Application of wavelet artificial neural networks in forecasting seasonal rainfall time series in Queensland, Australia

M Ghamariadyan\textsuperscript{1}, M A Imteaz\textsuperscript{2}

\textsuperscript{1,2}Department of Civil & Construction Engineering, Swinburne University of Technology, Hawthorn, VIC 3122, Australia.

\textsuperscript{2}Corresponding author email: mimteaz@swin.edu.au

Abstract. Predicting accurate future rainfall plays an essential role in managing the water resources as well as controlling the unforeseen phenomena like flooding or drought. In this study, an Artificial Intelligence (AI) method called wavelet artificial neural network is used to predict one-year ahead seasonal rainfall of Queensland, Australia. The technique used in this study is based on wavelet transform and artificial neural networks (ANN). Over the wavelet transforms analysis, the primary signal is split into some sub-time series called approximation and detail. The outcome of this process is some meticulous and valuable information that cannot be easily captured in the main time series. To assessment the skillfulness of the proposed method, five input sets combined of rainfall and three climate indices are defined. The comparisons showed that the wavelet-ANN (WANN) presented the lower root-mean-square-error (RMSE) with 64.8mm compared to ANN with 88.9mm, and it outperforms the ANN forecast with respect to statistical measures. Since fewer data in predicting seasonal rainfall is available compared to monthly rainfall data, the application of WANN provides promising insights into using this AI method for predicting seasonal rainfall one year in advance in Australia.

Keywords: Artificial Intelligence, WANN, rainfall

1. Introduction

In the past decades, artificial intelligence (AI) based methods were presented to overcome the intrinsic complications of the water resources and hydrological phenomena. Artificial Neural Networks (ANN) are widely used tools across the world in different fields, especially for modeling hydrological phenomena [1]. In recent years, to acquire better results in terms of accuracy and reliability, artificial neural networks have been combined with other methods. The wavelet transform is one of the techniques that can improve the accuracy of the prediction [2]. The wavelet-ANN helps researchers to investigate climatological events precisely and help them to have better insight into these phenomena as well.

Moreover, rainfall in Australia changes from high tropical rainfall to the lowest rainfall in the driest regions in the interior zones. Australian extreme rainfall in terms of frequency and intensity is greater than the rate of changes for average rainfall in the long-term [3]. It is believed that some parameters such as the Southern Annular Mode (SAM), El Nino-Southern Oscillation (ENSO), the subtropical...
ridge, the Pacific and the Indian Ocean sea surface temperature (SST), and the Madden-Julian Oscillation (MJO) affect the rainfall variation in Australia [4]. It is also noteworthy that investigating lagged relationships between rainfall and climate indices is very useful to have a better insight into future rainfall.

Many of the past conducted research has investigated the simultaneous connection between rainfall and different climate indices. However, there are fewer studies that investigated lagged relationships of the rainfall and climate indices [5], [6]. The various widely used forecasting empirical methods such as ANN, group method of data handling, fuzzy logic, and regression has been applied in many rainfall predictions [7], [8].

The application of the Wavelet transform (WF) was initially approved in the signal processing study. From that time on, many data scientists have employed this approach in their research due to the effectiveness of this method. The WF was presented rapidly as a substitution to the previous techniques like Fourier transform. This approach outperforms over traditional spectra analysis methods in non-stationary, multi-scaled, and localized phenomena. Therefore, it is highly suggested for analyzing irregularly distributed events [13]. In WF, the primary time series (signal) is split into some low-resolution sub-signals. In this process, many trivial signals are provided which hold hidden and meticulous information of the main signal. This trivial information cannot be detected in the main signal. Hence, this technique would be a useful approach, especially when it is coupled with an AI method such as ANN.

Many studies made use of wavelet analysis in different fields. Pnevmatikos and Hatziigeorgiou applied a discrete wavelet for detecting damage in a framed structure due to a strong earthquake [8]. They also proved that wavelet analysis could perform well in dealing with such problems. Mandrikova et al. (2016) successfully applied wavelet transform in astronomy by using the wavelet transform in the geomagnetic data processing [10]. The effectiveness of the approach was proved by experimental results. Kim et al. implemented a wavelet transform in fluid mechanics to ascertain the physical mechanism of the flow [11]. In hydrology, a wavelet-ANN technique for computing the river sedimentation load was developed by Kisi, 2010 [12]. For forecasting rainfall, Nourani et al., 2012 took advantage of wavelet analysis for forecasting rainfall [13].

This study aims to predict the seasonal rainfall for Pleystowe weather station in Queensland with a one-year lead time. The technique used in this study is a coupled wavelet ANN (WANN) method. The approach includes using the split sub-signals (called ‘approximation’ and ‘detail’) provided in the wavelet analysis process as the inputs of the ANN. The outcomes of the WANN are compared with the results of the ANN.

2. Methodology

2.1. Artificial Neural Networks

ANNs are effective tools that try to imitate human brain properties by learning and training rules. An artificial neural network is an information processing system comprises nodes connected by some weighted synaptic connections. The typical learning procedure for an ANN is a supervised learning method called back-propagation algorithm. This algorithm contains a feed-forward phase (Figure. 1). The complexity of the problem affects the hidden nodes’ size. It can be obtained through the trial and error process. The network is mathematically shown as follows [14]:

$$Z = f(\sum WX + \theta)$$  \hspace{1cm} (1)

where, $Z$ stands for output value of the hidden layer; $X$ shows hidden or input node value; $W$ and $\theta$ are weights and bias, respectively.

The training algorithm used in this study is the Levenberg–Marquardt algorithm (LMA). LMA is the most rapid and efficient training algorithm, among other training algorithms. It can be described as follows [14]:

$$X_{k+1} = X_k - [J^T J + \mu I]^{-1} J^T e$$  \hspace{1cm} (2)
where, $J$ represents the Jacobian matrix, $e$ called residual error, $X$ shows the weights of the neural network, and $\mu$ stands for the learning ratio, and it is used to control the process of learning.

In this study, to eliminate the overtraining, the early stop technique is employed. Put it differently; if the network is overly trained, it works well through the training data set. However, it provides poor outcomes in dealing with a fresh data set. Therefore, a verification data set is selected as a cross-validating the models.

![Figure 1](image-url)  
**Figure 1.** An artificial neural network structure.

### 2.2. Wavelet Transform Analysis

The advent of wavelet transforms in spectral analysis can be considered as a milestone due to its ability to provide a time-frequency description of a signal in the time domain. The concept of the wavelet transform is basically originated from the Fourier transform (FT). Through the FT analysis, the main signal is broken down into the sine waves. These sine waves are not spatially and temporally localized. However, in dealing with sudden changes in the signal, a new technique, which is localized in time and space, is required. This technique is called a wavelet transform. In wavelet analysis, the mother wavelet is applied to capture the relationship between the wavelet and the primary signal. There are several wavelet functions, such as Bior, Morlet, Dmey, Symlet, Db, etc.

For the time series $f(t)$, the time-scale discrete wavelet transform is defined as [15]:

$$f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \varphi \left(\frac{t-b}{a}\right) dt$$  \hspace{1cm} (3)

where, $\varphi(t)$ shows the wavelet, which is typically smaller than the primary signal $f(t)$, $a$ stands for frequency (scale) component, $b$ is the transformation variable, which determines proceeding the wavelet through $f(t)$. In the process of wavelet transform analysis, the primary signal is split into the sub-signals called ‘approximation’ and ‘detail’. The ‘approximation’ and ‘detail’ in every single level are considered as inputs of the ANN model (Figure. 2). In this study, it was found that the optimum decomposition level is three by trial and error. The approximations and details of the rainfall time series after decomposition through the calibration process are shown in Figure. 3. In this figure, ‘A’ and ‘D’ represent the approximation and detail time series. These sub-time series is employed as the input in the ANN modeling. Therefore, using these time series provides the network with more information, which is hidden and cannot be captured by the ANN solely.
3. Case study and data
The Pleystowe station in Queensland, Australia, is selected as a case study to evaluate the proposed model’s performance for calculating future seasonal rainfall (Figure 4). This station is chosen due to holding the long record of data as well as having a small number of missing values.

Moreover, this station is situated in the coastal area with high annual rainfall of 2154.9mm in Queensland, Australia.

In this study, the spring rainfall (September to November) is considered. The rainfall time series is obtained from the website of the Bureau of Meteorology (BOM)[16].

![Figure 3. Approximation and detail time series after the process of splitting the rainfall time series using three-level of decomposition over the calibration period.](image)

![Figure 4. Schematic map of the study area.](image)
In this study, the data of the years 1908 to 2016 (109 years) are employed for predicting the spring rainfall of Pleystowe weather station located in Queensland, Australia. For calibrating the model, the data of the years 1908 to 2005 (98 years) and the data of the remaining years are used for verification of the model.

Moreover, in the current study, the impact of various effective climate indices such as SOI, IPO, and Nino3.4 on future rainfall of the Pleystowe station is investigated. It is noteworthy that due to the effect of large-scale patterns and Southern Oscillation (SO) phases on the climate of Australia, these factors are the most effective parameters for forecasting the climate in Australia.

The pressure differences between Tahiti and Darwin is employed to measure the ENSO [4]. It should be noted that contrary to ENSO, the IPO shows an irregular inter-decadal cycle. Nonetheless, the third climate index used in this study (Nino 3.4 anomalies) measures the sea surface temperature (SST) in the equatorial Pacific Ocean. The data of Nino 3.4 is downloaded from the Royal Netherlands Meteorological Institute (KNMI) Climate Explorer [17]. The data related to SOI is obtained from the Australian BOM [16]. Besides, the IPO monthly values are collected from Met Office [18].

The five input variables using rainfall and climate indices are used to determine the impact of each variable on rainfall and to investigate the performance of each input set as well (Table 1).

| Model | Input variable                      |
|-------|-------------------------------------|
| I     | Rainfall                            |
| II    | IPO                                 |
| III   | SOI                                 |
| IV    | Nino 3.4                            |
| V     | Rainfall+IPO+SOI+Nino3.4            |

To assess the skills of each model, different statistical measures such as the root-mean-squared error (RMSE), refined index of agreement (dr) and correlation coefficient (R) are used. It should be noted that for calculating the correlation coefficient, the related correlation coefficient codes in the Matlab software is run. In this study, an improved version of Willmott’s index of agreement called the refined index of agreement is used. This index was presented by Willmott in 1980s. It is dimensionless and ranges between 1.0 and -1.0. It is noteworthy that this parameter reasonably measures the accuracy of the model compared to other statistical parameters [19]. The RMSE is a measurement of the deviation between predicted and observed values.

The following equations are the mathematical expressions of the statistical parameters used in this study.

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - O_i)^2}
\]  \hspace{1cm} (4)

\[
d_r = \begin{cases} 
1 - \frac{\sum_{i=1}^{n} |P_i - O_i|}{\sum_{i=1}^{n} |P_i - \overline{O}|}, & \text{when } \sum_{i=1}^{n} |P_i - O_i| \leq c \sum_{i=1}^{n} |O_i - \overline{O}| \\
\frac{\sum_{i=1}^{n} |P_i - O_i|}{\sum_{i=1}^{n} |O_i - \overline{O}|} - 1, & \text{when } \sum_{i=1}^{n} |P_i - O_i| > c \sum_{i=1}^{n} |O_i - \overline{O}| 
\end{cases}
\]  \hspace{1cm} (5)

where, \(d_r\) is the refined index of agreement, \(c\) is applied to scale the errors (\(c = 2\)), \(P\) and \(O\) are predicted and observed values, and \(\overline{P}\) and \(\overline{O}\) signify modeled and observed mean magnitude, correspondingly.
4. Results

To obtain the best model, the five models (input variables) consist of rainfall, and three selected climate indices are defined. The best input set is selected with respect to the RMSE values, which provides a clear and straightforward quantitative measure between modeled and observed values. It should be noted that ENSO has a more substantial impact on Queensland climate variability due to its location. Since the Pleystowe station is located in the Queensland state, it will be more affected by ENSO more than other climate indices. Due to the importance of Queensland in the agriculture industry of Australia, forecasting the one-year ahead seasonal rainfall is extremely useful for this region. The performance of the ANN and WANN are shown in Tables 2 and 3.

It is evident that the RMSE values for ANN range from 88.9mm to 109.2mm for different input sets through the verification process. However, the RMSE values for the WANN vary from 64.8mm to 106.7mm. The best result is obtained using the fourth and fifth input sets for ANN and WANN with RMSE values of 88.9mm and 64.8mm, respectively. In other words, the last input set, including all predictors, presents better results than other input sets. The WANN can consider the impact of different predictors for forecasting the seasonal rainfall in Pleystowe weather stations. In terms of the correlation coefficient, $R=0.60$ through the WANN, while it is 0.26 for ANN through the verification period.

The refined index of agreement ($d_r$) value is acquired 0.55 for the WANN compared to 0.36 for the ANN. It is seen that for WANN, the refined index of agreement values are higher than the values from the ANN and this indicated the prediction accuracy of the WANN is superior to the ANN. This index is logically related to the accuracy of the model compared to other existing indices.

Figure 5 compares the skillfulness of the ANN and WANN for cross-validation data. It is seen that the accuracy of the WANN for predicting the seasonal rainfall is higher than the ANN in the verification period.

Table 2. Best ANN input set performance.

| NO | Calibration | Verification |
|----|-------------|--------------|
|    | RMSE | $d_r$ | R | RMSE | $d_r$ | R |
| I  | 101.0 | 0.56 | 0.36 | 100.4 | 0.31 | 0.30 |
| II | 89.1  | 0.74 | 0.72 | 96.4  | 0.40 | 0.34 |
| III| 99.4  | 0.68 | 0.59 | 109.2 | 0.24 | 0.35 |
| IV | 88.2  | 0.59 | 0.58 | 88.9  | 0.36 | 0.26 |
| V  | 103.1 | 0.55 | 0.49 | 94.6  | 0.38 | 0.58 |

Table 3. Best WANN input set performance.

| NO  | Calibration | Verification |
|-----|-------------|--------------|
|     | RMSE | $d_r$ | R | RMSE | $d_r$ | R |
| I   | 65.1 | 0.81 | 0.8 | 104.1 | 0.41 | 0.36 |
| II  | 103.3 | 0.53 | 0.47 | 90.7  | 0.45 | 0.43 |
| III | 93.1  | 0.57 | 0.64 | 106.7 | 0.30 | 0.34 |
| IV  | 106.0 | 0.74 | 0.43 | 103.5 | 0.43 | 0.28 |
| V   | 59.7  | 0.72 | 0.82 | 64.8  | 0.55 | 0.60 |
5. Conclusion
The prediction of future seasonal rainfall one year in advance is remarkably significant for managing the water resources and mitigating the impact of harmful events such as drought and flooding. Artificial Neural Networks can be used to overcome these complexities. Sometimes for improving results and exactness of the models, hybrid models are used especially for modeling the complex phenomena. One of the hybrid models is wavelet-ANN, which is becoming a widely used approach in different fields throughout the world. In recent years, the application of hybrid techniques such as WANN has been increased given their capability to obtain excellent predictive outcomes. The WANN is a hybrid method, which is capable of improving the ability of the ANN using various input set. In this study, the WT was applied along with the ANN to build a hybrid method called WANN. It was shown that the WANN has better performance in comparison with the ANN in forecasting the next year's seasonal rainfall in the selected weather station. This is due to the fact that wavelet analysis, the hidden valuable information of the main signal are extracted by decomposing the signal into some components called approximation and detail. To the best of our knowledge, this technique can be useful in predicting future seasonal rainfall with excellent accuracy.

6. References
[1] Shukla RP, Tripathi KC, Pandey AC, Das IM 2011 Prediction of Indian summer monsoon rainfall using Niño indices: a neural network approach Atmos. Res. 102(1-2) 99-109.
[2] Ghamariadyan M, Imteaz MA, Mekanik F 2019 A hybrid wavelet neural network (HWNN) for forecasting rainfall using temperature and climate indices In IOP Conference Series: Earth and Environmental Science 351(1) 012003.
[3] Garnaut, R 2008 The Garnaut climate change review Cambridge.
[4] Risbey JS, Pook MJ, McIntosh PC, Wheeler MC, Hendon HH 2009 On the remote drivers of rainfall variability in Australia Mon. Weather Rev. 137(10) 3233-53.
[5] Abbot J, Marohasy J 2014 Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks Atmos. Res. 138 166-78.
[6] Schepen A, Wang QJ, Robertson D 2012 Evidence for using lagged climate indices to forecast Australian seasonal rainfall J. Clim. 25(4) 1230-46.
[7] El-Shafie A, Mukhlisin M, Najah AA, Taha MR 2011 Performance of artificial neural network and regression techniques for rainfall-runoff prediction Int. J. Phys. Sci. 6(8) 1997-2003.
[8] Mekanik F, Imteaz MA, Talei A 2016 Seasonal rainfall forecasting by adaptive network-based fuzzy inference system (ANFIS) using large scale climate signals Clim. Dyn. 46(9-10) 3097-111.
[9] Pnevmatikos NG, Hatzigeorgiou GD 2017 Damage detection of framed structures subjected to earthquake excitation using discrete wavelet analysis B. Earthq. Eng. 15(1) 227-48.
[10] Mandrikova OV, Solovjev IS, Khomutov SY, Baishev DG, Geppener VV, Klionskiy DM 2016 Estimation of geomagnetic field disturbance using the wavelet transform Pattern Recognit. Image Anal. 26(4) 773-81.
[11] Kim YH, Cierpka C, Wereley ST 2011 Flow field around a vibrating cantilever: coherent structure eduction by continuous wavelet transform and proper orthogonal decomposition J. Fluid Mech. 669:584.
[12] Kişi Ö 2010 Daily suspended sediment estimation using neuro-wavelet models Int. J. Earth Sci. 99(6) 1471-82.
[13] Nourani V, Komasi M, Alami MT 2012 Hybrid wavelet–genetic programming approach to optimize ANN modeling of rainfall–runoff process J. Hydrol. Eng. 17(6) 724-41.
[14] Haykin SS 2009 Neural networks and learning machines Pearson.
[15] Mallat, S 1999 A wavelet tour of signal processing Elsevier.
[16] Australia’s official weather forecasts and weather radar-Bureau of Meteorology. Available: http://www.bom.gov.au
[17] Royal Netherlands Meteorological Institute (KNMI) Climate Explorer. Available: http://www.climexp.knmi.nl
[18] Meteorology office weather forecasts for the UK. Available: https://www.metoffice.gov.uk
[19] Willmott CJ, Robeson SM, Matsuura K 2012 A refined index of model performance Int. J. Climatol. 32(13) 2088-94.