (December-February) had the weakest relationship. They found that in some parts of Australia seasonal rainfalls in all seasons were considerably correlated with the SO in the preceding season. The highest lag correlation occurred with spring rainfall. For some regions this correlation was also significant with the SO two seasons (six

months) earlier. McBride and Nicholls (1983) also calculated the correlations with the data divided into two subseries from 1932 to 1953 and from 1954 to 1974. It was found that there was a westward shift with time of the correlation pattern. This was associated with considerable changes in the magnitude of the correlations in some regions.

In both the observations and during June–December in the coupled general circulation model (CGCM) Power et al. (2006) found that the relationship of all-Australia rainfall and temperature with ENSO as measured by Nino4 or the SOI was found to be nonlinear which is responsible for rainfall and temperature changes over Australia. Australia usually becomes much wetter with a large La Nina SST anomaly or a large La Nina SOI excursion. However, the magnitude of an El Nino SST or SOI anomaly is not a good indicator of how dry Australia will essentially become. Although Australia certainly dries out during El Nino events; however, the degree of drying is not significantly associated with the magnitude of the El Nino SST anomaly. With respect to climate prediction, the surprising results of Power et al. (2006) has significant implications. In other words, for many parts of the world where the response is linear the magnitude of El Nino SST anomalies has considerable impact on climate predictions. However, in Australia the magnitude of El Nino SST anomalies could be less influential as a large El Nino SST signal does not increase the risk of severe drought (at least in terms of continental average rainfall) compared to the risk associated with a more modest El Nino SST anomaly (Power et al. 2006). The "El Nino of the century" during 1997–1998 had only a weak effect on SEA rainfall, while in the east Pacific Ocean it had very large SST anomalies. On the other hand, the worst drought in Australian's recorded history was possibly the relatively weak event occurred during 2002–2003, as temperatures were very warm and rainfall was recorded low (Watkins, 2002). According to Wang and Hendon (2007), during El Nino Australia usually experiences drought particularly during austral spring (September–November) across the eastern two-thirds of the continent. Nevertheless, there were some exceptions from this pattern. For instance, the near-record-strength El Nino during 1997 was linked with near-normal rainfall. On the contrary, during the modest El Nino in 2002 eastern Australia experienced near-record drought. This obvious contrast brings the attention to the issue that how the magnitude of the drought is associated with the magnitude and character of El Nino as measured by the broad-scale SST anomaly in the equatorial

eastern Pacific. One of the underlying reasons for this contrasting behaviour during these El Nino events is the internal and unpredictable atmospheric noise. According to Wang and Hendon (2007), rainfall in Australia is sensitive to the zonal distribution of SST anomalies during El Nino. Specifically, the highest sensitivity is to the SST variations on the eastern edge of the Pacific warm pool compared to that of the eastern Pacific where there are usually largest El Nino variations. In 1997 maximum anomalies were moved well into the eastern Pacific; thereby having less effect on rainfall in Australia. On the other hand, positive SST anomalies increased near the date line in 2002. They revealed that these results offer a possible physical basis that predicting the strength of El Nino is not adequate to precisely forecast rainfall variations across Australia. Murphy and Timbal (2008) also mentioned that the decade-long rainfall deficiencies seem to be independent of the ENSO-rainfall influence. A little beyond average rainfall during the 1998–1999 La Nina did little to relieve long-term deficiencies in southeast Australia. It was also understood that the main fall in rainfall has happened in autumn when the influence of ENSO on rainfall is weak in southeast Australia. Furthermore, although there is a relationship between Indian Ocean SSTs and rainfall in southeast Australia; however, it is not evident whether rainfall of southeast Australia is truly a response to SST forcing, mainly in autumn.

Verdon and Franks (2005) investigated the relationship between SST variability happening over the Indian Ocean and winter rainfall variability in eastern Australia. They compared Six indices of SST variability and determined their relationship to rainfall over eastern Australia. It was found that there is a strong relationship between a number of these indices and winter rainfall through an analysis of historical rainfall data for Queensland, New South Wales, and Victoria. Particularly, a good indication of winter rainfall variability in eastern Australia can be provided through anomalous SSTs over the Indonesian area. Further, it was shown that regardless of potential influences on rainfall by ENSO, this relationship is true. Verdon and Franks (2005) proposed a probable physical process in which the Indian Ocean SST anomalies might affect winter rainfall. This possible process involves the impact of the Indian Ocean SST anomalies on the nature of the northwest Australian cloud band. Verdon and Franks (2005)'s study revealed noticeable controls on winter climate variability similar to that induced in

summer by the better known ENSO processes, thereby, provides enhanced understanding of year-round seasonal climates.

Changes in the magnitude and spatial extent of inter-annual variations in Australian wheat yield was also found by Potgieter et al. (2005). They discovered that in between El Nino events this variation is significant and it is related to variation in rainfall with other variables based on the timing and location of SST anomalies. Power et al. (2006) further indicated that a simple proof of changes of rainfall in Australia by ENSO is that the east of the continent has the tendency to have rainfall in the lowest tercile virtually everywhere along the coast east of the Great Dividing Range. However, rainfall over the same area tends to be in the upper tercile during La Nina years. The mean El Nino response is weaker in autumn and summer. The pattern toward a drier southeast Australia (SEA) climate is possibly not associated with ENSO variations due to limited influence of ENSO on SEA in autumn. Nicholls et al. (1997) found that the strength of the SOI-rainfall link varies significantly with time. It was also found that the SOI–rainfall relationship had the greatest change in the southeast after the early 1970s. Power et al. (2006) also found that the effect of ENSO on rainfall in Australia changes significantly on inter-decadal time scales. In both the model and the observations the relationship between ENSO and climate of Australia as measured by correlation coefficients is strong in some decades, while it is weak in other decades. It was shown by a series of decadal-long perturbation experiments with the coupled general circulation model (CGCM) that the level of predictability is low. According to Mantua et al. (1997), the Interdecadal Pacific Oscillation (IPO) is an interdecadal El Nino–like SST trend which is closely associated with the interdecadal component of an index for the Pacific Decadal Oscillation (PDO). It was demonstrated by Power et al. (1999) that during the twentieth century an index for the PDO, IPO, is statistically connected with variations in ENSO's influence on the climate in Australia. Power et. al (1999) revealed that the influence of ENSO on Australia is increased when the IPO is in a negative phase. Power et al. (2006) used a simple nonlinear stochastic model to show that, at least in theory even in the case where it is not possible to forecast more than 1 year, the interdecadal excursions in ENSO indices will have the tendency to be statistically coherent with interdecadal variations in ENSO teleconnections measures. In Power et al. (2006)'s model, the interdecadal component of ENSO variability demonstrates a

residual ENSO-like SST trend due to random ENSO variations. For instance, a given interdecadal period might include a larger number of El Nino events compared to La Nina events. Or if for a given interdecadal period frequency of the El Nino and La Nina events is identical, then the SST anomalies of the El Nino events might have been larger compared to the SST anomalies related with the La Nina events that happened in the same interdecadal period. Obviously, this also implies that indices similar to low frequency excursions in ENSO indices seem to change ENSO teleconnections without the necessity of predictability of more than 1 year. When the impact of ENSO on Australia is usually greatest, the IPO of CGCM is statistically associated with interdecadal variations in the effect of ENSO on Australia during the period of June to December. On the other hand, in both the model and the observations IPO indices estimate interdecadal changes in Nino4. Thus, though the IPO or PDO just reflects random variations in ENSO statistics on interdecadal time scales they can appear to change ENSO teleconnections in the presence of nonlinear teleconnections. Although modulations can happen without nonlinearity, however, the chance of changes occurrence will be increased by nonlinearity. It should be pointed out that although in the Australian context nonlinearity is an significant characteristic; however, nonlinearity is not essentially of great importance for the appearance of clear changes of ENSO teleconnections by the IPO or PDO in all areas. In the linear case, during IPO positive phases, the influences of El Nino will be increased, while the effects of La Nina will be weakened. On the other hand, during IPO negative phases effects of La Nina will be improved, while the influence of El Nino will be decreased. Moreover, different forms of nonlinearity can cause increase in different forms of obvious modulation. For instance, IPO positive phases will increase the influences of ENSO in a particular area, if the El Nino SST anomalies compared to La Nina SST anomalies were closely associated with changes in that region. Nonlinear ENSO teleconnections are not limited to Australia. This pattern can be seen, as an important example, in the nonlinear nature of ENSO teleconnections in the southwestern United States and northern Mexico. In this area the magnitudes of La Nina SST anomalies are apparently less associated with the rainfall anomalies' magnitudes compared to those of the El Nino SST anomalies. Although the risk of dry conditions will be increased by La Nina SSTs; however, the degree of drying is not strongly associated with the SST anomaly's magnitude. Thereby, the above conclusions is relevant in this area as well (Power et al. 2006).

The rainfall probabilities produced by the Australian Bureau of Meteorology's operational seasonal prediction scheme also demonstrates the limited effect of ENSO in autumn (Drosdowsky & Chambers 2001). According to Murphy and Timbal (2008), the loading on the second mode in SSTs used, do regulate southeastern Australia rainfall probabilities in autumn. However, the loadings of the first mode of SST variability do little to regulate the rainfall probabilities for south east Australia in this season. The loadings of the first mode of SST variability reflect El Nino-like SST anomalies in the Pacific Ocean, while the loading on the second mode in SSTs used corresponds to Indian Ocean SST variability. Therefore, SST variability in the Indian Ocean has a superior influence on SEA rainfall in autumn. According to Saji et al. (1999), equatorial SSTs also illustrate Indian Ocean SST variability. As discussed earlier Indian Ocean Dipole (IOD) is defined as the development of cold SSTs in the eastern Indian Ocean near Indonesia and warm SSTs in the west which causes droughts over Indonesia and heavy rains in eastern Africa. The IOD also has influence on Australian rainfall. For the 6 years of extreme positive IOD events corresponding to cool eastern Indian Ocean SSTs Saji et al. (1999) identified rainfall deciles for the March–November period; it was revealed that in south east Australia, virtually all of western and central Victoria showed the maximum rainfall deficiencies corresponding to very much below average rainfalls. However, most of southern Australia with the exception of the east coast exhibited below average rainfall. These results are in good agreement with those of Ashok et al. (2003).

Ashok et al. (2003) studied the effect of the IOD on the Australian winter rainfall using an atmospheric general circulation model and observed datasets of SST and rainfall. It was found that over the western and southern regions of Australia the IOD has substantial negative partial correlations with rainfall which extend south-eastward from Indonesia all the way to south east Australia. According to Ashok et al. (2003)'s atmospheric general circulation model sensitivity experiments, during the positive IOD events cold SST anomalies dominate west of the Indonesian archipelago which introduces an anomalous anti-cyclonic circulation over much of the Australian continent and at lower levels over the eastern tropical and subtropical Indian Ocean. Moreover, they discovered that in this region the response of the atmosphere to the IOD is

baroclinic. Over the affected areas of Australia this baroclinic response leads to anomalous subsidence and anomalous decrease in the rainfall.

In further studies comprehensive classification methods were used by Meyers et al. (2007) to identify IOD positive, negative and neutral years and similar ENSO classes and the response of rainfall in each group over Australia. The method proposed by Meyers et al. (2007) identified when the positive or negative extrema of the El Nino Southern Oscillation and IOD happen. Each year from 1876 to 1999 was classified using this method. Although this method is statistical in nature; however, it is strongly based on the oceanic physical mechanisms which control the variability of the near-equatorial Indo-Pacific basin. Meyers et al. (2007) found that some years could not be evidently categorized as a result of strong decadal variation. Nevertheless, these years must be identified and the reason for their ambiguity must be recognized. Meyers et al. (2007) tested the sensitivity of the years classification through calculating composite maps of the Indo-Pacific SST anomaly and the probability of below median Australian rainfall for dissimilar groups of the El Nino–Indian Ocean relationship. Cai et al. (2009) showed a slight different classification; according to Meyers et al., (2007) classification, three (1972, 1982 and 1997) of the six years used in Cai et al. (2009) calculations were also El Nino years and three (1961, 1967 and 1994) were not. Nevertheless, other researchers have categorized 1994 as an El Nino year. It was found that 1982 was the driest year on record for southeastern Australia. The second driest year on record for southeastern Australia was 1967. Although there are some dissimilarities in the response of rainfall for these two subsets (i.e. El Nino and non-El Nino); however, the demonstrated trend is very similar over most of southeastern Australia for both of them with the exception of the far northeast of the region. Through existing observations and re-analyses, Cai et al. (2009) showed that the positive IOD (pIOD) events increase from about four per 30 years early in the 20th century to about 10 over the last 30 years. On the other hand, the number of negative Indian Ocean Dipole (nIOD) events drops from about 10 to two over the same periods, respectively. A systematic trend in this parameter can be seen commencing early in the 20th century using a skewness measure which is defined as the difference in happenings of pIODs and nIODs. It was found that there exist more pIODs than nIODs, with consistent mean circulation variations in the pIOD-prevalent seasons after 1950. According to Cai et al. (2009), these changes

potentially explain much of the observed austral winter and spring rainfall decrease since 1950 over southeastern Australia. Moreover, These features are in good agreement with expected future climate change and therefore with what is projected from global warming. (Cai et al. 2009)

Ummenhofer et al. (2009) also examined the effect of the phases of ENSO and IOD on southeast Australia. According to Ummenhofer et al. (2009) over southeast Australia the sign of the rainfall anomalies is very inconsistent during pure La Nina and positive IOD events. However, El Nino years consistently lead to dry conditions and negative IOD years consistently result in wet conditions. Further, an absence of negative IOD events was seen throughout most of the multi-year droughts of the 20th century, and especially was a noticeable feature of the present Big Dry. However, the same is not true about ENSO, with both negative and positive phases happening during all of the major 20th century droughts. Ummenhofer et al. (2009) demonstrated that over the past 120 years in the region of southeastern Australia, Indian Ocean variability more than ENSO is the key driver of the main droughts. Particularly the IOD has remained consistently 'positive' or 'neutral' during almost all of Australia's iconic droughts, comprising the Federation Drought (1895–1902), the World War II drought (1937–1945), and the present ''Big Dry'' (post-1995). An interaction between the tropics and the temperate zone increases regional moisture advection. Thereby unexpectedly wet conditions dominate across southern regions of Australia during the IOD negative phase. Therefore, during the major droughts the noticeable lack of the ''negative'' phase of the IOD prevents normal rainfall quota of southeast Australia. Although, the Indian Ocean has a noticeable role in driving southeastern droughts; however, the ''Big Dry'' has still an outstanding severity which seems to be related to recent large increases in air temperature (Ummenhofer et al. 2009).

Further, it was revealed that climatic influence of ENSO on middle latitudes west of the western Pacific, for example southeast Australia, during austral spring (i.e. September–November) is conducted through the tropical Indian Ocean (TIO) (Cai et al. 2012). Nevertheless, it is not clear whether this pathway is symmetric in regards to the positive and negative phases of ENSO and the IOD. Cai et al. (2012) showed in their study that a strong asymmetry exists. In regards to ENSO, only the influence of El Nino is conducted through the TIO pathway. Moreover, the effect of La Nina was

delivered through the Pacific–South America pattern. Regarding the IOD, a greater convection anomaly and wave train response happens during positive IOD (pIOD) events compared to the negative IOD (nIOD) events. This impact asymmetry is in agreement with the positive skewness of the IOD. This is mainly attributed to a negative skewness of SST anomalies in the east IOD (IODE) pole. According to Cai et al. (2012), convection anomalies in the IODE region are more sensitive to a per unit change of cold SST anomalies compared to the same unit change of warm SST anomalies. The study of Cai et al. (2012) demonstrated that despite the greater damping, the IOD skewness happens due to a breakdown of this damping as recommended by previous studies. Much of the spring rainfall decrease over southeast Australia during the 2000s can be explained via this IOD impact asymmetry. They concluded that the main cause of this decrease in rainfall is the increased happenings of pIOD events, rather than the lack of nIOD events (Cai et al. 2012).

Gallant et al. (2012) questions the independence of ENSO and IOD. They argue that Indian Ocean SSTs, in particular those in the far eastern tropics, undoubtedly have a relationship with rainfall in southeast Australia. However, it is not evident whether these SSTs actually lead to the rainfall anomalies. In general in a given year the IOD forms from May onwards, yet in early autumn of IOD positive years strong rainfall shortages are already evident over southeast Australia, which is before the IOD forms. Hence, the IOD may be a response to the same atmospheric forcing as the southeast Australia rainfall anomalies, so that both may be originated by a third, preceding mechanism. This claim is yet to be further investigated.

Other than southeast Australia, southwest Australia is also affected by IOD. High-pressure anomalies over Australia in which result in low rainfall in southwest Western Australia (SWWA) produce anomalous easterly winds over the eastern Indian Ocean which cause upwelling and, thereby, cool SSTs (England et al. 2006). Inter-annual rainfall extremes over SWWA through observations, reanalysis data, and a long-term natural integration of the global coupled climate system was investigated by England et al. (2006). A characteristic dipole pattern of Indian Ocean SST anomalies was revealed during extreme rainfall years. Although this pattern was surprisingly consistent between the reanalysis fields and the coupled climate model, however, it was different from most previous SST dipoles definitions in the area. Particularly, the dipole demonstrates peak

amplitudes in the eastern Indian Ocean next to the west coast of Australia. Anomalously cool waters appear during dry years in the tropical/subtropical eastern Indian Ocean which are close to an area of unusually warm water in the subtropics off SWWA. The sign of this dipole of anomalous SST keeps changing between dry and wet years. This dipole seems to happen in phase with a large-scale reorganization of winds over the tropical/subtropical Indian Ocean. The wind field changes SST by anomalous air–sea heat fluxes in the subtropics and by anomalous Ekman transport in the tropical Indian Ocean. The large-scale advection of moisture onto the SWWA coast is also altered by the winds. At the basin scale, the anomalous wind field can be understood as an acceleration of the Indian Ocean climatological mean anticyclone during dry years. Similarly, the anomalous wind field can also be inferred as deceleration of the Indian Ocean climatological mean anticyclone during wet years. Furthermore, they discovered that dry (wet) years see a strengthening (weakening) and coinciding southward (northward) shift of the sub-polar westerlies causing a comparable southward (northward) shift of the rain bearing fronts linked with the sub-polar front. England et al. (2006) further revealed that there is also a relationship between extreme rainfall years and the IOD. In other words, in some years the IOD acts to strengthen the eastern tropical pole of SST explained above, and to reinforce wind anomalies along the northern flank of the Indian Ocean anticyclone. In this way, in the Indian Ocean both tropical and extra-tropical processes produce SST and wind anomalies off SWWA that result in moisture transport and rainfall extremes in the area. An analysis of the seasonal evolution of the climate extremes demonstrated a progressive anomalies increase in SST and atmospheric circulation toward a wintertime maximum, which correspond to the season of highest SWWA rainfall. The SST anomalies can appear as early as the summertime months which may have significant effects on forecasting of SWWA rainfall extremes.

Chowdhury and Beecham (2013) investigated the effect of climate drivers such as SOI, DMI and Nino3.4 on South Australian rainfall. They analyzed recent records of monthly rainfall and climate driver index values from 1981 to 2010 for 53 rainfall stations. These stations were located across eight South Australian natural resources management (NRM) regions. They applied the Pearson, Kendall and Spearman correlation tests between rainfall and climate drivers and between the climate drivers

themselves. It was found that neither South Australian summer (December to February) nor autumn (March to May) rainfalls were considerably affected by climate indices. It was found that in the south and east parts of South Australia winter rainfall specifically in July and August was significantly influenced by both SOI and DMI. Moreover, they found that in winter both SOI and DMI were inter-correlated. Spring rainfall specifically in September and October was found considerably affected by DMI in the south and east parts of South Australia. With regards to ENSO phenomena, it was found that while both SOI and Nino3.4 were correlated; however, for South Australian winter and spring rainfall SOI was more effective on the region. The authors indicated that the results of the study are beneficial for producing stochastic rainfall and for developing downscaling techniques to produce rainfall projections in the region.

Recently it has been demonstrated that the Southern Hemisphere annular mode (SAM) is associated with Australian rainfall. According to Thompson and Wallace (2000), the SAM as the main mode of large scale variability in the southern hemisphere extra-tropical circulation, is a regionally symmetric varying mass exchange between the polar regions and the mid-latitudes. During the high phase of the SAM index corresponding to a southward contraction of the mid-latitude storm track, daily rainfall and surface temperatures over Australia were compared to that during the low phase corresponding to equator-ward expansion of the storm track. Using observations for the period 1979–2005, daily changes in Australian rainfall and surface temperature linked with the SAM were documented by Hendon et al. (2007). A poleward contraction of the mid-latitude westerlies categorizes the high index polarity of the SAM. During winter, the high index polarity of the SAM is linked with reduced daily rainfall over southeast and southwest Australia. However, during summer it is connected with intensified daily rainfall on the southern east coast of Australia and reduced rainfall in western Tasmania. Up to 15% of the weekly rainfall variance in these regions can be described by the changes in the SAM, which especially during winter is similar to the variance accounted for by the ENSO. The most prevalent temperature anomalies linked with the SAM take place during the spring and summer seasons, when the high index polarity of the SAM is related with anomalously low maximum temperature over most of central/eastern subtropical Australia. The areas of reduced maximum temperature are also linked with increased rainfall. Gillett et al. (2006) discussed possible effects of recent trends in

Australian rainfall and temperature. A positive pattern in the SAM index for summer was found that can consider up to half the observed positive rainfall pattern in southeast Australia in summer. Although there is a trend toward the positive phase in autumn, however, absence of relationship between the SAM and autumn rainfall in the southeast prevents any direct attribution. It was also found that during the positive phase of the SAM the monthly means of both rainfall and temperature in southeast Australia are below average. Previous studies on the effect of the SAM on surface climate have focused mostly on individual countries. In this study station observations of temperature and rainfall were used to recognize the effect of the SAM on land regions over the whole of the Southern Hemisphere. It was demonstrated that the positive phase of the SAM is related to a significant cooling over Antarctica and much of Australia, as well as a significant warming over the Antarctic Peninsula, Argentina, Tasmania and the south of New Zealand. Due to the southward shift of the storm track, the positive phase of the SAM is also connected with anomalously dry conditions over southern South America, New Zealand and Tasmania; it is also associated with anomalously wet conditions over much of Australia and South Africa. These effects on populated areas of the Southern Hemisphere may have effects on weather and seasonal forecasting and future climate change (Hendon et al. 2007).

Other researchers have also found strong trends in the SAM index towards higher values (i.e. more southerly storm track) in summer and autumn. It was demonstrated that the resulting changes to the circulation are commonly confined to the south polar areas and they attributed the trend bulk to variations in stratospheric ozone (Thompson and Solomon 2002). Evidence was presented that recent patterns in the Southern Hemisphere tropospheric circulation can be taken as a bias toward the high-index polarity of this trend, with stronger westerly flow encircling the polar cap. It was discovered that the largest and most considerable tropospheric patterns can be traced to recent patterns in the lower stratospheric polar vortex, that are mainly attributed to photochemical ozone losses. The pattern toward stronger circumpolar flow during the summer-fall season has contributed considerably to the observed warming over the Antarctic Peninsula and Patagonia and to the cooling over eastern Antarctica and the Antarctic plateau. Although Marshall et al. (2004) also found comparable trends; however, they concluded that increase of greenhouse gas has also played a significant

role. It was revealed that recent observed patterns in the annual and austral summer SAM are improbable to be attributed to internal climate variability as they exceed any equivalent-length patterns in a millennial General Circulation Model (GCM) control run with constant forcings. However, It was demonstrated that observed patterns in the SAM are consistent with the combined influences of anthropogenic and natural forcings in GCM simulations. The assertion that this process is mainly responsible for changes in the SAM was challenged since these patterns originate earlier than stratospheric ozone depletion. Furthermore, anthropogenic forcings have a larger influence on the austral summer SAM combined with natural forcings compared with when acting in isolation.

Meneghini et al. (2007) used an index of the pressure difference between 40oS and 65oS which was only covering the Australian area (90^{o}–180^{o}E) and was comparable to SAM index. It was found that in winter this index was more closely associated with rainfall than the SOI in the south of the country. Furthermore, in summer this index was more closely related with rainfall in the southeast Australia north of the Great Dividing Range. However, in autumn little relationship was found when the trend in their regional index was strongest. Their regional index was not inevitably directly similar to the hemispheric SAM index and was claimed to be closely aligned with the latitude of the Australian subtropical ridge (STR). In more details, Meneghini et al. (2007) explored the relationships between seasonal Australian rainfall and the SAM. Two seasonal indices of the SAM comprising the Antarctic Oscillation Index (AOI), and an Australian regional version (AOIR) were produced using ERA-40 mean sea-level pressure (MSLP) reanalysis data. The seasonal rainfall data were supplied by the Australian Bureau of Meteorology based on gridded monthly rainfall. An important inverse relationship was found between the SAM and rainfall in southern Australia for the period 1958-2002. However, in northern Australia a significant in-phase relationship was found between the SAM and rainfall. Moreover, in winter in southern Australia widespread significant inverse relationships were only observed with the AOIR. In comparison with the SOI, the AOIR is responsible for more of winter rainfall variation in southwest Western Australia, southern South Australia, western and southern Victoria, and western Tasmania. According to the Meneghini et al. (2007)'s results, the changes in SAM affect southern South Australia, Victoria, and Tasmania and may be partially accountable for the current decline in winter rainfall for these regions.

However, it is not responsible for the long-term decline in southwest Western Australian winter rainfall.

2.4 Seasonal Rainfall Forecasting in Australia

The seasonal rainfall predictions and streamflow forecasts are valuable for management of land and water resources, specifically in Australia since the streamflow variability is higher compared to most parts of the world (Chiew et al. 1998). Seasonal rainfall predictions are valuable for users such as irrigators and water managers to aid in developing risk-management strategies and to inform decisions. Both statistical and dynamical climate forecast systems are broadly used in practice to generate seasonal rainfall predictions up to a year in advance (Goddard et al. 2001). Statistical forecast systems are based on empirical relationships between observed variables and thereby depend on the availability of long data records and stationary relationships between the variables. On the other hand, dynamical forecast systems are based on numerical simulations which directly model physical processes, however they are more expensive to implement and operate compared to statistical climate forecast systems (Anderson et al. 1999; Schepen et al. 2012). In spite of considerable research studies and technological advances, complicated dynamical forecast systems are still not capable to consistently outperform simple statistical prediction systems for forecasting ENSO and other climate variables (Barnston et al. 1999; Halide & Ridd 2008; Quan et al. 2006). Until dynamical forecast systems progress considerably, statistical forecast systems will continue to be enhanced and play a role in seasonal rainfall estimation (Rajeevan et al. 2007; Schepen et al. 2012).

As discussed in section 2.3 several studies have tried to explain the relationships between oceanic and atmospheric circulation anomalies and Australian monthly or seasonal rainfalls (Meneghini et al. 2007; Murphy & Timbal 2008; Risbey et al. 2009; Wang & Hendon 2007). Simultaneous relationships are usually quantified through linear regression analyses between rainfalls and a variety of climate indices which represent anomalies in climate variables such as SST, upper ocean heat content, atmospheric pressure, and zonal wind. The strengths of the relationships change with season and location. For estimating next seasonal rainfall, climate indices which have a strong simultaneous relationship with seasonal rainfall are natural candidates to be used

as predictors in a statistical forecast system. However, for the purpose of prediction it is the lagged relationships between seasonal rainfall and climate indices that are of particular importance. A strong simultaneous relationship does not essentially result in a strong lagged relationship (Schepen et al. 2012). Hence, quantifying the evidence which supports the use of various lagged climate indices for seasonal rainfall predictions is unavoidable. Previous research studies have shown the usefulness of lagged indices of the Southern Oscillation for predicting Australian seasonal rainfall in some areas and seasons (Chiew et al. 1998; McBride & Nicholls 1983; Stone et al. 1996). McBride and Nicholls, (1983) found that in all seasons, in some parts of Australia seasonal rainfalls are considerably correlated with the SO in the preceding season. Based on the reported results the strongest lag correlations happen with spring rainfall, which for some regions is also considerably correlated with the SO previous two seasons (former six months). Drosdowsky (1993) examined the lagged relationships between winter rainfall over Australia and Indian Ocean SST anomalies. It was found that it was feasible to predict early winter (April to July) rainfall over parts of southern and eastern Australia from summer to early autumn (December to March) by using an index of SST anomalies in a region off the west Australian coast. Comparable results were found with the anomalous pressure gradient between the west coast and central Indian Ocean, which aids as an estimation to the meridional component of the geo-strophic wind. It was suggested that the SST and rainfall anomalies were forced by the anomalous atmospheric circulation.

Cheiw et al. (1998) presented a summary of the relationship between ENSO and rainfall, drought and stream flow in Australia. The teleconnection between ENSO and the hydroclimate of Australia was studied by evaluating the lag correlations between rainfall and streamflow and the ENSO indicators of several former months. According to the analyses, it was found that in Australia dry conditions are associated with El Nino. Moreover, it was revealed that in most parts of Australia the relationship between rainfall and streamflow and ENSO was statistically important, while it was not adequately strong to consistently and precisely forecast rainfall and streamflow. It was found that the teleconnection was stronger in the second part of the year, and spring rainfall in eastern Australia as well as summer rainfall in north-east Australia can be estimated with some success several months in advance using the ENSO indicators. It was indicated that the indicators of ENSO can similarly be used to assist in estimating

spring runoff in south-east Australia and summer runoff in north-east and east coasts of Australia. According to the analysis, it was revealed that in the streamflow data the serial correlation, unlike rainfall, is usually comparable or higher than the lag stream flow-ENSO correlation. Furthermore, it was noted that the former months' streamflow must be used together with the indicators of ENSO to develop stream flow prediction models.

For the prediction of Australian seasonal rainfall changes an operational system was introduced by Drosdowsky and Chambers (2001) which used SST anomaly pattern over the Indian and Pacific Oceans. Rotated principal components with individual monthly values at 1- and 3-month lead times as predictors represented the SST anomalies. For instance, in order to predict March to May seasonal rainfall, November and January SST anomalies were used. Rotated principal components of a gridded 18 rainfall dataset were also used to represent the historical seasonal rainfall with using the principal component loadings as weights to project the predictions back to the original 18 grid points. Linear discriminant analysis were used to produce the predictions of seasonal rainfall in two (above/below median) or three categories (terciles). The linear error in probability space (LEPS) skill score was used to measure hindcast skill which was assessed by cross validation. To select the best model or the best combination of predictors, experiments were also carried out by using a double or nested cross validation procedure. To maintain continuity of estimation probabilities between the overlapping 3-month seasons for every season and location, the first two rotated SSTA components lagged by 1 and 3 months as predictors were used by the model selected for operational seasonal forecasts. SST analysis of either the Bureau of Meteorology's or the National Centers for Environmental Prediction was projected onto the set of SST principal components to calculate current values of the principal component amplitudes. According to Drosdowsky and Chambers (2001), over the 5-yr period from January–March 1994 to December–February 1998/99 the hindcasts and experimental real-time predictions using the SST-based schemes demonstrate enhanced skill over parts of southern Australia during the autumn period compared to forecasts using the SOI alone.

Moreover, correlations between several other climate indices averaged over the preceding two months and Australian seasonal rainfalls for the four main seasons were investigated by Kirono et al. (2010). Based on the season and location, statistical

significant correlations were found between the lagged climate indices and seasonal rainfalls using Pearson correlation coefficient. Nevertheless, in depth quantitative results were only presented for southeast Australia. The lag relationships between rainfall across Australia and runoff across southeast Australia versus 12 atmospheric-oceanic predictors and how the relationships vary over time were investigated in the work of Kirono et al. (2010). Based on the rainfall data analysis, it was found that the greatest relationship is in spring and summer in northeast Australia and in spring in southeast Australia. It was found that regarding spring rainfall in eastern Australia the best predictors are Nino4 (SST in western Pacific) and thermocline (20^{o}C isotherm of the Pacific). With regards to summer rainfall in northeast Australia, the best predictors were proved to be Nino4 and SOI (pressure difference between Tahiti and Darwin). In northern Australia the greatest relationship is in spring and autumn with Nino4 as the best predictor. The relationship is significant in summer in western Australia, where SST2 (i.e. SST over the Indian Ocean) and II (SST over the Indonesian region) is the best predictor in the southwest and northwest, respectively. According to the runoff analysis across southeast Australia, the greatest predictability of runoff in the southern parts is in winter and spring, with antecedent runoff as the best predictor. The relationship between spring runoff and Nino4, thermocline and SOI is also relatively high which can be used together with antecedent runoff for spring runoff prediction. The study revealed that the atmospheric-oceanic variables in the northern parts of southeast Australia are superior runoff predictors to antecedent runoff. They also have considerable correlation with winter, spring and summer runoff. The runoff serial correlation is reduced for longer lead times particularly over the northern regions. Moreover, the atmospheric-oceanic variables are possibly superior predictors for runoff prediction. The correlations between runoff versus the predictors change with time, this has implication in developing prediction models which presume stationary in the historical data.

As can be seen from the reviewed studies, most seasonal rainfall forecasts with the use of antecedent months' climate indices as predictors use linear methods. Among the few nonlinear seasonal rainfall forecast available is the work of Schepen et al. (2012). Schepen et al. (2012) applied a rigorous Bayesian joint probability modelling approach to find the cross-validation predictive densities of gridded Australian seasonal rainfall

totals by means of lagged climate indices as predictors over the period of 1950-2009. They quantified the evidence which supports the use of each climate index as a predictor of seasonal rainfall by means of the pseudo-Bayes factor based on cross-validation predictive densities. The results revealed that the use of climate indices from the Pacific region was significantly supported by these evidences. However, there was weaker, but positive, evidence for the use of climate indices from the Indian region and the extra-tropical region. For each climate index they mapped and compared the spatial structure as well as seasonal variation of the evidence. With regards to seasonal variation, the strongest evidence was found from August–October to November–January. However, the weakest evidence was found from March–May to May–July. In terms of spatial structure, the strongest supporting evidence was found for prediction in northern and eastern Australia. Nevertheless, in some areas and seasons there was little evidence which supports the use of climate indices for estimating seasonal rainfall. It was found that climate indices resulting from SST anomalies in the Pacific region exhibit stronger relationship with Australian seasonal rainfall totals compared to climate indices resulted from SST anomalies in the Indian region. If climate indices resulting from atmospheric variables represent the large scale circulation, they are also strongly supported. According to Schepen et al. (2012), many climate indices demonstrate comparable supporting evidence for predicting Australian seasonal rainfall which result in the view of combining climate indices in multiple predictor models and/or model averaging.

Combining predictions from multiple models has the capacity to merge the strengths of individual models and to better represent prediction uncertainty compared to the use of a single model. Wang et al. (2012) developed a Bayesian model averaging (BMA) method for combining forecasts from multiple models which gives greater weights to better performing models. The objective of this study was to develop a BMA method that had the capacity to produce relatively stable weights in the presence of significant sampling variability, which result in robust predictions for future events. The BMA method was applied to combine predictions from multiple statistical models for seasonal rainfall predictions over Australia. Climate indices were used as predictors. It was revealed that the fully combined forecasts successfully merge the best models skills to maximize the spatial coverage of positive skill. Although the skill was generally low for

the first half of the year; however, it was more positive for the second half of the year. It was found that models in the Indian and extra-tropical groups produce suitable and occasionally different skills, while models in the Pacific group produce the most skill. Prediction uncertainty spread was found to be reliably presented through the fully combined probabilistic predictions. It was also found that when forecast lead time was increased from 0 to 1 month, the prediction skill held well. According to Wang et al. (2012), the BMA method was superior to the approach of using a model with two fixed predictors chosen a priori. The BMA method was also better than the approach of selecting the best model based on predictive performance.

Abbot and Marohasy (2012) reviewed the application of artificial neural networks to rainfall predictions. By inputting recognized climate indices, monthly historical rainfall data, and atmospheric temperatures into a prototype stand-alone, dynamic, recurrent, time-delay, artificial neural network, Abbot and Marohasy (2012) evaluated the application of artificial intelligence to monthly and seasonal rainfall prediction in Queensland, Australia. Monthly rainfall predictions 3 months in advance for the period 1993 to 2009 were selected as outputs. The outputs were compared with observed rainfall data in terms of time-series plots, root mean squared error (RMSE), and Pearson correlation coefficients. Comparison of RMSE values of the model with forecasts generated by Predictive Ocean Atmosphere Model for Australia (POAMA) revealed that the model achieved a lower RMSE for 16 of the 17 sites compared. POAMA is the general circulation model (GCM) used by Australian Bureau of Meteorology. Abbot and Marohasy (2012) also considered the preliminary model design with possible considerable enhancement such as inclusion of output from GCMs and experimentation with other input attributes. Later in 2014, Abbot and Marohasy used artificial neural networks to evaluate the utility of climate indices with regards to their capacity to estimate rainfall as a continuous variable. The significance of the Inter-decadal Pacific Oscillation as an index which have never been used in the official seasonal predictions for Queensland was highlighted through the ANN results. Before this study, the official seasonal forecasts for Queensland were based on statistical models. For three geographically different areas within Queensland, the ANN estimations were exhibited to be superior to predictions from the POAMA model in terms of lower Root Mean Square Errors (RMSE), Mean Absolute Error (MAE) and Correlation Coefficients (r).

Mekanik et al. (2013) investigated the application of Artificial Neural Networks (ANN) and Multiple Linear Regression analysis (MLR) to forecast seasonal spring rainfall in Victoria, Australia ENSO and IOD as potential predictors. The use of dual (combined lagged ENSO-IOD) input sets for calibrating and validating ANN and MLR Models was proposed to investigate the simultaneous effect of past values of these two major climate modes on long-term spring rainfall prediction. The MLR models that did not violate the limits of statistical significance and multicollinearity were selected for future spring rainfall forecast. The ANN was developed in the form of multilayer perceptron using Levenberg-Marquardt algorithm. Both MLR and ANN modelling were assessed statistically using mean square error (MSE), mean absolute error (MAE), Pearson correlation (r) and Willmott index of agreement (d).The developed MR and ANN models were tested on out-of-sample test sets; the MLR models showed very poor generalization ability for east Victoria with correlation coefficients of -0.99~ -0.90 compared to ANN with correlation coefficients of 0.42~0.93; ANN models also showed better generalization ability for central and west Victoria with correlation coefficients of 0.68~0.85 and 0.58~0.97 respectively. The ability of multiple regression models to forecast out-of-sample sets is compatible with ANN for Daylesford in central Victoria and Kaniva in west Victoria (r=0.92 and 0.67 respectively). The errors of the testing sets for ANN models are generally lower compared to multiple regression models. The statistical analysis suggest the potential of ANN over MLR models for rainfall forecasting using large scale climate modes.

2.5 Summery

Large-scale climate modes like ENSO and IOD affect the climate and particularly rainfall around the world. Australia is influenced by these phenomena as well. The nature of simultaneous relationships between climate indices and rainfall have been studied in numerous theoretical research for Australia. To date, majority of the studies on the relationship between simultaneous large scale climate modes and the Australian rainfall have adopted either the linear regression analysis (Cai et al. 2011, 2012; Risbey et al. 2009) or probability/categorical analysis (Power et al.1999; Verdon et al. 2004; Kiem & Verdon- Kidd 2009; and Gallant et al. 2012). The use of antecedent climate modes for the purpose of future predictions is proposed only in few studies and non-linear techniques have rarely been used. The aim of the present study is to address the

existing gap in seasonal rainfall prediction in Australia. On the other hand, research on Victoria's seasonal rainfall variability has not received enough attention. According to Verdon-Kidd and Kiem (2009b), in comparison to eastern Australia and particularly Queensland, the performance of past studies on Victorian seasonal rainfall predictability have been low and a maximum predictability of only 30% was achieved. Therefore, there is a need to further examine the relationship between seasonal rainfall in Victoria and the effective large scale climate modes in order to develop a reliable seasonal rainfall forecast model.

Chapter 3

Data and Study Area

3.1 Study Area

Victoria in Australia is the second smallest state, and covers 227,600 square kilometers. Roughly 36% of Victoria is forest and the main forest belt is located in the east. The highest mountains in Victoria are Mount Bogong and Mount Feathertop, at 1986 m and 1922 m respectively. Temperature in Victoria varies widely across the seasons, however most parts of Victoria are characterized by a warm, temperate climate, with the south-east corner of Australia having cool to mild wet winters and warm and dry summers. Normal daily summer temperatures vary between 14^{o}C to 23^{o}C in the coastal areas, 16^{o}C to 31^{o}C in the inland and 11^{o}C to 20^{o}C in the mountains. During winter, temperatures vary between 7^{o}C to 14^{o}C in coastal areas, 5^{o}C to 16^{o}C in the inland and 0^{o}C to 5^{o}C in the mountains. Snow settles on the Australian Alps from June to September in the north-east of Victoria.

In recent years, Australia has experienced dramatic flood events due to inclement climate conditions. From December 2010 until early 2011, major floods due to heavy rainfall occurred in several states in Australia including Victoria, New South Wales and Queensland (AbdulRauf & Zeephongsekul, 2014). According to the BoM, flooding is Australia's most costly natural disaster. Although floods do not have as wide an effect on Australia as droughts, they still cause damage to dwellings, transportation networks, and other infrastructure. In Victoria, flooding occurs mostly in winter and spring. In this study, three distinct regions of east, central and west Victoria are considered as case studies. Nine stations with maximum records of data were selected from these regions. Figure 3.1 shows the location details of the stations considered in this study. Monthly records of rainfall for 110 years (1900-2009) were obtained from the BoM for this study. Spring (September-November) rainfalls were obtained from the monthly rainfall data. Table 3.1 gives the geographical locations of the stations, their annual mean rainfall and the percentage of missing values for each station. Buchan, Daylesford, Heathcote and Kaniva each had a single missing value, Rainbow two missing values

and the rest of the stations had no missing values for spring rainfall. The missing values were in-filled using long-term averages of each station over the period 1900-2009.

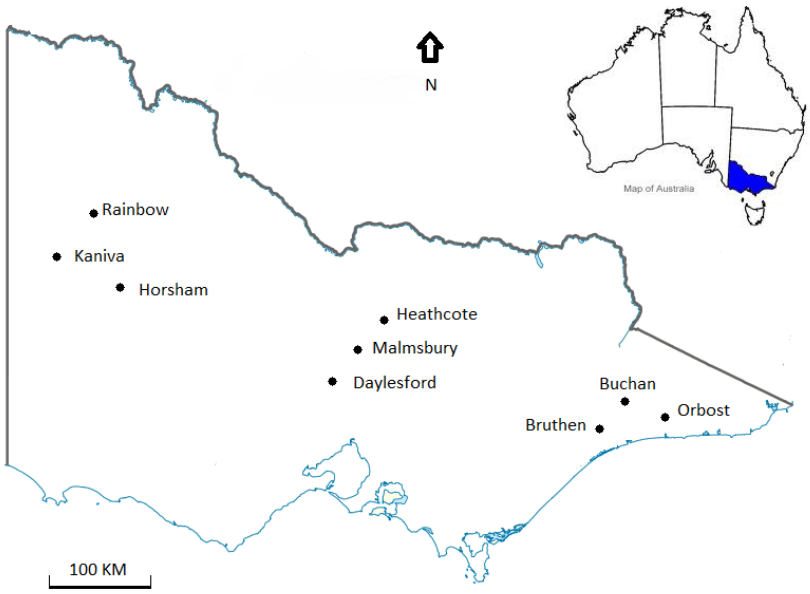

Figure 3-1. Schematic map of the study area

Table 3.1. Details of geographical location and recorded data of rainfall stations of the study

Region	Site No.	Site Name	Latitude	Longitude	Annual Mean Rainfall (mm)	% of Missing Observations
East	084003	Bruthen	37.71°S	147.83°E	759.5	0
	084005	Buchan	37.50°S	148.17° E	824.2	0
	084030	Orbost	37.69°S	148.46° E	845.1	0
Center	088042	Malmsbury	37.2 °S	144.37° E	726.0	0
	088020	Daylesford	37.34°S	144.16° E	879.3	1
	088029	Heathcote	36.96°S	144.69° E	575.9	1
West	079023	Horsham	36.66° S	142.07° E	447.3	0
	078078	Kaniva	36.37° S	141.24° E	452.1	1
	077051	Rainbow	35.94° S	141.94° E	349.0	2

The climate modes investigated in this study are chosen because they show simultaneous relationships with monthly or seasonal rainfalls across Australia. As discussed in Chapter 2, the dominant climate modes affecting Australia's climate are the El Nino Southern Oscillation (ENSO), the Indian Ocean Dipole (IOD), the Southern Annual Mode (SAM) and the Interdecadal Pacific Oscillation (IPO). The Southern Oscillation Index (SOI), which is a measure of Sea Level Pressure (SLP) anomalies between Darwin and Tahiti is a quantitative indicator of ENSO. ENSO is also represented by the Sea Surface Temperature (SST) anomalies in the equatorial Pacific Ocean. The SST anomalies vary depending on their region; the most important ones include Nino3 ($5^oS – 5^oN$, $150^o – 90^oW$), Nino3.4 ($5^o S – 5^oN$, $170^o – 120^oW$) and Nino4 ($5^oS – 5^oN$, $160^o – 150^oW$) (Risbey et al. 2009). SOI and Nino3.4 are the most common ENSO indicators used in analyzing this phenomenon.

IOD is the other coupled ocean-atmosphere phenomenon taking place in the Indian Ocean. A measure of IOD is the Dipole Mode Index (DMI), which is the difference in average SST anomalies between the tropical Western Indian Ocean (10^oS - 10^oN, 50^o - 70^oE) and the tropical Eastern Indian ocean (10^oS - Equator, 90^o - 110^oE) (Kirono et al. 2010).

The IPO, which is a low frequency (15–35 years) form of variability of the tropical and extra-tropical Pacific Ocean, modulates the relationship between ENSO and rainfall on a multi-decadal scale in some parts of Victoria. In this study, monthly values of Nino3.4, SOI and DMI were obtained from Royal Netherlands Meteorological Institute (KNMI) Climate Explorer website (http://climexp.knmi.nl/). The monthly IPO index was derived from Parker et al. (2007). Table 3.2 summarizes the climate indices used in this study, and Figure 3.2 to 3.4 show the intensity of the climate modes during the period of study.

Table 3.2. Climate indices investigated as potential predictors of Victoria's seasonal rainfall

Predictors	Predictor Definition	Region	Data source
SOI	Anomaly of mean sea level pressure (MSLP) difference between Tahiti and Darwin	Pacific	KNMI Climate Explorer
Nino3.4	Average SST anomaly over 5°S–5°N and 170° – 120°W	Pacific	KNMI Climate Explorer
DMI	West Pole Index - East Pole Index: (Average SST anomaly over 10°S - 10°N, 50° - 70°E)-(Average SST anomaly over 10°S - Equator, 90° - 110°E)	Indian	KNMI Climate Explorer
IPO	empirical orthogonal function (EOF) of SST anomalies in Pacific Ocean	Pacific	Parker et al., (2007)

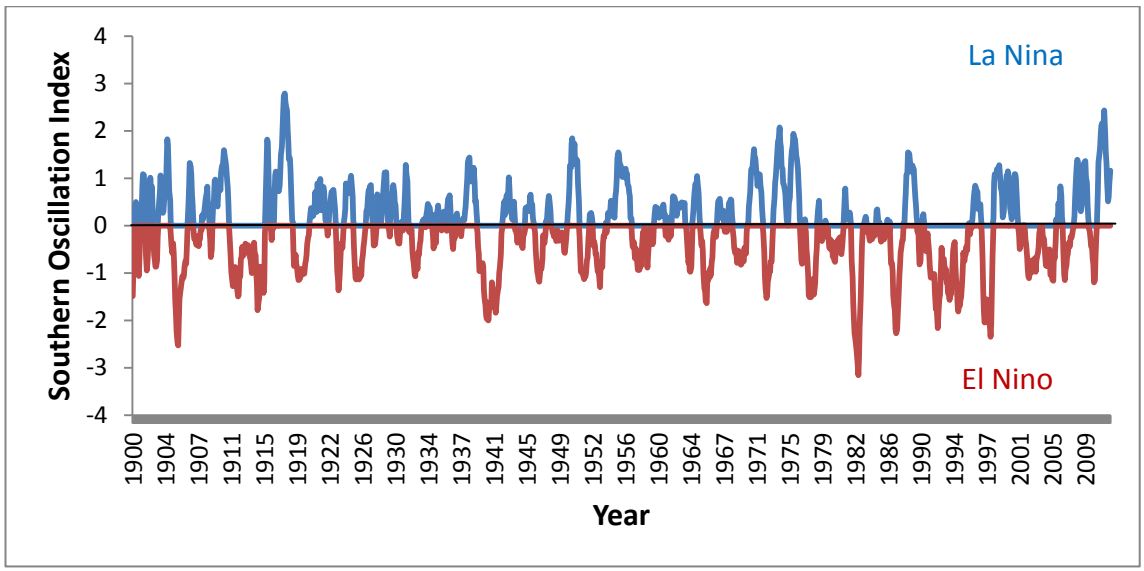

Figure 3-2. SOI five month running mean for the period 1900-2009 (Data source: http://climexp.knmi.nl/)

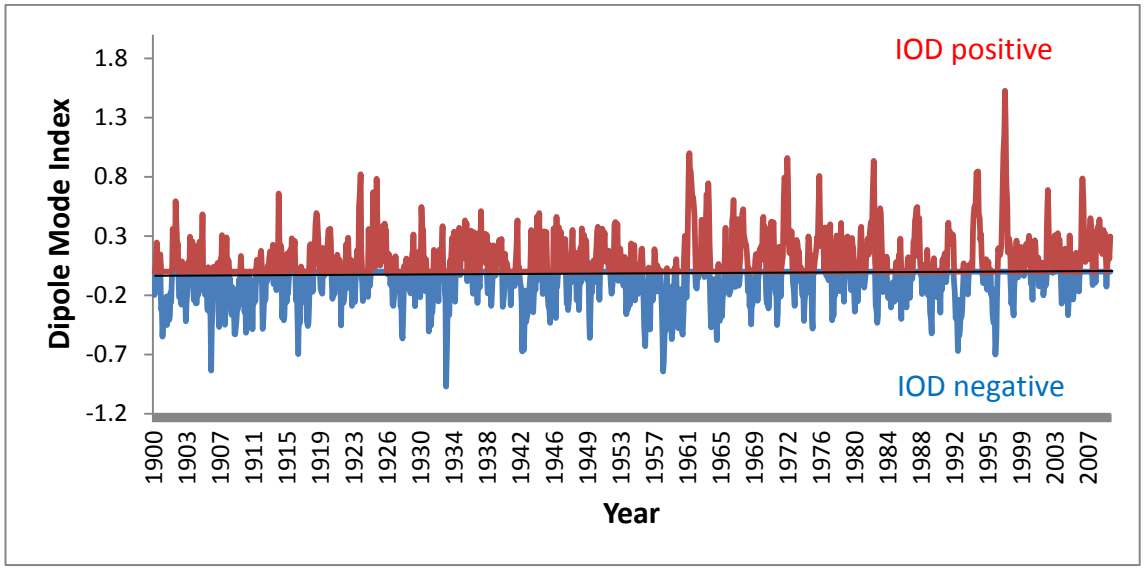

Figure 3-3. IOD monthly values for the period 1900-2009 (Data source: http://climexp.knmi.nl/)

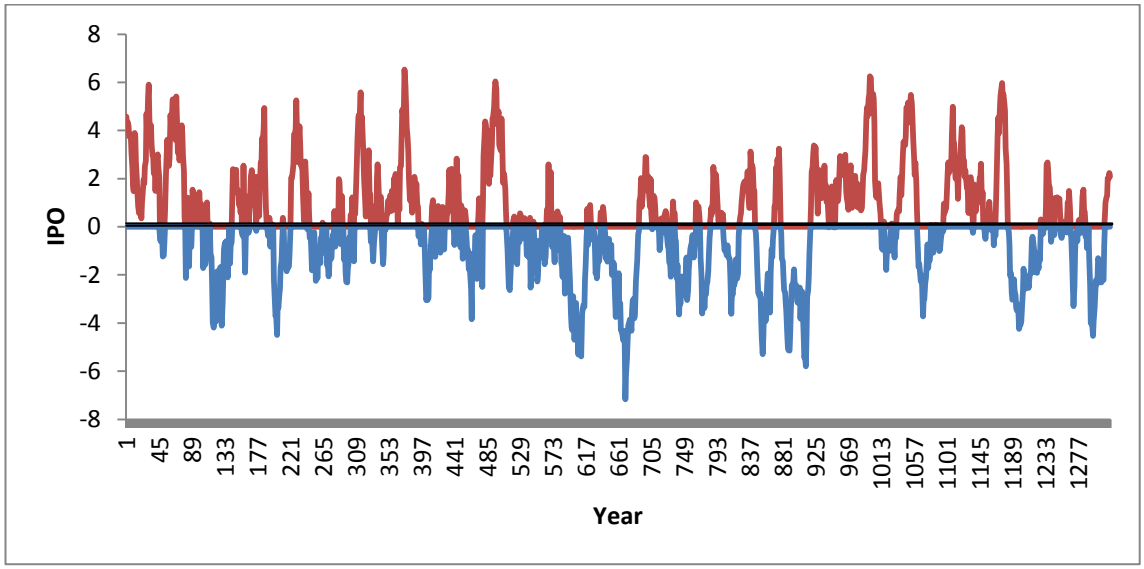

Figure 3-4. IPO monthly values for the period of 1900-2009 (Data source: Parker et al., 2007)

3.2 Data preprocessing

According to Maier and Dandy (2000) data pre-processing have significant effect on the model performance as different input variables could span different ranges and normalization of the variables insures that all variables receive equal attention during the calibration process of the modelling process. In general, in any model development process it is best to know the type of data considered in the model building stage. For example in the case of ANN models, some activation functions used in the output layer

of ANN have some limits in their range and the data must therefore be scaled within their range. For example, if the outputs of the logistic transfer function are between 0 and 1, the data are generally required to be scaled within the range of 0.1-0.9 or 0.2-0.8. It should be noted that if the transfer function of the output layer is unbounded (e.g. linear) scaling is not strictly required (Maier et al. 2010).

In the present study the data are normalized in the range of 0 to 1 using Eq. 3-1.

$$\hat{x}_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{3-1}$$

where, $\hat{x}_i$ is the normalized value, x_i is the observation value, and x_{max} and x_{min} are the maximum and minimum of the data respectively.

3.3 Model Verification

The performance criteria and error statistics used in this study are mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), and Pearson correlation (r).These criteria are described as follows:

a) Mean square error :

$$MSE = \frac{\sum_{i=1}^{n}(x_i - o_i)^2}{n} \tag{3-2}$$

b) Root mean square error or RMSE:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - o_i)^2}{n}} \tag{3-3}$$

where x_i and o_i are observed and simulated data of the *i*th observation, respectively; and *n* is the total number of observations. The RMSE accords extra importance to the outliers in the data set and is therefore biased towards errors in the simulation of high values (Dawson et al., 2006).

c) Mean Absolute Error or MAE:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |x_i - o_i|}{n} \tag{3-4}$$

MAE computes all deviations from the original data regardless of sign and is not weighted towards high values (Abrahart et al. 2004).

d) Correlation Coefficient (r):

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \times \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \tag{3-5}$$

where, r is the correlation coefficient between variables x and y; $\bar{x}$ and $\bar{y}$ are the average values of x and y, respectively and n is the number of data points. r ranges within the domain [-1, 1] where the values of 1 and -1 indicate positive and negative perfect linear correlation respectively, while r = 0 is an indication that there is no correlation between the two data series.

Chapter 4

Analysis of the Effect of Climate Mode Phases on rainfall

4.1 Classification Analysis

The El Nino Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD) are two ocean-atmospheric climate modes which change phase on a seasonal to inter-annual scale. The cool (positive) phase of ENSO, which is associated with cooler than average sea surface temperatures (SSTs) in the central and eastern tropical Pacific Ocean, is called La Nina. La Nina conditions (events) generally result in above-average rainfall over much of Australia (Verdon et al. 2004). In contrast, the warm (negative) phase of ENSO is associated with warmer than average SSTs in the central and eastern tropical Pacific Ocean. This phase of ENSO is known as El Nino, the occurrence of which generally result in below-average rainfall over much of eastern Australia (Wang & Hendon 2007). A neutral phase also exists, which is neither El Nino nor La Nina conditions.

The IOD also has three different conditions. The positive IOD (pIOD) is related to cooler than normal water in the tropical eastern Indian Ocean and warmer than normal water in the tropical western Indian Ocean. pIOD conditions are linked with decreased rainfall over the southern and central parts of Australia . On the other hand, a negative IOD (nIOD) brings warmer than normal water in the tropical eastern Indian Ocean and cooler than normal water in the tropical western Indian Ocean. nIOD conditions are linked to increased rainfall over parts of southern Australia (Gallant et al. 2012). The neutral IOD is the normal condition, where neither pIOD nor nIOD occur. It is believed that ENSO and IOD affect cool season (spring-winter) rainfall in Australia (Cai et al. 2011). Meyers et al. (2007) proposed a method to identify when the negative or positive phase of ENSO and IOD occur and applied the method to classify each year from 1876 to 1999 into the different phases of ENSO and IOD. Later, Ummenhofer et al. (2009) extended the classification to 2009; the classification is shown in Table 4.1. While some studies have investigated the effect of the phases of ENSO and IOD on some regions in

Australia (Ummenhofer et al. 2009; Cai et al. 2011, 2012), Victoria's rainfall has not been studied specifically.

Table 4.1. The years of ENSO and IOD based on the Ummenhofer et al. (2009) classification

	Negative IOD	Neutral IOD	Positive IOD
El Nino	1930	1877, 1888, 1899, 1905, 1911, 1914, 1918, 1925, 1940, 1941, 1965, 1972, 1986, 1987	1896, 1902, 1957,1963, 1982, 1991, 1997, 2009
Neutral ENSO	1915, 1958, 1968, 1974, 1980, 1985, 1989, 1992	1880, 1881, 1882, 1883, 1884, 1895, 1898, 1900, 1901, 1904, 1907, 1908, 1912, 1920, 1921, 1927, 1929, 1931, 1932, 1934, 1936, 1937, 1939, 1943, 1947, 1948, 1951, 1952, 1953, 1953, 1959, 1960, 1962, 1966, 1967, 1969, 1971, 1976, 1977, 1979, 1983, 1990, 1993, 1995, 2001, 2002, 2003, 2005, 2006,	1885, 1887, 1891, 1894, 1913, 1919, 1923, 1926, 1935, 1944, 1945, 1946, 1961, 1994, 2004, 2008
La Nina	1906, 1909, 1916, 1917, 1933, 1942, 1975	1878, 1879, 1886, 1889, 1890, 1892, 1893, 1897, 1903, 1910, 1922, 1924, 1928, 1938, 1949, 1950, 1954, 1955, 1956, 1964, 1970, 1973, 1978, 1981, 1984, 1988, 1996, 1998, 2000	1999, 2007

In order to evaluate the effect of different phases of ENSO and IOD on cool season rainfall in Victoria, seasonal rainfall anomalies are constructed for the three regions of the case study (east, central and west Victoria) based on Eq. 4-1:

$$x_{i(anom)} = x_i - \bar{x} \tag{4-1}$$

where $x_{i(anom)}$ is the anomaly of seasonal rainfall at year i , x_i is the value of seasonal rainfall at year i, and $\bar{x}$ is the long-term average rainfall for the particular season under study. The long-term average is calculated for the years 1900-2009 and the anomalies are created for this period.

The anomalies constructed for spring and winter rainfalls are categorized based on the six phases of ENSO and IOD (Neutral, El Nino and La Nina for ENSO and Neutral,

pIOD and nIOD for IOD). In this way, six groups of rainfall anomalies are constructed, based on the six phases of ENSO and IOD in order to evaluate whether the median of each group is statistically different from zero. A statistically significant zero median is referred to as normal climate condition and normal rainfall (BoM-2014). Figures 4.1 and 4.2 show the six categories of rainfall anomalies in colour-coded bars for spring and winter rainfall respectively. The neutral years of ENSO and IOD are colour-coded gray and are in the middle of each figure with El Nino and pIOD on the left and La Nina and nIOD on the right. The middle horizontal line in each bar shows the median rainfall. The Wilcoxon Signed Rank test is applied on each category to determine whether the median rainfall is statistically distinguishable from zero during the six phases. Tables 4.2 and 4.3 show the results of the Wilcoxon Signed Rank test for spring and winter rainfall anomalies respectively. Based on Table 4.2, in spring the median is significantly different from zero in east, central and west Victoria during El Nino and pIOD events; other than in east Victoria, the median is significantly different from zero during La Nina and nIOD events. This shows that the dry phases of ENSO and IOD have more effect on east Victoria than the wet phases. Table 4.3 shows that in winter the median is only distinguishable from zero in central Victoria during El Nino and nIOD events and in west Victoria during nIOD events. No other phases have significant effects in Victoria during winter. By comparing Table 4.2 and 4.3 it can be concluded that in Victoria spring rainfall is more under the influence of ENSO and IOD than winter rainfall. Therefore, more emphasis needs to be placed on predicting spring rainfall in Victoria based on large-scale climate modes.

Table 4.2. Testing for significant difference of median from zero for spring rainfall (Wilcoxon Signed Rank test)

Region	LN	EN	Neutral-ENSO	IOD+	IOD-	Neutral-IOD
East	---	0.027*	---	0.004*	---	---
Centre	0.015*	0.006*	---	0.000*	0.020*	---
West	0.028*	0.003*	---	0.002*	0.007**	---

*The significance level is 0.05

Table 4.3. Testing for significant difference of median from zero for winter rainfall (Wilcoxon Signed Rank test)

Region	LN	EN	Neutral-ENSO	IOD+	IOD-	Neutral-IOD
East	---	---	---	---	---	---
Centre	---	0.033*	---	---	0.013*	---
West	---	---	---	---	0.039*	---

*The significance level is 0.05

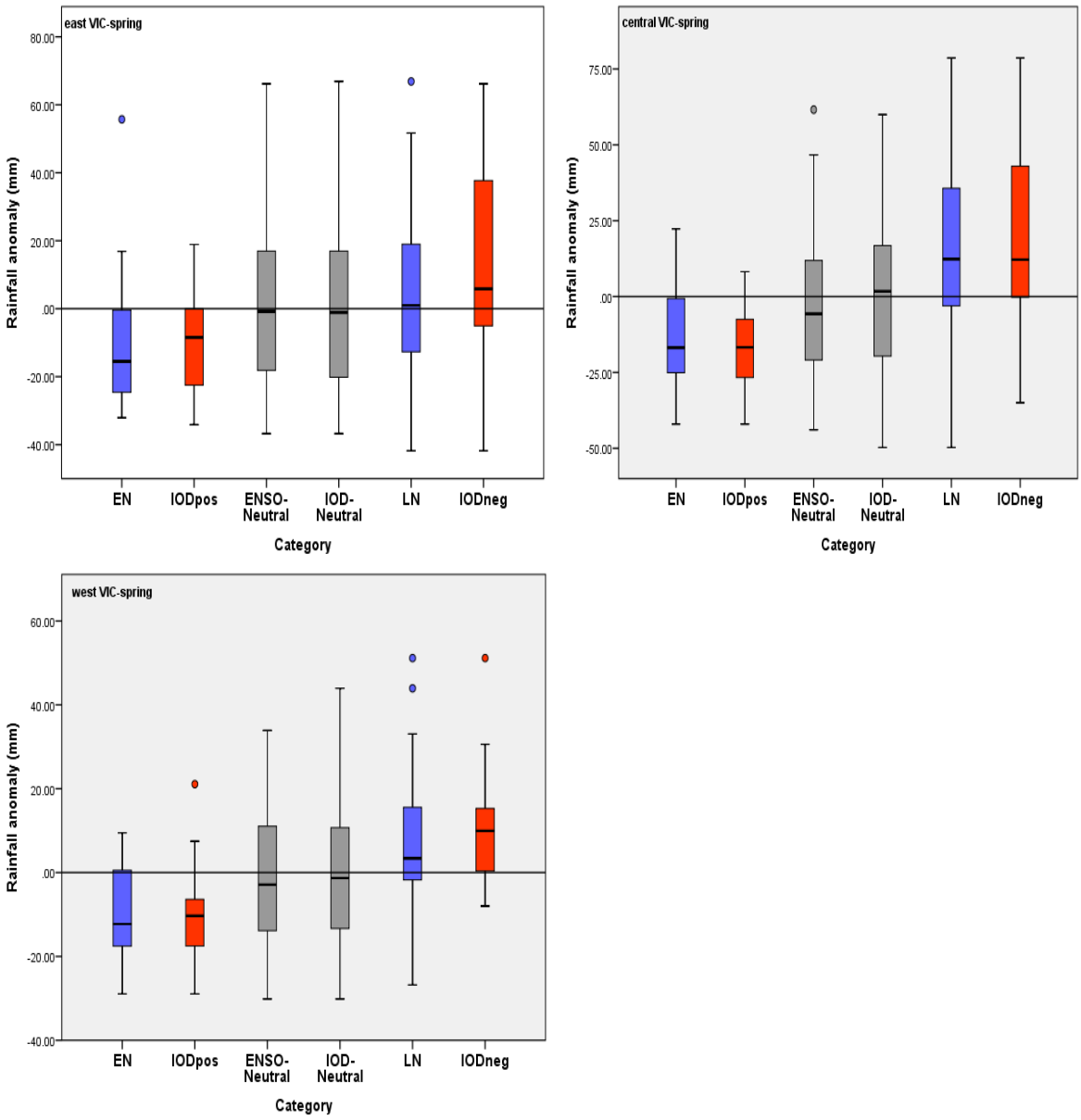

Figure 4-1. Spring rainfall anomalies for the different ENSO/IOD categories for Victoria. The neutral years of ENSO and IOD are colour-coded gray and are in the middle of each figure with El Nino and pIOD on the left and La Nina and nIOD on the right. The middle line in each bar shows the median rainfall.

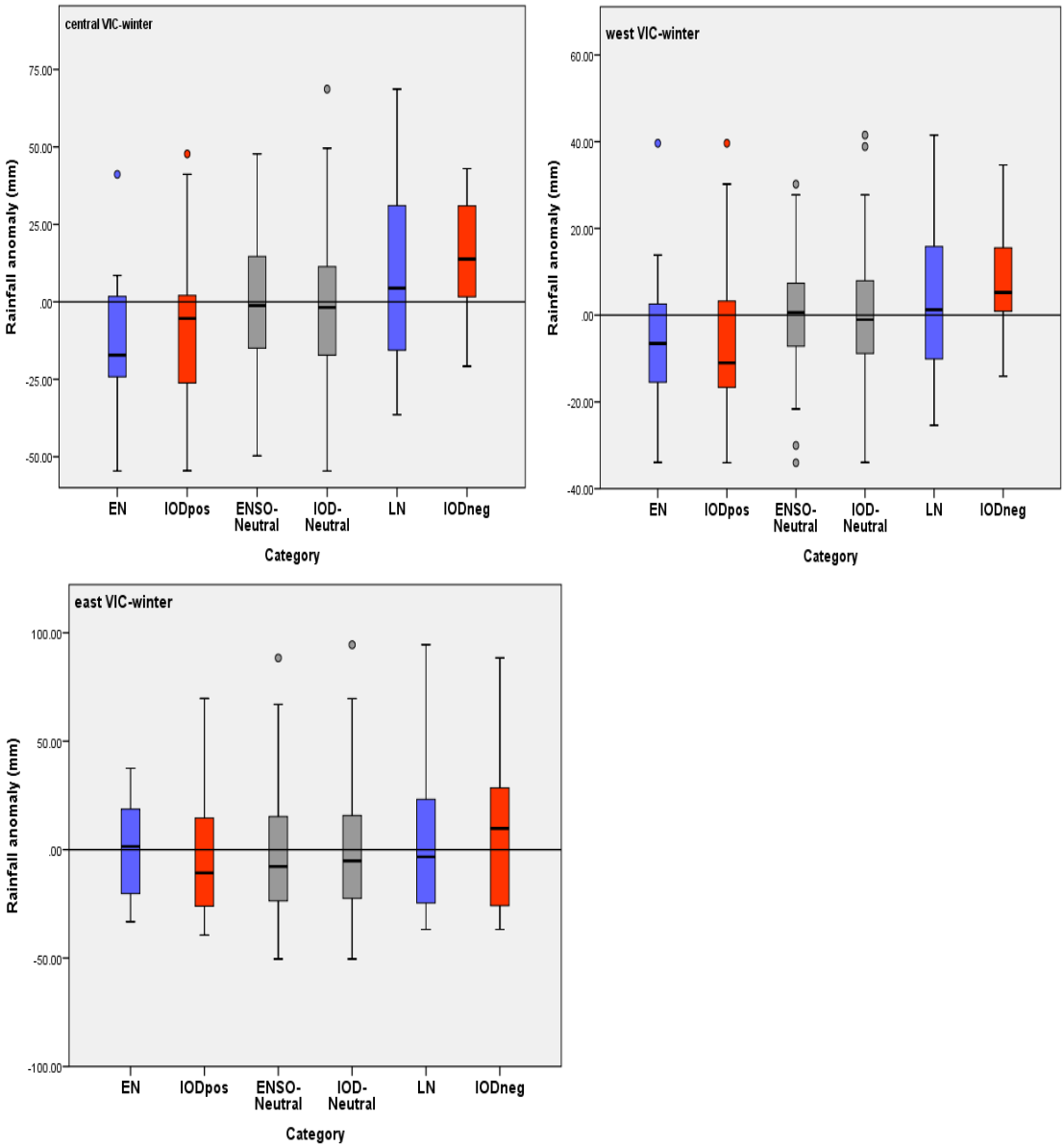

Figure 4-2. Winter rainfall anomalies for the different ENSO/IOD categories for Victoria. The neutral years of ENSO and IOD are colour-coded gray and are in the middle of each figure with El Nino and pIOD on the left and La Nina and nIOD on the right. The middle line in each bar shows the median rainfall.

After investigating the effect of separate phases of ENSO and IOD on cool season rainfall in Victoria, the combined effect of ENSO/IOD on Victoria's seasonal rainfall was investigated. Since IOD is strongly correlated with ENSO in austral spring and less correlated with ENSO in winter, the independent and dependent impacts of IOD and ENSO are felt in both seasons in Australia (Cai et al. 2011). The dependence and independence effect of IOD and ENSO in Victoria in particular are investigated in this section. In order to show the strength of the relationship between IOD and ENSO in different seasons, Pearson correlations between Nino3.4, SOI (ENSO indicators) and DMI (IOD indicator) were calculated for the four seasons (Figure 4.3). It can be seen from this figure that there is a moderate to weak relationship between IOD and ENSO in spring and winter, but in summer and autumn no relationship exists between these two climate modes. Based on Figure 4.3, the relationship between ENSO and IOD in spring is stronger than in winter, with a maximum correlation coefficient of 0.59 compared to -0.40. According to Meyers et al. (2007), the IOD and ENSO can occur together such that they reinforce each other, but this need not happen; conditions may

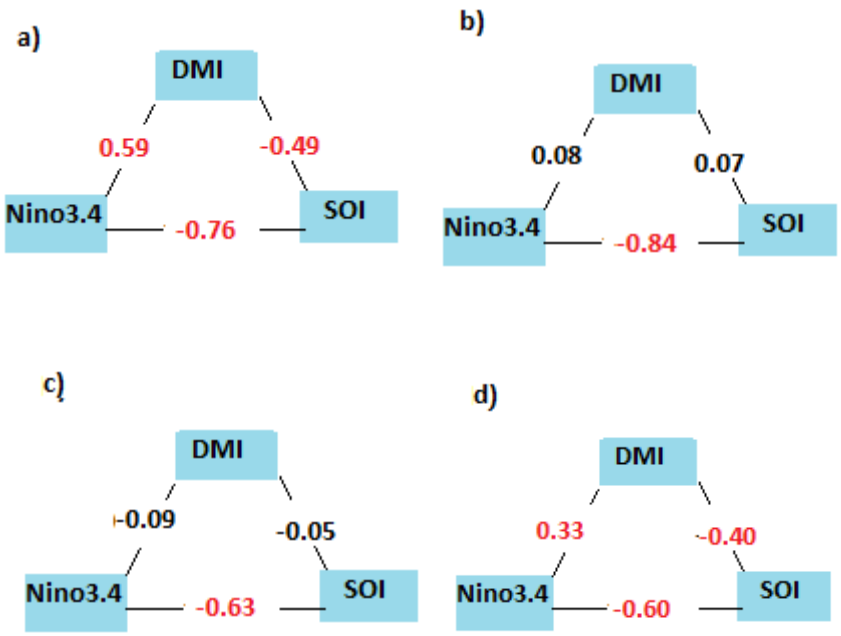

Figure 4-3. Pearson correlations (r) of climate indices a) spring b) summer c) fall d) winter. Significant correlations are shown in red

exist that a phase of ENSO is accompanied by a phase of IOD. For example, an El Nino event can occur while IOD is in its positive phase. This condition will reinforce the effect of ENSO and IOD. In the classification of Meyers et al. (2007) and Ummenhofer et al. (2009), this situation is considered by categorizing ENSO and IOD into nine categories, of which two seldom occur (Meyers et al. 2007). These categories are El Nino –pIOD, pure El Nino, pure pIOD, Neutral, pure La Nina, La Nina-nIOD and pure nIOD. The term "pure" is used in the present study to identify that the only active phase of climate is the phase where the term "pure" is associated with and the other mode is in neutral condition. In general, El Nino–pIOD, pure El Nino, and pure pIOD are accompanied by dry (below-average rainfall) conditions, while pure La Nina, La Nina-nIOD and pure nIOD are accompanied by wet (above-average rainfall) conditions. To investigate the reinforcement effect of ENSO and IOD on rainfall, this classification was used to categorize Victoria's spring and winter rainfall anomalies (see Figures 4.4 and 4.5). The Wilcoxon Signed Rank test is applied to determine whether the median of each category of anomalies is statistically significant from zero (Table 4.4). It can be seen from Table 4.4 that for east Victoria El Nino-pIOD co-occurrence is statistically significant; a median rainfall anomaly of -20 mm is recorded for east Victoria during the co-occurrence of El Nino-pIOD with all but one event out of seven showing below average rainfall. For central and west Victoria, a median rainfall anomaly of -22 and -15 mm respectively is recorded during pure El Nino events with all but two events out of eleven having below average rainfall. Also, a median of -20 mm and -12 mm is recorded for central and west Victoria respectively, during pure pIOD events with all but one event out of twelve showing below average rainfall. For central Victoria, with a median of + 12 mm during pure La Nina, sixteen out of twenty-one events have above average rainfalls. The only region in Victoria with six events out of seven events having above average rainfall in spring during the co-occurrence of La Nina-nIOD is west Victoria with a median of +15 mm.

Table 4.4. Testing for significant differences of median from zero, spring rainfall (Wilcoxon Signed Rank test)

Region	EN,IOD+	EN	IOD+	Neutral	IOD-	LN	LN,IOD-
East	0.028*	---	---	---	---	---	---
Centre	---	0.026*	0.005*	---	---	0.035*	---
West	---	0.013*	0.028*	---	---	---	0.043*

*The significance level is 0.05

Table 4.5. Testing for significant difference of median from zero, winter rainfall (Wilcoxon Signed Rank test)

Region	EN,IOD+	EN	IOD+	Neutral	IOD-	LN	LN,IOD-
East	---	---	---	---	---	---	---
Centre	---	---	---	---	---	---	0.018*
West	---	---	---	---	---	---	---

*The significance level is 0.05

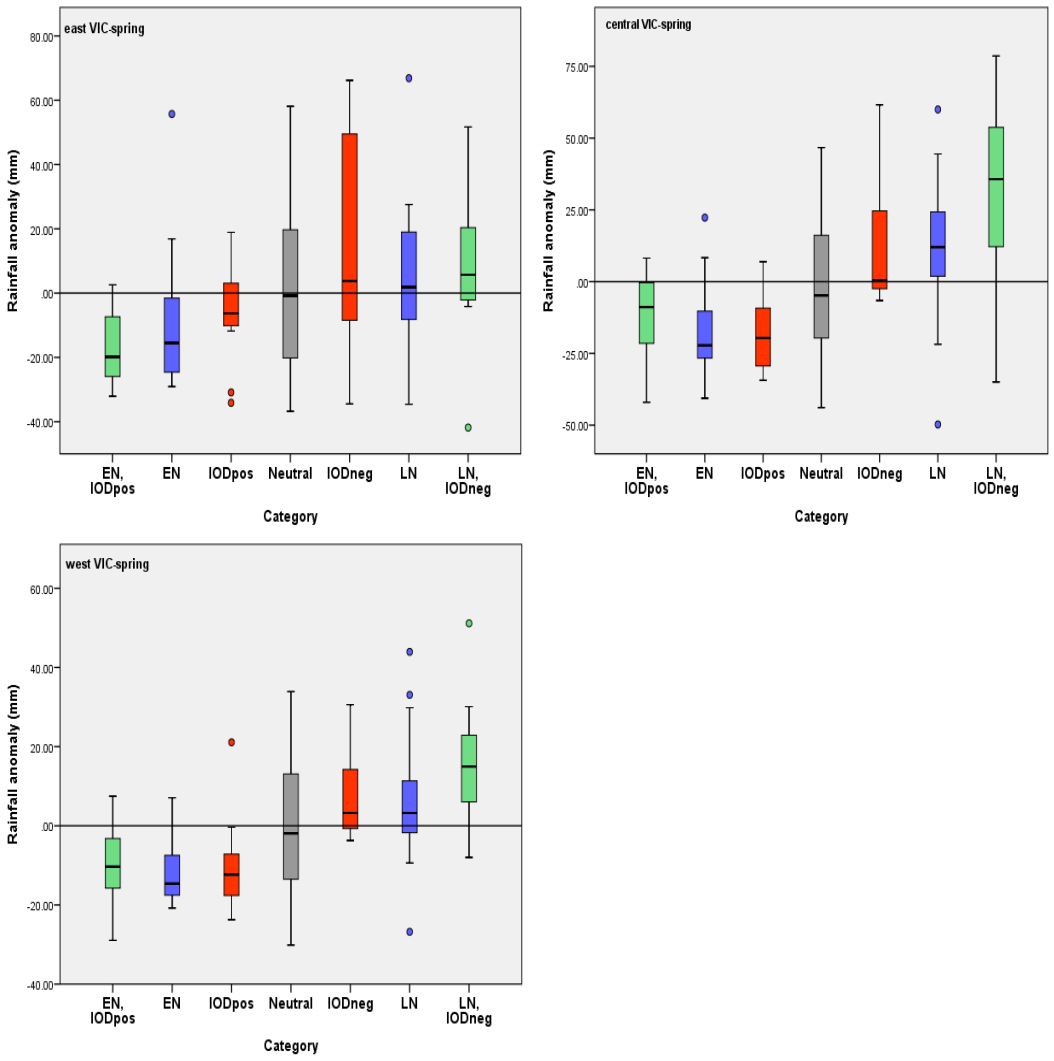

Figure 4-4. Spring rainfall anomalies for the different ENSO/IOD categories for Victoria. The neutral years are colour-coded gray and are in the middle of each figure with El Nino-pIOD, El Nino and pIOD on the left and La Nina-nIOD, La Nina and nIOD on the right. The middle line in each bar shows the median rainfall.

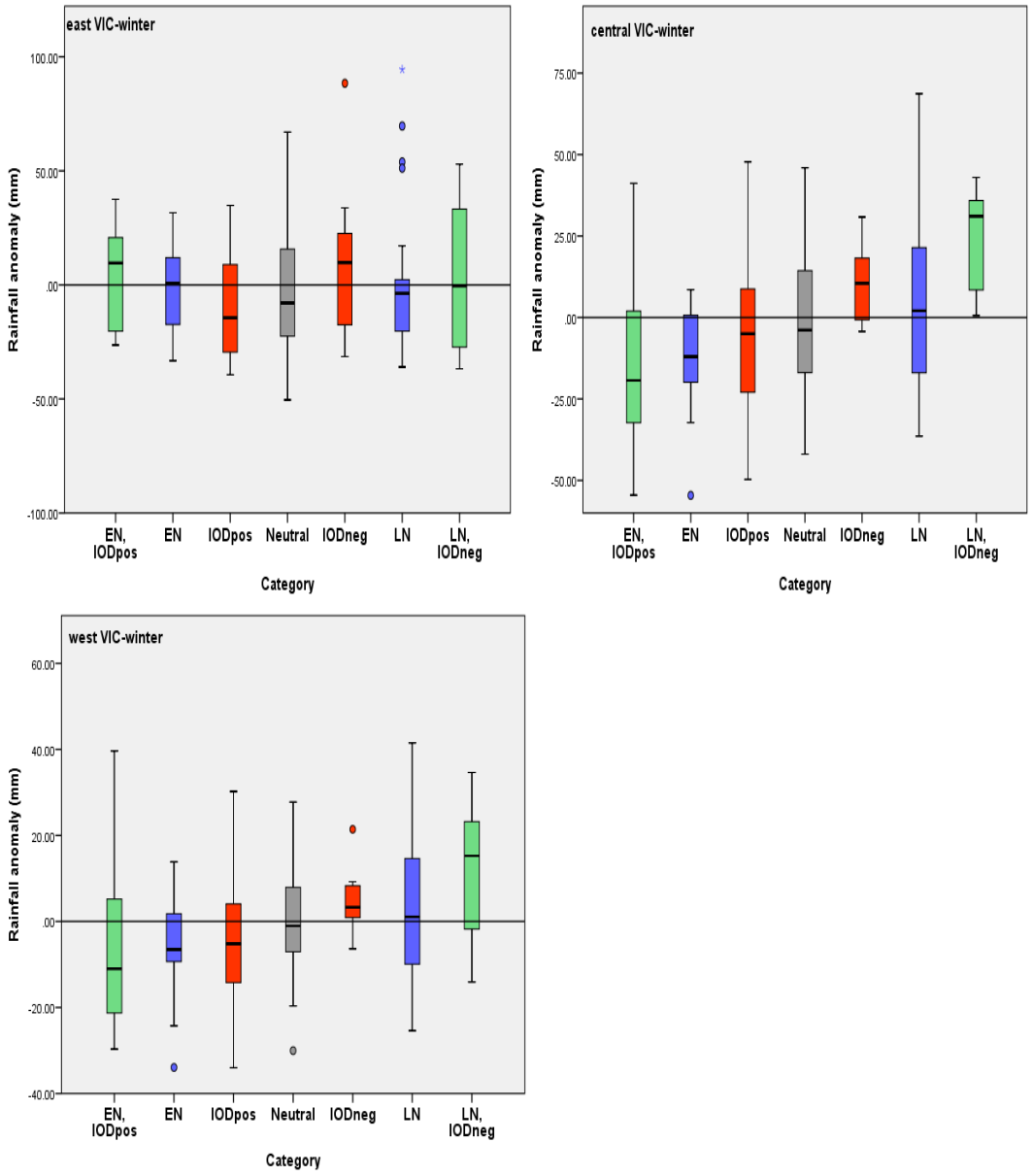

Figure 4-5. Winter rainfall anomalies for the different ENSO/IOD categories for Victoria. The neutral years are colour-coded gray and are in the middle of each figure with El Nino-pIOD, El Nino and pIOD on the left and La Nina-nIOD, La Nina and nIOD on the right. The middle line in each bar shows the median rainfall.

In general, the results show that rainfall anomalies during pure nIOD events, pure La Nina events and when they both occur at the same time (i.e. wet phase) are mostly variable; however, pure El Nino years, pure pIOD events and the co-occurrence of El Nino-pIOD (i.e. dry phase) consistently result in dry conditions during spring in Victoria. Based on Table 4.5, during winter none of the ENSO and IOD events affect Victoria, except for central Victoria where La Nina-nIOD occurring at the same time

affects the rainfall. As discussed earlier, it appears that winter rainfall in Victoria is not as much affected by different phases of ENSO and IOD and their co-occurrence compared to spring rainfall. The results derived here highlight the importance of spring rainfall prediction based on large-scale climate modes for Victoria. Further, the result indicate that, although the general conception indicates that ENSO and IOD affect south-east Australia during spring and winter, Victoria is most influenced by the large-scale climate modes in spring.

4.1.1 Summary of classification analysis

In this section the effect of different phases of ENSO and IOD on Victoria's cool seasonal rainfall was investigated. Rainfall anomalies were constructed and classified based on the classification of Meyers et al. (2007) and Ummenhofer et al. (2009). The Wilcoxon Signed Rank test was applied in two steps to investigate the significance of the median of each category from zero. In the first step, the anomalies were tested based on the three phases of ENSO and three phases of IOD. It was discovered that spring rainfall in Victoria is more under the influence of these phases than winter rainfall. In the next step, the co-occurrence of ENSO and IOD was taken into account and it was discovered that the dry phases of ENSO and IOD and their co-occurrence (i.e. pure El Nino years, pure pIOD events and the co-occurrence of El Nino-pIOD) consistently result in dry conditions, while the wet phases of these phenomena are highly variable. It was also revealed that it is only when El Nino and pIOD occur at the same time that the events affect east Victoria's rainfall. The results highlight the importance of dry phases of ENSO/IOD on spring rainfall and the necessity of forecasting spring rainfall based on large-scale climate modes as potential predictors.

4.2 Pearson Correlation Analysis

Categorizing rainfall based on the phases of ENSO and IOD in Section 4.1 revealed the effect of each phase of these climate modes on Victoria's cool season rainfall. It was discovered that the dry phases of these climate modes have more significant effects on rainfall variations than the wet phases. While classification analysis gives a good understanding of the natural causes of the increase or decrease in rainfall over Victoria, it cannot show the magnitude of the influence of the climate modes on seasonal rainfall. In order to have an understanding of the strength of different climate modes on seasonal

rainfall in Victoria, Pearson correlation analysis is used in two separate approaches. In the first approach the concurrent relationship between seasonal rainfall and climate modes is investigated; the second approach takes into account the relationship between seasonal rainfall and antecedent climate modes. The results of the second approach is used later as the basis of selecting the appropriate lagged climate modes as inputs for seasonal rainfall forecasting models.

4.2.1 Concurrent relationships

In order to evaluate the linear relationships between spring, summer, autumn and winter rainfalls and spring, summer, autumn and winter ENSO/IOD, a Pearson correlation analysis was conducted for the three regions under study. The results are demonstrated in Tables 4.6 to 4.9 for spring, summer, autumn and winter respectively. It can be seen from Table 4.6 that the relationship between ENSO and IOD with east Victoria's spring rainfall is very weak (r_{max}= -0.27) compared to the other regions. In central and west Victoria, a significant increase can be seen in the correlation coefficient values between ENSO and IOD with spring rainfalls. The maximum correlation coefficient in central and west Victoria can be seen with regard to IOD with r_{max}= -0.44 and -0.45 respectively. In summer (Table 4.7) the influence of ENSO is quite weak in Victoria and can only be seen in central Victoria with r_{max}= 0.20. IOD has no significant influence on summer rainfall in this state. In autumn, ENSO has no influence on the region and IOD has a minor effect on central and west Victoria with r_{max}= -0.29 and -0.20 respectively (Table 4.8). Although the effect of ENSO in winter is increased compared to summer and autumn, this effect cannot be seen for east Victoria. ENSO affects central and west Victoria in winter with a correlation coefficient of r_{max}= 0.43 and 0.38 respectively compared to IOD with r_{max}= -0.33 and -0.29 respectively. In general, it appears that ENSO and IOD influence the cool season rainfalls (spring and winter) in Victoria, with IOD being the dominant predictor in spring and SOI in winter; the effects of ENSO and IOD are very weak for east Victoria.

Table 4.6. Pearson correlation (r) of spring climate indices and spring rainfall

Region	$Nino34_{(S-O-N)}$	$SOI_{(S-O-N)}$	$DMI_{(S-O-N)}$
East	-0.27**	0.21*	-0.25**
Centre	-0.38**	0.43**	-0.44**
West	-0.39**	0.43**	-0.45**

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

Table 4.7. Pearson correlation (r) of summer climate indices and fall rainfall

Region	$Nino34_{(S-O-N)}$	$SOI_{(S-O-N)}$	$DMI_{(S-O-N)}$
East	----	----	----
Centre	-0.20*	0.20*	----
West	----	----	----

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

Table 4.8. Pearson correlation (r) of autumn climate indices and fall rainfall

Region	Nino34	SOI	DMI
East	----	----	----
Centre	----	----	-0.29**
West	----	----	-0.20*

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

Table 4.9. Pearson correlation (r) of winter climate indices and winter rainfall

Region	Nino34	SOI	DMI
East	----	----	----
Centre	-0.21*	0.43**	-0.33**
West	----	0.38**	-0.29**

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

Analysis of the correlation coefficient between ENSO/IOD and seasonal rainfall in Victoria gives some insight into the strength of these relationships in general, but this analysis does not reveal the strength of the relationship between rainfall and specifically the phases of climate modes (i.e. El Nino/ La Nina and pIOD/nIOD). In order to examine the strength of the relationship between the effective phases of ENSO (El Nino/La Nina) and IOD (pIOD/nIOD) and cool season rainfall (spring/winter), further analysis is required. Therefore, seasonal rainfalls associated with the specific years of El Nino and La Nina classified by Ummenhofer et al. (2009) were separated from the data for each region. The extracted rainfalls were then put in order based on the related years. In this way the neutral years are removed from the rainfall series and only the El Nino and La Nina years remain. We will call this method merging as the years of El Nino and La Nina were merged to produce a time series of related seasonal rainfalls. The same method was applied in order to produce merged rainfalls based on pIOD/nIOD. The related Nino3.4 and SOI values were also merged based on the years of El Nino and La Nina to produce the series of merged Nino3.4 and SOI; DMI was also aggregated based on the years of pIOD/nIOD to produce the series of merged IOD. When the merged series of seasonal rainfall (winter and spring separately) and merged climate modes were produced, the relationship between these significant phases of climate modes and their related seasonal rainfall could be examined. For this purpose Pearson correlations between the merged ENSO indices and rainfalls were calculated, and the results are shown in Table 4.10. The results of previous analysis with un-merged rainfalls are also shown for comparison. It can be seen from Table 4.10 that merged rainfalls show much higher correlation coefficients (r_{max}= 0.35 for east, r_{max}= 0.58 for central and west) for spring rainfall compared to the correlation

coefficients of normal rainfalls (r_{max}= - 0.27 for east, r_{max}= 0.43 for central and west) with ENSO indices. In this way the strength of the two phases of ENSO (El Nino/ La Nina) on spring rainfall in Victoria is shown. For winter rainfalls, Table 4.10 shows that merged rainfalls show much higher correlation coefficients (r_{max}= 0.55 for central and r_{max}= 0.44 for west) than the correlation coefficients of normal rainfalls (r_{max}= 0.43 for central and r_{max}= 0.38 for west) with ENSO indices. It can be seen that even by combining the years of El Nino/La Nina, east Victoria does not show any significant relationship with the two phases of ENSO in winter.

The same analysis was conducted for the phases of IOD and the results are shown in Table 4.11. It can be seen from this table that merged rainfalls based on the two IOD phases show higher correlation coefficients (r_{max}= -0.33 for east, r_{max}= -0.58 for central and r_{max}= -0.54 for west) for spring rainfall than the un-merged rainfalls, with r_{max}= -0.25 for east, r_{max}= -0.44 for central and r_{max}= -0. 45 for west. For winter, the results show higher correlation coefficients for merged rainfalls compared to normal rainfalls (r_{max}= -0.54 for central and r_{max}= -0.43 for west compared to r_{max}= -0.33 for central and r_{max}= -0.29 for west). It appears that even by combining the winter rainfall for east Victoria based on pIOD and nIOD, no relationship can be found between rainfall and winter IOD in this region.

Table 4.10. Comparison between Pearson correlations (r) of merge and un-merge rainfalls with ENSO indices

	Merged El Nino/La Nina rainfalls				Un- Merged rainfalls			
Region	**Spring**		**Winter**		**Spring**		**Winter**	
	Nino3.4	**SOI**	**Nino3.4**	**SOI**	**Nino3.4**	**SOI**	**Nino3.4**	**SOI**
East	-0.33*	0.35*	----	----	-0.27**	0.21*	----	----
Centre	-0.50**	0.58**	-0.32*	0.55**	-0.38**	0.43**	-0.21*	0.43**
West	-0.50**	0.58**	----	0.44**	-0.39**	0.43**	----	0.38**

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

Table 4.11. Comparison between Pearson correlations (r) of merged and un-merge rainfalls with IOD indices

	Merged pIOD/nIOD rainfalls		Un-merged rainfalls	
Region	**Spring**	**Winter**	**Spring**	**Winter**
	IOD	**IOD**	**IOD**	**IOD**
East	-0.33*	----	-0.25**	----
Centre	-0.58**	-0.54**	-0.44**	-0.33**
West	-0.54**	-0.43**	-0.45**	-0.29**

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

4.2.2 Lagged relationships

As discussed in Sections 4.1 and 4.2.1, the season that is most influenced by ENSO and IOD in Victoria is spring. Therefore, the development of a forecast model to accurately predict spring rainfall based on large-scale climate modes is crucial for Victoria. Any modelling process requires the identification of the predictors; in rainfall forecasting models it is the past values of the large scale climate modes that are essential for forecasting rainfall. In order to evaluate the significant predictors of rainfall, Pearson correlation analysis was conducted between monthly climate modes prior to spring and spring rainfall in three regions of the case study; Pearson correlations between spring rainfall at year *n* and the Dec_{n-1}-Aug_n monthly values of ENSO and IOD indicators (Nino3.4, SOI and DMI) were calculated ("*n"* being the year for which spring rainfall is being predicted). As discussed in Chapter 3, for each region three stations were chosen based on the data records. The results of correlation analyses between rainfall and lagged climate modes are shown in Table 4.12. It can be seen from this table that only the three months (June, July and August) Nino3.4, SOI and DMI have statistically significant correlations with spring rainfalls; this result is consistent with the findings of Chiew et al. (1998) and Verdon et al. (2004), substantiating that not only the highest correlations between rainfall and climate indicators are obtained up to three month lags i.e. there is no further significant relationship after lag 3; in addition, these correlations are very weak for Victoria ($|r_{max}|$=0.30 for east Victoria, $|r_{max}|$=0.39 for central Victoria and $|r_{max}|$=0.36 for west Victoria). As these relationships are weak based on linear

analysis, a sophisticated modelling approach needs to be applied in order to forecast rainfall.

Table 4.12. Pearson correlations (r) of lagged climate indices and spring rainfall

Region	Station	Lagged climate indices								
		$Ni34_{(Jun)}$	$Ni34_{(Jul)}$	$Ni34_{(Aug)}$	$SOI_{(Jun)}$	$SOI_{(Jul)}$	$SOI_{(Aug)}$	$DMI_{(Jun)}$	$DMI_{(Jul)}$	$DMI_{(Aug)}$
	Bruthen	-0.20*	-0.25**	-0.28**	---	---	---	-0.25**	---	---
East	**Buchan**	-0.22*	-0.26**	-0.24*	-0.20*	---	---	-0.30**	---	---
	Orbost	---	-0.24*	-0.26**	---	---	---	-0.29**	-0.21*	---
	Malmsbury	-0.22*	-0.22*	-0.29**	---	0.32**	0.30**	---	-0.30**	-0.31**
Centre	**Daylesford**	-0.30**	-0.28**	-0.33**	0.20*	0.37**	0.34**	---	-0.29**	-0.28**
	Heathcote	-0.30**	-0.30**	-0.38**	---	0.36**	0.39**	---	-0.25**	-0.28**
	Horsham	-0.22*	-0.23*	-0.31**	----	0.26**	0.25**	---	-0.28**	-0.31**
West	**Kaniva**	-0.32**	-0.32**	-0.36**	0.23*	0.33**	0.31**	---	-0.30**	-0.31**
	Rainbow	-0.31**	-0.31**	-0.36**	0.20*	0.33**	0.33**	---	-0.25**	-0.26**

**: correlation is significant at the 0.01 level (2-tailed).
* : correlation is significant at the 0.05 level (2-tailed).

In addition to ENSO and IOD, IPO and SAM also affect some parts of Victoria. SAM was not used in this study due to the limited data records. It is believed that IPO enhances the effect of ENSO during the La Nina phase for some parts of Victoria (Verdon et al. 2004). Pearson correlation analysis was conducted to find the relationships between monthly IPO values prior to spring and spring rainfall for the three regions. Table 4.13 shows the result of the correlation analyses; as can be seen in Table 4.13, of the nine stations, only four have significant, although weak, lagged relationships between IPO and rainfall. Daylesford in central Victoria correlates with IPO $_{(Jun, Jul, Aug)}$ with ($|r_{max}|$) of 0.24, while Heathcote and Kaniva only have significant correlation with $IPO_{(Jul, Aug)}$ with ($|r_{max}|$) of 0.25. Rainbow has only a significant correlation with IPO $_{(Aug)}$. It can be seen from Table 4.13 that not only these lagged correlations weak, they also do not cover a common and wide range of months and stations. In general, it appears that due to the multi-decadal variability of IPO, a weak

relationship exists between monthly IPO and seasonal rainfall for only some parts of Victoria.

Table 4.13. Pearson correlations (r) of lagged IPO and spring rainfall

Region	Station	Lagged IPO		
		June	**July**	**August**
East	Bruthen	----	----	----
	Buchan	----	----	----
	Orbost	----	----	----
Centre	Malmsbury	----	----	----
	Daylesford	-0.19*	-0.24*	-0.23*
	Heathcote	----	-0.25*	-0.20*
West	Horsham	----	----	----
	Kaniva	----	-0.23*	-0.22*
	Rainbow	----	----	-0.21*

**: correlation is significant at the 99% level
* : correlation is significant at the 95% level

4.2.3 Summary of Pearson correlation analysis

Pearson correlation analysis was conducted for concurrent seasonal rainfalls in Victoria and seasonal climate modes. It was discovered that only the cool season rainfalls (spring and winter) have statistically significant relationships with climate modes; it was also found that these relationships are moderate to weak. Furthermore, by merging the years of El Nino/La Nina, and in a separate analysis the years of pIOD/nIOD, the strength of the relationship between rainfall and the phases of ENSO and IOD was examined. The results showed increased correlation coefficients considering the merged phases of ENSO and IOD with spring and winter rainfall for Victoria which revealed the magnitude of the strength of the relationships. The relationship between antecedent climate modes and spring rainfall was also examined using Pearson correlation analysis. It was discovered that only the climate modes of the three months June, July and August

have statistically significant relationships with spring rainfall, although these relationships are quite weak.

Chapter 5

Multiple Linear Regression (MLR) Analysis

5.1 Introduction

Multiple linear regression (MLR) models are commonly used for the prediction of water resources and hydrological variables. Many studies of rainfall forecasting (He et al. 2014; Nicholson 2014; Mekanik et al. 2013) and flood forecasting (Latt et al. 2014; Chavoshi et al. 2013) have used MLR as a tool for prediction purposes. Rossel and Cadier (2009) applied multiple regression models for the prediction of monthly rainfall in the Ecuador, using as predictors precipitation, sea surface temperature (SST), meridional and zonal wind in the eastern equatorial Pacific. The developed models were used to predict rainfall anomalies in the out-of-sample test set. They discovered that there is significant predictive power for the rainy months of the year with the best predictability being for the period from March to May. Their revealed that the developed multiple linear models explained 60–82% of the monthly precipitation variance. The relationship between topography and Korean precipitation was examined by MLR models (Um et al. 2011). Sadhuram and Murthy (2008) used MLR to forecast Indian summer monsoon rainfall using SST anomalies in Indian Ocean. They noted that the developed model was able to forecast the rainfall with good accuracy. Ihara et al. (2007) examined the relationship between ENSO and Indian Ocean indices with Indian summer monsoon rainfall using MLR. They discovered that the combination of Nino3 and the zonal wind anomalies over the equatorial Indian Ocean are good predictors of rainfall in the region. In the present study, MLR models are developed as a benchmark for comparison with the developed ANN models.

5.2 Methodology

MLR is a linear statistical technique that enables finding the best relationship between a variable (dependent, predicant) and several other variables (independent, predictor) through the least square method. Multiple regression models can be presented by the following equation:

$$O_i = a + b_1 X_{1i} + b_2 X_{2i} + c \quad (5\text{-}1)$$

where, O_i is the dependent variable , X_{1i} and X_{2i} are first and second independent variable respectively , b_1 and b_2 are model coefficients of first and second independent variable respectively, a is constant, and c is the error.

regression weights (b_1 and b_1) are computed in a way to minimize the sum of squared deviations defined by the following equation:

$$\sum_{i=1}^{N} (x_i - O_i)^2 \quad (5\text{-}2)$$

MLR models are linear statistical methods that are based on established statistical assumptions. In the process of building regression models, the related assumptions must be evaluated and satisfied before the models can be reported as reliable. Among the most important statistical assumptions is the evaluation of the goodness-of-fit of the model and the statistical significance of the estimated parameters of the constructed regression model. The techniques commonly used to verify the goodness-of-fit of regression models are hypothesis testing, R-squared and analysis of the residuals. In the present study, for the purpose of evaluating the goodness-of-fit of the models, the F-test was used to verify the statistical significance of the overall fit.

The next statistical criterion that needs to be satisfied while developing a MLR model is evaluation of the statistical significance of the individual parameters of the model. Each parameter in Eq. 5-1 needs to be statistically significant within the 95% confidence level; the test used for this purpose is the t-test, which evaluates the significance of the individual parameters. While the t-test tests the importance of individual coefficients, the F-test is used to compare different models to evaluate the model that best fits the population of the sample data (Um et al. 2011).

Verifying the multicollinearity is also an important stage in MLR modelling; Multicollinearity occurs when the predictors are highly correlated, which will result in dramatic change in parameter estimates in response to small changes in the data or the

model. The indicators used to identify multicollinearity among predictors are tolerance (T) and variance inflation factor (VIF):

$$\text{Tolerance} = 1 - R^2 \quad \text{and} \quad \text{VIF} = \frac{1}{\text{Tolerance}} \tag{5-3}$$

where, R^2 is the coefficient of multiple determination :

$$R^2 = \frac{\text{ssr}}{\text{sst}} = 1 - \frac{\text{sse}}{\text{sst}} \tag{5-4}$$

where, sst is the total sum of squares, ssr is the regression sum of squares and sse is the error sum of squares. According to Lin (2008), a tolerance of less than 0.20–0.10 or a VIF greater than 5–10 indicates a multicollinearity problem.

After the model is developed, the residuals should be examined in order to evaluate the independence of the errors of the models. The statistical test used to evaluate residual independence is the Durbin-Watson test (DW), which tests the existence of serial correlations between the model errors. Field (2009) states that values less than 1 or greater than 3 are a matter of concern when the DW test is applied.

5.3 Results and Discussion

According to Schepen et al. (2012) potentially suitable predictors for seasonal rainfall totals are lagged oceanic and atmospheric climate indices. In the present study, MLR was used in order to find the relationship between the significant lagged rainfall predictors (i.e. Nino3.4, SOI and DMI) discussed in Section 4.2.2 and spring rainfall. IPO will be discussed in Chapter 6. ENSO-IOD input sets were organized based on the statistically significant months (June, July and August) as potential predictors of spring rainfall for multiple regression analysis. For the purpose of calibration of the models, the years 1900-2006 were considered, and the reliable models were then tested on an out-of-sample test set of 2007-2009. The models were developed using SPSS software (IBM SPSS-2012). Table 5.1 shows a sample of the models that were developed. It can be seen from this table that the combination of ENSO-IOD is based on a single month for each climate indicator (e.g. $\text{Nino3.4}_{(\text{Jun})}$ –$\text{DMI}_{(\text{July})}$) and multiple lagged models (e.g.

Nino3.4($_{\text{Jun-Jul-Aug}}$) – DMI($_{\text{Jun-July-Aug}}$) are not shown in this table. After evaluating all the potential combinations of multiple lagged climate modes, it was discovered that only single lagged climate mode models (e.g. Nino3.4$_{(Jun)}$ –DMI$_{(July)}$) produced statistically significant results. The other models were not statistically reliable and are therefore not reported here. The non-significant models' results of this study are stored at Swinburne University of Technology library and can be retrieved (http://hdl.handle.net/1959.3/355557). A sample of a single lagged MLR is as follows:

$$Rainfall_{(spring)} = -0.24 \times Nino3.4_{(Aug)} + -0.20 \times DMI_{(July)} + 0.55 \quad (5\text{-}5)$$

Table 5.1. Multiple regression model sets developed for each station

	Nino3.4$_x$-DMI$_y$	**SOI$_x$-DMI$_y$**
Bruthen	Jun-Jun, Jul-Jun, Aug-Jun	--------
Buchan	Jun-Jun, Jul-Jun, Aug-Jun	Jun-Jun
Orbost	Jul-Jun, Jul-Jul, Aug-Jun, Aug-Jul,	--------
Malmsbury	Jun-Jul, Jun-Aug, Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug	Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug
Daylesford	Jun-Jul, Jun-Aug, Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug	Jun-Jul, Jun-Aug, Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug
Heathcote	Jun-Jul, Jun-Aug, Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug	Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug
Horsham	Jun-Jul, Jun-Aug, Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug	Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug
Kaniva	Jun-Jul, Jun-Aug, Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug	Jun-Jul, Jun-Aug, Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug
Rainbow	Jun-Jul, Jun-Aug, Jul-Jul, Jul-Aug, Aug-Jul, Aug-Aug	Jun-Jul, Jun-Aug, Jul-Jul, Jun-Aug, Aug-Jul, Aug-Aug

Note:" x-y" are the lagged months of the climate indices

As discussed earlier, the MLR technique has some assumptions that need to be satisfied before any model can be reported as reliable. In order to verify the statistical significance of the overall fit, the F-test needs to be conducted. The F-test is used to compare different models in order to evaluate the model that best fits the population of the sample data (Um et al. 2011). After the model significance is tested, the significance

of individual predictors and their related coefficients in Eq. 5-4 needs to be investigated using a t-test; in this way the importance of individual coefficients can be evaluated. After investigating the statistical significance of the parameters and the goodness-of-fit of the models, the indicators used to identify multicollinearity among predictors, tolerance (T) and variance inflation factor (VIF), were examined as discussed in Section 5.2. To evaluate the independence of the errors of the models, the Durbin-Watson test (DW), which tests the serial correlations between errors, was applied. According to Lin (2008), a tolerance of less than 0.20–0.10 or a VIF greater than 5–10 indicates a multicollinearity problem. According to Field (2009), values less than 1 or greater than 3 are of concern when the DW test is applied. It was discovered that, of the models with more than a single predictor for each climate mode (e.g. $Nino3.4_{(Jun\text{-}Jul\text{-}Aug)}$-$DMI_{(Jun\text{-}Jul\text{-}Aug)}$), the statistical assumptions were not satisfied, i.e. the t-test and VIF were not statistically significant. In addition, it was discovered that, of the models with single predictors (e.g. $Nino3.4_{(Aug)}$-$DMI_{(Aug)}$), only some of the models satisfied the statistical criteria. After examining all four major statistical factors (i.e. F-test, t-test, multicollinearity and DW statistics) the models that did not violate the limits of statistical significance were selected and models with lowest errors were chosen as the best model for each station. The regression coefficients, variance inflation factors (VIF), Durbin-Watson statistics (DW) and the Pearson correlations (r) of the best models are shown in Table 5.2. It can be seen from this table that the VIFs for the selected models are near one, i.e. there is no multicollinearity among the predictors; in addition, the DW statistics show that the residuals of the models have no autocorrelation, confirming the goodness-of-fit of the models. The Nino3.4-DMI based models proved to be statistically significant and to have better forecasting ability than the SOI-DMI models for Victoria, with a maximum Pearson correlation coefficient of 0.35 for east Victoria, 0.37 for central Victoria and 0.39 for west Victoria. Table 5.3 shows the MSE, MAE, RMSE and Pearson correlations (r) of the best MLR models for the three regions. It can be seen from Table 5.3 that the errors are relatively low for all the stations. The statistically significant models were then tested on an out-of-sample test set, and the results are shown in Table 5.4. This table indicates that the models are able to forecast rainfall for three consecutive years with good correlation coefficients for Malmsbury, Daylesford

and Kaniva. For the rest of the stations, the models do not show good r values. The regression models developed are shown in Figures 5.1to 5.3. These models will be further compared with ANN models.

Table 5.2. Summary of the best regression models

Region	Station	Models	Coefficient										r	VIF	DW
			Const.	$Nino34_{(Jun)}$	$Nino34_{(Jul)}$	$Nino34_{(Aug)}$	$SOI_{(Jun)}$	$SOI_{(Jul)}$	$SOI_{(Aug)}$	$DMI_{(Jun)}$	$DMI_{(Jul)}$	$DMI_{(Aug)}$			
East	Bruthen	$Ni34_{(Jul)}$-$DMI_{(Jun)}$	0.65	---	-0.24	---	---	---	---	-0.24	---	---	0.32	1.10	1.90
	Buchan	$Ni34_{(Jul)}$-$DMI_{(Jun)}$	0.51	---	-0.17	---	---	---	---	-0.23	---	---	0.35	1.10	2.10
	Orbost	$Ni34_{(Aug)}$-$DMI_{(Jun)}$	0.56	---	---	-0.20	---	---	---	-0.27	---	---	0.35	1.10	2.00
Centre	Malmsbury	$Ni34_{(Aug)}$-$DMI_{(Jul)}$	0.55	---	---	-0.20	---	---	---	---	-0.22	---	0.36	1.12	1.90
	Daylesford	$Ni34_{(Jun)}$-$DMI_{(Jul)}$	0.62	-0.25	---	---	---	---	---	---	-0.29	---	0.37	1.10	1.81
	Heathcote	$Ni34_{(Jun)}$-$DMI_{(Aug)}$	0.60	-0.29	---	---	---	---	---	---	---	-0.24	0.37	1.10	1.80
West	Horsham	$Ni34_{(Aug)}$-$DMI_{(Jul)}$	0.55	---	---	-0.24	---	---	---	---	-0.20	---	0.36	1.12	2.00
	Kaniva	$Ni34_{(Jun)}$-$DMI_{(Jul)}$	0.67	-0.32	---	---	---	---	---	---	-0.27	---	0.39	1.10	2.00
	Rainbow	$Ni34_{(Aug)}$-$DMI_{(Jun)}$	0.56	-0.29	---	---	---	---	---	---	---	-0.20	0.36	1.10	2.25

Table 5.3. Performance of the regression models: calibration set

Region	Station	r	RMSE	MSE	MAE
East	Bruthen	0.32	0.22	0.048	0.171
	Buchan	0.35	0.16	0.026	0.171
	Orbost	0.35	0.19	0.038	0.157
Centre	Malmsbury	0.36	0.17	0.030	0.140
	Daylesford	0.37	0.20	0.039	0.155
	Heathcote	0.37	0.19	0.035	0.153
West	Horsham	0.36	0.18	0.033	0.149
	Kaniva	0.39	0.20	0.041	0.163
	Rainbow	0.36	0.18	0.031	0.142

Table 5.4. Performance of the regression models: test set

Region	Station	r	RMSE	MSE	MAE
East	Bruthen	-0.99	0.13	0.016	0.085
	Buchan	-0.90	0.15	0.023	0.180
	Orbost	-0.99	0.15	0.024	0.150
Centre	Malmsbury	0.48	0.11	0.013	0.100
	Daylesford	0.92	0.21	0.043	0.205
	Heathcote	-0.50	0.16	0.026	0.158
West	Horsham	0.25	0.17	0.030	0.149
	Kaniva	0.67	0.23	0.051	0.163
	Rainbow	-0.74	0.17	0.029	0.142

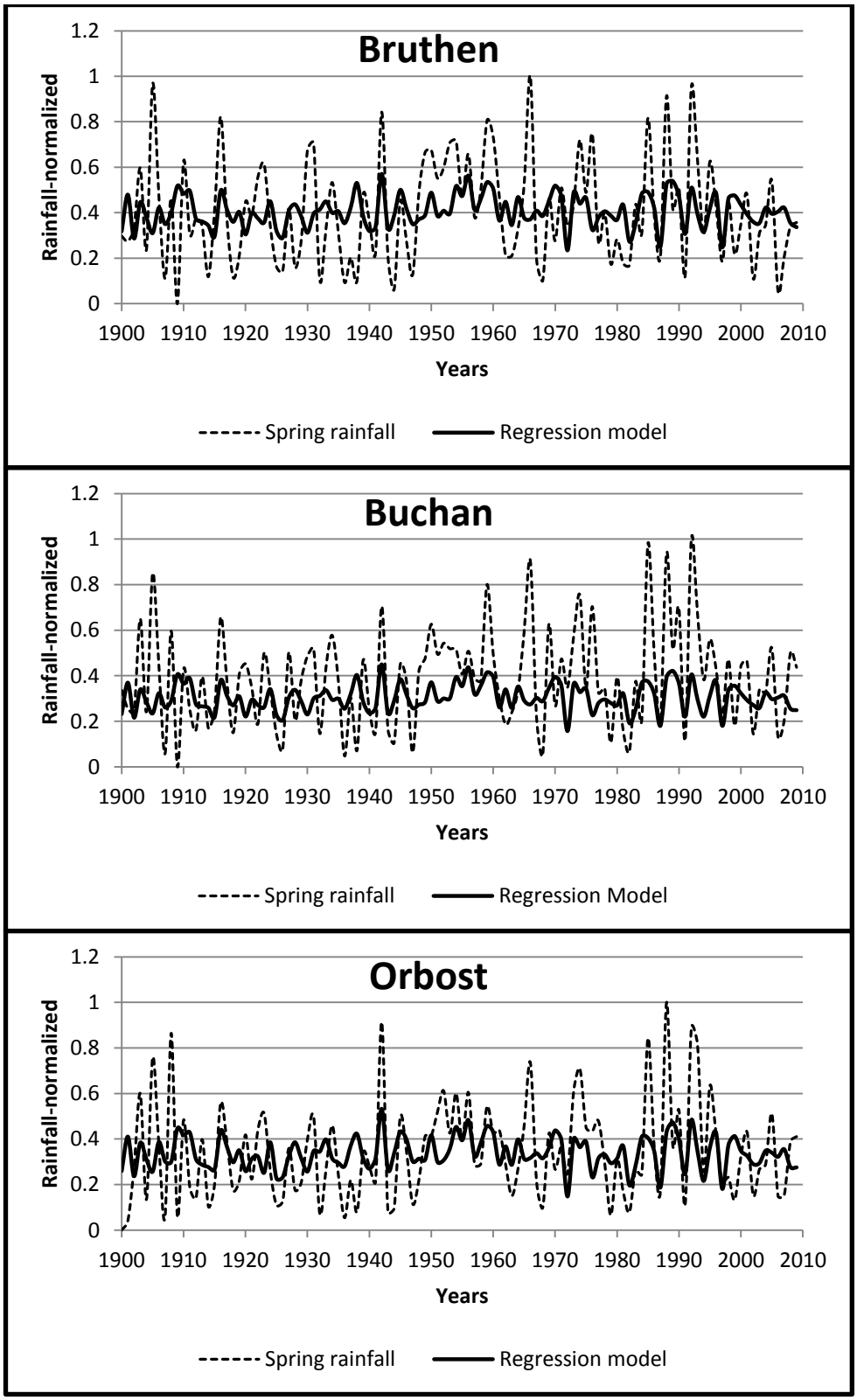

Figure 5-1. Multiple regression models for east Victoria

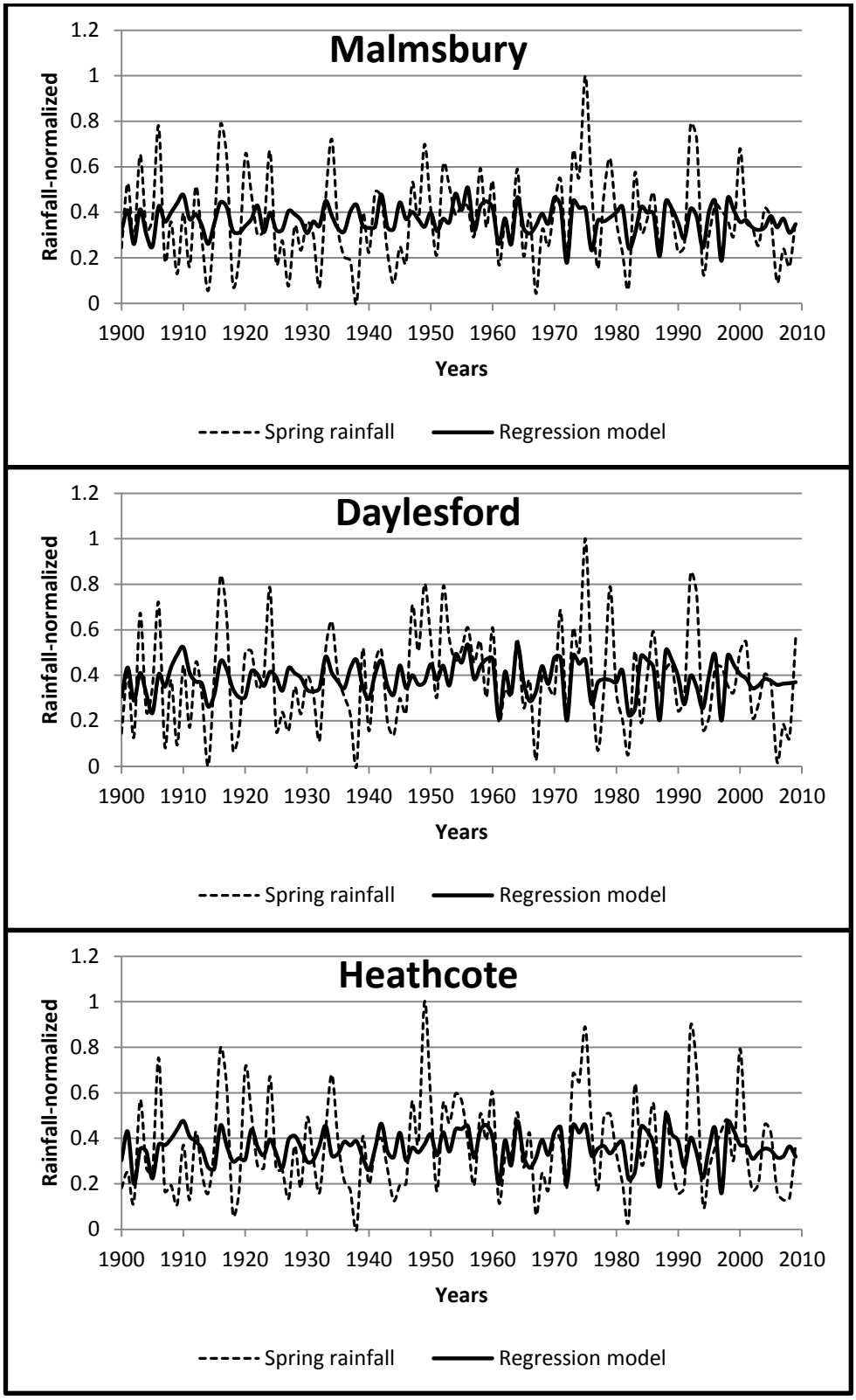

Figure 5-2. Multiple regression models for central Victoria

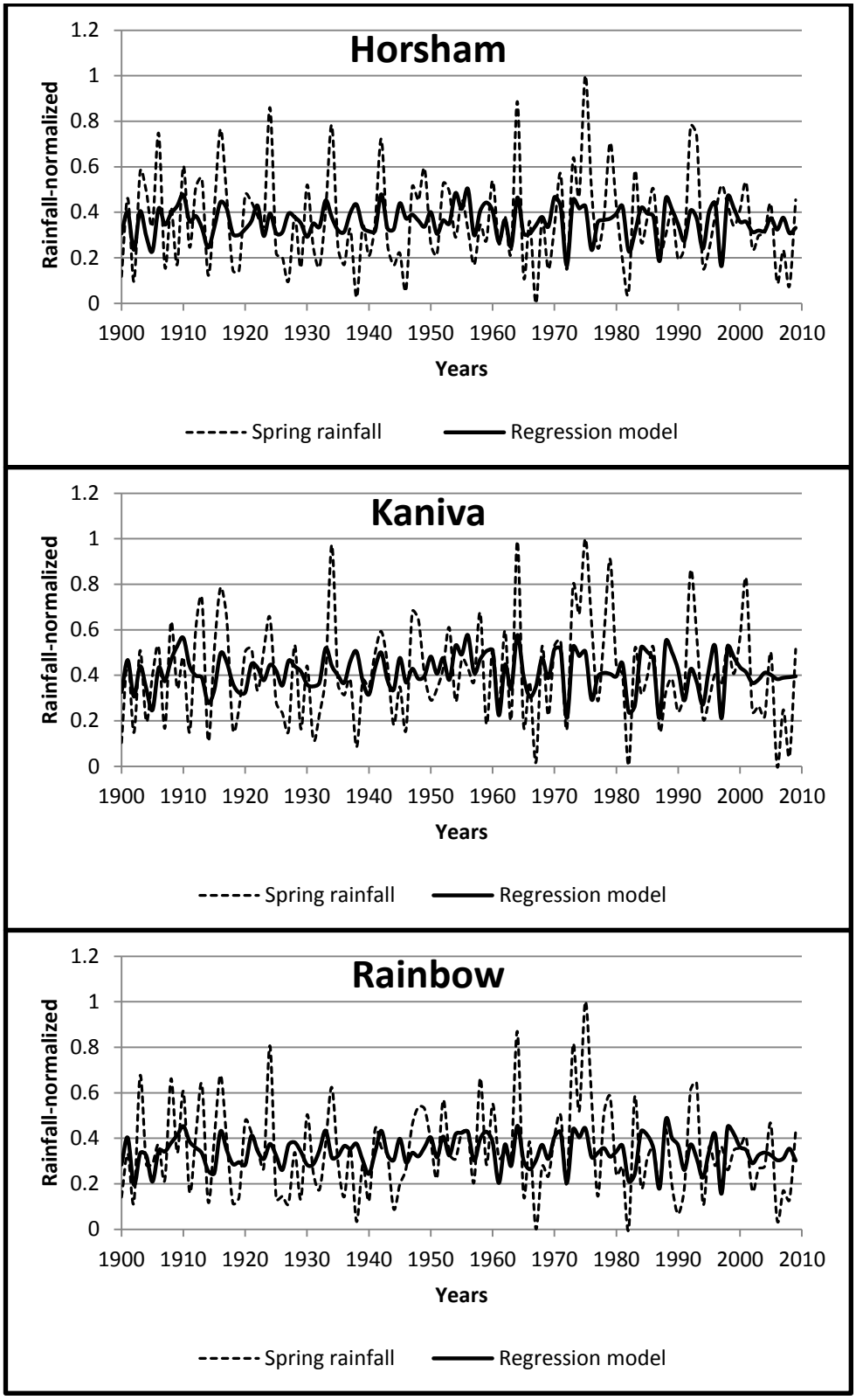

Figure 5-3. Multiple regression models for west Victoria

5.4 Summary of MLR analysis

In this section, multiple linear regression (MLR) analysis was conducted in order to produce rainfall forecast models based on lagged climate modes as predictors. It was discovered that only a few of the single lagged models were statistically significant and were used to forecast rainfall. The multiple lagged climate mode models were not reliable, as they did not meet the statistical assumptions of the MLR technique. Due to the limitations of the MLR technique, care needs to be taken while developing forecasting models; if the statistical assumptions of the MLR technique are not examined and are not satisfied, the models developed could be misleading. The MLR models developed and tested were able to forecast spring rainfall three consecutive years in advance for some stations with acceptable correlation coefficients, while showing less accurate results for the others. It should be noted that MLR models are linear techniques that cannot model nonlinear complex relationships. This is the main drawback of MLR models. Therefore, the use of nonlinear techniques with fewer limitations, such as Artificial Neural Networks (ANNs), is proposed and will be discussed in the following sections.

Chapter 6

Artificial Neural Networks

6.1 Introduction

Many probabilistic and deterministic modelling approaches have been used by hydrologists and climatologists in order to capture rainfall characteristics. Conceptual and physically-based models require an in-depth knowledge of complex atmospheric phenomena; these models need a large amount of calibration data and they have to deal with over-parameterisation effects and parameter redundancy (De Vos & Reintjas, 2005). On the other hand, due to inaccurate initial conditions, limited spatial resolution, and parameterization schemes of subscale phenomena rainfall forecasts with use of numerical weather prediction (NWP) models is far from being satisfactory (Ramirez et al. 2005). As discussed earlier, the use of numerical models has not demonstrated useful performance in rainfall prediction. The NWP forecasts obtained from these models are an average over grid point values, and therefore is a function of the model's spatial resolution only. in this case neglecting the temporal variation will lead to consistent inaccuracies since rainfall is highly variable both in space and time. (Ramirez et al., 2005). These problems lead to investigating the potential of data-driven techniques. A variety of new approaches in hydrological modelling has opened since the introduction of artificial intelligence-based models such as Artificial Neural Networks (ANNs). A comprehensive review on application of ANN in hydrology has been provided by Maier and Dandy (2000) and Maier et al., (2010); the reviews highlights the potential of ANN as alternative modelling tools worthy of further exploration.

ANNs are mathematical models that has the ability to find nonlinear relationships between input and output parameters without the need to solve complex partial differential equations (Yilmaz et al. 2011). ANNs have been used in many hydrological and meteorological applications; for rainfall-runoff modelling (Akhtar et al. 2009; Chiang & Chang, 2009; Chiang et al. 2004; De Vos & Rientjes, 2005; Sudheer et al. 2002; Tokar & Johnson, 1999); for stream-flow forecasting (Campolo et al. 1999; Firat & Gungor, 2007; Kisi 2007; Riad et al. 2004; Turan & Yurdusev, 2009), and for

groundwater modelling (Coulibaly et al. 2001; Daliakopoulos et al. 2005; Rogers & Dowla, 1994). They have also been used for rainfall forecasting (Hsu et al. 1995; Luk et al. 2000 and 2001; Mekanik et al. 2011; Rami'rez et al. 2005; Toth et al. 2000).

Two major advantages of neural networks are (1) their ability to represent both linear and nonlinear relationships and (2) to learn these relationships directly from the data which they are modelling (Abbot & Marohasy, 2014). The use of ANN can be seen frequently in modelling and forecasting Indian monsoon rainfall. Iyengar and Raghu Kanth (2005) showed that the nonlinear characteristic of Indian monsoon rainfall can be modeled using ANN techniques, whereas the linear part is amenable for modelling through simple regression concepts. It was found that the proposed model explained between 75 to 80% of the interannual variability of eight regional rainfall series considered in the study. It was demonstrated that the model was capable of foreshadowing the drought of 2002, with the help of only antecedent data. Chakraverty and Gupta (2008) also predicted Indian summer monsoon rainfall 6 years in advance with the use of ANN. Different networks was constructed and compared with the results of previous studies. It was revealed that the developed ANN model had better results than those of previous studies.

While many studies have used antecedent rainfall values to forecast rainfall using ANN modelling technique, some studies around the world have taken into account the large scale climate variables as ANN inputs. Chattopadhyay (2007) used a three layer ANN with backpropagation learning for predicting the average summer monsoon rainfall over India. Nine predictors consisting of the monthly summer monsoon rainfall totals, tropical rainfall indices and sea surface temperature anomalies were used as inputs for the ANN model. The results showed supremacy to the persistence forecast and MLR prediction. Chattopadhay et al. (2010) predicted the average winter monsoon rainfall of a given year using exponential regression equation and ANN model. Substantial variables were selected and the rainfall amounts and the SST anomalies in the winter monsoon months of the previous year were chosen as predictors. Levenberg-Marquardt algorithm was used to produce the regression coefficients for the multiple exponential regression equation. The ANN in the form of a multilayer perceptron with sigmoid non-

linear function and genetic-algorithm based variable selection was developed. The Willmott's index, percentage error of prediction, and prediction yields were used to evaluate the performance of the predictive models; the potential of ANN over exponential regression was then demonstrated.

Shukla et al. (2011) further tried to improve the seasonal forecast skill of the Indian Summer Monsoon Rainfall Index (ISMRI) using ANN approach. In this regard, the influence of SST indices of Nino1+2, Nino3, Nino3.4 and Nino4 regions on ISMRI were evaluated using correlation analysis with a lag period of 1–8 seasons. Considerable positive correlations were found between ISMRI and Nino3 and Nino3.4 indices with a lag of 4 (June–July–August) and 5 (March–April–May) seasons and Nino4 index with a lag of 5 seasons before the start of monsoon. They reported that the level of confidence for the correlations was above 99%. Multiple linear regression models was also used to predict ISMRI using SST indices. Comparison of the results revealed the superior prediction skills of the ANN models to all the linear regression models. Based on the results of the developed ANN models, it revealed that the relationship between the Nino indices and the ISMRI is non-linear in nature.

Other than India, many studies has focused on the application of ANN for rainfall prediction around the world. French et al. (1992) represented an ANN model which consisted of a three layer neural network to forecast rainfall intensity in space and time. They compared the results of their model with a space-time mathematical rainfall simulation model. They discovered that neural networks are capable of learning the complex nonlinear relationships of rainfall and they perform well in multi-site rainfall forecasting. Later in 1998, Lee et al. (1998) proposed a divide-and-conquer approach for predicting the daily rainfall at 367 locations in Switzerland based on the daily rainfall at nearby 100 locations. Based on their approach, the region was split into four sub-areas and for each sub-area different techniques were proposed. Based on the location information, radial basis function (RBF) neural networks were used for the two larger areas. Since it was assumed that for the two smaller areas precipitation is implemented by the orographic effect, therefore a simple linear regression model which used elevation as its only information was applied. For the two larger areas RBF

networks produced acceptable predictions while for the smaller areas the linear regression models produced large errors. Hartmann et al. (2008) forecasted summer rainfall in the Yangtze River basin using ANN technique. The input variables for the ANN included the SOI, the East Atlantic/Western Russia (EA/WR) pattern, the Scandinavia (SCA) pattern, the Polar/Eurasia (POL) pattern and several indices calculated from SST, sea level pressures (SLP) and snow data from December to April of 1993 to 2002. The precipitation from May to September of 1994 to 2002 was selected as the output variable of the ANN model. Using a principal component analysis (PCA), the output variable was classified into six different regions and rainfall was forecasted from May to September 2002. The results revealed that winter SST and SLP indices are the most significant predictors of summer rainfall in the Yangtze River basin. However, it seemed that the Tibetan Plateau snow depth, the SOI and the other teleconnection indices did not have significant effect on accurate rainfall forecast. The authors indicated that this could be attributed to the length of the available time series, which does not permit a deeper analysis of the influence of multi-annual oscillations. The results proved the capability of the ANN algorithms in forecasting most of the rainfall variability in the Yangtze River basin. The ANN model was able to show more than 77% of the total variance of the measured rainfall for five out of the six investigated regions.

Multilayer perceptron is not the only type of ANN being used in rainfall forecasting. Lin and Wu (2009) proposed a hybrid Self Organizing Map (SOM) and multilayer perceptron network to forecast typhoon rainfall. SOM technique was used for input selection while multilayer perceptron was used for the purpose of training and forecasting. The model was applied to Tanshui River Basin for the purpose of forecasting one hour a head typhoon rainfall for ten rain gauges. It was concluded that the proposed model has advantage over the conventional model used to forecast typhoon rainfall. Luk et al. (2001) conducted three types of ANN namely multi layer feed-forward neural network (MLFN), Elman partial recurrent neural network (Elman) and time delay neural network (TDNN) to forecast rainfall for an urban catchment. They showed that all the mentioned networks have the ability to forecast rainfall 15

minutes ahead for 16 rain gauges. The results showed that there is a relation between the optimal complexity and the number of the hidden nodes and the lag of the series. It was concluded that networks with lower lags outperforms the ones with higher lags which reveals the short term memory characteristics of rainfall at the mentioned rain gauges.

As discussed in Chapter 2, the application of nonlinear techniques for rainfall forecasting using antecedent climate modes has not widely been explored for Australia. Among the few studies focusing on this aspect of rainfall forecasting only three studies have used ANN modelling approach; Abbot and Marohasy (2012) developed ANN models to forecast monthly and seasonal rainfall in Queensland, Australia. The input sets considered for their model were climate indices, monthly historical rainfall data, and atmospheric temperatures. They compared their forecast results in regards to RMSE values with forecasts generated by the Australian Bureau of Meteorology's Predictive Ocean Atmosphere Model for Australia (POAMA)-1.5 general circulation model (GCM). It was concluded that the prototype achieved a lower RMSE for 16 of the 17 sites compared. In 2014, Abbot and Marohasy expanded their study and introduced the IPO into their model. After input selection and developing the ANN models, the results of the models were compared with POAMA and was found that for Queensland, ANN models outperform the forecasts obtained by POAMA model (Abbot & Marohasy 2014). Mekanik et al. (2013) also developed ANN models for spring rainfall forecast in Victoria, southeast Australia. The results of ANN models were compared with Multiple Linear Regression (MLR) models and they were proved to be superior to the developed linear models.

6.2 Methodology

ANN has been inspired by biological neural networks; it consists of simple neurons and connections that process information in order to find a relationship between inputs and outputs. ANN is generally classified based on the direction of the flow of information into two classes: feedforward networks and recurrent networks. Abbot and Marohasy (2012) classified neural network rainfall models into three approaches: function models, time series models and classification models. Function models are static models, which

use sets of inputs with the aim of forecasting a corresponding rainfall value. Time series networks forecast rainfall over discrete intervals of time, and classification models are less used in rainfall prediction. In the present study, the developed models are a combination of static models in terms of model development and time series models in terms of input organization.

6.2.1 Determination of Network Architecture

Determining the network architecture is the most difficult task in the model building process, and includes determining the number of connection weights and the way information flows through a network. Feedforward networks have been used in a wide range of forecasting and prediction applications (Maier et al. 2012). In this type of network, neurons (nodes) in one layer are only connected to the neurons of the next layer without any backward connections. Another type of network is the recurrent network, where the connections of the neurons are not only to the next layer's neurons, but also to the previous layer, to the same layer and even to themselves. In feedforward networks, dynamic systems need to be treated explicitly by including lagged inputs, whereas in recurrent networks the system can be analyzed implicitly. The advantage of a feedforward network over a recurrent network is that it is capable of capturing long-term dependencies, and this characteristic is crucial when inputs at high lags have a significant effect on the network output (Maier & Dandy 2000). Multi Layer Perceptron (MLP) is now one of the most popular network architectures used by researchers. Figure 6.1 shows a schematic feedforward MLP network. As Figure 6.1 shows, three different layers construct a simple MLP. The first layer, known as the input layer, simply serves to introduce the values of input variables (in the case of this study the inputs are lagged climate modes) to the network. The number of nodes in this layer is fixed based on the number of inputs. The last layer is the output layer. The number of nodes in the output layer represents the number of the output of the model. The hidden layer, which serves as the nonlinear part of the network, is the layer between the input and the output layer. A network without a hidden layer constructs a linear relationship among the available input and output.

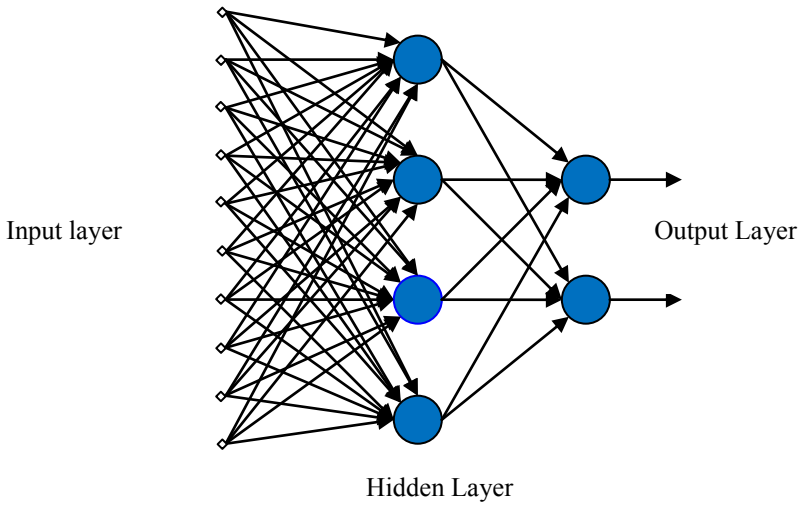

Figure 6-1. A typical feedforward MLP network with one hidden layer.

The number of hidden layers can vary according to the type of problem being solved. It has been shown that ANN with one hidden layer can approximate any function, given that sufficient degrees of freedom are provided (Hornik 1991).The number of hidden nodes is a critical aspect of the network geometry. There should be enough hidden nodes in order to enable representation of the function to be approximated. Too many hidden nodes on the other hand can result in over-fitting the network.

Traditionally, optimal network geometry has been found by trial and error, but recently some systematic approaches have been evolved, including pruning and constructive algorithms. The basic idea of a pruning algorithm is to start with a network that is large enough to capture the desired input-output relationship and to subsequently remove or disable unnecessary weights and/or nodes. A constructive algorithm approaches the problem of optimizing the number of hidden layer nodes from the opposite direction to pruning algorithm. The smallest possible network is used at the start of training; hidden layer nodes and connections are then added one at a time in an attempt to improve model performance. In this study, a constructive algorithm has been used.

6.2.2 Training (optimization)

Training a network is the process of optimizing the connection weights. This phase is equal to parameter estimation in conventional statistical models like regression. The mean square error (MSE) is the common error function used. Training can be done using either global or local methods. Local methods fall into two major categories: first-order and second-order. First-order methods are based on a linear model (Gradient descent), whereas second–order models are based on a quadratic model (e.g Newton's method). In both cases, iterative techniques are used to minimize the error function. ANN-based models in hydrology over the past years have shown that there is great interest in the use of multi-layer feedforward neural networks (FNNs) trained by the standard back-propagation (BP) algorithm. One of the major disadvantages of BP is its slow convergence, which leads to sub-optimal solution. This shortcoming makes FNN trained with BP fail to find solutions to even rather simple pattern classifications. Second-order nonlinear techniques are usually faster and more reliable than any BP variant (Coulibaly et al. 2000). Therefore in the present study, Levenberg-Marquardt back-propagation training (LMBP) for multi-layer feedforward neural network training was chosen in the model development phase. More details of the Levenberg-Marquardt algorithm are given in the next section.

6.2.3 Levenberg-Marquardt Back-Propagation Training

When prediction is of interest, it is common practice to provide input-output examples to a multi-layer feedforward neural network (FNN) and minimize the error function using either a first-order or second-order optimization method. This is called supervised training, and can be formulated as the sum of nonlinear least squares between the observed and the predicted outputs defined by:

$$E = \frac{1}{2}\sum_{p=1}^{n}\sum_{k=1}^{m}(y_{pk} - \hat{y}_{pk})^2 \qquad (6\text{-}1)$$

where, n is the number of observations (patterns) and m is the total output units, y is the target output and $\hat{y}$ is the predicted output. In our case, where there is only one output unit (m=1), Eq. 6-1 reduces to:

$$E = \frac{1}{2}\sum_{p=1}^{n}(y_p - \hat{y}_p)^2 \quad (6\text{-}2)$$

This is the common function which is minimized in least square regression. In the BP, minimization of E is done by the use of the steepest descent method. The gradient of the error function is then computed by applying the chain rule on the hidden layers of the FNN. Consider the multilayer feedforward neural network depicted in Figure 6.1. If the hidden layer consists of M neurons the network can be described mathematically by the following equations:

$$net_{pj} = \sum W_{ji} y_{pi} + W_{jo} \quad (6\text{-}3)$$

$$f(net_{pj}) = \frac{2}{(1+\exp(-2net_{pj}))} - 1 \quad (6\text{-}4)$$

where, net_{pj} is the weighted input into the jth hidden unit, n is the total number of input nodes, W_{ji} is the weight from input unit *i* to the hidden unit *j*, y_{pi} is the value of the ith input for pattern p, W_{jo} is the threshold or bias for neuron *j*, and $f(net_{pj})$ is the jth neuron's activation function, assuming that $f(.)$ is the tansigmoid function. Note that the input unit serves only for the receiving of the information and passing it to the hidden nodes, and does not perform any operation on the information. After the inputs have flown from the input layer to the hidden layer and the activation function in the hidden layer is applied, the results are flown to the output layer using the following equations:

$$net_{pk} = \sum_{j=1}^{M} W_{kj} \times f(net_{pj}) + W_{ko} \quad (6\text{-}5)$$

$$\hat{y}_{pk} = h(net_{pk}) \quad (6\text{-}6)$$

$$h(net_{pk}) = net_{pk} \quad (6\text{-}7)$$

where, M is the number of hidden units, W_{kj} represents the weight connecting the hidden node j to the output k, W_{ko} is the treshold value for neuron *k*, $\hat{y}_{pk}$ is the kth predicted output, and *h (.)* is the purelin activation function for the output layer.

Recall that the ultimate goal of the network training is to find the set of weights W_{ji} connecting the input units *i* to the hidden units *j* and W_{kj} connecting the hidden units *j* to output *k*, that minimize the objective function (Eq. 6-1). Since Eq. 6-1 is not an explicit function of the weights in the hidden layer, the first partial derivation of E is evaluated with respect to the weights using the chain rule and the weights are moved in the steepest-descent direction. This can be represented mathematically as:

$$\Delta W_{kj} = -\psi \frac{\partial E}{\partial W_{kj}} \tag{6-8}$$

The ψ is the learning rate, which simply scales the step size. ψ is chosen according to the relationship of $0 < \psi < 1$ and this is a common approach in BP training. Eq. 6-8 reveals that BP has the ability to suffer from inherent slowness and the local search nature of the first-order optimization method. However, BP is still the most widely used supervised training method for FNN. Overall, second-order nonlinear optimization techniques are generally more reliable. As a result, algorithms like the Levenberg-Marquardt (LMBP), which use the second derivative of E, are of more interest. The Levenberg-Marquardt algorithm is designed to achieve second-order training speed without the need to compute the Hessian matrix. When the performance function has the form of a sum of squares (as Eq. 6-1), then the Hessian matrix can be approximated as

$$H = J^T J \tag{6-9}$$

and the gradient can be computed as

$$g = J^T \mathbf{e} \tag{6-10}$$

where, *J* is the Jacobian matrix that comprises the first derivatives of the network errors with regard to the weights and biases, and **e** is the vector of network errors. The Jacobian matrix can be calculated based on a standard back-propagation technique that is much less complicated than computing the Hessian matrix. The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

$$w_{k+1} = w_k - [J^T J + \mu I]^{-1} J^T e \quad (6\text{-}11)$$

where, w is the weight vector and k is the index of iterations.

The Levenberg-Marquardt algorithm is faster and can deal better with a variety of problems than other usual methods. Therefore, in the present study LMBP was used for training the network.

6.2.4 Transfer (activation) function

Transfer functions are the main mathematical functions of a neural network. The use of linear transfer function is equivalent to using a single layer network. To gain the advantage of a multilayer network, a nonlinear function must be used. The sum of weighted inputs and bias produces the inputs to a transfer function. The most commonly used transfer functions are sigmoidal-type functions, such as the logistic or hyperbolic tangent function. However, it should be mentioned that, while it is necessary to extrapolate beyond the range of the training data, it is more suitable to use sigmoidal-type transfer functions in the hidden layer and linear transfer functions in the output layer. The types of transfer functions used in this study are tansigmoid function in the hidden layer and purelin function in the output layer.

6.2.5 Epoch size

The epoch size is equal to the number of training samples presented to the network between weigh updates. There are two types of epoch size: batch mode and on-line mode. On-line mode is when the epoch size is equal to one, and batch mode is when the epoch size is equal to the size of the training set. Generally, batch mode is preferable to on-line mode, as it forces the search to move in the direction of the true gradient at each weight update. A batch mode is used in this study.

6.2.6 Error function

The error function is the function that is minimized during training. Generally, the mean square error (MSE) is used as the error function. According to Masters (1993), there are

four main advantages of using MSE: "First of all it is calculated easily. Secondly, it penalizes large errors. Third, its partial derivative with respect to the weights can be calculated easily and lastly, it lies close to the heart of the normal distribution". In the present study the MSE was used as the performance criteria for deciding when to stop the training. To prevent over-training, the error of training and validation set was monitored throughout the modelling process. Initially the errors of both training and validation set decrease; however, at a certain point during the modelling process the error of validation set starts to increase, while the error of the training set is still decreasing (Figure 6.2). This is where the model starts over-fitting the data. The minimum point of the validation set is where the optimal weights of the network are obtained. This technique is called early-stop training, and in this way the network avoids over-fitting (Luk et al. 2000; Sarle 1995).

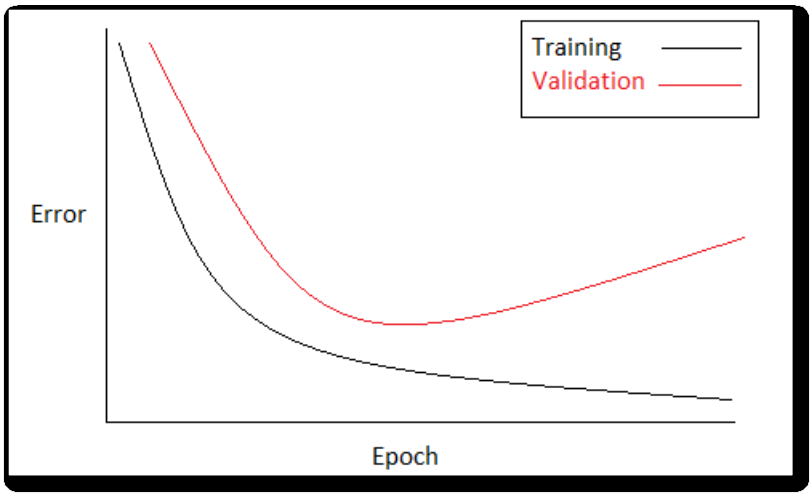

Figure 6-2. Schematic presentation of early-stop technique: avoiding over-fitting by monitoring validation error

6.3 Model development

Our model is a combination of static models in terms of model development and time series models in terms of input organization. As described in Section 6.2, a multi-layer perceptron neural network with one hidden layer was developed, and a constructive

algorithm (dynamic creation) was used in order to select the best number of hidden neurons for each network. The Levenberg-Marquardt algorithm was used to train the networks. The transfer function used in the hidden layer is the tansigmoid nonlinear function, and the purelin linear transfer function is used in the output layer.

Climate indicator and rainfall data were normalized within the range of [0,1] as discussed in Chapter 3. The data were divided into two sets: calibration (training and validation) and testing (out-of-sample test) sets. The calibration samples contained the data from 1900-2006, where 85% of the data were allocated to training and 15% were allocated to validating the models. The testing set contained data from 2007-2009.

The parameters for ANN modelling are basically network topology, node characteristics, training and learning rules. These rules specify an initial set of weights and indicate how weights should be adapted to improve performance. Multi-layered perceptrons are feed-forward nets (FNNs) with one or more hidden layers between the input and output nodes. The hidden layers are the most important part in an MLP, since they provide the nonlinearity between the input and output sets. More complex problems can be solved by increasing the number of hidden layers or neurons in the hidden layer. As discussed earlier, the Levenburg-Marquardt back-propagation algorithm was used for training the FNN model. The process of developing a suitable network is as follows:

i. Fixing the architecture
ii. Calibrating the network
iii. Testing the network

The general steps in ANN modelling, the definition and improvement of FNNs, the Levenburg-Marquardt back-propagation algorithm and its techniques were broadly explained in Section 6.2. A model based on feedforward neural network, with one hidden layer using tansigmoid function as the transfer function of the hidden layer, and purelin transfer function as the transfer function of the output layer was developed. The use of the tansigmoid transfer function was to enable nonlinearity of the network.

Choosing the number of hidden layers and the hidden neurons in the hidden layers is a critical task. These hidden neurons are responsible for mapping the complex nonlinear relationship between the inputs and the output. The major concern in developing an ANN structure is the determination of the appropriate hidden neurons in the hidden layers. There is no systematic way of selecting the best number of hidden neurons while developing an ANN and it is basically problem-dependent. Hornik (1991) proved that a single hidden layer network with a sufficiently large number of neurons can be used to approximate any measurable functional relationship between input data and the output variable to any desired accuracy. As discussed earlier, in the present study one hidden layer was used. The number of hidden neurons in the hidden layer was determined using the algorithm of dynamic creation. Initially, one hidden neuron is used and the training process is carried out. The final error is calculated. Then the algorithm progressively adds on hidden neurons. The number of hidden neurons which gives the minimum error among all other neurons is then accepted.

6.3.1 Calibrating the Network

ANN training is a nonlinear optimization process. Basically, the error between the network output and the target output is minimized by a predetermined algorithm, which repeatedly changes the values of ANN's connection weights. The ANN model implementation was carried out using MATLAB (Mathworks 2012). The model parameters were the connection weights, the momentum, learning rate, and the number of neurons in the hidden layer. These parameters were adjusted during the training process through the minimization of the mean square error (MSE). The *trainlm* function in MATLAB is used for this purpose. This function updates weights and bias values according to the Levenberg-Marquardt algorithm. A program in MATLAB was developed in order to use MATLAB functions to use the input files, apply the neural network, apply the early-stop technique, and calibrate and test the models. A written program in MATLAB is called an M-file and the first page of this program is shown in Figure 6.3.

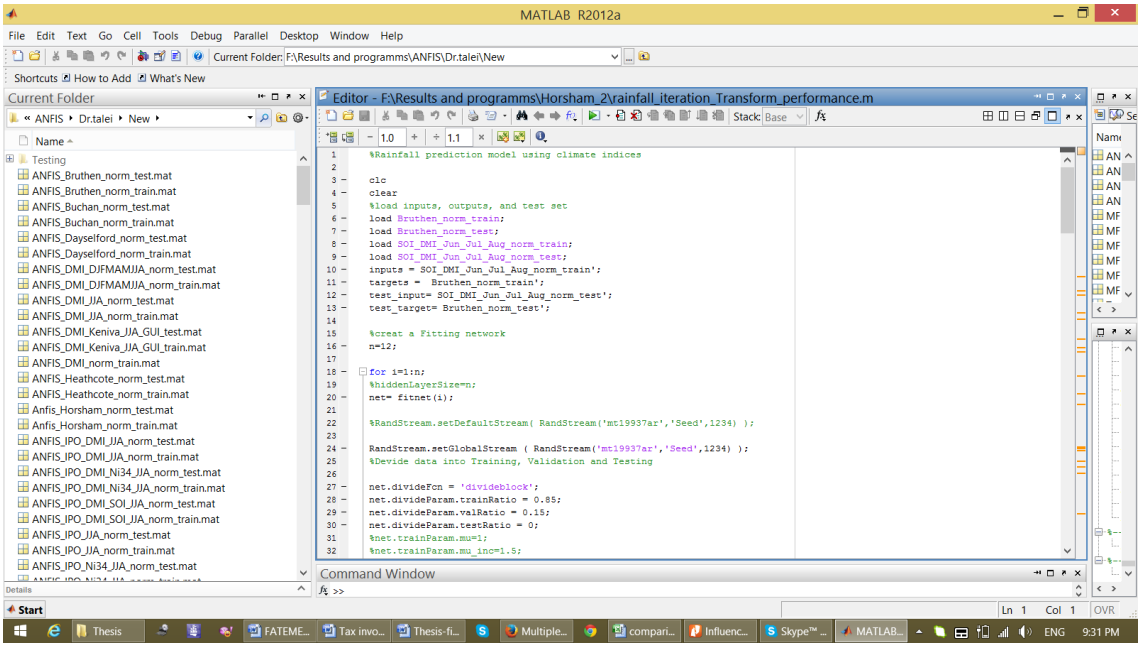

Figure 6-3. The first page of the M-File developed in MALTAB workspace

The training was stopped using the early-stop technique to decide the optimal learning. As discussed in Section 6.2.6, the training was stopped when the MSE over the validation set was found to be rising instead of reducing, even though the MSE over the training set was still reducing. This technique is used to stop the network from over-fitting. An over-fitted ANN would perform very well in the training set but fail to maintain the same level of accuracy when applied to the test set. Figure 6.4 shows an example of the early-stop technique process in MATLAB when developing the models.

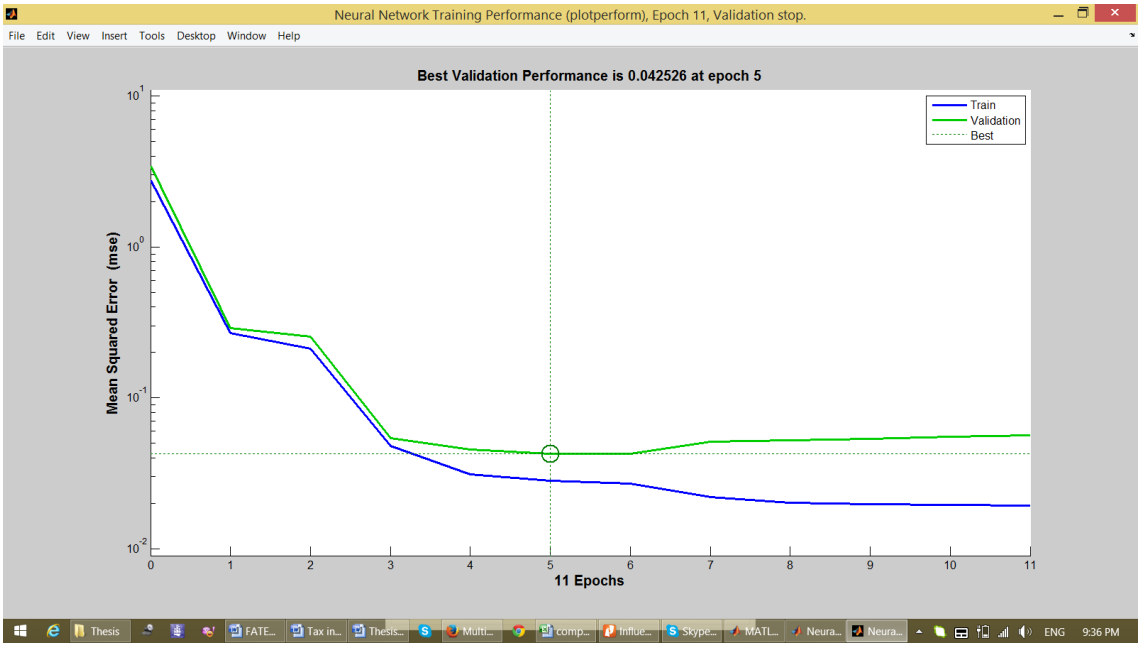

Figure 6-4. Sample of MSE curve for ANN training in MATLAB using early-stop technique

6.3.2 Testing the Network

The best model selected was applied to the test set to investigate the model's ability to work with an independent data set that had not been used in the calibration process of the model. The *sim* function in MATLAB software was incorporated into the developed program (written M-file) for this purpose.

6.3.3 Input Selection

Although neural networks are capable of handling the extra challenge of choosing both the variables and the size of the network simultaneously, prior input selection based on physical knowledge is beneficial for a more reliable neural network modelling process. ANN belongs to the class of data-driven approaches and can derive the relationship between the input and output from the presented data. However, choosing the inputs based on the existing knowledge of the physical process that governs the input-output set can be beneficial and lead to better modelling results. The input sets needed to calibrate and test the models were organized based on two hypotheses (scenarios). In the

first scenario only the individual climate modes (ENSO/IOD indicators) were considered as rainfall predictors; this scenario consisted of two different sections which will be further described in more detail. In the second scenario, the combined climate predictors were considered as rainfall predictors; this scenario also contained two separate sections. Table 6.1 shows the details of the inputs selected for this study.

Table 6.1. Input sets selected for developing ANN models

Scenarios	Input matrix for each category	Input category		
Scenario 1		**Single climate modes**		
Scenario 1-a	3×107	Nino3.4 $_{(Jun\text{-}July\text{-}Aug)}$	SOI $_{(Jun\text{-}July\text{-}Aug)}$	DMI $_{(Jun\text{-}July\text{-}Aug)}$
Scenario 1-b	9×107	Nino3.4 $_{(Dec\text{-}Aug)*}$	SOI $_{(Dec\text{-}Aug)*}$	DMI $_{(Dec\text{-}Aug)*}$
Scenario 2		**Dual climate modes**		
Scenario 2-a	6×107	Nino3.4 $_{(Jun\text{-}Jul\text{-}Aug)}$- DMI$_{(Jun\text{-}Jul\text{-}Aug)}$	SOI $_{(Jun\text{-}Jul\text{-}Aug)}$- DMI$_{(Jun\text{-}Jul\text{-}Aug)}$	
Scenario 2-b	6×107	IPO$_{(Jun\text{-}July\text{-}Aug)}$-SOI$_{(Jun\text{-}July\text{-}Aug)}$		

*Values of months Dec-Jan-Feb-Mar-Apr-May-Jun-Jul-Aug

More details of the two scenarios and the hypothesis upon which each scenario is based are described in the following section. The output of the ANN model for each scenario was intended to be the spring rainfall over each corresponding year.

1.Scenario 1: single climate mode predictors

In this scenario, the potential of each individual climate mode as spring rainfall predictor was investigated. ENSO and IOD indicators were chosen as single predictors of rainfall. This scenario comprised two different sets of predictors:

1-a) Three antecedent months' climate modes

As discussed in Section 4.2.2, the strongest statistically-significant relationship between

spring rainfall and climate indicators prior to spring occurs in the months of June, July and August (Table 4.12). These three months' climate modes are potential predictors of spring rainfall, as they show a statistically-significant linear relationship with the seasonal rainfall. Therefore, the input sets developed for scenario 1-a comprised three antecedent months (June-July-August) Nino3.4, SOI, and DMI, creating a matrix of 3 × 107 data for each climate index (i.e. Nino3.4 $_{(Jun\text{-}July\text{-}Aug)}$, SOI $_{(Jun\text{-}July\text{-}Aug)}$ and DMI $_{(Jun\text{-}July\text{-}Aug)}$) as shown in Table 6.1.

1-b) Nine antecedent months' climate modes

Based on Pearson correlation analysis, statistically-significant relationships between spring rainfall and climate indicators prior to spring occur in the months of June, July and August. However, Pearson correlation is a linear technique that is not capable of capturing complex nonlinear relationships. ENSO and IOD episodes start long before spring rainfalls occur; therefore, there could be a possibility that spring rainfall is influenced by climate modes further than the months of June, July and August. According to Risbey et al. (2009), the IOD occurs in May to November; the ENSO cycle on the other hand starts in April-May of the first year and continues until March-April of the following year (Verdon et al. 2004); when spring arrives it is in the middle of the ENSO and IOD cycles. Therefore, the second scenario is proposed, based on the work of Risbey et al. (2009) and Verdon et al. (2004); i.e. for the purpose of predicting spring rainfall, the author decided to examine further months of IOD and ENSO prior to spring rainfall. In this way, a broader range of antecedent monthly IOD and ENSO values were explored in order to investigate their effectiveness as rainfall predictors. The second scenario allocates nine months Dec_{n-1}-Aug_n monthly values of ENSO and IOD indicators (Nino3.4, SOI and DMI), where "n" is the year for which spring rainfall is being predicted. Input data for the second scenario for each index is a 9×107 matrix for each climate index, as shown in Table 6.1.

2. Scenario 2: **combined (dual) climate modes predictors**

2-a) Combined ENSO-IOD

The performance of seasonal predictability for southeast Australia was projected at only 30% compared to the successful seasonal forecasting schemes of other regions of Australia, such as Queensland. According to Verdon-Kidd and Kiem (2009b), the low predictability of southeast Australia's rainfall is prone to the compound interactions among the numerous climatic phenomena that affect this region's weather. In order to investigate the combined antecedent effect of ENSO and IOD on spring rainfall predictions, the dual input sets of $\text{Nino3.4}_{(\text{Jun-Jul-Aug})}$- $\text{DMI}_{(\text{Jun-Jul-Aug})}$ and $\text{SOI}_{(\text{Jun-Jul-Aug})}$-$\text{DMI}_{(\text{Jun-Jul-Aug})}$ were organized (Table 6.1).

2-b) Combined ENSO-IPO

According to Power et al., (1999), IPO modulates ENSO and its effect on Australian rainfall. Verdon et al. (2004) also found that there is a relationship between ENSO and IPO on a multi-decadal time scale; they discovered that the negative phase of the IPO amplifies the already increased La Nina rainfall and streamflow. A student t-test was used by Verdon et al. (2004) to determine whether rainfalls during an IPO negative-La Nina years were significantly higher than all other La Nina years. Based on Figure 6.5, adapted from Verdon et al. (2004), it can be seen that in Victoria, some of the considered stations have significantly higher rainfalls during the IPO negative phase in La Nina years, while others do not show a statistically significant difference. Figure 6.5 reinforces the results obtained earlier for correlation analysis between lagged IPO and rainfall shown in Table 4.13, i.e. only in some parts of Victoria does a relationship exist between IPO and rainfall, and this also applies to the lagged relationships. Therefore, in order to examine the effectiveness of IPO on seasonal rainfall forecasting, the use of combined antecedent ENSO and IPO in ANN modelling was investigated. A matrix of 6×107 inputs containing three antecedent months of June-July-August of combined ENSO and IPO was developed ($\text{IPO}_{(\text{Jun-Jul-Aug})}$ - $\text{SOI}_{(\text{Jun-Jul-Aug})}$) for the four stations discussed in Section 4.2.2 which are under the influence of IPO.

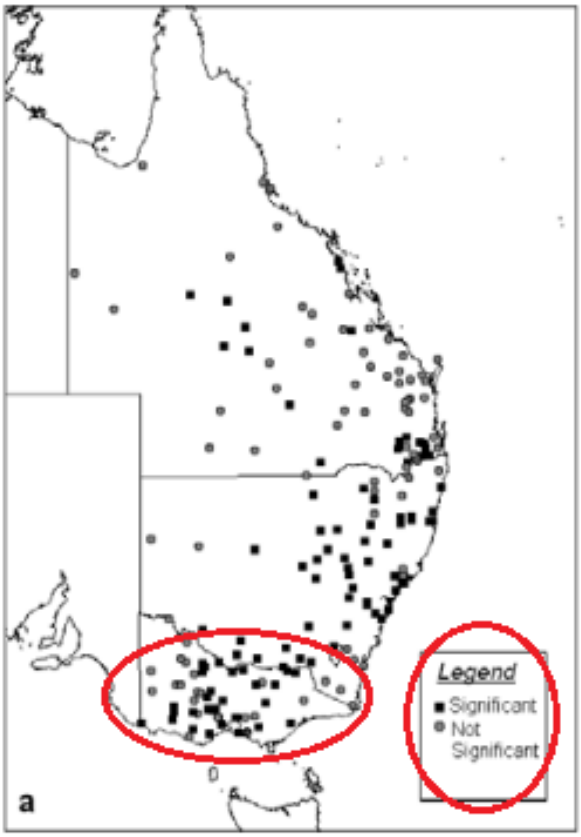

Figure 6-5. Results of significance test showing stations where the rainfall in IPO negative-La Nina years is significantly higher than all other La Nina years (adapted from Verdon et al. 2004).

6.4 Results and discussion

The results of single and combined climate mode models are discussed in this section. The best models from each scenario are selected for each station based on lowest forecast errors. Further, the best models are compared in order to find which scenario (i.e. single or combined climate modes) produces more accurate results.

6.4.1 First scenario: Single climate mode predictors

After arranging the input matrix according to Table 6.1, the inputs were used to develop multi-layer perceptron neural network models for each site. After calibrating (training

and validating) the models, the best models in regard to the lowest errors were chosen. The best ANN model performances for the first scenario are presented in Table 6.2.

Table 6.2. Performance of the best ANN models: first scenario

Region	Station	Model	r	RMSE	MAE	MSE
East	Bruthen	$DMI_{(Dec\ to\ Aug)}$	0.50	0.22	0.19	0.048
	Buchan	$SOI_{(Jun\text{-}July\text{-}Aug)}$	0.50	0.20	0.16	0.040
	Orbost	$DMI_{(Jun\text{-}July\text{-}Aug)}$	0.66	0.17	0.14	0.029
Centre	Malmsbury	$DMI_{(Dec\ to\ Aug)}$	0.53	0.17	0.12	0.029
	Daylesford	$DMI_{(Dec\ to\ Aug)}$	0.42	0.17	0.14	0.029
	Heathcote	$DMI_{(Dec\ to\ Aug)}$	0.51	0.20	0.16	0.040
West	Horsham	$DMI_{(Dec\ to\ Aug)}$	0.74	0.10	0.09	0.010
	Kaniva	$DMI_{(Dec\ to\ Aug)}$	0.65	0.17	0.13	0.029
	Rainbow	$DMI_{(Dec\ to\ Aug)}$	0.30	0.17	0.12	0.029

It can be seen from Table 6.2 that in east Victoria in terms of the extent in time of climate modes, three antecedent months' ENSO/IOD produce more accurate prediction of spring rainfall than nine antecedent months' ENSO/IOD. In this region two different climate indices, SOI and DMI, produce the best models, possibly due to the mixed effect of ENSO and IOD on east Victorian rainfalls, as east Victoria is close to the Pacific Ocean where ENSO occurs. The models for this region show correlation coefficients ranging from 0.50 to 0.66. For central Victoria, the models based on nine antecedent months' IOD show superior model performance compared to the models based on three antecedent months' climate modes. The correlation coefficients for models in central Victoria vary from 0.42 to 0.53. The same pattern can be seen for west Victoria, where the models based on nine antecedent months' IOD produce better forecasts compared to three antecedent months' IOD or ENSO indicators. It appears

that in general the models based on antecedent months' IOD indicator out-perform the models based on ENSO indicators. This result is consistent with the findings of Lim et al. (2011), indicating that the effect of IOD on rainfall can be seen across the southeast and southwest of Australia. A physical interpretation of the models could be that not only is IOD the dominant lagged predictor of spring rainfall for Victoria compared to ENSO, but the wider information of IOD also has more effect on the accuracy of the models compared to the lower ones.

After calibrating (training and validating) the models, the selected best models of all the regions were used to forecast spring rainfalls for three consecutive years in advance (2007-2009). In this way, the models' generalization ability was assessed. Table 6.3 shows the performances of the models on the testing dataset.

Table 6.3. Performance of the best ANN models (test-set): first scenario

Region	Station	Model	r	RMSE	MAE	MSE
East	Bruthen	$DMI_{(Dec\ to\ Aug)}$	0.49	0.14	0.11	0.020
	Buchan	$SOI_{(Jun\text{-}July\text{-}Aug)}$	0.40	0.20	0.16	0.040
	Orbost	$DMI_{(Jun\text{-}July\text{-}Aug)}$	0.12	0.17	0.13	0.029
Centre	Malmsbury	$DMI_{(Dec\ to\ Aug)}$	0.77	0.22	0.00	0.048
	Daylesford	$DMI_{(Dec\ to\ Aug)}$	0.95	0.17	0.07	0.029
	Heathcote	$DMI_{(Dec\ to\ Aug)}$	0.97	0.14	0.12	0.020
West	Horsham	$DMI_{(Dec\ to\ Aug)}$	0.99	0.00	0.05	0.00
	Kaniva	$DMI_{(Dec\ to\ Aug)}$	0.98	0.10	0.06	0.010
	Rainbow	$DMI_{(Dec\ to\ Aug)}$	0.96	0.00	0.05	0.00

It can be seen from Table 6.3 that IOD-based models are able to predict spring rainfall in the test set with a very low error (RMSE, MAE, and MSE) and a very high correlation for most of Victoria. The generalization ability of the models for east Victoria in regard to correlation coefficients is not as good as for other parts of Victoria; the correlation coefficients for east Victoria vary from 0.12 to 0.49 compared to central

and west Victoria with correlation coefficients from 0.77 to 0.99. The models' errors are also lower for central and west Victoria compared to east Victoria. The best models are shown in Figures 6.6 to 6.8. It can be seen from the figures that the models perform the best in west Victoria. In general the models are able to follow the pattern of the observation, however, in some cases the peaks and troughs are not captured.

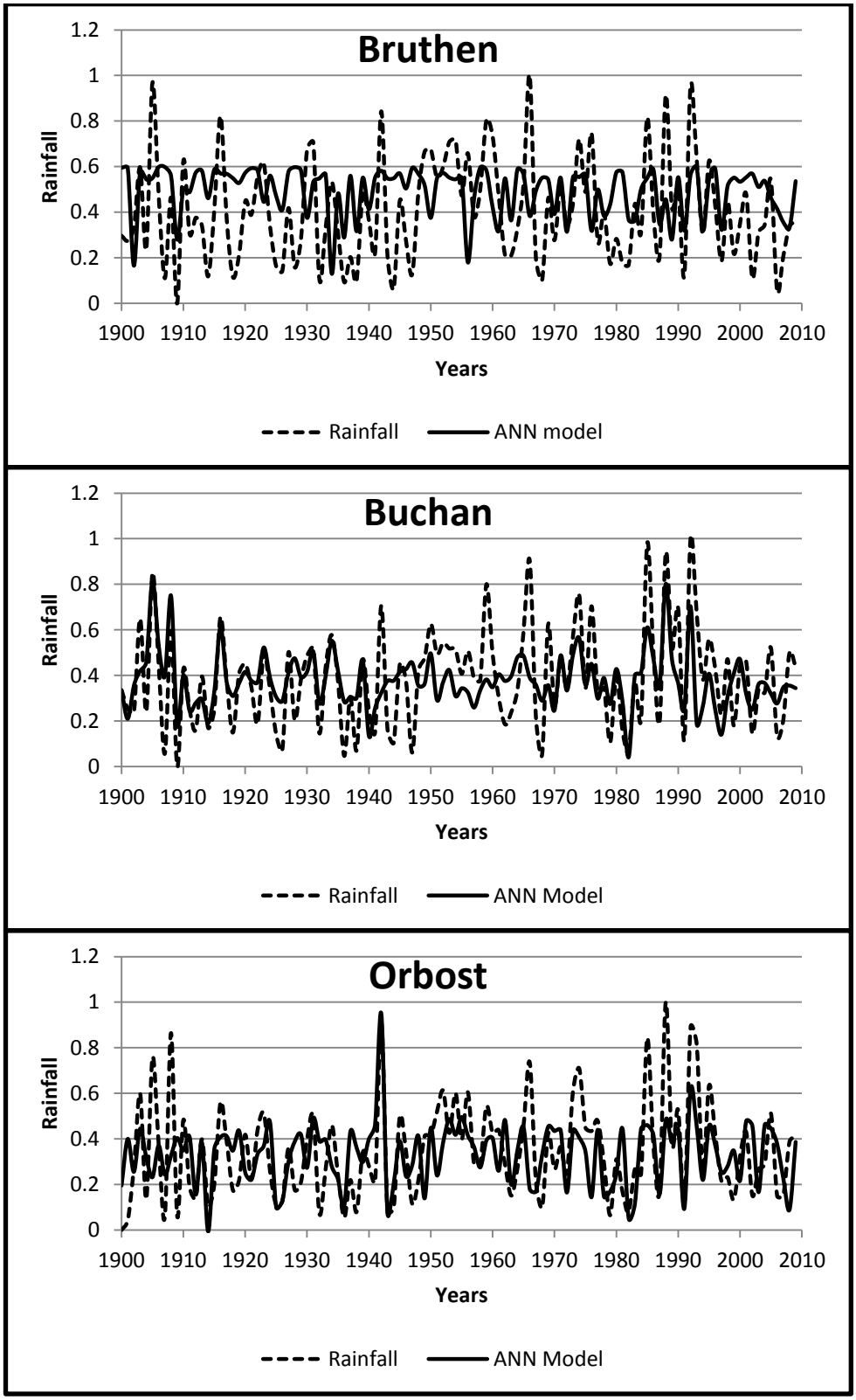

Figure 6-6. Single ANN modelling for east Victoria

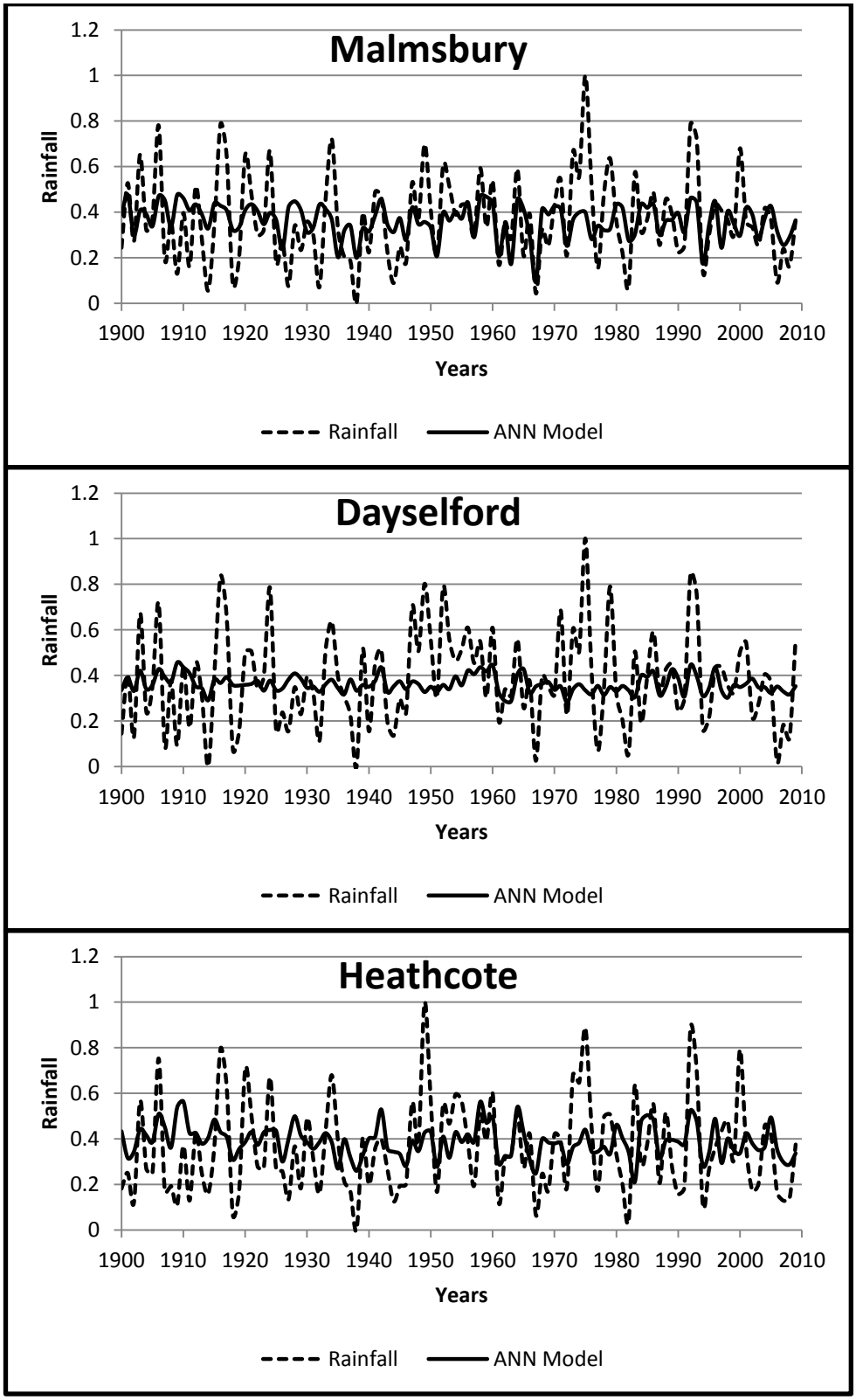

Figure 6-7. Single ANN modelling for central Victoria

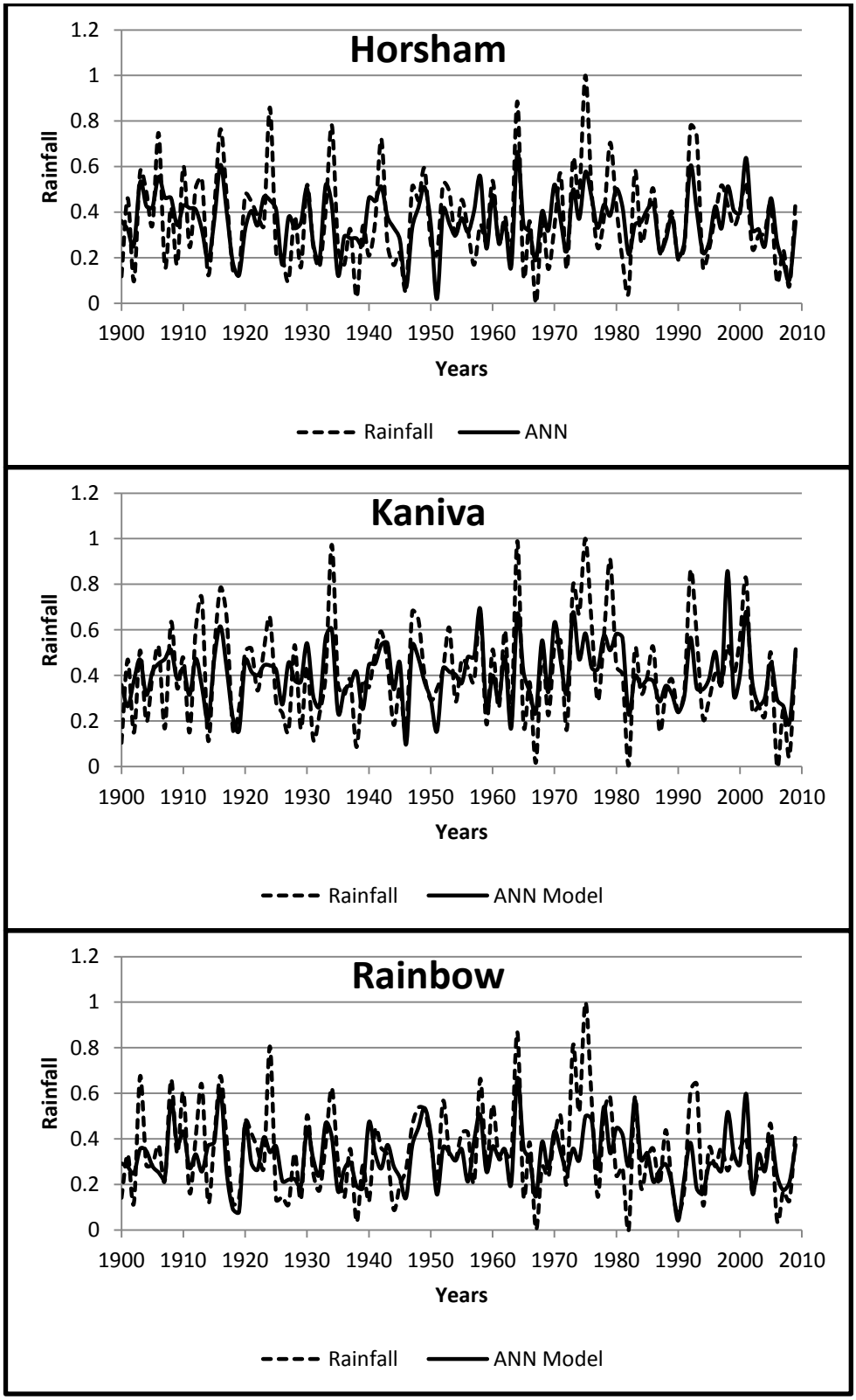

Figure 6-8. Single ANN modelling for west Victoria

6.4.2 Second scenario: combined climate mode predictors

2-a) ENSO-IOD results

Two sets of Nino3.4 $_{(Jun\text{-}July\text{-}Aug)}$-DMI $_{(Jun\text{-}July\text{-}Aug)}$ and SOI $_{(Jun\text{-}July\text{-}Aug)}$-DMI $_{(Jun\text{-}July\text{-}Aug)}$ were used as inputs for developing ANN models for the three regions, as shown in Table 6.1. Table 6.4 summarizes the prediction skills of the combined ENSO-IOD models regarding RMSE, MAE, MSE and Pearson correlations (r). The results were then compared with MLR models as benchmarks, since combined ENSO-IOD indices were also used as inputs for the MLR models. It can be seen from Table 6.4 that the correlation coefficients of ANN models for east and west Victoria are significantly higher compared to the MLR models, and the errors (RMSE, MAE, and MSE) are generally lower. For central Victoria, the correlation coefficients are generally higher than the MLR models; however, the performance of the MLR models regarding MSE and MAE is better for this region. The higher correlation coefficients of ANN models indicate that ANN is more capable of finding the pattern and trend of the observations than MLR models.

Table 6.4. Comparison of performance of combined ENSO-IOD ANN and MLR models

			ANN models				MLR models			
Region	**Station**	**ANN Model**	**r**	**RMSE**	**MAE**	**MSE**	**r**	**RMSE**	**MSE**	**MAE**
East	Bruthen	Ni34-DMI	0.75	0.15	0.120	0.023	0.32	0.22	0.048	0.171
	Buchan	Ni34-DMI	0.65	0.17	0.154	0.028	0.35	0.16	0.026	0.171
	Orbost	SOI-DMI	0.64	0.18	0.145	0.034	0.35	0.19	0.038	0.157
Centre	Malmsbury	Ni34-DMI	0.54	0.18	0.130	0.034	0.36	0.17	0.030	0.140
	Daylesford	Ni34-DMI	0.36	0.20	0.168	0.039	0.37	0.20	0.039	0.155
	Heathcote	SOI-DMI	0.52	0.21	0.158	0.044	0.37	0.19	0.035	0.153
West	Horsham	Ni34-DMI	0.64	0.15	0.193	0.023	0.36	0.18	0.033	0.149
	Kaniva	SOI-DMI	0.56	0.20	0.158	0.042	0.39	0.20	0.041	0.163
	Rainbow	SOI-DMI	0.53	0.15	0.115	0.023	0.36	0.18	0.031	0.142

After calibrating the models, in order to evaluate the generalization ability of the developed ANN models, out-of-sample tests were carried out for the years 2007-2009 (Table 6.5). It can be seen that MLR models show very poor generalization ability for

east Victoria (r = -0.99, -0.90 and -0.99 for Bruthen, Buchan and Orbost respectively) compared to ANN, with correlation coefficients of 0.93, 0.76 and 0. 42; the ANN models also show better generalization ability for central and west Victoria, with correlation coefficients of 0.68 to 0.85 and 0.58 to 0.97 respectively compared to the MLR models. However, the ability of MLR models to forecast out-of-sample sets is comparable with ANN for Daylesford in central Victoria and Kaniva in west Victoria (r=0.92 and 0.67 respectively). In addition, the errors of the testing sets for ANN models are generally lower compared to those for the multiple regression models. Figures 6.9 to 6.11 show the comparison between combined ANN models and combined regression models. For most of Victoria, the superiority of ANN over MLR models in regard to error and pattern recognition can be seen from the figures.

Table 6.5. Performance of dual ANN models and MLR models for the test set

Region	Station	ANN				Regression			
		r	RMSE	MAE	MSE	r	RMSE	MAE	MSE
East	Bruthen	0.93	0.134	0.120	0.018	-0.99	0.126	0.085	0.016
	Buchan	0.76	0.089	0.080	0.008	-0.90	0.152	0.180	0.023
	Orbost	0.42	0.122	0.107	0.015	-0.99	0.155	0.150	0.024
Centre	Malmsbury	0.68	0.084	0.080	0.007	0.48	0.114	0.100	0.013
	Daylesford	0.85	0.182	0.164	0.033	0.92	0.207	0.205	0.043
	Heathcote	0.71	0.134	0.125	0.018	-0.50	0.161	0.158	0.026
West	Horsham	0.80	0.095	0.080	0.009	0.25	0.173	0.149	0.030
	Kaniva	0.97	0.114	0.110	0.013	0.67	0.226	0.163	0.051
	Rainbow	0.58	0.130	0.128	0.017	-0.74	0.170	0.142	0.029

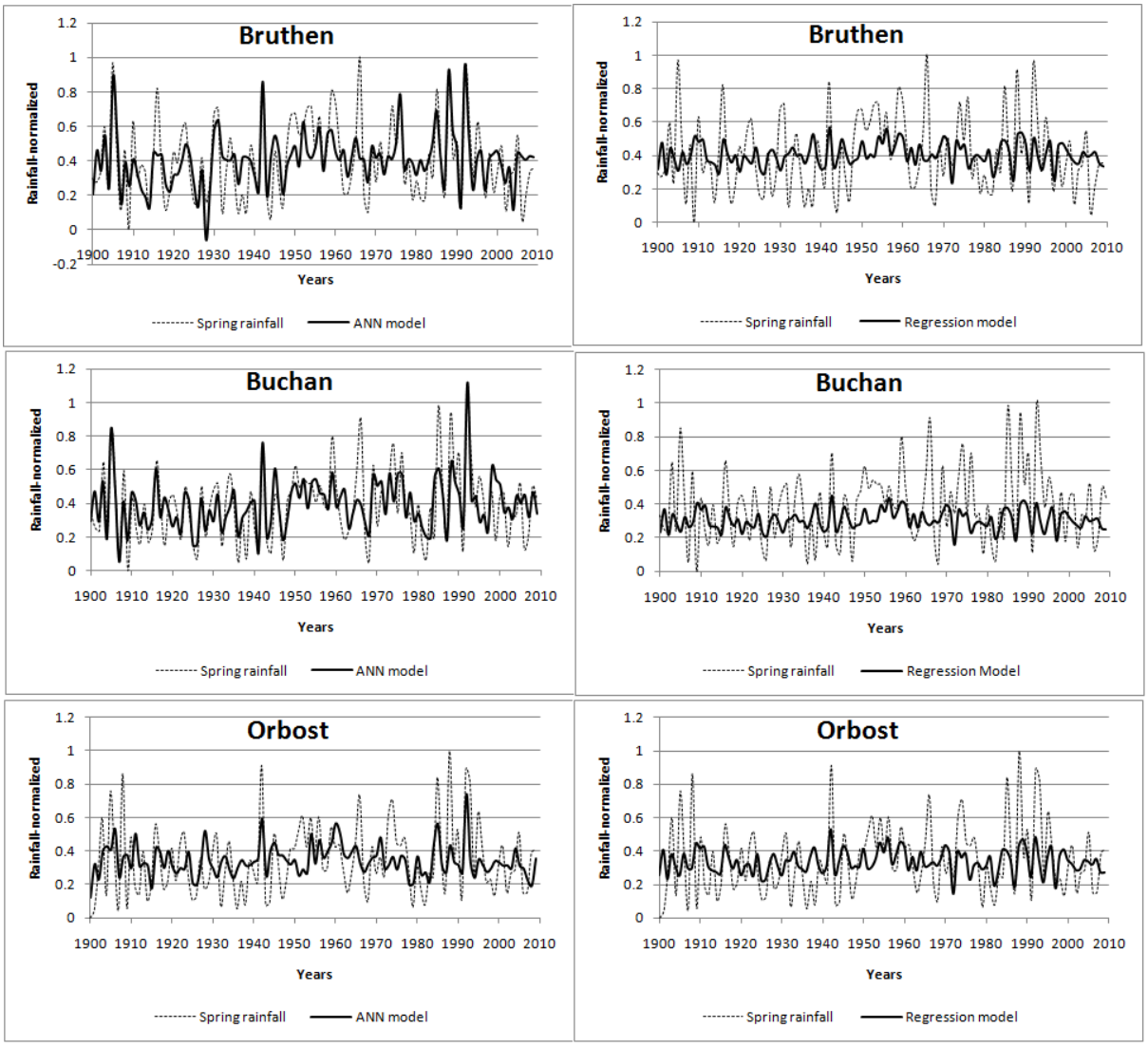

Figure 6-9. Comparing combined ANN modelling with combined MLR modelling for east Victoria

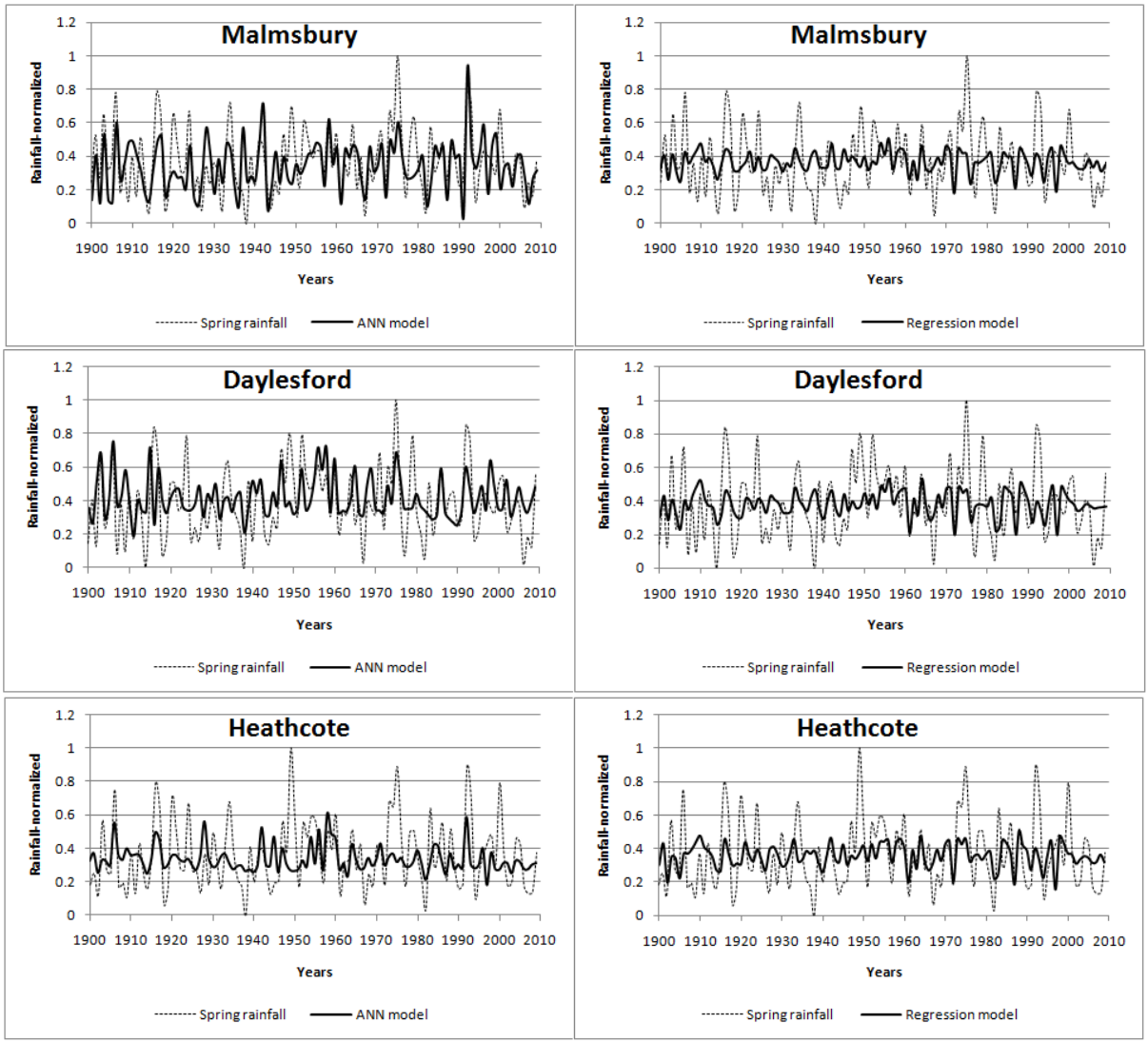

Figure 6-10. Comparing combined ANN modelling with combined MLR modelling for central Victoria

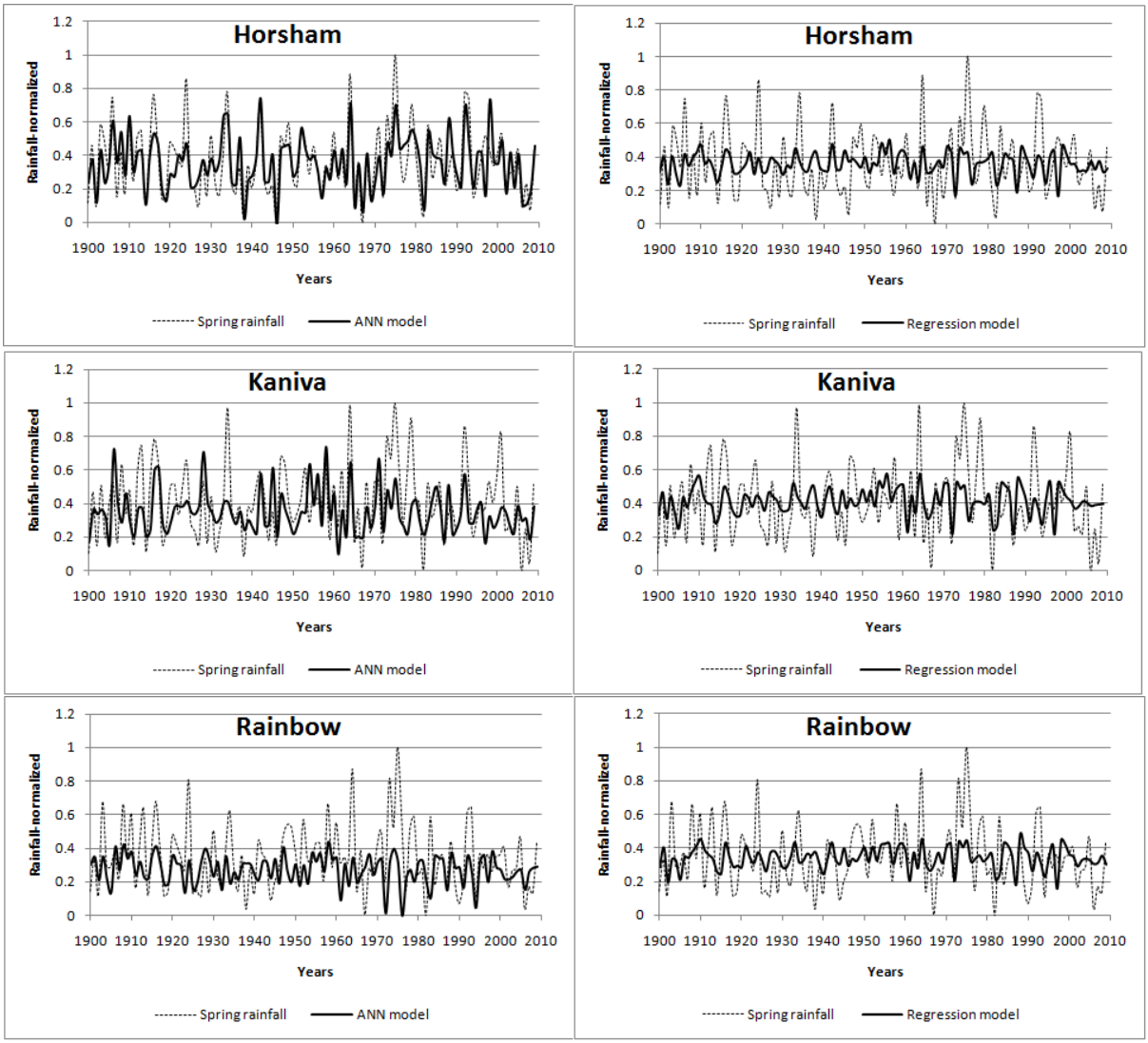

Figure 6-11. Comparing combined ANN modelling with combined MLR modelling for west Victoria

2-b) ENSO-IPO results

As discussed in Table 4.13, the lag relationship between IPO and spring rainfall is quite weak. This relationship is not consistent throughout Victoria, and only some stations show statistically-significant relationships. In order to evaluate the effect of IPO and ENSO on rainfall predictability using neural networks, a set of inputs containing combined $IPO_{(Jun\text{-}July\text{-}Aug)}$-$SOI_{(Jun\text{-}July\text{-}Aug)}$ was constructed (Table 6.1) and used to train and validate ANN models for the stations for which IPO is effective. The results of the

combined IPO-SOI models are shown in Tables 6.6 and 6.7 for calibration and testing sets respectively. It can be seen from Tables 6.6 and 6.7 that combining IPO with ENSO does not improve the rainfall forecasting performance compared to the previously discussed models (Tables 6.2 to 6.5). It can be seen from Table 6.6 that the testing results are poor, considering the correlation coefficients and error criteria, possibly because negative IPO affects only some parts of Victoria and enhances the rainfall during La Nina events specifically. Therefore, the use of IPO in seasonal rainfall forecasting in Victoria is not recommended.

Table 6.6. Performance of the combined models (IPO-SOI)

Region	Station	r	RMSE	MAE	MSE
East	Bruthen	----	----	----	----
	Buchan	----	----	----	----
	Orbost	----	----	----	----
Centre	Malmsbury	----	----	----	----
	Daylesford	0.29	0.20	0.15	0.04
	Heathcote	0.41	0.21	0.15	0.04
West	Horsham	----	----	----	---
	Kaniva	0.61	0.18	0.14	0.03
	Rainbow	0.43	0.15	0.11	0.02

Table 6.7. Performance of the combined models (IPO-SOI) for the test set

Region	Station	r	RMSE	MAE	MSE
East	Bruthen	----	----	----	----
	Buchan	----	----	----	----
	Orbost	----	----	----	----
Centre	Malmsbury	----	----	----	----
	Daylesford	-0.21	0.27	0.23	0.07
	Heathcote	0.64	0.18	0.15	0.03
West	Horsham	----	----	----	
	Kaniva	-0.87	0.37	0.30	0.14
	Rainbow	-0.36	0.22	0.19	0.05

To further evaluate the ability of ANN to model spring rainfall, the peaks and troughs of ANN-predicted spring rainfall and actual spring rainfall were cross-plotted (Figures

6.12-6.14). Table 6.8 shows the correlation coefficient values for the peaks and troughs. It can be seen from the figures and Table 6.8 that ANN is able to capture the peaks with a correlation coefficient of 0.41~0.59 for east Victoria. Apart from Bruthen with a weak correlation coefficient of -0.03, the Buchan and Orbost models were able to forecast the troughs with a correlation coefficient of r=0.46. For central Victoria (Figure 6.13), the models were able to forecast the troughs better than the peaks (r=0.42~0.53). For west Victoria, the peaks and troughs were modelled better compared to the rest of Victoria (r=0.037~0.69) (Figure 6.14).

Table 6.8. Correlation coefficients of the models for the peaks and troughs

Station	Peak	Trough
Bruthen	0.41	-0.03
Buchan	0.42	0.46
Orbost	0.59	0.46
Malmsbury	0.52	0.53
Daylesford	0.06	0.42
Heathcote	-0.02	0.42
Horsham	0.69	0.68
Kaniva	0.55	0.46
Rainbow	0.37	0.00

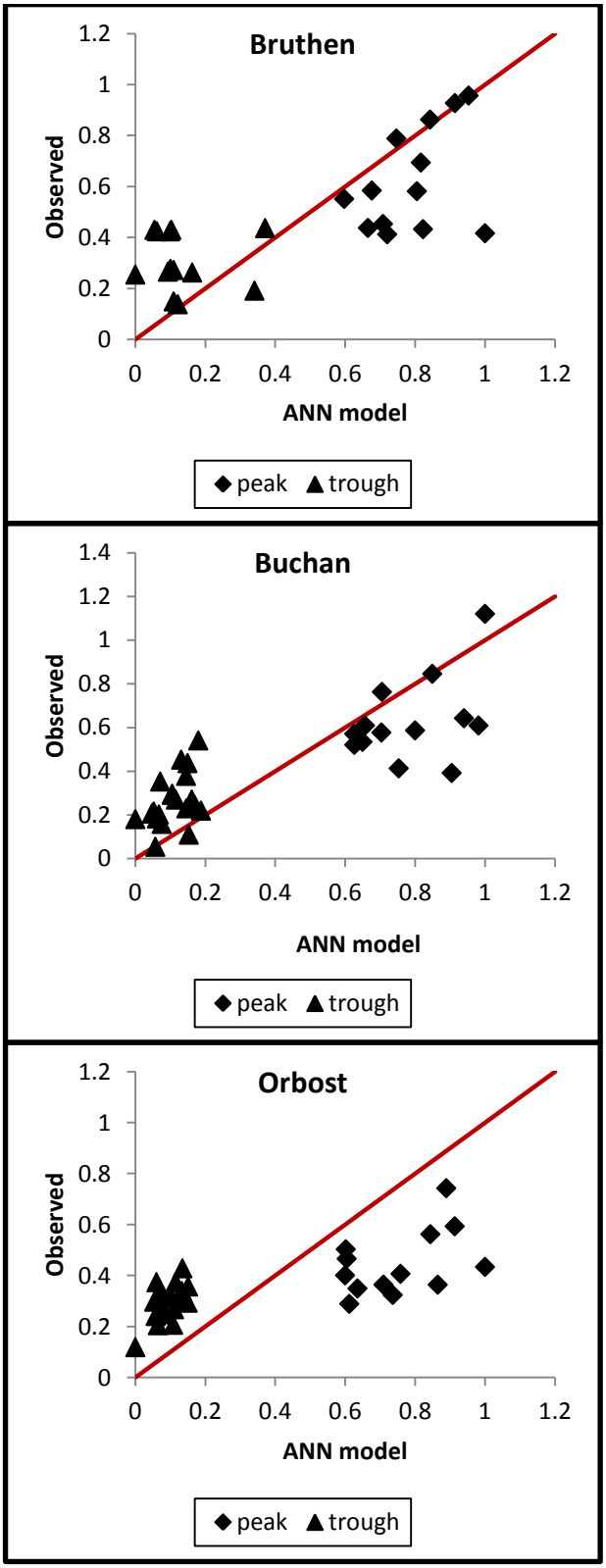

Figure 6-12. Evaluating the performance of ANN models for the peaks and troughs-East Victoria

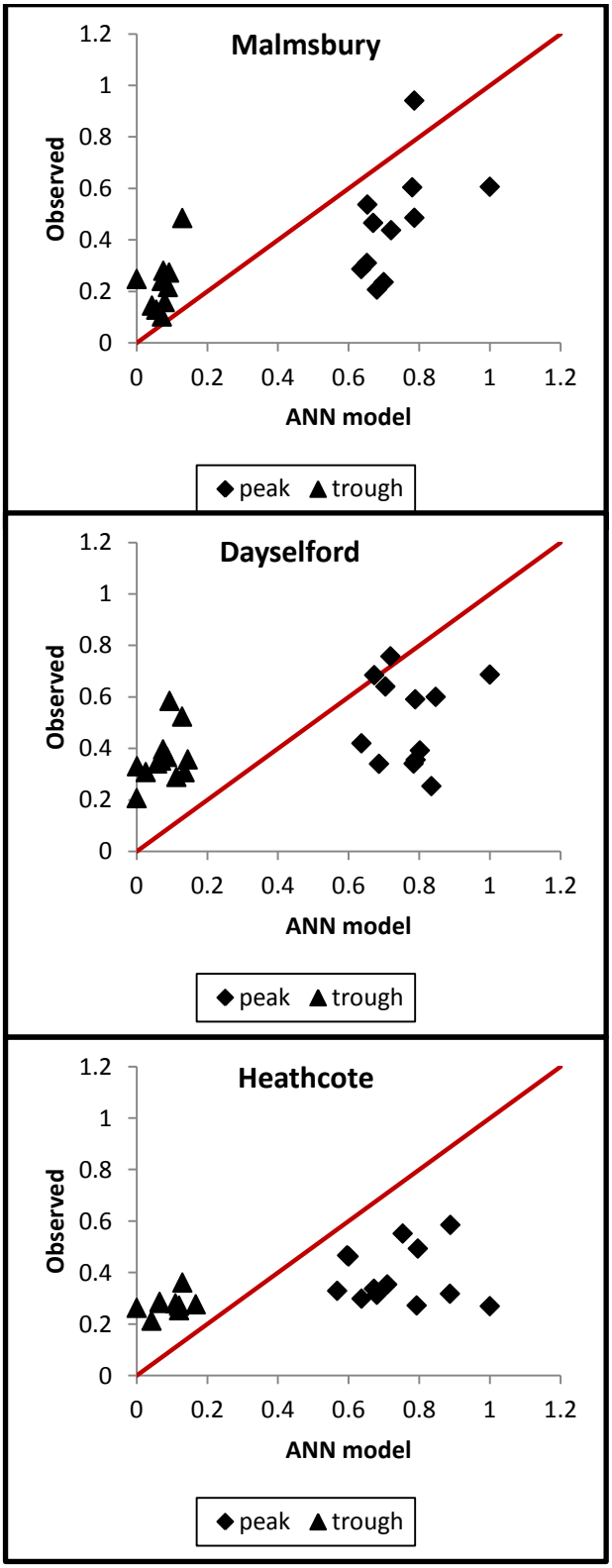

Figure 6-13. Evaluating the performance of ANN models for the peaks and troughs-Central Victoria

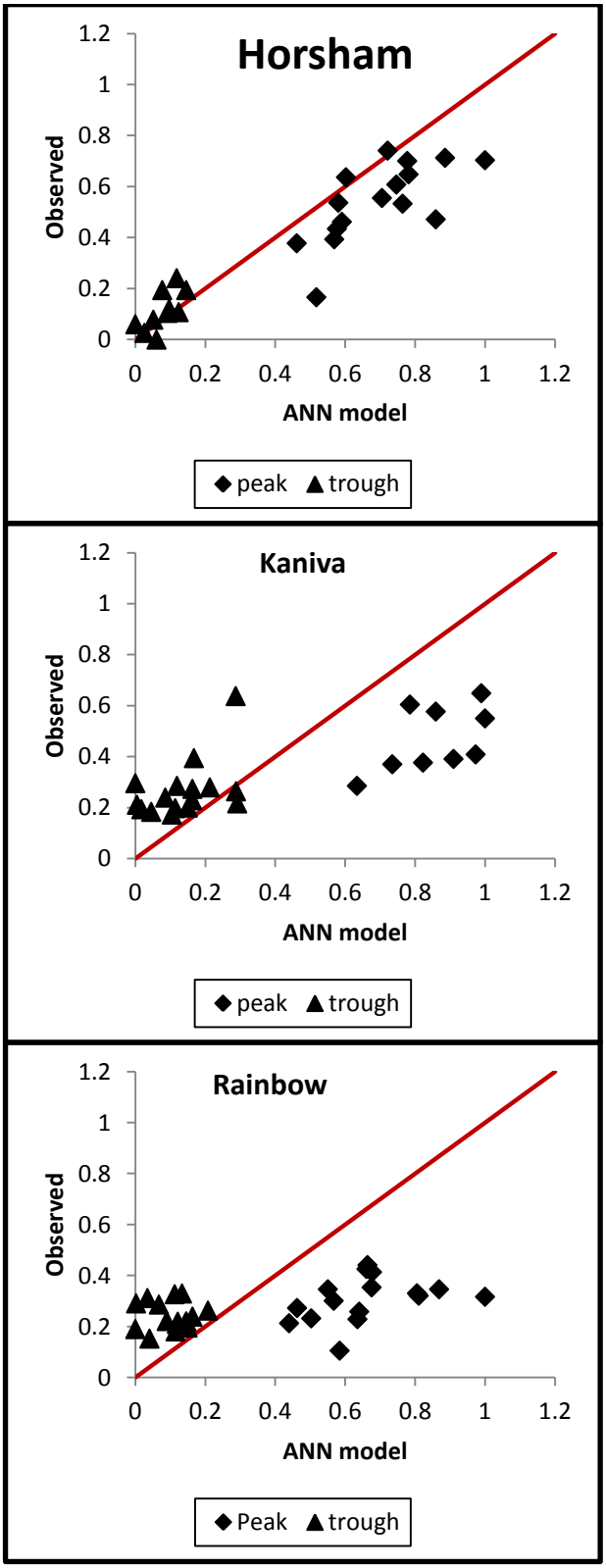

Figure 6-14. Evaluating the performance of ANN models for the peaks and troughs-West Victoria

6.4.3 Comparison of the models for single and combined climate modes scenarios

The error criteria (MSE, MAE) and trend criterion (r) differ. While the error criteria show the average error of the models, the trend criterion shows how well the model is capable of following the pattern of observation. The results of the best single and combined (dual) models are shown next to each other for the sake of comparison for the validation and test sets in Tables 6.9 and 6.10 respectively. The best model performances are shown in bold. It can be seen from Tables 6.9 and 6.10 that for east Victoria, the combined models produce forecasts with lower errors and higher correlation coefficients compared to the single models. For central Victoria, the performance of the validation sets for both single and combined models are similar, although the single models show a better generalization ability in the test set; i.e. the correlation coefficients of the single models in the test set are higher than those of the dual models. However, the errors are almost the same. For west Victoria, a similar pattern can be seen. In some cases, these two types of criteria (error and trend) might lead to confusion, as a model might show low errors but have a very poor trend criterion. For example, Table 6.10 shows that for Malmsbury, the single model has a performance of RMSE=0.22, r= 0.77 while the dual model shows a performance of RMSE=0.08 and r= 0.68. It can be seen that, while the trend criterion (r) of the single model is higher, the error criterion (RMSE) of the dual model is better. Therefore, deciding on which models are better for a station based on both criteria is challenging. Therefore, the author proposes the concept of error trend ratio (ETR) criteria, defined as follows:

$$ETR = \frac{error}{trend} \tag{6-12}$$

As lower errors and higher correlation is desirable, the closer the ETR to zero, the better the models are.

Table 6.9. Comparison of the performance of the best single and combined (dual) ANN models: validation

		Single models				Dual models			
Region	**Station**	**r**	**RMSE**	**MAE**	**MSE**	**r**	**RMSE**	**MAE**	**MSE**
East	Bruthen	0.50	0.22	0.19	0.048	**0.75**	**0.15**	0.12	0.023
	Buchan	0.50	0.20	0.16	0.040	**0.65**	**0.17**	0.15	0.028
	Orbost	0.66	0.17	0.14	0.029	**0.64**	0.18	0.15	0.034
Centre	Malmsbury	0.53	0.17	0.12	0.029	0.54	0.18	0.13	0.034
	Daylesford	**0.42**	**0.17**	0.14	0.029	0.36	0.20	0.17	0.039
	Heathcote	0.51	0.20	0.16	0.040	0.52	0.21	0.16	0.044
West	Horsham	**0.74**	**0.10**	0.09	0.010	0.64	0.15	0.19	0.023
	Kaniva	**0.65**	**0.17**	0.13	0.029	0.56	0.20	0.16	0.042
	Rainbow	0.30	0.17	0.12	0.029	**0.53**	**0.15**	0.12	0.023

Table 6.10. Comparison of the performance of the best single and combined (dual) ANN models: test

		Single models				Dual models			
Region	**Station**	**r**	**RMSE**	**MAE**	**MSE**	**r**	**RMSE**	**MAE**	**MSE**
East	Bruthen	0.49	0.14	0.11	0.020	**0.93**	0.13	0.120	0.018
	Buchan	0.40	0.20	0.16	0.040	**0.76**	**0.09**	0.080	0.008
	Orbost	0.12	0.17	0.13	0.029	**0.42**	**0.12**	0.107	0.015
Centre	Malmsbury	**0.77**	0.22	0.00	0.048	0.68	**0.08**	0.080	0.007
	Daylesford	**0.95**	0.17	0.07	0.029	0.85	0.18	0.164	0.033
	Heathcote	**0.97**	0.14	0.12	0.020	0.71	0.13	0.125	0.018
West	Horsham	**0.99**	**0.00**	0.05	0.00	0.80	0.10	0.080	0.009
	Kaniva	0.98	0.10	0.06	0.010	0.97	0.11	0.110	0.013
	Rainbow	**0.96**	**0.00**	0.05	0.00	0.58	0.13	0.128	0.017

The ETR based on MSE and MAE is shown in Tables 6.11 and 6.12 for validation and testing sets, respectively. Based on ETR shown in Tables 6.11 and 6.12, the dual models for east Victoria have better performance in regard to error and trend criteria than the single models. Recall from Section 6.4.1, in east Victoria both single ENSO and IOD indices were shown to affect east Victoria in the single modelling process. Therefore, the results of both single and dual models reinforce the assumption that

combined ENSO and IOD have more influence on east Victoria and the use of their combined antecedent values for rainfall forecasting provides better forecasting results. In relation to central and west Victoria, the results of the wider time frame of IOD models (nine antecedent months) are better in terms of the forecast accuracy compared to the dual models. A possible explanation for this is that, since east Victoria is closer to the Pacific Ocean where ENSO occurs, the effect of both ENSO and IOD can be felt in this region; however, moving towards central and west Victoria and closer to the Indian Ocean, IOD becomes the dominant predictor of rainfall.

Table 6.11. Comparison of model performance based on the error/trend ratio (ETR). Better models are indicated in bold.

Region	Station	Dual climate mode		Single climate mode	
		MSE/r	MAE/r	MSE/r	MAE/r
East	Bruthen	**0.031**	**0.160**	0.096	0.380
	Buchan	**0.043**	**0.237**	0.080	0.320
	Orbost	0.053	0.227	**0.044**	**0.212**
Centre	Malmsbury	0.063	0.241	**0.055**	**0.226**
	Daylesford	0.108	0.467	**0.069**	**0.333**
	Heathcote	0.085	**0.304**	**0.078**	0.314
West	Horsham	0.036	0.302	**0.014**	**0.122**
	Kaniva	0.075	0.282	**0.045**	**0.200**
	Rainbow	**0.043**	**0.217**	0.097	0.400

Table 6.12. Comparison of model performance based on the error/trend ratio (ETR) (test set). Better models are indicated in bold.

Region	Station	Dual climate mode		Single climate mode	
		MSE/r	MAE/r	MSE/r	MAE/r
East	Bruthen	**0.019**	**0.129**	0.041	0.224
	Buchan	**0.011**	**0.105**	0.100	0.400
	Orbost	**0.036**	**0.255**	0.242	1.083
Centre	Malmsbury	**0.010**	0.118	0.062	**0.000**
	Daylesford	0.039	0.193	**0.031**	**0.074**
	Heathcote	0.025	0.176	**0.021**	**0.124**
West	Horsham	0.011	0.100	**0.000**	**0.051**
	Kaniva	0.013	0.113	**0.010**	**0.061**
	Rainbow	0.029	0.221	**0.000**	**0.052**

6.5 Summary of Artificial Neural Network Analysis

The Artificial Neural Network technique was used for the purpose of seasonal spring rainfall forecasting using antecedent large-scale climate modes. ENSO, IOD and IPO were selected as inputs of ANN in two major scenarios. In the first scenario, each individual climate mode was considered as a rainfall predictor, while in the second scenario combined (dual) climate modes were considered. The first scenario contained two sub-sections; the first sub-section of the first scenario was based on three antecedent months of each individual climate mode, while the second sub-section contained nine antecedent months of individual climate modes. It was discovered that the nine antecedent months of IOD are better predictors of seasonal spring rainfall for most of Victoria and produce better forecasts compared to three antecedent months of ENSO or IOD. It was also discovered that IOD is a better predictor than ENSO for seasonal spring rainfall forecasting.

The second scenario also contained two sub-sections; the first sub-section contained combined antecedent three months of ENSO-IOD, while the second sub-section was based on combined three antecedent months of ENSO-IPO. It was found that combined

three antecedent months of ENSO-IOD are better predictors for spring rainfall compared to combined antecedent three months of ENSO-IPO.

Comparing the results of the two scenarios revealed that the models developed based on nine months antecedent IOD have better performance than the combined models in central and west Victoria. However, the combined climate mode models perform better for east Victoria. The physical reason for this may be the fact that east Victoria is closer to the Pacific Ocean where ENSO occurs, and the effect of both ENSO and IOD can therefore be felt in this region. However, towards central and west Victoria and closer to the Indian Ocean, IOD becomes the dominant predictor of rainfall in Victoria.

Chapter 7

Adaptive Network-based Fuzzy Inference system

7.1 Introduction

Fuzzy and ANN theories have been developed to imitate the thinking process of human brain to learn analogous strategies or experiences in order to make optimal decisions. However, the fundamental mechanisms of these two theories are different. ANN offers an advanced capability to extract significant features from complex databases and are capable of learning the relationship between any data pairs, whereas the fuzzy logic is based on the way how brains deal with inexact information. As fuzzy theories lack the ability of learning, it is difficult to tune the fuzzy rules and membership functions based on training data. capturing the advantages and strengths of both ANNs and fuzzy logic in a single framework. In order to take advantage of both ANN and fuzzy logic in a single framework, the neuro-fuzzy system was developed (Chang et al. 2014). In 1993, Jang proposed the adaptive network-based fuzzy inference system (ANFIS), which is one of the popular neuro-fuzzy systems. To date, ANFIS has been applied to a wide range of hydrological modelling including rainfall-runoff modelling (Nayak et al. 2004; Talei et al. 2010; Talei & Chua 2012), flood forecasting (Nayak et al. 2005; Talei et al., 2013), water resources management (Abolpour et al. 2007; Chang and Chang, 2006), water quality modelling (Yeon et al. 2008), and rainfall forecasting (El-Shafie et al. 2011). To the best of the author's knowledge, ANFIS has not been used for rainfall forecasting with the use of large-scale climate modes; therefore, this study examines the capability of ANFIS in this regard.

7.2 Methodology

Fuzzy logic is based on the idea of fuzzy sets. A set with no crisp or clear boundary is defined as a fuzzy set which contains elements with only partial membership. Unlike the two-valued Boolean logic (e.g. yes-no, one-zero, true-false, etc), fuzzy logic is a multi-valued logic and it deals with degrees of membership and degrees of truth. A membership function (MF) is defined as any curve that identifies how each point in the

input space is mapped to a membership value (or degree of membership) between 0 (completely false) and 1(completely true). Membership functions can be chosen arbitrarily based on simplicity, convenience, efficiency and speed. If a classical set is expressed as Eq. 7-1, then a fuzzy set A in X when X is the universe discourse with elements that are noted as x, are defined by Eq. 7-2:

$$A= [x| x > 10] \quad (7\text{-}1)$$

$$A= [x, \mu_A(x) \mid x \in X] \quad (7\text{-}2)$$

where $\mu_A(x)$ is called the membership function of x in A. Each element of X is mapped by the membership function to a value between 0 and 1.

Different membership functions are built based on basic functions like piecewise linear functions, the Gaussian distribution function, the sigmoid curve, quadratic and cubic polynomial curves. The simplest membership functions are triangular and trapezoidal functions. In Fuzzy logic, logical operations such as AND, OR, NOT, etc have their own definition based on the membership value concept. Generally fuzzy intersection is related to the operator AND, fuzzy union to OR and fuzzy complement to NOT.

7.3 Fuzzy rules

Another important component of Fuzzy logic is the Fuzzy rules; in order to express knowledge in a fuzzy-based system, conditional statements that comprise logic are used as Fuzzy rules. Fuzzy rules relate the fuzzy sets to each other. In the rule:

IF X is x THEN Y is y

x and y are linguistic or actual values of the variables X and Y respectively. The if-part and then-part of such conditional statements are called the “antecedent” and “consequent” respectively. Through fuzzy if-then rules, the inexact and uncertain modes of reasoning that have a significant impact on human decision-making ability can be captured (Jang 1993). Interpreting the if-then rules requires fuzzifying inputs such that all fuzzy statements in the antecedent are mapped to a membership between 0 and 1. If

a fuzzy set is only used in the antecedent part of a rule, and the consequent has a precise (non-fuzzy) value, then this rule is known as the Takagi-Sugeno fuzzy if-then rule (Takagi & Sugeno 1985) This form of if-then rules has an important role in fuzzy inference systems and they are used in control and modelling.

7.4 Fuzzy inference system

The process of mapping from a given input to an output using fuzzy logic is called fuzzy inference. A fuzzy inference system (FIS) is composed of four components: a) a set of fuzzy if-then rules; b) fuzzifying inputs which transform crisp values into fuzzy values based on the degree to which they belong to appropriate fuzzy sets through membership functions; c) an interface engine which performs inference operations on the rules; d) deffuzzification, which transforms the fuzzy results into a crisp output (Talei 2013). Deffuzzification can be done using different methods including centre of area, middle of maximum, largest maximum, and smallest maximum. A basic fuzzy inference system is shown in Figure 7.1.

Fuzzy inference systems are categorized into linguistic and precise models based on the method chosen to determine the output. The Mamdani fuzzy inference system (Mamdani & Assilian, 1975) is the most widely-used linguistic system. In the Mamdani fuzzy inference system the rules in both antecedent and the consequent are defined by fuzzy sets.

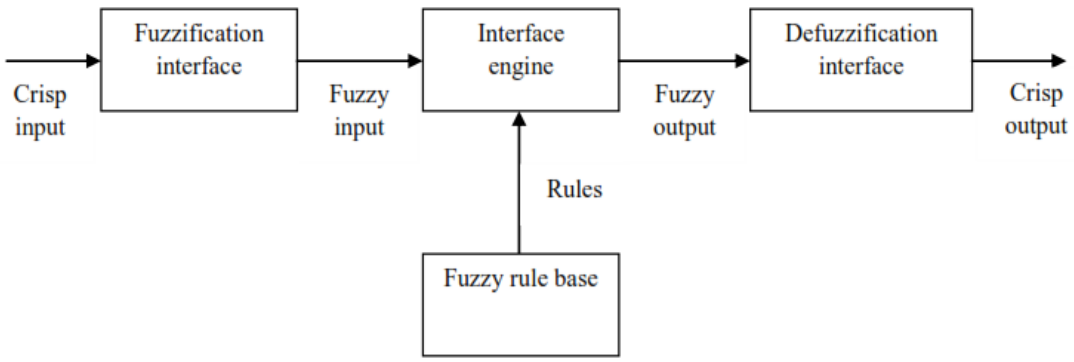

Figure 7-1. Basic structure of a fuzzy inference system (adapted from Talei, 2013)

Of the precise fuzzy models, the Takagi-Sugeno FIS is one of the most frequently used models (Takagi & Sugeno 1985). In the Takagi-Sugeno FIS, a fuzzy rule is composed of a weighted linear combination of crisp inputs rather than a fuzzy set. A typical Takagi-Sugeno FIS with two fuzzy if-then rules is the first-order Takagi–Sugeno FIS denoted as:

IF (x is A_1) AND (y is B_1) THEN ($f_1 = p_1x + q_1y + r_1$) (7-3)

IF (y is A_2) AND (y is B_2) THEN ($f_2 = p_2x + q_2y + r_2$) (7-4)

where A_1, A_2 and B_1, B_2 are membership values of the input variables x and y, respectively, and p_1, q_1, r_1 and p_2, q_2, r_2 are the parameters of the output functions f_1 and f_2, respectively. A typical Takagi-Sugeno inference system with a triangular membership functions producing an output function f from input variables x and y is shown in Figure 7.2.

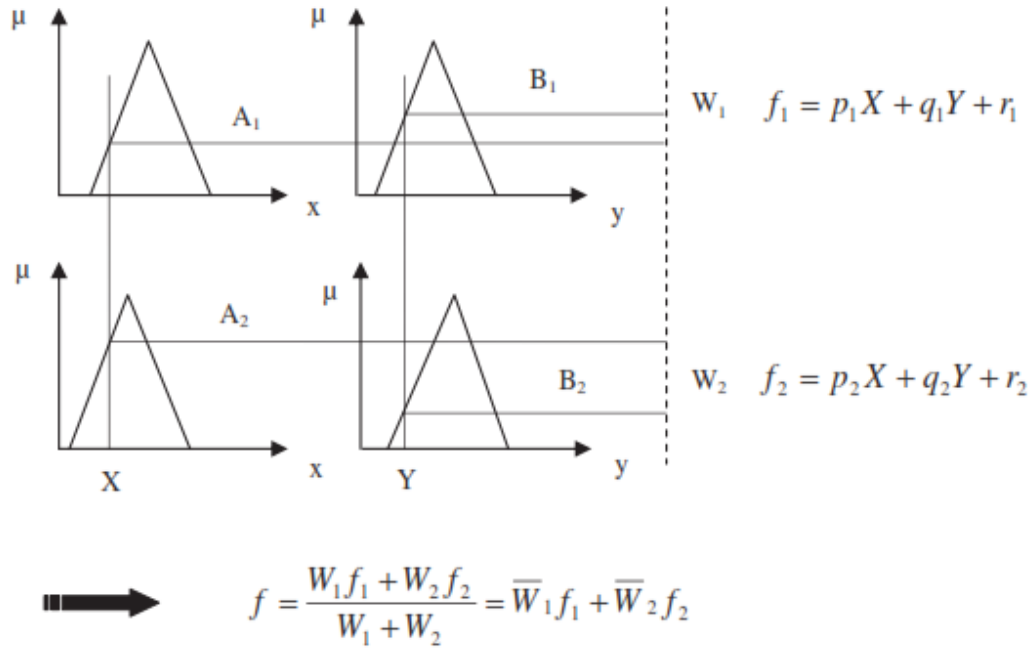

Figure 7-2. A typical Takagi-Sugeno fuzzy inference system (adapted from Talei et al. 2010)

7.5 Adaptive Network-Fuzzy Inference System (ANFIS)

The hybridization of artificial neural networks and fuzzy logic has resulted in adaptive neuro fuzzy inference systems where the membership functions of a FIS are tuned using a back-propagation algorithm. ANFIS has been successfully applied in various problems, especially on control processes such as chemical reactors, automatic trains and nuclear reactors. However, the main problem with fuzzy logic is that there is no systematic procedure to design a fuzzy controller (Chang & Chang, 2006). On the other hand, a neural network has the ability to learn from the input-output pairs, self-organize its structure, and adapt to it in an interactive manner. Thus, use of the Adaptive Network-based Fuzzy Inference system (ANFIS) methodology was proposed to self-organize the network structure and to adapt parameters of the fuzzy system (Chang & Chang 2006). Neuro-fuzzy systems (NFSs) are categorized into two groups: linguistic NFSs and precise NFSs. Linguistic NFSs use a Mamdani-type inference system in their structure while precise NFSs use a Sugeno-type (Takagi-Sugeno) inference system. ANFIS is the most widely used precise NFS and was developed by Jang (1993). In this model the global parameter tuning is done by minimizing the global error of the model (Jang 1993). ANFIS has the ability of extracting fuzzy rules from numerical data and adaptively constructing a rule base (Chang et al. 2006). A typical ANFIS structure with two inputs x and y and one output z is presented in Figure 7.2. The architecture of ANFIS consists of a five-layer MLP network. The structure and description of the layers are as follows (Talei et al. 2010):

Layer (1):

In this layer, the fuzzy membership values for an input variable are estimated. The output of the node *i* is defined as:

$$O_i^1 = \mu_{Ai}(x) \qquad \text{for } i = 1, 2 \tag{7-5}$$

$$O_i^1 = \mu_{Bi-2}(y) \qquad \text{for } i = 3, 4 \tag{7-6}$$

where x (or y) is input and μ_{Ai} (or μ_{Bi-2}) is the fuzzy set associated with this node.

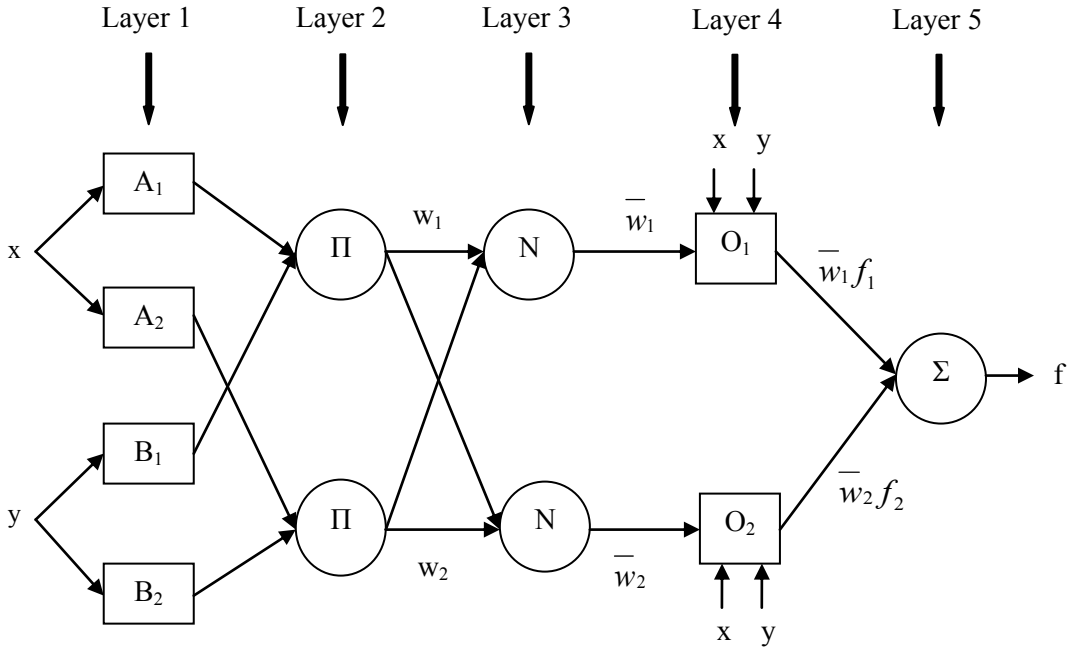

Figure 7-3. A typical ANFIS structure

This fuzzy set is characterized by the shape of a continuous and piecewise differentiable function called membership function. Assuming a triangular membership function, the output O_i^1 can be computed as:

$$O_i^1 = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right) \tag{7-7}$$

where *a* and *b* are the parameters that locate the feet of the triangle, while *c* is the parameter that locates the peak of the triangle.

Layer (2):

Multiply the incoming signals from the previous layer and calculate the firing strength of the rule. The output O_i^2 of the node i can be computed as:

$$O_i^2 = w_i = \mu_{Ai}(x)\mu_{Bi-2}(y) \qquad \text{for } i = 1, 2 \tag{7-8}$$

Layer (3):

Each node in this layer (denoted by N in Figure 7.2) computes the normalized firing strength as:

$$O_i^3 = \overline{w}_i = \frac{w_i}{w_1 + w_2} \qquad \text{for } i = 1, 2 \tag{7-9}$$

Layer (4):

The node i in this layer calculates the contribution of ith rule in the model output function, which is defined based on the first-order Takagi-Sugeno method as:

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i \left(p_i x + q_i y + r_i\right) \qquad \text{for } i = 1, 2 \tag{7-10}$$

Layer (5):

The single node of this layer calculates the weighted global output of the system as:

$$O^5 = f = \frac{\sum_i w_i f_i}{\sum_i w_i} = \sum_i \overline{w}_i f_i \tag{7-11}$$

In this study a back-propagation (BP) algorithm was used to modify the initially chosen membership functions and the least mean square (LMS) algorithm was used to determine the coefficients of the linear output functions (Jang 1993).

7.6 Results and Discussion

While ANFIS is applied to many cases of long-term and short-term rainfall forecasts, the use of large-scale climate indicators as potential rainfall predictors has not been considered in rainfall modelling using the ANFIS process in any study in Australia. The aim of developing ANFIS in the present research is to firstly, have a nonlinear bench

mark to compare with the developed ANN models; and secondly, to investigate whether the fuzzy logic characteristic of ANFIS, which has not been used previously in rainfall modelling using large-scale climate modes, can help improve the forecast accuracy of seasonal predictions.

The input sets chosen for ANFIS models are slightly different from the input sets developed for ANN models. ANFIS modelling requires a large amount of hardware memory. Owing to the large amount of data and constraints in hardware requirements, inputs with nine antecedent single climate modes were not able to be used in ANFIS model development, as more sophisticated computers are required. The data were classified into the calibration set (1900-1999) and the testing set (2000-2009). ANFIS and ANN models were developed using the calibration set and were later tested on the testing set. The input sets used in ANFIS model development comprise single (Nino3.4, SOI, and DMI) and combined (Nino3.4 $_{(Jun\text{-}Jul\text{-}Aug)}$-DMI$_{(Jun\text{-}Jul\text{-}Aug)}$, SOI $_{(Jun\text{-}Jul\text{-}Aug)}$-DMI$_{(Jun\text{-}Jul\text{-}Aug)}$, IPO$_{(Jun\text{-}July\text{-}Aug)}$-SOI$_{(Jun\text{-}July\text{-}Aug)}$, IPO$_{(Jun\text{-}July\text{-}Aug)}$-Nino3.4$_{(Jun\text{-}July\text{-}Aug)}$) antecedent climate modes and are shown in Table 7.1.

Table 7.1. Input sets for developing ANFIS models

Scenarios	Input matrix for each category	Input category	
Scenario 1-a	3×107	Nino3.4 $_{(Jun\text{-}July\text{-}Aug)}$	SOI $_{(Jun\text{-}July\text{-}Aug)}$ DMI $_{(Jun\text{-}July\text{-}Aug)}$
Scenario 2-a	6×107	Nino3.4 $_{(Jun\text{-}Jul\text{-}Aug)}$- DMI$_{(Jun\text{-}Jul\text{-}Aug)}$	SOI $_{(Jun\text{-}Jul\text{-}Aug)}$- DMI$_{(Jun\text{-}Jul\text{-}Aug)}$
Scenario 2-b	6×107	IPO$_{(Jun\text{-}July\text{-}Aug)}$-SOI$_{(Jun\text{-}July\text{-}Aug)}$	IPO$_{(Jun\text{-}July\text{-}Aug)}$-Nino3.4 $_{(Jun\text{-}July\text{-}Aug)}$

The ANFIS models developed in this study used the BP algorithm to modify the initially chosen membership functions and the LMS algorithm was used to determine the coefficients of the linear output functions. Different analyses were carried out to identify the appropriate number and type of membership functions for ANFIS. The

results showed that using 2 or 3 triangular membership functions is appropriate to achieve the best modelling results during calibration of the models.

After developing the models based on the input sets discussed above, the best models were chosen and tested on an out-of-sample test set. The performance criteria of the best models are shown in Table 7.2. Table 7.2 shows that in most of the stations, ANFIS models with antecedent single IOD set outperform both single ENSO predictors and combined ENSO-IOD predictors. The only region where ENSO index (Nino3.4) produces better models than IOD is central Victoria. For the same testing period, ANN models were also examined. The performance of ANN and ANFIS on out-of-sample test sets is shown in Table 7.3. The best ANN models are based on combined ENSO-IOD for east and central Victoria and single IOD for west Victoria. As discussed earlier, the nine antecedent months were not used in model building for ANFIS due to hardware limitations. In order to have the same benchmark in input selection for both ANFIS and ANN modelling, ANN models based on nine antecedent months were also not considered. It can be seen from Table 7.3 that, for east Victoria, ANN generally has a better performance than ANFIS. For central and west Victoria the performance of ANN and ANFIS are comparable, and both models show almost the same results.

Table 7.2. Performance of the ANFIS models

Region	Station	Model	MF*	r	RMSE	MAE
East	Bruthen	DMI	3	0.84	14.4	10.3
	Buchan	DMI	3	0.84	13.6	9.2
	Orbost	DMI	3	0.85	13.7	9.9
Centre	Malmsbury	Ni34	3	0.86	14.4	10.2
	Daylesford	DMI	2	0.54	23.8	17.5
	Heathcote	DMI	2	0.63	19.4	15.3
West	Horsham	DMI	2	0.60	15.0	11.1
	Kaniva	DMI	3	0.89	7.3	5.2
	Rainbow	DMI	2	0.56	13.0	9.7

*MF is the number of membership functions

Table 7.3. Comparison of the ANFIS and ANN models: testing set

		ANFIS				ANN			
Region	**Station**	**Model**	**r**	**RMSE**	**MAE**	**Model**	**r**	**RMSE**	**MAE**
East	Bruthen	DMI	0.53	16.7	14.0	SOI-DMI	0.51	16.6	11.7
	Buchan	DMI	0.29	24.0	22.3	Ni34-SOI	0.30	15.2	12.3
	Orbost	DMI	0.39	18.9	16.9	SOI-DMI	0.32	15.2	12.0
Centre	Malmsbury	Ni34	0.55	25.0	23.5	Ni34-DMI	0.23	21.8	15.2
	Daylesford	DMI	0.45	22.6	19.3	DMI	0.47	22.5	19.4
	Heathcote	DMI	0.45	21.9	18.2	SOI-DMI	0.57	21.2	17.3
West	Horsham	DMI	0.57	11.4	8.5	DMI	0.62	11.0	9.0
	Kaniva	DMI	0.66	13.6	10.1	DMI	0.56	15.3	12.1
	Rainbow	DMI	0.57	10.9	7.8	DMI	0.62	9.5	7.0

As discussed in Chapter 6, combined ENSO-IPO did not improve the performance of ANN models. In order to evaluate whether fuzzy logic is capable of extracting the complex relationship among ENSO, IPO and Victoria's spring rainfalls, ANFIS models were calibrated and tested for the four stations (Daylesford, Heathcote, Kaniva and Rainbow) where IPO is effective using ENSO-IPO input sets. Tables 7.4-7.7 show the performance of the models based on IPO for the four stations. It can be seen that in all four stations the combination of ENSO-IPO does not produce acceptable forecast results, and the performance of the ANFIS models is very poor in the test set. This result is in accordance with the ANN modelling results, where ENSO-IPO input sets did not improve the models' accuracy. In general, monthly IPO values combined with monthly ENSO indicators are not suitable predictors of spring rainfall in Victoria.

Table 7.4. Performance of the ANFIS-based models for Daylesford

	Train			Test		
Predictors	**r**	**RMSE**	**MAE**	**r**	**RMSE**	**MAE**
IPO-SOI	0.99	2.35	1.20	-0.26	54.5	39.6
IPO-Nino34	0.98	6.33	3.4	-0.13	48.7	42.7

Table 7.5. Performance of the ANFIS-based models for Heathcote

	Train			Test		
Predictors	**r**	**RMSE**	**MAE**	**r**	**RMSE**	**MAE**
IPO-SOI	0.99	1.84	1.10	-0.19	42.64	31.77
IPO-Nino34	0.97	5.80	3.24	0.44	27.25	20.20

Table 7.6. Performance of the ANFIS-based models for Kaniva

	Train			Test		
Predictors	**r**	**RMSE**	**MAE**	**r**	**RMSE**	**MAE**
IPO-SOI	0.99	1.13	0.57	-0.63	43.23	36.00
IPO-Nino34	0.96	4.26	2.14	-0.13	28.21	23.10

Table 7.7. Performance of the ANFIS-based models for Rainbow

	Train			Test		
Predictors	**r**	**RMSE**	**MAE**	**r**	**RMSE**	**MAE**
IPO-SOI	0.99	1.54	0.79	-0.47	40.89	35.0
IPO-Nino34	0.96	4.15	2.025	0.14	23.39	18.78

7.7 Summary of Adaptive Network-based Fuzzy Inference System (ANFIS) Analysis

In order to compare the results of ANN modelling with a nonlinear technique, ANFIS models were developed for the rainfall stations under study. This is the first time ANFIS approach is applied for Victoria's seasonal rainfall prediction with the use of large-scale climate modes. It was discovered that ANFIS models based on three antecedent months' IOD outperform those based on ENSO or combined ENSO-IOD. It was also discovered that combining ENSO and IPO does not improve the forecast results. Compared to ANFIS, it was found that ANN models have lower errors in east Victoria; however, the results of ANN and ANFIS models are almost comparable for central and west Victoria. The results show the ability of both ANN and ANFIS in forecasting spring rainfall in Victoria.

Chapter 8

Model Comparisons

8.1 Comparison of ANN and ANFIS forecasts with POAMA

The Predictive Ocean Atmosphere Model for Australia (POAMA) was developed by the Australian Bureau of Meteorology (BoM) together with the Commonwealth Scientific and Industrial Research Organization (CSIRO). POAMA is a coupled atmosphere–ocean climate prediction model that is based on atmosphere and ocean general circulation models (Lim et al., 2011). The first version of POAMA (POAMA-1) went operational in 2002 and continued to produced routine forecasts until 2007. In September 2007, POAMA 1.5 replaced POAMA-1 and later in 2011 POAMA-2 was adapted by BOM to produce routine seasonal forecasts two times a month. (BoM-2014). Forecasts are produces every week by running 33 scenarios for the coming 9 month. For example if 30 of the 33 ensemble members propose that the condition would be dry ahead, it is said that there is about a 90% chance of dry conditions in the next season. POAMA forecasts are initialized with observed atmospheric and oceanic conditions. According to Lim et al. (2011) the atmospheric initial conditions are provided by a new Atmosphere and Land Initialization (ALI) scheme. In 2013, a new version of POAMA-2 which is considered by BoM to be a state-of-the-art seasonal to inter-annual forecast system has become the new BoM official seasonal forecast model. The results of the rainfall forecasts are presented for grid pints of 250×250 km^2. POAMA produces quantitative seasonal and monthly rainfall predictions and present its official forecasts as the probability of exceeding the long-term average value (Abbot and Marohasy, 2014). The POAMA forecasts are used as benchmark for comparison with the results of the developed models in this study.

The POAMA rainfalls obtained from BoM are in the form of anomalies; these anomalies were used to calculate the magnitude of POAMA-forecast rainfall for every spring by adding the seasonal rainfall climatology of 1981-2010 to each anomaly (E. Lim, personal communication, 30 Sep, 2014). The proposed method can also be found

in Cottrill et al. (2013).

The POAMA forecast results were obtained for the period 2000-2009; the ANN and ANFIS results for the out-of-sample test set for the same period are compared with those of the POAMA. Table 8.1 shows the comparison between ANN and POAMA forecasts. As shown in the table, ANN forecasts are comparable with POAMA forecasts in regard to error criteria for most of the stations and show superior results, especially in west Victoria. It can be seen from Table 8.1 that ANN models outperform POAMA in terms of correlation coefficients in all stations except Malmsbury. The differences between the correlation coefficients of ANN models and POAMA are very significant in central and west Victoria. The r_{max} for the POAMA model in central Victoria is 0.30, as opposed to 0.57 for ANN, in west Victoria r_{max} is 0.19 for the POAMA as opposed to 0.62 for ANN. It can be seen from Table 8.1 that both POAMA and ANN models show higher errors in central Victoria compared to east and west Victoria. It seems that the process governing central Victoria's rainfall is very complicated compared to east and west Victoria, and more research needs to be done in this region. In order to have a more detailed comparison between the models the observed and simulated rainfall values by ANN and POAMA for the period 2000-2009 are plotted and shown in Figures 8.1 to 8.3.

Table 8.1. Comparison of POAMA and ANN models

		ANN			POAMA		
Region	**Station**	**MAE**	**RMSE**	**r**	**MAE**	**RMSE**	**r**
East	Bruthen	11.7	16.6	0.51	12.1	15.9	0.21
	Buchan	12.3	15.2	0.30	14.0	15.70	0.26
	Orbost	12.0	15.2	0.32	11.6	15.31	0.20
Central	Malmsbury	15.2	21.8	0.23	19.5	21.2	0.30
	Daylesford	19.4	22.5	0.47	20.7	23.87	0.051
	Heathcote	17.3	21.2	0.57	19.3	23.29	0.23
West	Horsham	9.0	11.0	0.62	11.8	13.86	0.015
	Kaniva	12.1	15.3	0.56	13.5	18	0.19
	Rainbow	7.0	9.5	0.62	13.2	13.67	0.065

Table 8.2 shows the comparison between POAMA and ANFIS models. It can be seen that POAMA perform better in east Victoria compared to ANFIS models in terms of error, however the correlation coefficient of ANFIS models are better in this region. For central and west Victoria, ANFIS models are superior to POAMA.

Table 8.2. Comparison of POAMA and ANFIS models

Station	ANFIS			POAMA		
	RMSE	MAE	r	RMSE	MAE	r
Bruthen	21.0	14.0	0.53	15.9	12.1	0.21
Buchan	24.7	22.3	0.29	15.70	14.0	0.26
Orbost	18.9	16.9	0.39	15.31	11.6	0.20
Malmsbury	25.0	23.5	0.55	21.2	19.5	0.30
Daylesford	22.6	19.3	0.45	23.9	20.7	0.051
Heathcote	22	18.2	0.45	23.3	19.3	0.23
Horsham	11.4	8.5	0.57	13.9	11.8	0.015
Kaniva	13.6	10.1	0.66	18.0	13.5	0.19
Rainbow	10.9	7.8	0.57	13.7	13.2	0.065

Figures 8.4 to 8.6 show the results of ANFIS and POAMA forecasts. As can be seen in the figures, the POAMA tends to give more flat rainfall forecasts when facing extreme cases compared to ANN and ANFIS. Since ANN and ANFIS learn from the associations they are more responsive against extreme cases. In general, by comparing the results of the ANN, ANFIS and POAMA it can be concluded that until POAMA further develops, ANN and ANFIS approach are reliable nonlinear statistical options which can produce similar and in some cases better forecasts which could be beneficial for agriculture and water management.

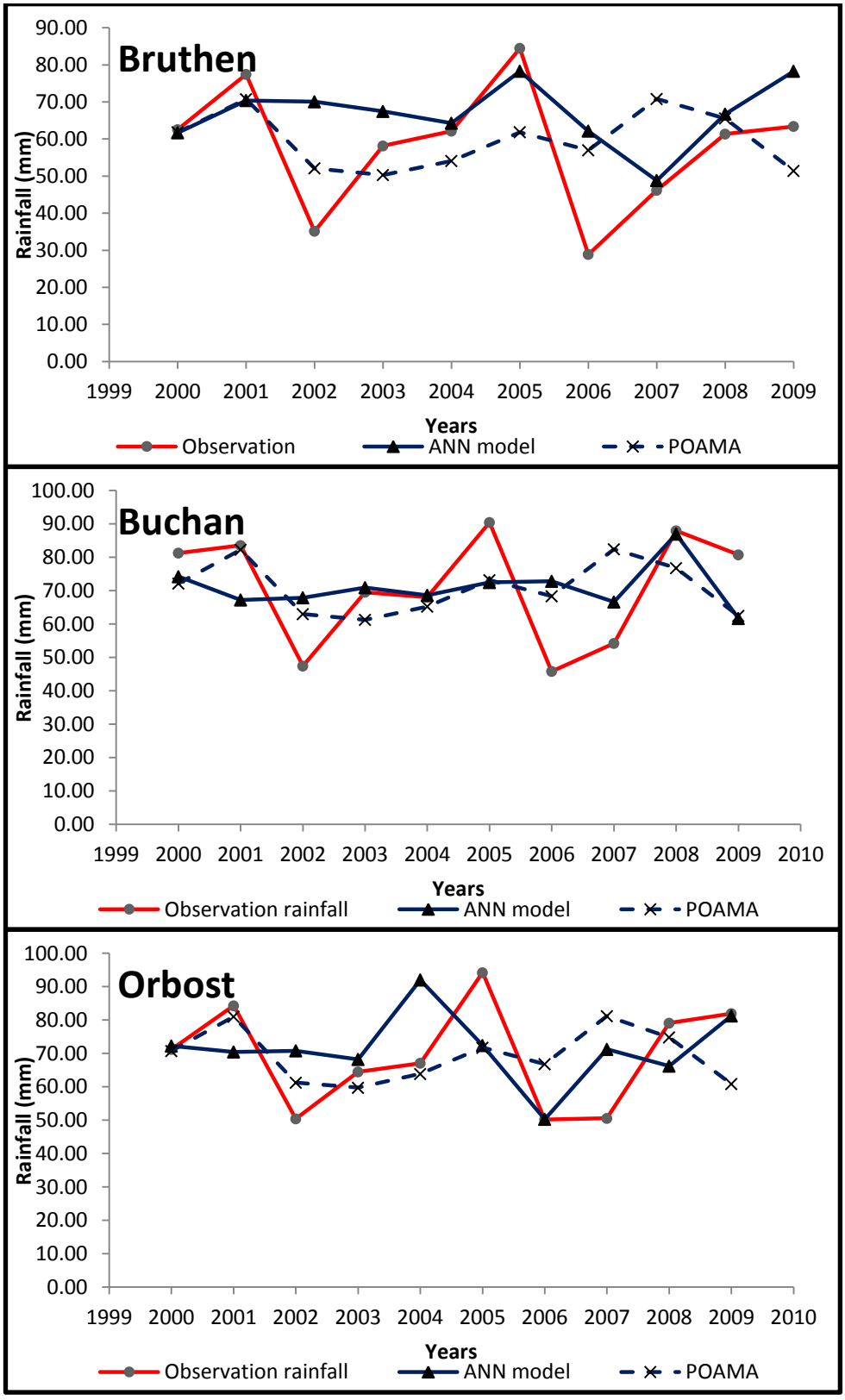

Figure 8-1. Comparison of ANN models with POAMA for east Victoria

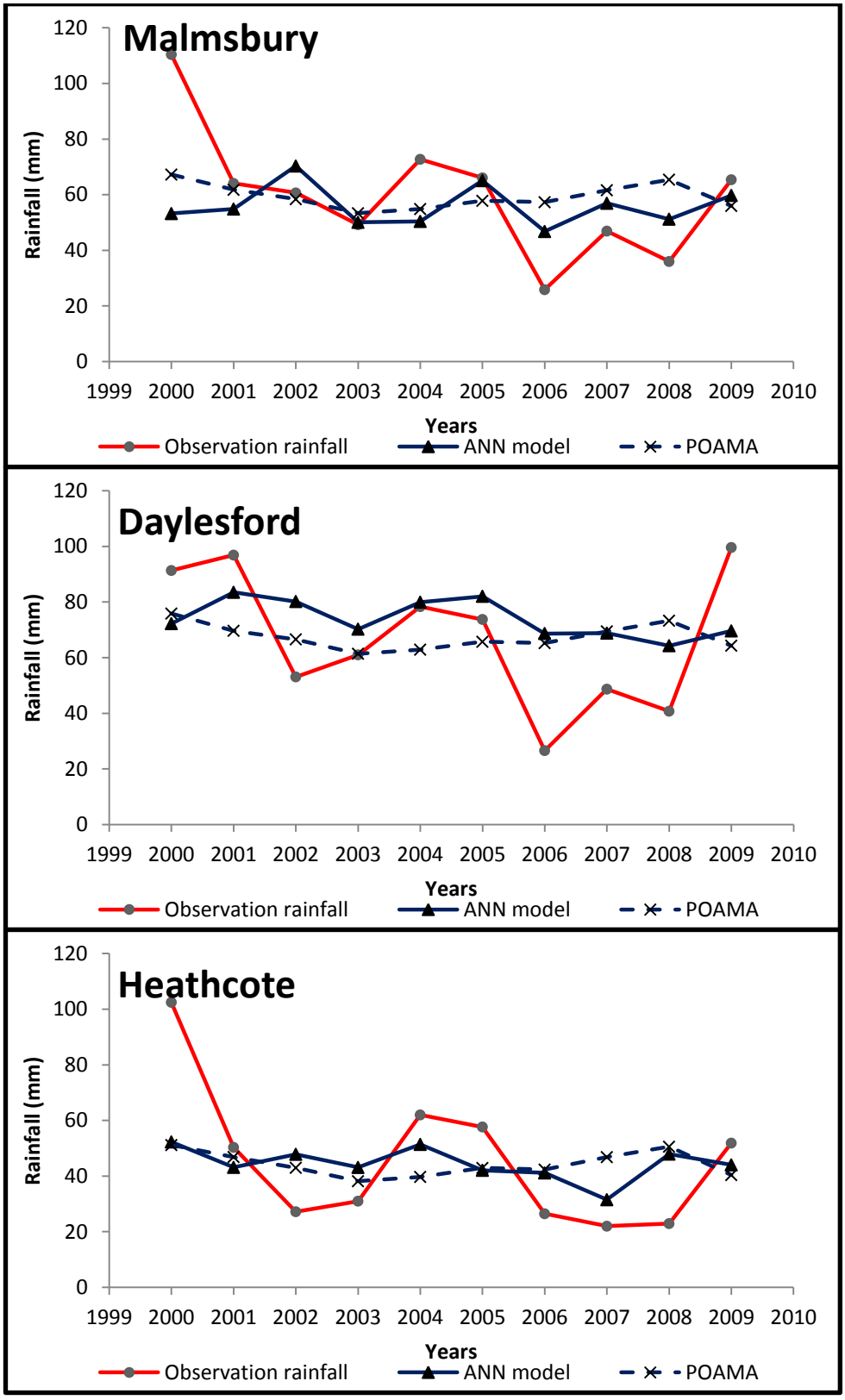

Figure 8-2. Comparison of ANN models with POAMA for central Victoria

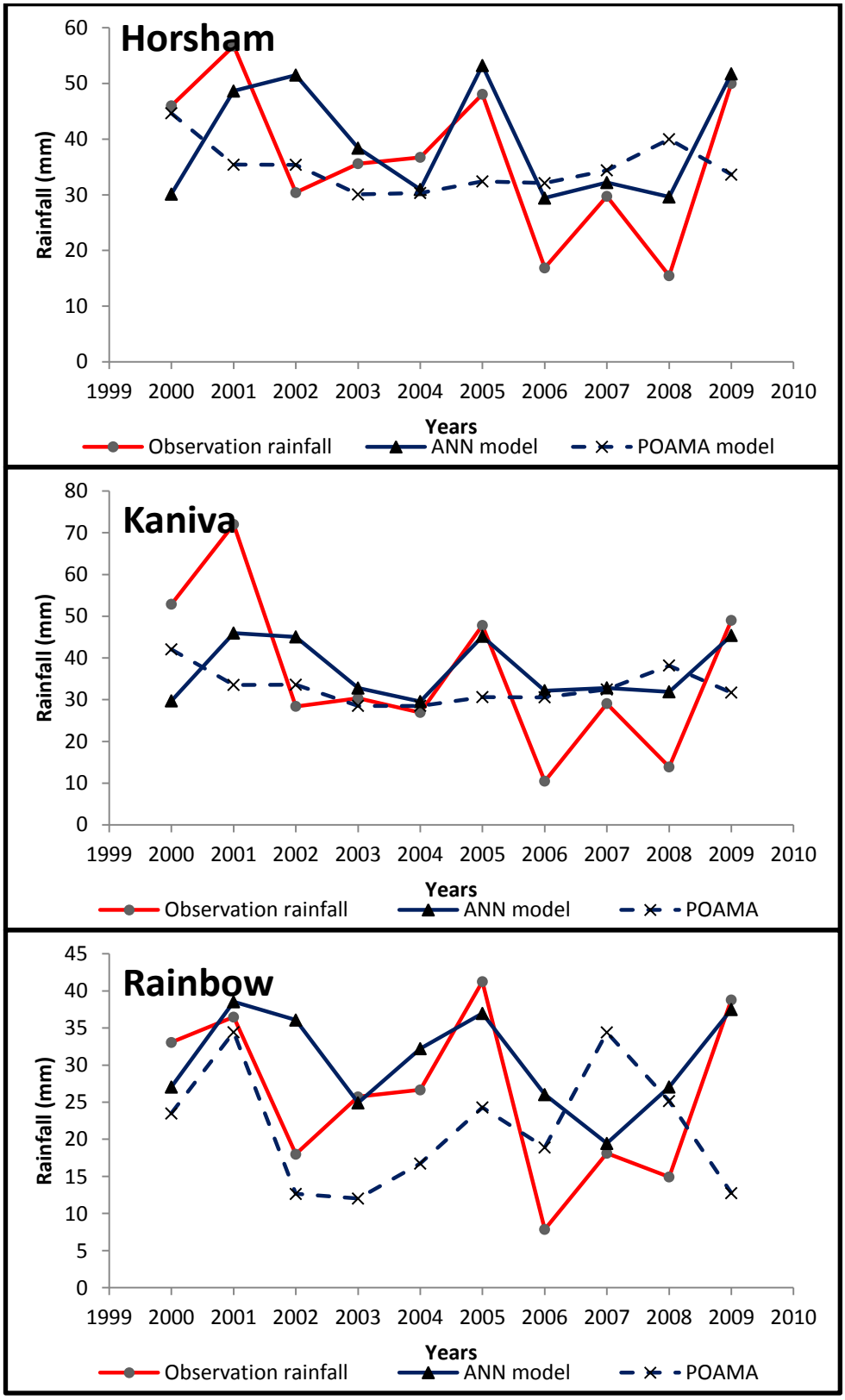

Figure 8-3. Comparison of ANN models with POAMA for west Victoria

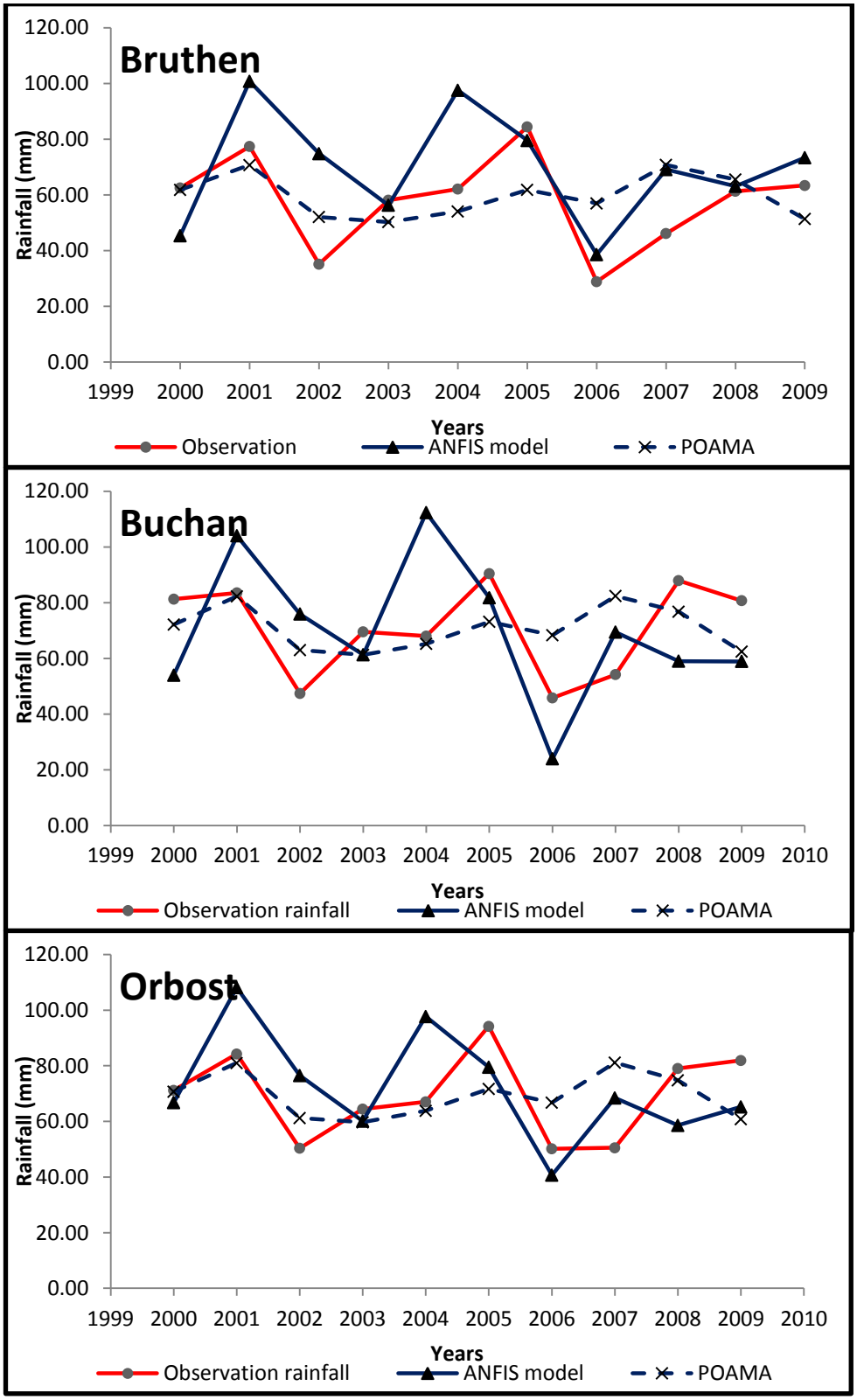

Figure 8-4. Comparison of ANFIS models with POAMA for east Victoria

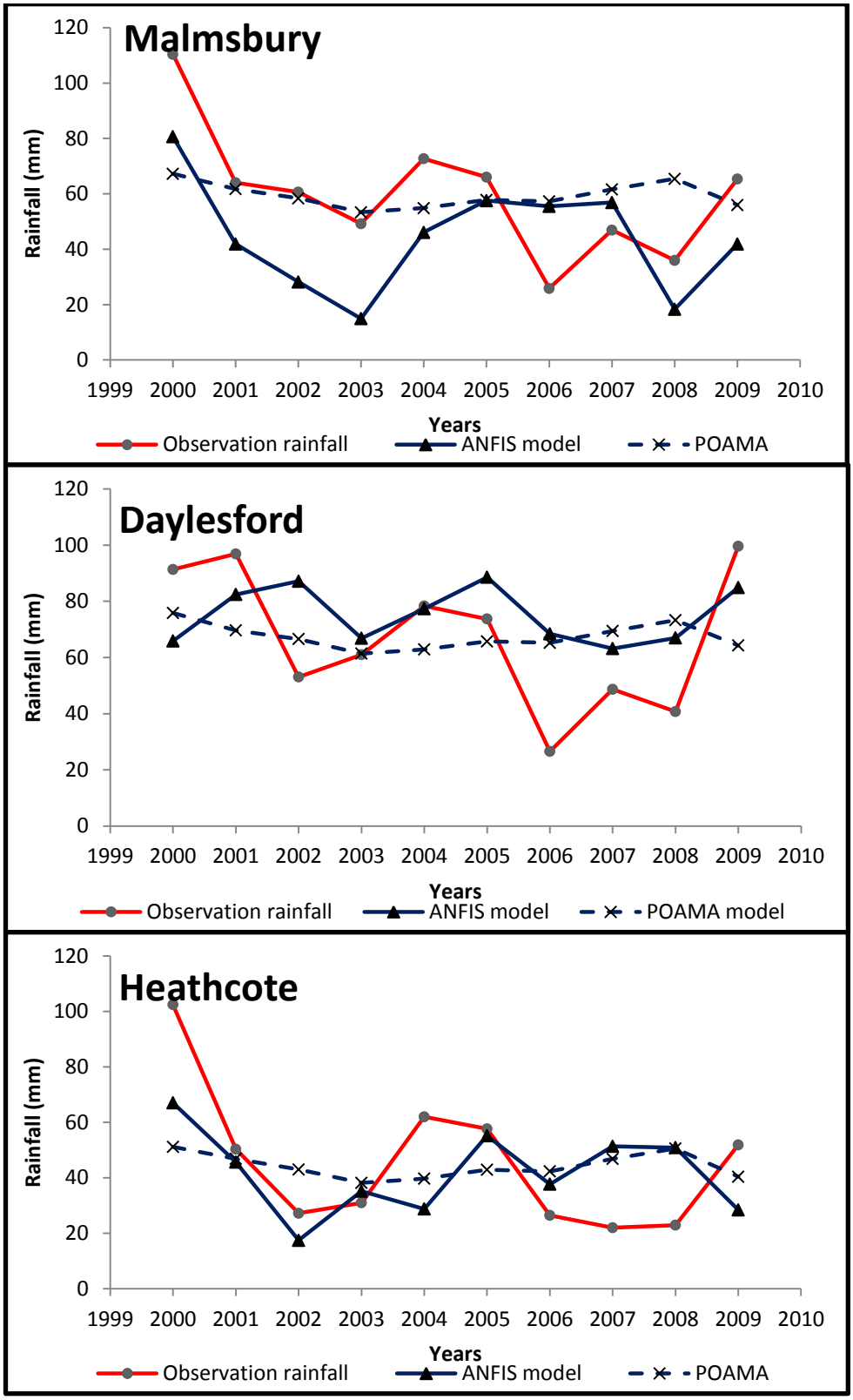

Figure 8-5. Comparison of ANFIS models with POAMA for central Victoria

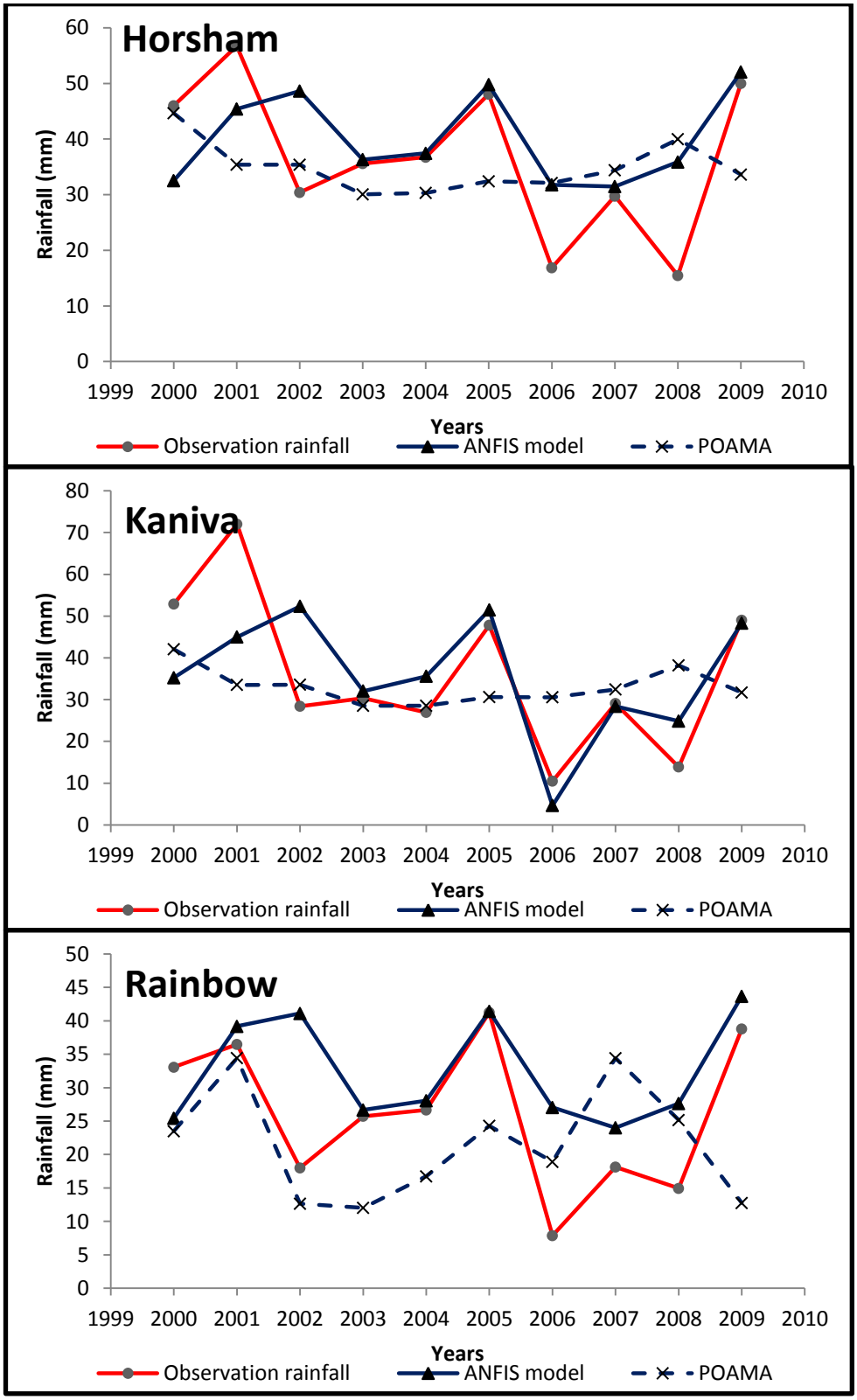

Figure 8-6. Comparison of ANFIS models with POAMA model for west Victor

8.2 Summary

In this study the results of the developed nonlinear models were compared with the Predictive Ocean Atmosphere Model for Australia (POAMA) which is the official forecast model used by the BoM. The results revealed that ANN models are comparable with POAMA in all of the stations in the case study in terms of error criteria and are superior in terms of trend criteria. The results also revealed that ANFIS models are better estimators of future rainfall in central and west Victoria compared to POAMA however, POAMA performs better in east Victoria. The model comparison suggests that nonlinear modelling techniques such as ANN and ANFIS are promising tools for forecasting Victoria's seasonal rainfall.

Chapter 9

Summary, Conclusion and Recommendations

9.1 Summary

The main focus of this study is the development of a non-linear rainfall forecast model for Victoria, Australia using antecedent large-scale climate predictors. Artificial Neural Network (ANN) was chosen as the primary modelling technique due to its capability to extract complex relationships from the data. ANN has been rarely used in rainfall forecasting in Australia in conjunction with the use of large-scale climate modes. In order to compare the forecast results of ANN models with those of linear and non-linear models, Multiple Linear Regression (MLR) models and Adaptive Network-based Fuzzy Inference System (ANFIS) were developed respectively. The large-scale climate modes taking place in the Pacific and Indian Oceans were considered in this study as potential rainfall predictors; the El Nino Southern Oscillation (ENSO) and the Interdecadal Pacific Oscillation (IPO), which occur in the Pacific Ocean, and the Indian Ocean Dipole (IOD) which occurs in the Indian Ocean were examined in this study. Furthermore, this study investigated the concurrent and antecedent relationship between seasonal rainfall and large-scale climate modes; classification analysis and Pearson correlation analysis were used in this regard. Three distinct regions in Victoria, Australia were considered as case studies; from each region three rainfall stations were selected. Monthly rainfall and climate mode data were obtained from the Australian Bureau of Meteorology (BoM) and the Royal Netherlands Meteorological Institute (KNMI) Climate Explorer website, respectively. The Nino3.4 and Southern Oscillation Index (SOI) were chosen as ENSO indicators and the Dipole Mode Index (DMI) was chosen as the IOD indicator. The IPO index was obtained from Parker et al., (2007).

In the classification analysis, seasonal rainfall anomalies were constructed and classified based on the phases of ENSO and IOD in two different steps. In the first step, the anomalies were classified based on the years of El Nino, La Nina and Neutral for ENSO and the years of positive IOD (pIOD), negative IOD (nIOD) and Neutral for IOD.

Wilcoxon Signed Rank tests were applied to the classified anomalies in order to examine the significance of the median of each category from zero. Categories with zero median are known as normal rainfalls. It was revealed that spring rainfall is affected by the phases of ENSO and IOD; however, winter rainfall is less influenced by these phases. It was also revealed that east Victoria is only influenced by El Nino and pIOD, which are both related to decreased rainfall and dry conditions in Australia.

In the next step of classification analysis, the co-occurrence of the phases of ENSO and IOD and their effect on Victoria's seasonal rainfall was investigated. Seasonal rainfall anomalies were classified based on the years of El Nino–pIOD, pure El Nino, pure pIOD, Neutral, pure La Nina, La Nina-nIOD and pure nIOD. In the present study, the term "pure" is used to identify that the only active phase of climate is the phase with which the term "pure" is associated and the other climate mode is in neutral condition. The Wilcoxon Signed Rank test was applied to examine the statistically significant median from zero in each category. It was revealed that the dry phases of ENSO and IOD and their occurrence (i.e. pure El Nino events, pure pIOD events and the co-occurrence of El Nino-pIOD) consistently result in dry conditions, while the wet phases of these phenomena (La Nina, nIOD and the co-occurrence of La Nina-nIOD) are highly variable. It was also revealed that the only phase of climate affecting east Victoria's rainfall is the co-occurrence of El Nino and pIOD. The results highlight the importance of the dry phases of ENSO/IOD for spring rainfall and the necessity of forecasting spring rainfall based on large-scale climate modes as potential predictors.

While classification analysis gave a good understanding of the natural causes of the increase or decrease in rainfall over Victoria, it could not show the magnitude of the influence of the climate modes on seasonal rainfall. In order to gain an understanding of the magnitude of strength of different climate modes on seasonal rainfall in Victoria, Pearson correlation analysis was further used in two separate approaches. In the first approach, the concurrent relationship between seasonal rainfall and climate modes was investigated. The first approach consisted of two sections; in the first section the correlation coefficient between seasonal rainfall and seasonal climate modes was calculated. It was revealed that only the cool season rainfalls (winter and spring) have

statistically significant relationships with the climate modes, and these relationships proved to be moderate to weak. In the second section of the concurrent correlation analysis, the rainfalls based on the years of El Nino/La Nina were merged in order to remove the neutral years; rainfalls based on pIOD/ nIOD were also merged and Neutral IOD years were removed. Nino3.4, SOI and DMI were also merged separately based on the years of El Nino/La Nina and pIOD/nIOD. Pearson correlation analysis between the merged rainfall and merged ENSO and IOD indices was conducted in order to examine the strength of the relationship between rainfall and the effective phases of ENSO and IOD. It was discovered that the correlation coefficients between the merged rainfalls and merged climate modes increased significantly, which revealed the magnitude of the strengths of these relationships.

In addition to examining the concurrent relationship between rainfall and climate modes, the relationship between spring rainfall and antecedent climate modes was also examined using Pearson correlation analysis. It was discovered that only the climate modes in the three months of June, July and August have statistically significant relationships with spring rainfall, although these relationships are quite weak. It was also found that IPO is only effective in some parts of Victoria and not all stations show significant lagged relationships with IPO.

The statistically significant lagged climate modes were used as the basis of input selection for seasonal rainfall forecasting models using MLR, ANN, and ANFIS techniques. MLR models were constructed based on combined antecedent ENSO-IOD to forecast rainfall for the three regions under study. It was revealed that, due to statistical limitations, the MLR models with the multiple lagged climate mode (e.g. $Nino3.4_{(Jun\text{-}Jul\text{-}Aug)}$-$DMI_{(Jun\text{-}Jul\text{-}Aug)}$) were not reliable, as they did not meet the statistical assumptions of the MLR technique; only some of the single month ENSO-IOD (e.g. $Nino3.4_{(Aug)}$-$DMI_{(Aug)}$) models were statistically significant. After calibrating the models, the models were tested and used to forecast spring rainfall three consecutive years in advance in the stations under study; the models were able to forecast spring rainfall with acceptable errors for some stations, while showing less accurate results for the others. As MLR is a linear technique, it is not able to capture complex non-linear

relationships. ANN models were therefore further developed in order to forecast spring rainfall based on antecedent large-scale climate modes.

In developing ANN models, two scenarios were considered in selecting the inputs. The first scenario consisted of two different input sets; the first set was developed based on the results of lagged correlation analysis, i.e. antecedent single climate modes (Nino3.4, SOI, DMI and IPO) were considered, which took into account the three months of June, July and August (e.g. $Nino3.4_{(Jun\text{-}Jul\text{-}Aug)}$). As Pearson correlation is a linear technique and does not capture the non-linear relationships, the second input set was developed in order to examine a broader range of climate modes. The second set consisted of single climate modes for nine antecedent months from December in the previous year until August of the current year (e.g. $Nino3.4_{(Dec\text{-}\ Aug)}$). It was revealed that ANN models based on the wider time frame of climate modes (nine antecedent months) showed superior results compared to models based on three months' climate modes. It was also found that models based on nine months' IOD outperformed models based on ENSO indicators.

The second scenario of input development was based on combinations of climate modes. Two sets of combined antecedent ENSO-IOD and ENSO-IPO based on the three months of June, July and August (e.g. $Nino3.4_{(Jun\text{-}Jul\text{-}Aug)}$-$DMI_{(Jun\text{-}Jul\text{-}Aug)}$) were developed and used in ANN modelling for forecasting spring rainfall in the stations under study. It was discovered that the combined ENSO-IOD models outperform the combined ENSO-IPO models. Comparing the results of single climate mode models with combined climate mode models revealed that other than in east Victoria, where combined climate mode models show superior results, for the rest of Victoria single IOD models based on nine antecedent months outperform the combined climate mode models. It was concluded that, for most of Victoria, IOD is the dominant predictor of spring rainfall. The physical reason for this may be the fact that east Victoria is closer to the Pacific Ocean where ENSO occurs, and the effect of both ENSO and IOD can therefore be felt in this region. However, towards central and west Victoria and closer to the Indian Ocean, IOD becomes the dominant predictor of rainfall in Victoria.

In order to compare the results of ANN models with another non-linear technique, ANFIS models were developed. As ANFIS modelling requires sufficient hardware memory, it was not possible to use the nine antecedent climate modes in the modelling process due to hardware restrictions. The results of ANFIS models based on single three antecedent month climate modes and combined antecedent months ENSO-IOD and ENSO-IPO revealed that models based on three antecedent months' IOD outperform those based on ENSO. It was also revealed that combining ENSO-IOD and ENSO-IPO does not improve the forecast results. By comparing ANN and ANFIS models, it was found that ANN models have lower errors in east Victoria, and the results of ANN and ANFIS models are comparable for central and west Victoria.

In order to evaluate the accuracy of the developed models with an official benchmark, the Predictive Ocean Atmosphere Model for Australia (POAMA) was considered. POAMA is the official forecast model used in the Australian Bureau of Meteorology (BoM) for daily to seasonal rainfall forecasts. POAMA is a dynamic model which uses a variety of ocean and atmosphere variables in order to forecast rainfall. The POAMA seasonal forecasts for the stations under study were obtained from the BoM and compared with the results of ANN and ANFIS models. It was revealed that ANN forecasts are comparable with POAMA forecasts in regard to error criteria for most of the stations, and show superior results especially in west Victoria. It was found that ANN models outperform POAMA in terms of correlation coefficients in all stations except one. The superiority of the correlation coefficients of ANN models compared to POAMA is very significant in central and west Victoria. The results show that both POAMA and ANN models produce higher errors in central Victoria than in east and west Victoria. It appears that the process governing central Victoria's rainfall is very complicated compared to east and west Victoria, and more research needs to be done in this region. In general, by comparing the results of the ANN and POAMA models it can be concluded that until POAMA further develops, with a simpler non-linear statistical technique like ANN with fewer input variables similar and even much better results can be obtained compared to using a complex dynamic model such as POAMA.

The results of ANFIS models were also compared to those of the POAMA. It was discovered that in terms of error criteria, POAMA performs better in the east and part of central Victoria than ANFIS models. In west and part of central Victoria, ANFIS models are comparable with if not better than the POAMA. In terms of the correlation coefficients of the models, ANFIS models are better than POAMA in all the stations studied. In general, the POAMA tends to give more flat rainfall forecasts when facing extreme cases compared to ANN and ANFIS. Since these two artificial intelligence techniques learn from the associations, they are more responsive to extreme cases. The forecast of extreme cases is crucial for flood management and drought mitigation; ANFIS and ANN models are more reliable in this regard than POAMA.

9.2 Conclusion and Recommendations

The cool season rainfall (winter and spring) in Victoria, Australia is under the influence of large-scale climate modes such as ENSO and IOD. The dry phases of large-scale climate modes have more effect on Victoria's rainfall than the wet phases. The linear relationships between Victoria's cool season rainfall and large-scale climate modes are not strong. Seasonal rainfall can be modeled and forecast by the use of artificial intelligence techniques such as Artificial Neural Network (ANN) and Adaptive Network-based Fuzzy Inference System (ANFIS) with good accuracy in Victoria, Australia. In the process of forecasting seasonal rainfall in this region, antecedent large-scale climate modes as potential predictors of rainfall can be used in ANN and ANFIS modelling approaches for accurate seasonal rainfall forecasting. The Indian Ocean climate mode produces more accurate rainfall forecasts compared to the Pacific Ocean climate modes in most of Victoria when used in the developed non-linear models. The results of the developed ANN models are comparable with and even better than the forecast results of the official model (POAMA) used by the Australian Bureau of Meteorology. As a pioneer study, this research has revealed the potential of artificial intelligence techniques in seasonal rainfall forecasting in Australia. The study has revealed that the use of antecedent large-scale climate modes in non-linear modelling techniques can be beneficial for more accurate rainfall forecasting around Australia. This research may be extended in future by developing forecasting models for other

parts of Australia, improving the models by incorporating genetic algorithm techniques for more accurate input selection, and expanding and improving the models for grid rainfall forecasts as opposed to station rainfall forecasts. The approach taken in this study can be used in official rainfall forecasts.

References

ABBOT, J. & MAROHASY, J. 2012. Application of artificial neural networks to rainfall forecasting in Queensland, Australia. *Advances in Atmospheric Sciences,* 29**,** 717-730.

ABBOT, J. & MAROHASY, J. 2014. Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks. *Atmospheric Research,* 138**,** 166-178.

ABDUL RAUF, U. F. & ZEEPHONGSEKUL, P. 2014. Copula based analysis of rainfall severity and duration: A case study. *Theoretical and Applied Climatology,* 115**,** 153-166.

ABOLPOUR, B., JAVAN, M. & KARAMOUZ, M. 2007. Water allocation improvement in river basin using adaptive neural fuzzy reinforcement learning approach. *Applied Soft Computing,* 7**,** 265-285.

ABRAHART, R.J., KNEALE, P.E., SEE, L.M., 2004. Neural networks for hydrological modelling. Taylor & Francis

AKHTAR, M., CORZO, G., VAN ANDEL, S. & JONOSKI, A. 2009. River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin. *Hydrology and Earth System Sciences,* 13**,** 1607.

ANDERSON, J., VAN DEN DOOL, H., BARNSTON, A., CHEN, W., STERN, W. & PLOSHAY, J. 1999. Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bulletin of the American Meteorological Society,* 80**,** 1349-1361.

ASHOK, K., GUAN, Z., SAJI, N. & YAMAGATA, T. 2004. Individual and combined influences of ENSO and the Indian Ocean dipole on the Indian summer monsoon. *Journal of Climate,* 17**,** 3141-3155.

ASHOK, K., GUAN, Z. & YAMAGATA, T. 2001. Impact of the Indian Ocean dipole on the relationship between the Indian monsoon rainfall and ENSO. *Geophysical Research Letters,* 28**,** 4499-4502.

ASHOK, K., GUAN, Z. & YAMAGATA, T. 2003. Influence of the Indian Ocean Dipole on the Australian winter rainfall. *Geophysical Research Letters,* 30**,** CLM 6-1 - 6-4.

BAMSTON, A. G., CHELLIAH, M. & GOLDENBERG, S. B. 1997. Documentation of a highly ENSO-related sst region in the equatorial pacific: Research note. *Atmosphere-ocean*, 35, 367-383.

BARNSTON, A. G., HE, Y. & GLANTZ, M. H. 1999. Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997-98 El Niño episode and the 1998 La Niña onset. *Bulletin of the American Meteorological Society,* 80**,** 217-243.

BARSUGLI, J. J. & SARDESHMUKH, P. D. 2002. Global atmospheric sensitivity to tropical SST anomalies throughout the Indo-Pacific basin. *Journal of Climate,* 15**,** 3427-3442.

BEECHAM, S. & CHOWDHURY, R. 2010. Temporal characteristics and variability of point rainfall: a statistical and wavelet analysis. *International Journal of Climatology,* 30**,** 458-473.

Bom.gov.au, (2013). *Australia's official weather forecasts & weather radar - Bureau of Meteorology*. [online] Available at: http://www.bom.gov.au/ [Accessed 25 June 2013].

BOOKHAGEN, B. & STRECKER, M. R. 2010. Modern Andean rainfall variation during ENSO cycles and its impact on the Amazon drainage basin. *Amazonia, landscape and species evolution: a look into the past, 1st edn. Blackwell, Oxford***,** 223-243.

CAI, W., COWAN, T. & SULLIVAN, A. 2009. Recent unprecedented skewness towards positive Indian Ocean Dipole occurrences and its impact on Australian rainfall. *Geophysical Research Letters,* 36.

CAI, W., VAN RENSCH, P., COWAN, T. & HENDON, H. H. 2011. Teleconnection pathways of ENSO and the IOD and the mechanisms for impacts on Australian rainfall. *Journal of Climate,* 24**,** 3910-3923.

CAI, W., VAN RENSCH, P., COWAN, T. & HENDON, H. H. 2012. An asymmetry in the IOD and ENSO teleconnection pathway and its impact on Australian climate. *Journal of Climate,* 25**,** 6318-6329.

CAMPOLO, M., ANDREUSSI, P. & SOLDATI, A. 1999. River flood forecasting with a neural network model. *Water Resources Research,* 35**,** 1191-1197.

CHAKRAVERTY, S. & GUPTA, P. 2008. Comparison of neural network configurations in the long-range forecast of southwest monsoon rainfall over India. *Neural Computing & Applications,* 17**,** 187-192.

CHAMBERS, L. E. 2003. *South Australian rainfall variability and trends*, Bureau of Meteorology Research Centre.

CHANG, F.-J. & CHANG, Y.-T. 2006. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Advances in Water Resources,* 29**,** 1-10.

CHANG, F. J., CHIANG, Y. M., TSAI, M. J., SHIEH, M. C., HSU, K. L. & SOROOSHIAN, S. 2014. Watershed rainfall forecasting using neuro-fuzzy networks with the assimilation of multi-sensor information. *Journal of Hydrology*, 508, 374-384.

CHATTOPADHYAY, G., CHATTOPADHYAY, S. & JAIN, R. 2010. Multivariate forecast of winter monsoon rainfall in India using SST anomaly as a predictor: Neurocomputing and statistical approaches. *Comptes Rendus Geoscience,* 342**,** 755-765.

CHATTOPADHYAY, S. 2007. Feed forward Artificial Neural Network model to predict the average summer-monsoon rainfall in India. *Acta Geophysica,* 55**,** 369-382.

CHAVOSHI, S., SULAIMAN, W. N. A., SAGHAFIAN, B., SULAIMAN, M. N. B. & MANAF, L. A. 2013. Flood prediction in southern strip of Caspian Sea watershed. *Water Resources*, 40, 593-605.

CHENG, K., HSU, H., TSAI, M., CHANG, K. & LEE, R. 2004. Test and analysis of trend existence in rainfall data. *In:* Asian Pacific Association of Hydrology and Water Resources 2nd Conference, Singapore,

CHIANG, Y. M. & CHANG, F. J. 2009. Integrating hydrometeorological information for rainfall-runoff modelling by artificial neural networks. *Hydrological Processes,* 23**,** 1650-1659.
CHIANG, Y. M., CHANG, L. C. & CHANG, F. J. 2004. Comparison of static-feedforward and dynamic-feedback neural networks for rainfall–runoff modelling. *Journal of Hydrology,* 290**,** 297-311.
CHIEW, F., PIECHOTA, T. C., DRACUP, J. & MCMAHON, T. 1998. El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting. *Journal of Hydrology,* 204**,** 138-149.
CHOWDHURY, R. & BEECHAM, S. 2010. Australian rainfall trends and their relation to the southern oscillation index. *Hydrological Processes,* 24**,** 504-514.
CHOWDHURY, R. K. & BEECHAM, S. 2013. Influence of SOI, DMI and Niño3. 4 on South Australian rainfall. *Stochastic Environmental Research and Risk Assessment,* 27**,** 1909-1920.
COTTRILL, A., HENDON, H. H., LIM, E. P., LANGFORD, S., SHELTON, K., CHARLES, A., MCCLYMONT, D., JONES, D. & KULESHOV, Y. 2013. Seasonal forecasting in the pacific using the coupled model POAMA-2. *Weather and Forecasting,* 28**,** 668-680.
COULIBALY, P., ANCTIL, F., ARAVENA, R. & BOBÉE, B. 2001. Artificial neural network modelling of water table depth fluctuations. *Water Resources Research,* 37**,** 885-896.
COULIBALY, P., ANCTIL, F. & BOBEE, B. 2000. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *Journal of Hydrology,* 230**,** 244-257.
DALIAKOPOULOS, I. N., COULIBALY, P. & TSANIS, I. K. 2005. Groundwater level forecasting using artificial neural networks. *Journal of Hydrology,* 309**,** 229-240.
DAWSON, C. W., SEE, L. M., ABRAHART, R. J. & HEPPENSTALL, A. J. 2006. Symbiotic adaptive neuro-evolution applied to rainfall–runoff modelling in northern England. *Neural Networks,* 19**,** 236-247.
DE VOS, N. & RIENTJES, T. 2005. Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation. *Hydrology and Earth System Sciences Discussions,* 2**,** 365-415.
RIO, S., PENAS, Á. & FRAILE, R. 2005. Analysis of recent climatic variations in Castile and Leon (Spain). *Atmospheric Research,* 73**,** 69-85.
DROSDOWSKY, W. 1993. Potential predictability of winter rainfall over southern and eastern Australia using Indian Ocean sea-surface temperature anomalies. *Aust. Met. Mag,* 42**,** 6.
DROSDOWSKY, W. & CHMABERS, L. E. 2001. Near-global sea surface temperature anomalies as predictors of Australian seasonal rainfall. *Journal of Climate,* 14**,** 1677-1687.
EL-SHAFIE, A., JAAFER, O. & SEYED, A. 2011. Adaptive neuro-fuzzy inference system based model for rainfall forecasting in Klang River, Malaysia. *International Journal of Physical Sciences,* 6**,** 2875-2888.

ENGLAND, M. H., UMMENHOFER, C. C. & SANTOSO, A. 2006. Interannual rainfall extremes over southwest Western Australia linked to Indian Ocean climate variability. *Journal of Climate,* 19**,** 1948-1969.
EVANS, A. D., BENNETT, J. M. & EWENZ, C. M. 2009. South Australian rainfall variability and climate extremes. *Climate Dynamics,* 33**,** 477-493.
FENG, M. & MEYERS, G. 2003. Interannual variability in the tropical Indian Ocean: a two-year time-scale of Indian Ocean Dipole. *Deep Sea Research Part II: Topical Studies in Oceanography,* 50**,** 2263-2284.
FIELD, A. P. 2009. *Discovering statistics using SPSS*, SAGE publications Ltd.
FIRAT, M. & GÜNGÖR, M. 2007. River flow estimation using adaptive neuro fuzzy inference system. *Mathematics and Computers in Simulation,* 75**,** 87-96.
FOLLAND, C., RENWICK, J., SALINGER, M. & MULLAN, A. 2002. Relative influences of the interdecadal Pacific oscillation and ENSO on the South Pacific convergence zone. *Geophysical Research Letters,* 29**,** 21-1-21-4.
FRENCH, M. N., KRAJEWSKI, W. F. & CUYKENDALL, R. R. 1992. Rainfall forecasting in space and time using a neural network. *Journal of Hydrology,* 137**,** 1-31.
GALLANT, A., KIEM, A., VERDON-KIDD, D., STONE, R., KAROLY, D. & DE MICHELE, C. 2012. Understanding hydroclimate processes in the Murray-Darling Basin for natural resources management. *Hydrology & Earth System Sciences,* 16.
GILLETT, N. P., KELL, T. D. & JONES, P. D. 2006. Regional climate impacts of the Southern Annular Mode. *Geophysical Research Letters,* 33.
GODDARD, L., MASON, S. J., ZEBIAK, S. E., ROPELEWSKI, C. F., BASHER, R. & CANE, M. A. 2001. Current approaches to seasonal to interannual climate predictions. *International Journal of Climatology,* 21**,** 1111-1152.
GONG, D. & WANG, S. 1999. Definition of Antarctic oscillation index. *Geophysical Research Letters,* 26**,** 459-462.
GRIMM, A. M. 2011. Interannual climate variability in South America: impacts on seasonal precipitation, extreme events, and possible effects of climate change. *Stochastic Environmental Research and Risk Assessment,* 25**,** 537-554.
GUO, Y., WANG, J. & ZHAO, Y. 2004. Numerical simulation of the 1999 Yangtze River Valley heavy rainfall including sensitivity experiments with different SSTA. *Advances in Atmospheric Sciences,* 21**,** 23-33.
HALIDE, H. & RIDD, P. 2008. Complicated ENSO models do not significantly outperform very simple ENSO models. *International Journal of Climatology,* 28**,** 219-233.
HARTMANN, H., BECKER, S. & KING, L. 2008. Predicting summer rainfall in the Yangtze River basin with neural networks. *International Journal of Climatology,* 28**,** 925-936.
HE, X., GUAN, H., ZHANG, X. & SIMMONS, C. T. 2014. A wavelet-based multiple linear regression model for forecasting monthly rainfall. *International Journal of Climatology,* 34**,** 1898-1912.
HENDON, H. H., THOMPSON, D. W. J. & WHEELER, M. C. 2007. Australian rainfall and surface temperature variations associated with the Southern Hemisphere annular mode. *Journal of Climate,* 20**,** 2452-2467.

HOERLING, M. P. & KUMAR, A. 2000. Understanding and predicting extratropical teleconnections related to ENSO, Cambridge University Press.
HORNIK, K. 1991. Approximation capabilities of multilayer feedforward networks. *Neural Networks,* 4**,** 251-257.
HSU, K., GUPTA, H. V. & SOROOSHIAN, S. 1995. Artificial neural network modelling of the rainfall-runoff process. *Water Resources Research,* 31**,** 2517-2530.
IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp. Accessed on 5 June, 2011
IHARA, C., KUSHNIR, Y., CANE, M. A. & DE LA PEÑA, V. H. 2007. Indian summer monsoon rainfall and its link with ENSO and Indian Ocean climate indices. *International Journal of Climatology,* 27**,** 179-187.
IYENGAR, R. & KANTH, S. R. 2005. Intrinsic mode functions and a strategy for forecasting Indian monsoon rainfall. *Meteorology and Atmospheric Physics,* 90**,** 17-36.
JANG, J.-S. 1993. ANFIS: adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on,* 23**,** 665-685.
KIEM, A. S., FRANKS, S. W. & KUCZERA, G. 2003. Multi-decadal variability of flood risk. *Geophysical Research Letters,* 30.
KIEM, A. S. & VERDON-KIDD, D. C. 2009. Climatic drivers of victorian streamflow: Is ENSO the dominant influence. *Australian Journal of Water Resources,* 13**,** 17-29.
KIM, T.-W., YOO, C. & AHN, J.-H. 2008. Influence of climate variation on seasonal precipitation in the Colorado River Basin. *Stochastic Environmental Research and Risk Assessment,* 22**,** 411-420.
KIRONO, D. G. C., CHIEW, F. H. S. & KENT, D. M. 2010. Identification of best predictors for forecasting seasonal rainfall and runoff in Australia. *Hydrological Processes,* 24**,** 1237-1247.
KIŞI, Ö. 2007. Streamflow forecasting using different artificial neural network algorithms. *Journal of Hydrologic Engineering,* 12**,** 532-539.
LATT, Z. Z. & WITTENBERG, H. 2014. Improving Flood Forecasting in a Developing Country: A Comparative Study of Stepwise Multiple Linear Regression and Artificial Neural Network. *Water Resources Management,* 28**,** 2109-2128.
LAU, K. & WENG, H. 2001. Coherent modes of global SST and summer rainfall over China: An assessment of the regional impacts of the 1997-98 El Nino. *Journal of Climate,* 14**,** 1294-1308.
LEE, S., CHO, S. & WONG, P. M. 1998. Rainfall prediction using artificial neural networks. *Journal of Geographic Information and Decision Analysis,* 2**,** 233-242.
LIM, E. P., HENDON, H. H., ANDERSON, D. L. T., CHARLES, A. & ALVES, O. 2011. Dynamical, statistical-dynamical, and multimodel ensemble forecasts of Australian spring season rainfall. *Monthly Weather Review,* 139**,** 958-975.
LIN, F. J. 2008. Solving multicollinearity in the process of fitting regression model using the nested estimate procedure. *Quality & Quantity,* 42**,** 417-426.
LIN, G.-F. & WU, M.-C. 2009. A hybrid neural network model for typhoon-rainfall forecasting. *Journal of Hydrology,* 375**,** 450-458.

LOUGH, J. M. 2007. Tropical river flow and rainfall reconstructions from coral luminescence: Great Barrier Reef, Australia. *Paleoceanography,* 22.
LUK, K., BALL, J. & SHARMA, A. 2000. A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting. *Journal of Hydrology,* 227**,** 56-65.
LUK, K. C., BALL, J. & SHARMA, A. 2001. An application of artificial neural networks for rainfall forecasting. *Mathematical and Computer modelling,* 33**,** 683-693.
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 and software,* 15**,** 101-124.
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 and Software,* 25**,** 891-909.
MAMDANI, E. H. & ASSILIAN, S. 1975. An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of man-machine studies,* 7**,** 1-13.
MANATSA, D., CHIPINDU, B. & BEHERA, S. K. 2012. Shifts in IOD and their impacts on association with East Africa rainfall. *Theoretical and Applied Climatology,* 110**,** 115-128.
MANTUA, N. J. & HARE, S. R. 2002. The Pacific decadal oscillation. *Journal of Oceanography,* 58**,** 35-44.
MANTUA, N. J., HARE, S. R., ZHANG, Y., WALLACE, J. M. & FRANCIS, R. C. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the American Meteorological Society,* 78**,** 1069-1079.
MARIOTTI, A., ZENG, N. & LAU, K. M. 2002. Euro-Mediterranean rainfall and ENSO—a seasonally varying relationship. *Geophysical Research Letters,* 29**,** 59-1-59-4.
MARSHALL, G. J. 2003. Trends in the Southern Annular Mode from observations and reanalyses. *Journal of Climate,* 16**,** 4134-4143.
MARSHALL, G. J., STOTT, P. A., TURNER, J., CONNOLLEY, W. M., KING, J. C. & LACHLAN-COPE, T. A. 2004. Causes of exceptional atmospheric circulation changes in the Southern Hemisphere. *Geophysical Research Letters,* 31**,** L14205 1-4.
MASON, S. J. & GODDARD, L. 2001. Probabilistic precipitation anomalies associated with ENSO. *Bulletin of the American Meteorological Society,* 82**,** 619-638.
Masters, T. (1993). Practical neural network recipes in C++: Morgan Kaufmann
Mathworks.com, (2012). MathWorks - MATLAB and Simulink for Technical Computing. [online] Available at: http://www.mathworks.com [Accessed 15 Nov 2012].
MCBRIDE, J. L. & NICHOLLS, N. 1983. Seasonal relationships between Australian rainfall and the Southern Oscillation. *Monthly Weather Review,* 111**,** 1998-2004.
MEKANIK, F. Lee, T.S. Imteaz, M. 2011. *Rainfall time series modelling for a mountainous region in west Iran.* Master Thesis, University Putra Malaysia.

MEKANIK, F., IMTEAZ, M. A., GATO-TRINIDAD, S. & ELMAHDI, A. 2013. Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. *Journal of Hydrology,* 503**,** 11-21.
MENEGHINI, B., SIMMONDS, I. & SMITH, I. N. 2007. Association between Australian rainfall and the Southern Annular Mode. *International Journal of Climatology,* 27**,** 109-121.
MEYERS, G. 1996. Variation of Indonesian throughflow and the El Niño-Southern Oscillation. *Journal of Geophysical Research: Oceans (1978–2012),* 101**,** 12255-12263.
MEYERS, G., MCINTOSH, P., PIGOT, L. & POOK, M. 2007. The years of El Niño, La Niña and interactions with the tropical Indian Ocean. *Journal of Climate,* 20**,** 2872-2880.
MURPHY, B. F. & TIMBAL, B. 2008. A review of recent climate variability and climate change in southeastern Australia. *International Journal of Climatology,* 28**,** 859-879.
MURPHY, M. J., GEORGAKAKOS, K. P. & SHAMIR, E. 2014. Climatological analysis of December rainfall in the Panama Canal Watershed. *International Journal of Climatology,* 34**,** 403-415.
NAZEMOSADAT, M. J. & CORDERY, I. 1997. The influence of geopotential heights on New South Wales rainfall. Meteorology and Atmospheric Physics, 63, 179-193.
NAYAK, P., SUDHEER, K., RANGAN, D. & RAMASASTRI, K. 2004. A neuro-fuzzy computing technique for modelling hydrological time series. *Journal of Hydrology,* 291**,** 52-66.
NAYAK, P., SUDHEER, K., RANGAN, D. & RAMASASTRI, K. 2005. Short-term flood forecasting with a neurofuzzy model. *Water Resources Research,* 41.
NICHOLLS, N. 1989. Sea surface temperatures and Australian winter rainfall. *Journal of Climate,* 2**,** 965-973.
NICHOLLS, N., DROSDOWSKY, W. & LAVERY, B. 1997. Australian rainfall variability and change. *Weather,* 52**,** 66-72.
NIU, J. 2013. Precipitation in the Pearl River basin, South China: scaling, regional patterns, and influence of large-scale climate anomalies. *Stochastic Environmental Research and Risk Assessment,* 27**,** 1253-1268.
NOURANI, V., ALAMI, M. T. & AMINFAR, M. H. 2009. A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Engineering Applications of Artificial Intelligence,* 22**,** 466-472.
POTGIETER, A. B., HAMMER, G. L., MEINKE, H., STONE, R. C. & GODDARD, L. 2005. Three putative types of El Nino revealed by spatial variability in impact on Australian wheat yield. *Journal of Climate,* 18**,** 1566-1574.
POWER, S., CASEY, T., FOLLAND, C., COLMAN, A. & MEHTA, V. 1999. Inter-decadal modulation of the impact of ENSO on Australia. *Climate Dynamics,* 15**,** 319-324.
POWER, S. B., HAYLOCK, M., COLMAN, R. & WANG, X. 2006. The predictability of interdecadal changes in ENSO activity and ENSO teleconnections. *Journal of Climate,* 19**,** 4755-4771.

QUAN, X., HOERLING, M., WHITAKER, J., BATES, G. & XU, T. 2006. Diagnosing sources of US seasonal forecast skill. *Journal of Climate,* 19**,** 3279-3293.
RAJEEVAN, M., PAI, D., KUMAR, R. A. & LAL, B. 2007. New statistical models for long-range forecasting of southwest monsoon rainfall over India. *Climate Dynamics,* 28**,** 813-828.
RAMÍREZ, M. C. P., VELHO, H. F. C. & FERREIRA, N. J. 2005. Artificial neural network technique for rainfall forecasting applied to the Sao Paulo region. *Journal of Hydrology,* 301**,** 146-162.
RAO, S. A., BEHERA, S. K., MASUMOTO, Y. & YAMAGATA, T. 2002. Interannual subsurface variability in the tropical Indian Ocean with a special emphasis on the Indian Ocean dipole. *Deep Sea Research Part II: Topical Studies in Oceanography,* 49**,** 1549-1572.
RIAD, S., MANIA, J., BOUCHAOU, L. & NAJJAR, Y. 2004. Predicting catchment flow in a semi-arid region via an artificial neural network technique. *Hydrological Processes,* 18**,** 2387-2393.
RISBEY, J. S., POOK, M. J., MCINTOSH, P. C., WHEELER, M. C. & HENDON, H. H. 2009. On the remote drivers of rainfall variability in Australia. *Monthly Weather Review,* 137**,** 3233-3253.
ROGERS, L. L. & DOWLA, F. U. 1994. Optimization of groundwater remediation using artificial neural networks with parallel solute transport modelling. *Water Resources Research,* 30**,** 457-481.
ROSSEL, F. & CADIER, E. 2009. El Niño and prediction of anomalous monthly rainfalls in Ecuador. *Hydrological Processes,* 23**,** 3253-3260.
SADHURAM, Y. & MURTHY, T. R. 2008. Simple multiple regression model for long range forecasting of Indian summer monsoon rainfall. *Meteorology and Atmospheric Physics,* 99**,** 17-24.
SAJI, N. H., GOSWAMI, B. N., VINAYACHANDRAN, P. N. & YAMAGATA, T. 1999. A dipole mode in the tropical Indian ocean. *Nature,* 401**,** 360-363.
SARKAR, S., SINGH, R. P. & KAFATOS, M. 2004. Further evidences for the weakening relationship of Indian rainfall and ENSO over India. *Geophysical research letters,* 31.
SARLE, W. S. 1995. Stopped training and other remedies for overfitting. *In:* Proceedings of the 27th Symposium on the Interface of Computing Science and Statisfi (.'. _\'. pp. 352-360. Interface Foundation of North America, Fairfax Station. VA, USA.
SCHEPEN, A., WANG, Q. J. & ROBERTSON, D. 2012. Evidence for using lagged climate indices to forecast Australian seasonal rainfall. *Journal of Climate,* 25**,** 1230-1246.
SHUKLA, R. P., TRIPATHI, K. C., PANDEY, A. C. & DAS, I. 2011. Prediction of Indian summer monsoon rainfall using Niño indices: a neural network approach. *Atmospheric Research,* 102**,** 99-109.
SIMMONDS, I. & HOPE, P. 1997. Persistence characteristics of Australian rainfall anomalies. *International Journal of Climatology,* 17**,** 597-613.
STONE, R. C., HAMMER, G. L. & MARCUSSEN, T. 1996. Prediction of global rainfall probabilities using phases of the Southern Oscillation Index.

SUDHEER, K., GOSAIN, A. & RAMASASTRI, K. 2002. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrological Processes,* 16**,** 1325-1330.
TAKAGI, T. & SUGENO, M. 1985. Fuzzy identification of systems and its applications to modelling and control. *Systems, Man and Cybernetics, IEEE Transactions on***,** 116-132.
TALEI, A. & CHUA, L. H. 2012. Influence of lag time on event-based rainfall–runoff modelling using the data driven approach. *Journal of Hydrology,* 438**,** 223-233.
TALEI, A., CHUA, L. H. C. & QUEK, C. 2010. A novel application of a neuro-fuzzy computational technique in event-based rainfall–runoff modelling. *Expert Systems with Applications,* 37**,** 7456-7468.
TALEI, A., CHUA, L. H. C., QUEK, C. & JANSSON, P.-E. 2013. Runoff forecasting using a Takagi–Sugeno neuro-fuzzy model with online learning. *Journal of Hydrology,* 488**,** 17-32.
TALEI, A. 2013. Rainfall-Runoff Modelling With Neuro-Fuzzy Systems, PhD Thesis, Nanyang Technological University of Singapore.
THOMPSON, D. W. J. & SOLOMON, S. 2002. Interpretation of recent Southern Hemisphere climate change. *Science,* 296**,** 895-899.
THOMPSON, D. W. J. & WALLACE, J. M. 2000. Annular modes in the extratropical circulation. Part I: Month-to-month variability. *Journal of Climate,* 13**,** 1000-1016.
THOMPSON, D., Atmos.colostate.edu, (2014). *Annular Modes - Introduction*. [online] Available at: http://www.atmos.colostate.edu/~davet/ao/introduction.html [Accessed 15 May 2014].
TOKAR, A. S. & JOHNSON, P. A. 1999. Rainfall-runoff modelling using artificial neural networks. *Journal of Hydrologic Engineering,* 4**,** 232-239.
TOTH, E., BRATH, A. & MONTANARI, A. 2000. Comparison of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology,* 239**,** 132-147.
TURAN, M. E. & YURDUSEV, M. A. 2009. River flow estimation from upstream flow records by artificial intelligence methods. *Journal of Hydrology,* 369**,** 71-77.
UM, M. J., YUN, H., JEONG, C. S. & HEO, J. H. 2011. Factor analysis and multiple regression between topography and precipitation on Jeju Island, Korea. *Journal of Hydrology,* 410**,** 189-203.
UMMENHOFER, C. C., ENGLAND, M. H., MCINTOSH, P. C., MEYERS, G. A., POOK, M. J., RISBEY, J. S., GUPTA, A. S. & TASCHETTO, A. S. 2009. What causes southeast Australia's worst droughts? *Geophysical Research Letters,* 36.
US Department of Commerce, N. (2014). NOAA - National Oceanic and Atmospheric Administration. [online] Noaa.gov. Available at: http://www.noaa.gov/ [Accessed 10 Nov 2014].
VENTURA, F., ROSSI PISA, P. & ARDIZZONI, E. 2002. Temperature and precipitation trends in Bologna (Italy) from 1952 to 1999. *Atmospheric Research,* 61**,** 203-214.

VERDON-KIDD, D. C. & KIEM, A. S. 2009a. Nature and causes of protracted droughts in southeast Australia: Comparison between the Federation, WWII, and Big Dry droughts. *Geophysical Research Letters,* 36.

VERDON-KIDD, D. C. & KIEM, A. S. 2009b. On the relationship between large-scale climate modes and regional synoptic patterns that drive Victorian rainfall. *Hydrology and Earth System Sciences,* 13**,** 467-479.

VERDON, D. C. & FRANKS, S. W. 2005. Indian Ocean sea surface temperature variability and winter rainfall: Eastern Australia. *Water Resources Research,* 41**,** 1-10.

VERDON, D. C., WYATT, A. M., KIEM, A. S. & FRANKS, S. W. 2004. Multidecadal variability of rainfall and streamflow: Eastern Australia. *Water Resources Research,* 40**,** W10201.

WANG, G. & HENDON, H. H. 2007. Sensitivity of Australian Rainfall to Inter–El Niño Variations. *Journal of Climate,* 20.

WANG, Q., SCHEPEN, A. & ROBERTSON, D. E. 2012. Merging Seasonal Rainfall Forecasts from Multiple Statistical Models through Bayesian Model Averaging. *Journal of Climate,* 25.

WATKINS, A. 2002. Australian climate summary: dry and warm conditions dominate. *Bull. Aust. Met. Ocean. Soc,* 15**,** 109-14.

WIJFFELS, S. & MEYERS, G. 2004. An intersection of oceanic waveguides: Variability in the Indonesian Throughflow region. *Journal of Physical Oceanography,* 34**,** 1232-1253.

YAMAGATA, T., BEHERA, S. K., LUO, J. J., MASSON, S., JURY, M. R. & RAO, S. A. 2004. Coupled ocean-atmosphere variability in the tropical Indian Ocean. *Earth's Climate***,** 189-211.

YEON, I., KIM, J. & JUN, K. 2008. Application of artificial intelligence models in water quality forecasting. *Environmental technology,* 29**,** 625-631.

YILMAZ, A. G., IMTEAZ, M. A. & JENKINS, G. 2011. Catchment flow estimation using Artificial Neural Networks in the mountainous Euphrates Basin. *Journal of Hydrology,* 410**,** 134-140.

YUFU, G., YAN, Z. & JIA, W. 2002. Numerical simulation of the relationships between the 1998 Yangtze River valley floods and SST anomalies. *Advances in Atmospheric Sciences,* 19**,** 391-404.

ZHANG, Y., WALLACE, J. M. & BATTISTI, D. S. 1997. ENSO-like interdecadal variability: 1900-93. *Journal of Climate,* 10**,** 1004-1020.