# A Multivariate Artificial Neural Network Approach for Rainfall Forecasting: Case Study of Victoria, Australia

F. Mekanik and M. A. Imteaz

*Abstract*— **El Nino southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) have enormous effects on the precipitations around the world. Australian rainfall is also affected by these key modes of complex climate variables. Many studies have tried to establish the relationships of these large-scale climate indices among the rainfalls of different parts of Australia, particularly Western Australia, New South Wales, Queensland and Victoria. Unlike the other regions, no clear relationship can be found between each individual large-scale climate mode and Victorian rainfall. Past studies considering southeast Australian rainfall predictability could achieve a maximum of 30% correlation. This study looks into the lagged-time relationships of these modes on Victorian spring rainfall. On the other hand, few attempts have been made to establish the combined effect of these indices on rainfall in order to develop a better understanding and forecasting system. Since rainfall is a complicated atmospheric phenomenon, linear techniques might not be sufficient enough to capture its characteristics. This research attempts to find a nonlinear relationship between the Victorian rainfall and the lagged-indices affecting the region using Artificial Neural Networks (ANN). It was discovered that ANN modelling is able to provide higher correlations using the lagged-indices to forecast spring rainfall in compared to linear methods. Using these indices in an ANN model increased the model correlation up to 99%, 98% and 43% for the three case study stations of Horsham, Melbourne and Orbost in Victoria, Australia respectively. it seems that IOD has a higher effect on the centre and west of Victoria more than the ENSO, while ENSO seems to have a stronger effect on the east side. This method can be used for other parts of the world where a relationship exists between rainfall and large scale climate modes which could not be established by linear methods.**

*Index Terms*—**Rainfall, ENSO, IOD, forecasting, ANN**

## I. Introduction

Forecasting rainfall several months or seasons in advance can be beneficial for the management of water resources. Many studies have tried to find the relationships between large-scale climate modes and rainfall in different parts around the world using different linear and nonlinear methods [1-6].

Manuscript received July 23, 2012; revised August 9, 2012.

F. Mekanik is a PhD student at Swinburne University of Technology (phone: +61-449634593; e-mail: fmekanik @ swin.edu.au).

M.A. Imteaz is a senior lecturer at Swinburne University of Technology (email: mimteaz@swin.edu.au).

It is believed that Australian rainfall is affected by several major climate patterns. The major drivers bringing rainfall over Australia which have been investigated by many researchers are, El Nino Southern Oscillation (ENSO), Indian Ocean Dipole (IOD) and Southern Annular Mode (SAM) [7]. ENSO is represented by two different type of indices; the Southern Oscillation Index (SOI) which is a measure of Sea Level Pressure (SLP) anomalies between Darwin and Tahiti and Sea surface temperature (SST) anomalies in equatorial Pacific Ocean noted as NINO3 (5 $^{o}$S – 5$^{o}$N, 150$^{o}$– 90$^{o}$W), NINO3.4 (5$^{o}$ S – 5$^{o}$N, 170$^{o}$ – 120$^{o}$W) and NINO4 (5$^{o}$S – 5$^{o}$N, 160$^{o}$ – 150$^{o}$W) [8]. The IOD is also a coupled ocean-atmosphere phenomenon in the equatorial Indian Ocean [9]. SAM is the major mode of atmospheric variability on the mid and high latitude of southern hemisphere. SAM is not considered in this study due to its shorter range of data. Many researchers have conducted different studies in different parts of Australia trying to establish the relationship between these climate modes and Australian rainfall [7, 8, 10-16]. Victoria is one of the regions that so far did not show good correlation of its rainfall and the climate modes. According to [17] in comparison to eastern Australia and particularly Queensland, past studies considering southeast Australian rainfall predictability could achieve a maximum of 30% correlation. Other than the work of [18] which analysed the combined impact of ENSO and SAM on Victorian rainfall, other studies focused only on finding the relationship between rainfall and a single driver. In the work of [16] Murphy the maximum correlation of 37% was achieved for spring rainfall with spring NINO4. On the other hand, the studies did not take into account the relationship between previous time lags of these drivers as a potential predictor for future rainfall. Kirono et al, [7] is the only accessible publication which considered the relationship between Australian rainfall and two months average lag of different climate indices.

According to [18] Victorian rainfall variability is not driven by a single climate mode. Further to investigating the effect of different lagged-time climate indices on Victorian spring rainfall, this paper also contributes to finding the lag relationship of combined climate indices and Victorian rainfall in order to have a better understanding of the variability of Victorian rainfall in regards to large scale climate modes. Thus, this study is distinguished from previous studies by

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

forecasting spring rainfall three consecutive years in advance by using the lag relationship of separate and combined climate indices. The combined ENSO-IOD sets include lagged NINO3-DMI, NINO4-DMI, NINO3.4-DMI and SOI-DMI, since there is no agreement on which of the ENSO indices can better represent this ocean-atmospheric phenomenon.

## II. DATA

### *A. Rainfall Data*

Historical monthly rainfall data was obtained from the Australian Bureau of Meteorology for Horsham, Melbourne and Orbost, Victoria, Australia as a case study. Fig.1 shows the location details of the station considered in this study. Spring (September - November) rainfall was obtained from monthly rainfall data from January 1900 to December 2009 (www.bom.gov.au/climate/data/) .

### *B. Climate Indices*

In this study, monthly values of NINO3, NINO4, and NINO3.4 were used as representation of ENSO. In addition to this SST related indices, Southern Oscillation Index (SOI) which is the SLP representation of ENSO was also considered in this study. A measure of IOD is the Dipole Mode Index (DMI) which is the difference in average SST anomalies between the tropical Western Indian Ocean ($10^{O}$S - $10^{O}$N, $50^{O}$ - $70^{O}$E) and the tropical Eastern Indian ocean ($10^{O}$S - Equator, $90^{O}$ - $110^{O}$E)[7]. ENSO and IOD indices were obtained from Climate Explorer website http://climexp.knmi.nl/ .

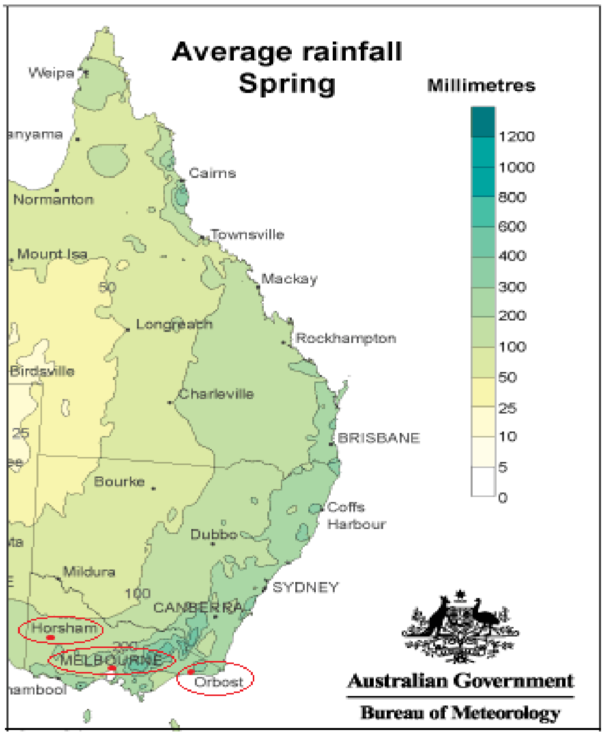


Fig.1. Map of the study area
(Adopted from www.bom.gov.au/jsp/ncc/climate_averages/rainfall)

## III. ARTIFICIAL NEURAL NETWORK.

The parameters for ANN modelling are basically network topology, neurons characteristics, training and learning rules. Multi-Layered Perceptrons (MLP) are feed-forward nets with one or more hidden layers between the input and output neurons (Fig. 2). The number of input and output neurons is based on the number of input and output data. Basically, the input layer only serves as receiving the input data for further processing in the network. The hidden layers are a very important part in a 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. The output neuron is the desired output of the model. The process of developing an ANN model is to find a) suitable input data set, b) determine the number of hidden layers and neurons, and c) training and testing the network. Mathematically, the network depicted in Fig. 2 can be expressed as follow:

$$Y_t = f_2[\sum_{j=1}^{l} w_j f_1(\sum_{i=1}^{l} w_i x_i)] \quad (1)$$

where $Y_t$ is the output of the network, $x_i$ is the input to the network, $w_i$ and $w_j$ are the weights between neurons of the input and hidden layer and between hidden layer and output respectively; $f_1$ and $f_2$ are the activation functions for the hidden layer and output layer respectively. According to [19] if extrapolating beyond the range of the training data is needed it is recommended to use sigmoiadal-type transfer functions in the hidden layers and linear transfer functions in the output layer. In this study $f_1$ is considered tansigmoid function which is a nonlinear function and $f_2$ is considered the linear purelin function defined as follow:

$$f_1 = \frac{2}{(1+\exp(-2x))} - 1 \quad (2)$$

$$f_2(x) = x \quad (3)$$

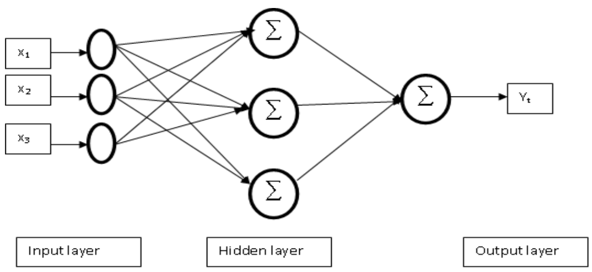


Fig. 2. A typical ANN architecture

Early stop technique was used to stop the network from over fitting. Number of hidden neurons was chosen based on trial and error considering 5,10,15,20,25,30,35 and 40 hidden nodes. In this study, 1900-2006 was selected as the training and validation period and 2007-2009 was used as test period.

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

The data were normalized between the range of 1 and 0 with Eq. 4. The models were evaluated using mean square error (MSE), Pearson correlation (R) and index of agreement *(d).*

$$\bar{x}_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (4)$$

## IV. RESULTS AND DISCUSSIONS

ANN modeling was carried out for separate and combined climate indices considering 5, 10, 15, 20, 25, 30 and 35 hidden neurons for three stations in Victoria. The best model of each input set was chosen based on the least MSE of training, validation and testing set; the best hidden neuron varied for different models and a unity was not observed among the models. Tables I to III show the MSE of the models for the three stations. The models are named after their input set. The best models are bolded in each table. It can be seen from Table I that for Horsham the best model is the model with lagged DMI as the input with a MSE of validation-test of 0.019, the next best model is NINO3-DMI with a MSE of 0.053. Lagged DMI seems to be the best predictor of Melbourne spring rainfall as well with a MSE 0.037, followed by NINO3-DMI with a MSE of 0.052 (Table II).

However, for Orbost which is located at far east lagged NINO4 is the better predictor (Table III). The MSE for Orbost are higher compared to Horsham and Melbourne; the results show that IOD has a higher effect on the centre and west of Victoria more than ENSO, while ENSO seems to have a stronger effect on the east side. This needs to be further investigated by considering more stations in the study.

TABLE I. MSE of the best models for Horsham

| Model | Train | Valid | Test | Valid-Test |
|---|---|---|---|---|
| NINO3 | 0.020 | 0.109 | 0.008 | 0.117 |
| NINO4 | 0.009 | 0.043 | 0.062 | 0.105 |
| NINO3.4 | 0.026 | 0.043 | 0.072 | 0.115 |
| SOI | 0.019 | 0.055 | 0.033 | 0.088 |
| **DMI** | **0.024** | **0.015** | **0.004** | **0.019** |
| NINO3-DMI | 0.047 | 0.037 | 0.018 | 0.055 |
| NINO4-DMI | 0.017 | 0.046 | 0.009 | 0.055 |
| NINO3.4-DMI | 0.00 | 0.032 | 0.021 | 0.053 |
| SOI-DMI | 0.032 | 0.030 | 0.030 | 0.060 |

TABLE II. MSE of the best models for Melbourne

| Model | Train | Valid | Test | Valid-Test |
|---|---|---|---|---|
| NINO3 | 0.027 | 0.030 | 0.050 | 0.080 |
| NINO4 | 0.025 | 0.032 | 0.042 | 0.074 |
| NINO3.4 | 0.074 | 0.027 | 0.029 | 0.056 |
| SOI | 0.006 | 0.064 | 0.006 | 0.070 |
| **DMI** | **0.022** | **0.032** | **0.005** | **0.037** |
| NINO3-DMI | 0.016 | 0.041 | 0.011 | 0.052 |
| NINO4-DMI | 0.053 | 0.046 | 0.009 | 0.055 |
| NINO3.4-DMI | 0.020 | 0.032 | 0.021 | 0.053 |
| SOI-DMI | 0.058 | 0.054 | 0.003 | 0.057 |

Combining ENSO and OID did not improve the model performances. One explanation is that ENSO and IOD have different effective periods on Victorian rainfall; using the same lagged months for both indices in order to predict rainfall might not capture their nature efficiently. Thus, further study is underway to investigate the use of different lags of ENSO and IOD for rainfall prediction. Pearson correlation of these models with the rainfall of the three stations is shown in Table IV. By using lagged predictors as inputs in the ANN modeling, the authors were able to predict spring rainfall 3 consecutive years in advance for Horsham, Melbourne and Orbost with correlations of 99%, 98% and 43% respectively.

TABLE III. MSE of the best models for Orbost

| Model | Train | Validation | Test | Valid-Test |
|---|---|---|---|---|
| NINO3 | 0.023 | 0.044 | 0.057 | 0.101 |
| **NINO4** | **0.015** | **0.055** | **0.019** | **0.074** |
| NINO3.4 | 0.057 | 0.049 | 0.026 | 0.075 |
| SOI | 0.011 | 0.078 | 0.036 | 0.114 |
| DMI | 0.027 | 0.064 | 0.190 | 0.254 |
| NINO3-DMI | 0.083 | 0.110 | 0.018 | 0.128 |
| NINO4-DMI | 0.00 | 0.103 | 0.019 | 0.122 |
| NINO3.4-DMI | 0.054 | 0.062 | 0.030 | 0.092 |
| SOI-DMI | 0.018 | 0.061 | 0.024 | 0.085 |

Fig. 3shows the best models for the three stations. It can be seen from Fig. 3 spring rainfall of the region follows a very noisy pattern. Using the lagged climate indices as predictors ANN was able to model the series in a way that not only follows the pattern but also is able to predict long-term future; ANN is smoothly fitting the series capturing all the peaks and minimum, however, it can be seen that validation and test sets are not well-modeled as the training set. This can be overcome by doing a K-fold cross-validation which is going to be done in the future work.

TABLE IV.Pearson correlation of the best models

| Model | Train | Validation | Test |
|---|---|---|---|
| Horsham | 66 | 74 | 99 |
| Melbourne | 56 | 34 | 98 |
| Orbost | 84 | 38 | 43 |

For better assessment of the model performance, an additional criterion, the Index of agreement (*d*) [20] has been chosen for model comparison. A *'d'* value close to 1 indicates a better fitted model. Table V shows '*d*' values for the three stations. Horsham is having a high '*d*' value both in validation and testing set while for Melbourne only the test has an acceptable '*d*'. The '*d*' value for Orbost is showing an average mode regarding predictibility.

TABLE V. 'd' values for the best models

| Model | Train | Validation | Test |
|---|---|---|---|
| Horsham | 0.75 | 0.81 | 0.94 |
| Melbourne | 0.62 | 0.37 | 0.92 |
| Orbost | 0.90 | 0.63 | 0.52 |

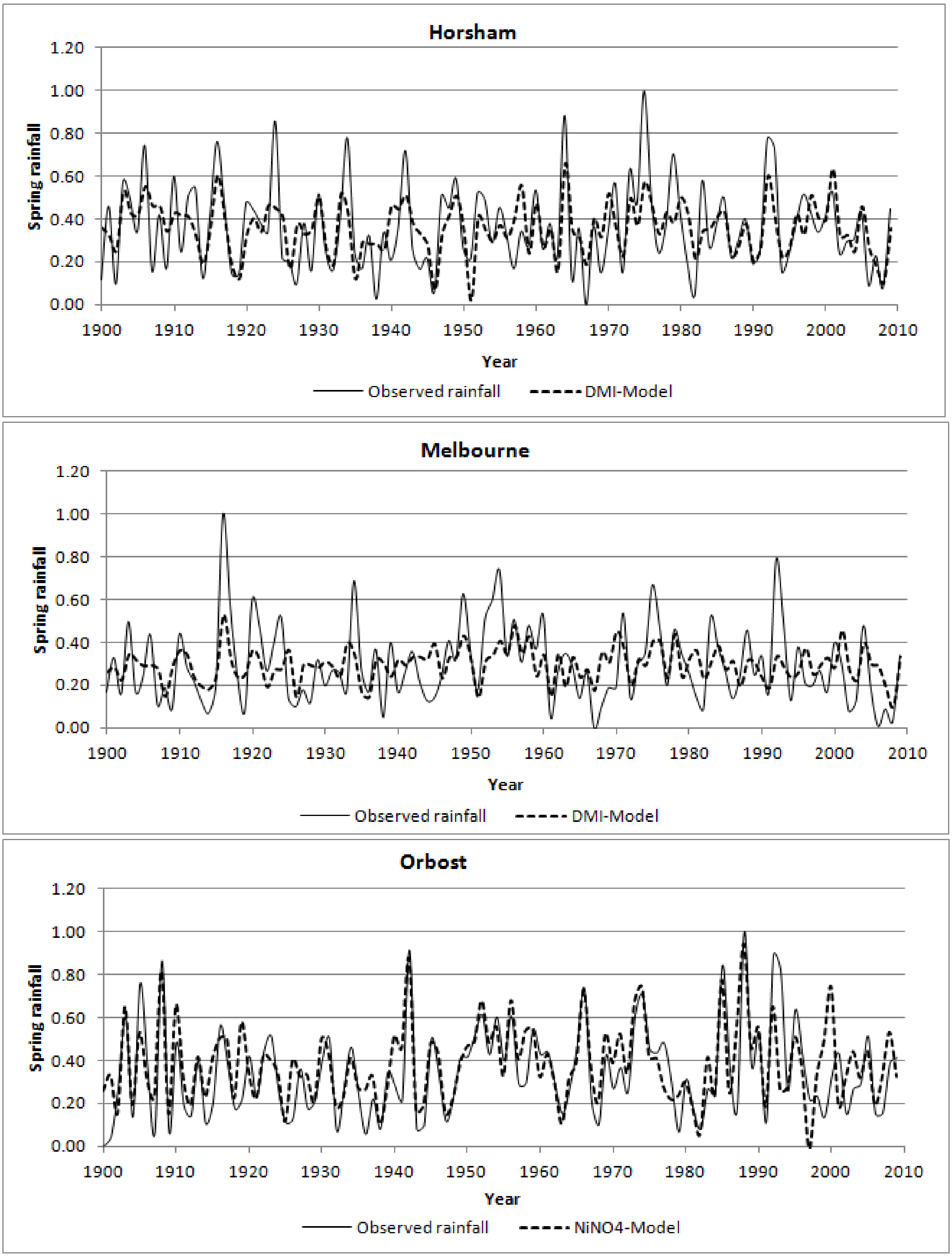


Fig. 3. Comparison of ANN model for spring rainfall for the three stations, (1900-1990=training period, 1991-2006= validation period, 2007-2009=testing period for ANN).

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

## V. CONCLUSIONS

This study attempted to predict spring rainfall three consecutive years in advance by considering single and combined lagged climate indices as potential predictors. A nonlinear Artificial Neural Networks method was performed in order to investigate the predictability of spring rainfall using lagged ENSO, IOD and ENSO-IOD representatives. NINO3, NINO4, NINO3.4 and SOI were chosen as ENSO representatives and DMI was chosen as IOD representative, the previous studies were focusing on finding the effect of these indices separately on Victorian rainfall but could not achieve a correlation of more than 30%. This study discovered that the single lagged climate indices have more effect on rainfall predictability than the combined lagged climate indices. Using these indices in an ANN model increased the model correlation up to 99%, 98% and 43% for the three case study stations of Horsham, Melbourne and Orbost respectively. It seems that IOD has a higher effect on the centre and west of Victoria more than the ENSO, while ENSO seems to have a stronger effect on the east side.

There is a need to further investigate this method on other different rainfall stations which will be covered in future studies. Also the effect of each lag on the predictability of rainfall needs further attention in a sensitivity analysis procedure.

## REFERENCES

[1] K. Lau and H. Weng, "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,* vol. 14, pp. 1294-1308, 2001.

[2] G. Yufu, Z. Yan, and W. Jia, "Numerical simulation of the relationships between the 1998 Yangtze River valley floods and SST anomalies," *Advances in Atmospheric Sciences,* vol. 19, pp. 391-404, 2002.

[3] J. J. Barsugli and P. D. Sardeshmukh, "Global atmospheric sensitivity to tropical SST anomalies throughout the Indo-Pacific basin," *Journal of Climate,* vol. 15, pp. 3427-3442, 2002.

[4] H. Hartmann, S. Becker, and L. King, "Predicting summer rainfall in the Yangtze River basin with neural networks," *International Journal of Climatology,* vol. 28, pp. 925-936, 2008.

[5] G. Chattopadhyay, S. Chattopadhyay, and R. Jain, "Multivariate forecast of winter monsoon rainfall in India using SST anomaly as a predictor: Neurocomputing and statistical approaches," *Comptes Rendus Geoscience,* vol. 342, pp. 755-765, 2010.

[6] R. P. Shukla, K. C. Tripathi, A. C. Pandey, and I. M. L. Das, "Prediction of Indian summer monsoon rainfall using Niño indices: A neural network approach," *Atmospheric Research,* vol. 102, pp. 99-109, 2011.

[7] D. G. C. Kirono, F. H. S. Chiew, and D. M. Kent, "Identification of best predictors for forecasting seasonal rainfall and runoff in Australia," *Hydrological Processes,* vol. 24, pp. 1237-1247, 2010.

[8] J. S. Risbey, M. J. Pook, P. C. McIntosh, M. C. Wheeler, and H. H. Hendon, "On the remote drivers of rainfall variability in Australia," *Monthly Weather Review,* vol. 137, pp. 3233-3253, 2009.

[9] N. H. Saji, B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, "A dipole mode in the tropical Indian ocean," *Nature,* vol. 401, pp. 360-363, 1999.

[10] B. Meneghini, I. Simmonds, and I. N. Smith, "Association between Australian rainfall and the Southern Annular Mode," *International Journal of Climatology,* vol. 27, pp. 109-121, 2007.

[11] C. C. Ummenhofer, A. Sen Gupta, M. J. Pook, and M. H. England, "Anomalous rainfall over southwest Western Australia forced by Indian Ocean sea surface temperatures," *Journal of Climate,* vol. 21, pp. 5113-5134, 2008.

[12] M. H. England, C. C. Ummenhofer, and A. Santoso, "Interannual rainfall extremes over southwest Western Australia linked to Indian Ocean climate variability," *Journal of Climate,* vol. 19, pp. 1948-1969, 2006.

[13] A. D. Evans, J. M. Bennett, and C. M. Ewenz, "South Australian rainfall variability and climate extremes," *Climate Dynamics,* vol. 33, pp. 477-493, 2009.

[14] N. Nicholls, "Local and remote causes of the southern Australian autumn-winter rainfall decline, 1958–2007," *Climate Dynamics,* vol. 34, pp. 835-845, 2010.

[15] D. C. Verdon, A. M. Wyatt, A. S. Kiem, and S. W. Franks, "Multidecadal variability of rainfall and streamflow: Eastern Australia," *Water Resources Research,* vol. 40, p. W10201, 2004.

[16] B. F. Murphy and B. Timbal, "A review of recent climate variability and climate change in southeastern Australia," *International Journal of Climatology,* vol. 28, pp. 859-879, 2008.

[17] D. C. Verdon-Kidd and A. S. Kiem, "On the relationship between large-scale climate modes and regional synoptic patterns that drive Victorian rainfall," *Hydrology and Earth System Sciences,* vol. 13, pp. 467-479, 2009.

[18] A. S. Kiem and D. C. Verdon-Kidd, "Climatic drivers of victorian streamflow: Is ENSO the dominant influence," *Australian Journal of Water Resources,* vol. 13, pp. 17-29, 2009.

[19] H. R. Maier and G. C. Dandy, "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications," *Environmental Modelling and software,* vol. 15, pp. 101-124, 2000.

[20] Willmott, C. J. (1982). Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society, 63,* 1309-1369

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)