Chapter 3

Forecasting of Medium-term Rainfall Using Artificial Neural Networks: Case Studies from Eastern Australia

John Abbot and Jennifer Marohasy

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.72619

Abstract

The advent of machine learning, of which artificial neural networks (ANN) are a component, has provided an opportunity for improved rainfall forecasts, which is of value for water infrastructure management, agriculture, mining and other industries. In this chapter, ANNs are shown to provide more skillful monthly rainfall forecasts for locations in south-eastern Queensland, Australia, for lead-times of 3–12 months. The skill of the forecasts from the ANNs is highest when the models are individually optimized for each month, and when longer-duration series are used as input. The ANN technique has application where there is temperature and rainfall data extending back at least 50 years. Such datasets exist for much of Europe and North America, though a review of the available literature indicates most research into the application of ANN has focused on China, India and Australia.

Keywords: rainfall, forecast, monthly, neural network, Australia

1. Introduction

Until relatively recently, simple statistical models were used by meteorological agencies around the world to forecast seasonal and monthly rainfall. Typically, these models use relationships between large scale climate indices, such as the Southern Oscillation Index, and rainfall at some future time, generally utilizing a small number of input variables, perhaps only one or two. For example, until May 2013 the Australian Bureau of Meteorology (BOM) generated seasonal rainfall forecasts based on a statistical scheme using an El Niño Southern Oscillation (ENSO) index as a primary predictor in a relatively simple statistical model [1, 2].

These traditional statistical models are limited in the number of input variables that can be effectively combined, while advances in machine learning has now significantly expanded this potential [3]. Machine learning has close relationships with artificial intelligence, pattern recognition and data mining. Data mining focuses on the discovery of previously unknown properties embedded in data [4], whereas machine learning focuses on prediction based on known properties learned from exposure to data sets during a process known as "training". A principle objective of the learning process is to construct a model that can generalize from experience [5]. Subsequently, the performance of the trained model can be tested using data not utilized in the training set. Performance of the model in testing gives confidence, but not certainty, that the model would provide reliable forecasts if deployed operationally.

Artificial neural networks (ANNs), a form or machine learning, provide several important advantages over simple statistical models. ANNs can accommodate non-linear relationships, and test multiple inputs, and this is particularly important when the influence of climate indices may vary geographically and temporarily in poorly understood ways [6].

Rather than progressing from simple statistical models to artificial neural networks, however, meteorological agencies around the world now tend to rely almost exclusively on general circulation modeling for rainfall forecasting. These are physical models, that attempt to simulate real-world oceanic and atmospheric circulation patterns. For example, the Australian Bureau of Meteorology uses the Predictive Oceanic Atmospheric Model for Australia (POAMA) as its operational model for forecasting daily rainfall, and also as the basis of its monthly and seasonal forecasts.

The skill of the forecasts from ANNs can be compared with POAMA (and other general circulation models) through a comparison of root mean square errors (RMSE), mean absolute errors (MAE) and correlation coefficients – with such comparisons a focus of this chapter. RMSE is commonly applied to compare skill between different rainfall forecasting models, as it gives a simple, transparent, quantitative measure of difference between input and target and is easily understood across disciplines [7]. RMSE is more sensitive than MAE to the occasional large error (the squaring process gives higher proportionate weight to the large errors), and is therefore arguably more useful when skill at forecasting floods is particularly relevant.

Given the importance of skillful monthly rainfall forecasts for most inhabited regions, for activities as diverse as crop harvesting, mine scheduling, and dam management, it is surprising that there are so few comparative studies and so little discussion about how advances in machine learning might aid medium-term rainfall forecasting – including the forecasting of flood events.

Queensland is a state in the north-east of Australia, facing the South Pacific. Brisbane is the fast-growing capital of Queensland, and is located in the south-east. Brisbane has a long history of flooding and the Wivenhoe dam was built specifically for flood mitigation following a flood in 1974.

The flooding of Brisbane in January 2011 has been termed a "dam release flood" [8] with the sudden release of water from the Wivenhoe storage a principal cause of flooding. The extent of the rainfall in 2010/2011 was *not* unprecedented relative to historical records that extend back to 1864; but the heavy rainfall was not forecast [9]. Because the heavy rain in December 2010 and January 2011 followed a long period of drought, and was not forecast, the Wivenhoe dam was not properly managed for flood mitigation – the purpose for which it was originally built. Because of the extent of the economic losses the flood event has led to a class action lawsuit against the dam operator SEQEB and the State of Queensland, with the trial expected to commence in late 2017 [10].

2. Data, method, and a first study

Monthly data were obtained from the Bureau's Climate Data Online. Data was downloaded for all 62 sites in south eastern Queensland that are considered in this chapter. The sites were chosen on the basis of their geographic spread and also the quality of the data: that is the desirability of long data series with few missing values.

South-eastern Queensland is defined very broadly in this chapter, and does not correspond with the administrative region of South East Queensland.

This chapter is a review of various studies undertaken since 2012 focused on this general area, with specific information on data and methodology in the published technical papers that are referenced [11–21]. However, in this first section, the method used in an early study [17] is provided in more detail, by way of background into how an ANN can be practically deployed to generate a rainfall forecast.

Many neural network applications incorporate multilayer perceptrons (MLPs) as fundamental processing elements (PEs) trained with a standard backpropagation algorithm. These neural networks can perform well in solving static problems but are limited in solving temporal problems, ones where the previous value of the input affects the current output. Recurrent networks, such as Jordan networks, extend the basic MLP architecture by also including context units, PEs that remember past activity. In the Jordan network, the output of the network is copied to the context unit.

In addition, the context units are locally recurrent, that is, they feedback onto themselves. The local recurrence decreases the values by a multiplicative time constant (τ) as they are fed back. This constant determines the memory depth, that is, how long a given value fed to the context unit will be "remembered". The context unit acts as a simple lowpass filter, creating an output ($y(n)$), calculated as a weighted average value of some of its more recent past inputs. In the case of the Jordan context unit, the output is obtained by summing the past values multiplied by the scalar τn, where:

$$y(n) = \sum x(n)\, \tau n \tag{1}$$

Genetic optimization provides an efficient way of selecting those inputs that are significant in determining target rainfall and eliminating those inputs with very low information content. Essentially, genetic optimization enables elimination of inputs that carry mainly noise rather than useful signal, so that the number of input considered in the optimized model might typically be reduced from over 40 to less than 10.

In the early study [17] forecasts were made for Lowood and two other sites within the Brisbane catchments (upstream of the Wivenhoe dam) using the Jordan network with optimization. The initial forecasts were made using only lagged input parameters – that is any variable measured in the past. The initial focus was on:

i. Benchmarking output from the ANN against POAMA (the operational model used by the Bureau);

ii. Evaluating the impact of 1-, 2- and 3-month lags, and accordingly the extent to which a rainfall forecast potentially loses skill moving from forecasting 1, 2 and 3 months in advance;

iii. Determining the capacity of the ANN to deal with unary, binary and ternary data sets, that is, with an increasing number of input variables;

iv. Determining if the ANN model could forecast the extreme rainfall of December 2010 and January 2011 for these sites; and

v. A 'best forecast' using only lagged values was constructed to explore the potential to improve forecasts through post-processing, as shown in **Figure 1**.

Figure 1 shows the skill of this forecast, as an orange line, relative to observed rainfall, blue line. Clearly the ANN was able to forecast that December 2011 was likely to be unusually wet at Lowood. The rainfall of over 600 mm for the month of December contributed to the flooding of Brisbane in early January 2011.

In this early study [17], several climate indices were also successfully forecast and then inputted as lead variables. A 'lead', also known as a forecast value, are un-measurable from the reference point of the current period, but potentially predictable in a forecast model.

Variations in rainfall in many parts of the world, including south-eastern Queensland, are associated with large-scale climate phenomena which can be described by climate indices typically measuring changes in temperatures and pressures across oceans [22–24]. ENSO, a Pacific Ocean phenomenon, can be represented by both the Southern Oscillation Index (SOI) and a combination of four different Niño values (Niño 4, Niño 3.4, Niño 3, Niño 1.2). The Interdecadal Pacific Oscillation (IPO) also measures pressure and temperature changes in the Pacific Ocean. The Indian Ocean Dipole measured by the Dipole Mode Index (DMI), is a measure of pressure and temperature changes in the Indian Ocean.

These large-scale climate indices were used as input data, together with local data for each site specifically rainfall, and temperatures. This data was always divided into training (75%), evaluation (15%) and testing sets (10%) – with absolutely no overlap.

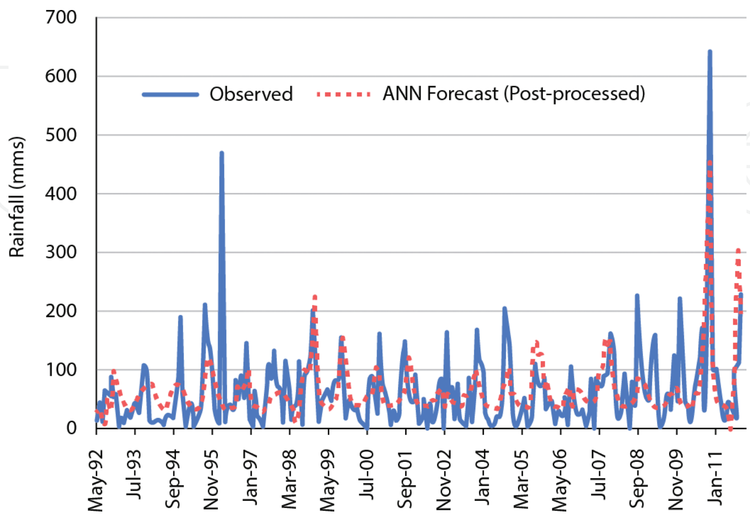

Figure 1. Rainfall at Lowood, south-eastern Queensland, as observed and forecast by the ANN [17].

3. Progressing to automation and single-month optimization

A wide range of ANN architectures have been applied in forecasting rainfall [11, 13, 25–27]. The selection of an ANN architecture is commonly achieved through a trial and error process [11–13]. This can, however, be very time-consuming: that is choosing network topology, and architecture based on trial and error – with the selected model with lowest error score, subsequently applied with all the data input sets.

In contrast, with the Neurosolutions Infinity software that we used since 2015 [19–21] the selection of network architecture and configuration was automated. This offered a great advantage in terms of arriving at an optimum forecast model for each data set of interest without prohibitive time outlay. The Infinity program uses a pre-set formula incorporating RMSE, MAE and correlation coefficient values to evaluate the accuracy for each ANN model and a corresponding set of selected inputs. Based on this formula, the program determines which ANN model and set of inputs is optimal.

In the early study using the Jordan ANN model (Lowood example), and also subsequently with the Infinity software automating the process of selection of network architecture, the default was for optimization of "all-months" within each calendar year at the same time. That is, it was assumed that the same climate indices would be important for each of the months of the year. This did not prove valid. Indeed, it is well known within the climate science literature, that the magnitude of the linear correlation between SOI and annual rainfall are highly variable both temporarily and geographically particularly for Australia [24, 28] has shown that other 'remote drivers' of Australian rainfall dominant at different times of the year. Furthermore, this temporal variability has been a reason why forecasting beyond autumn in the southern hemisphere, and spring in the northern Hemisphere spring has been considered particularly problematic [29–31].

In a study published in 2015 [15] we showed for the first time that the prominent rainfall peak in December 2010 for the location of Harrisville, which is near Lowood, could be much more skillfully forecast using a single-month optimization technique. Further, the very heavy rainfall could be forecast at the long-lead times of at least 12 months, as shown in **Table 1**. We have since extended the method of single month optimization to show its relevance to locations along the Queensland coast [20].

The process is much more time consuming as 12 optimizations are carried out to produce monthly rainfall forecasts for the entire year. Improved skill, however, is generally achieved as demonstrated by reduced MAE and RMSE and increased correlation coefficient r.

While single-month optimization clearly gives a lower RMSE and MAE than all-month optimization and therefore is a preferred forecast method, it should be noted that both methods of ANN optimization provided a superior forecast to climatology and POAMA as shown in **Table 1**.

The results shown in **Table 1** where achieved from a general regression neural network (GRNN). This same ANN architecture was used to generate forecasts for Bingera, which is in a coastal catchment to the north of Brisbane. Both single-month and all-month optimization methods gave skillful forecasts for Bingera, as shown in **Figure 2**.

Gatton			
	RMSE (mm)	MAE (mm)	r
All-month optimization	49.9	34.7	0.70
Single-month optimization	32.4	22.2	0.91
POAMA	59.2	40.0	0.51
Climatology	59.7	40.2	0.50
Harrisville			
All-month optimization	39.4	29.8	0.78
Single-month optimisation	28.4	19.1	0.88
Climatology	49.8	34.6	0.57
POAMA	49.3	34.2	0.59

Table 1. Skill parameters for monthly rainfall forecasting for Gatton and Harrisville on an annual basis for the test period July 2004 to August 2011 [15].

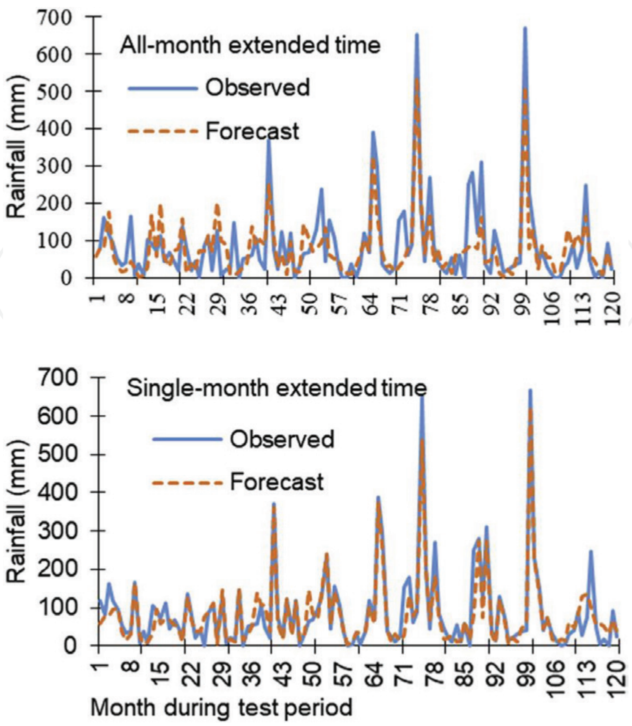

Figure 2. Test results for Bingera showing difference between single-month and all-month NN optimizations [20].

4. Best forecasts from longer datasets

Our very first paper on forecasting rainfall using ANNs [11] focused on 17 sites, chosen specifically because they had the longest continuous rainfall records for anywhere in Queensland. In that paper, we used a temperature series from the relatively far away destination of Sydney – with the comment that there is no comparable temperature dataset available for any Queensland location. We have since used local temperature series, even if this requires the infilling of missing values through interpolation and/or linear regression [32].

In our early work, we also limited the selection of climate indices to those of 'long duration' – specifically ENSO (SOI and the four Ninos), IPO and DMI. These climate indices were downloaded from the Royal Netherlands Meteorological Institute (KNMI) Climate Explorer.

Input	Data series duration
Local	
Rainfall	Long
MaxT	Long
MinT	Long
Regional	
SOI	Long
IPO	Long
DMI	Long
Nino 1.2	Long
Nino 3	Long
Nino 4	Long
Nino 3.4	Long
SAM	Short
QBO	Short
MEI	Short
SEIO	Short
WIO	Short

Table 2. Attributes used as input for ANN models.

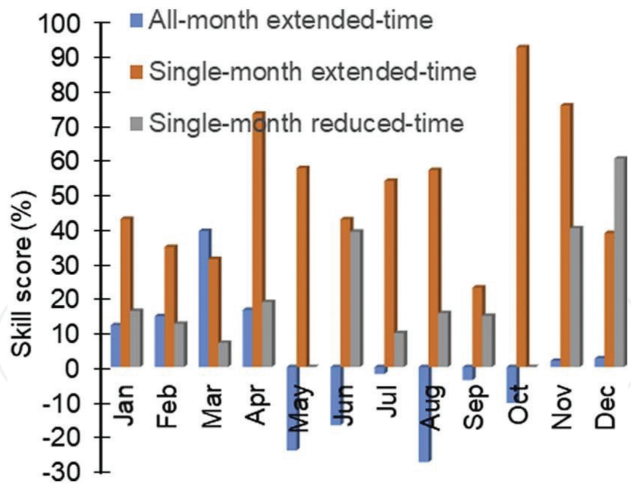

Figure 3. Skill scores for monthly forecasts for Victoria Mill (central Queensland), including the comparison between single-month optimization for short duration (reduced-time) and long duration (extended-time) forecasts [20].

There are many more climate indices available, but most are of much shorter published duration, often only beginning in the 1950s. These 'short' duration indices are listed in **Table 2**, specifically the south-eastern Indian Ocean index (SEIO), Western Indian Ocean Index (WIO), The Southern Annular Mode (SAM) and the Quasi-Biannual Oscillation (QBO). Data for these indices were variously sourced from the Bureau, KNMI Climate Explorer, and also the UK Met Office for the recent publication [20] in which we tested the trade-off between more climate indices and shorter series as input.

The skill of an ANN is dependent on finding patterns or relationships in data, and we intuitively thought that input series of less than 100 years would be unlikely to provide adequate information. In a recent study for a region in central Queensland it was apparent that the skill of the forecast depended on the particular month for which the forecast was being made, as shown in **Figure 3**. Nevertheless, it was apparent that of more importance than inputting long-duration series, was single-month optimization. When single month optimization was applied to both long and short duration series, the long duration series gave superior forecasts, as shown in **Figure 3**. More information on the definition of, and methodology used, to calculation the skill score is provided in the technical paper [20] and under the following section: Deterministic versus Probabilistic Forecasts, and Skill Scores.

5. Regional rainfall forecasts

The rainfall forecasts discussed so far correspond to specific locations within a region. The limitation in selection of those sites depends on the availability of sufficient historical data for adequately training and testing. In some cases, shorter data sets may be adequate, and shorter data duration may be offset by inclusion of additional climate indices [20].

Provided there are sufficient well-distributed sites within a geographical region, it may be possible to construct a regional rainfall forecast map, as shown in **Figures 4** and **5**. This

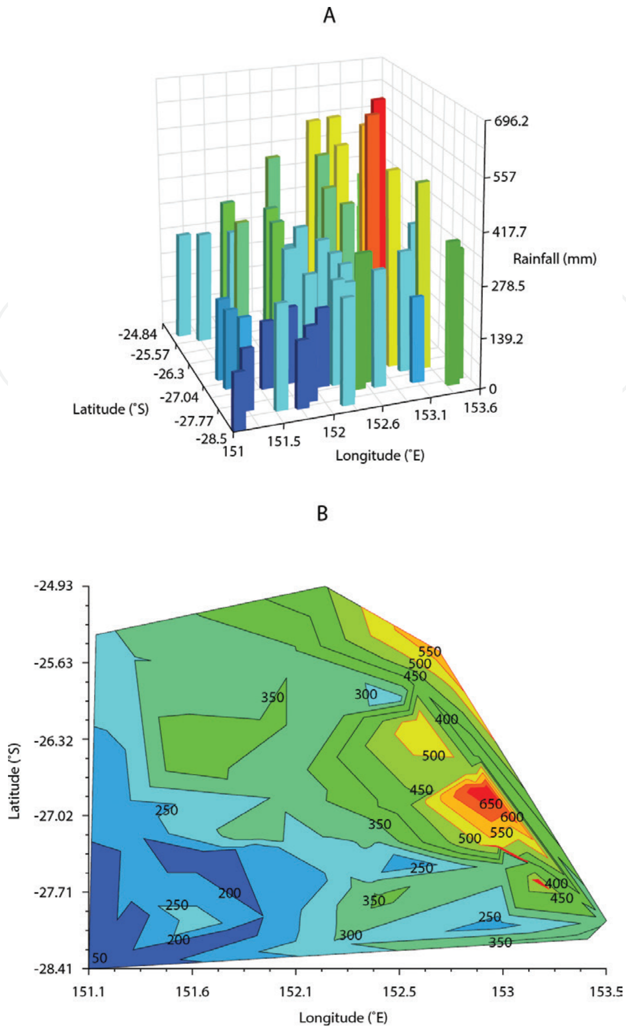

Figure 4. Forecast rainfall (mm) for December 2010 for the south-east Queensland region. A: bar chart for individual sites; and B: isohyet map with 50 mm interval spacing [21].

enables forecasts to be generated for specific locations within the geographical region that do not correspond to sites with rainfall data. It also enables rainfall forecasts to be made corresponding to a defined geographical area within the region, such as a water catchment area.

As an example, **Figure 4**, is a specific sub-region of south-eastern Queensland defined by longitudes 151.0°E and 153.5°E, and latitudes 24.4°S and 28.5°S based on rainfall data from 54 sites. The region extends about 300 km southwards along the Queensland Coast from Bundaberg to the Gold Coast, and extended inland from the coast to include the towns of Dalby, Inglewood and Mundubbera, a distance of about 200 km. This sub-region therefore approximates the same area as a single grid forecast area used by POAMA (250 km × 250 km).

This example is provided to show that it is possible to develop an isohyet map for such a region derived from the individual series to predict rainfall for a period of intense rainfall, as occurred in December 2010, and also a relatively dry month such as December 2005. The isohyet map was constructed using *Teraplot* software – and the forecast are 12 months in advance using single-month optimization [21].

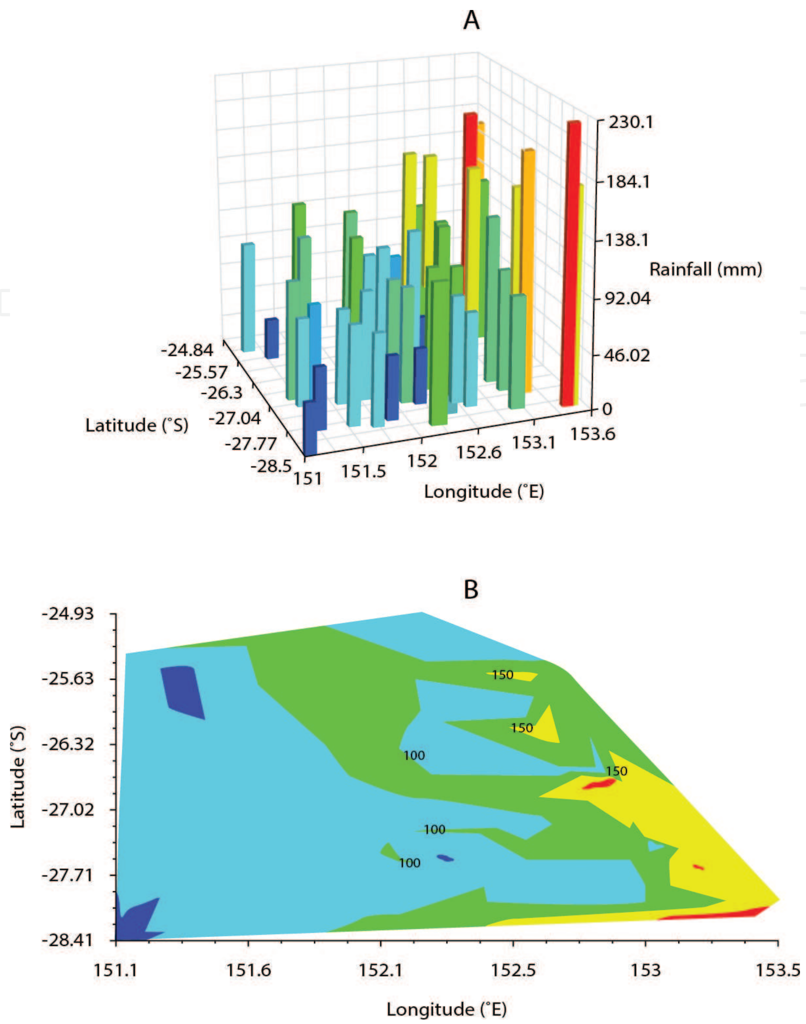

Figure 5. Forecast rainfall (mm) for December 2005 for the south-east Queensland region. A: bar chart for individual sites; and B: isohyet map with 50 mm interval spacing [21].

In **Figures 4** and **5**, it is evident that rainfall varies with topography, and was highest at the places of some altitude in both the relatively dry December of 2005, and also the flood month of December 2010. It is also evident that the ANN correctly predicted December 2010 as experiencing relatively higher rainfall across the sub-region, peaking at over 700 mm.

6. Deterministic versus probabilistic rainfall forecasts, and skill scores

In this chapter, we have shown that forecasts generated through application of an ANN are deterministic, that is numerical values are produced corresponding to the expected rainfall for a specific month. In contrast, medium-term rainfall forecasts issued to the public by the Australian Bureau of Meteorology are in the form of probabilities relative to the median seasonal rainfall, and do not differentiate between an anticipated rainfall slightly above the median and an extreme rainfall event, such as occurred in Queensland during the period December 2010 and January 2011 [14].

This is illustrated in **Figure 6** showing the seasonal forecast issued by the BOM for the period December 2010 to February 2011 for the entire continent of Australia (Bureau of Meteorology, Archive of rainfall forecasts). Although the forecast indicates that there is a probability in the range of 60–70% above median rainfall for the south-east Queensland region, there was no warning of the magnitude of the impending heavy rainfall. Furthermore, the extensive flooding that affected much of costal Queensland beyond the south-eastern region where the forecast is a 50% probability of above median. This contrasts with the specific quantities forecast for the isohyet maps shown in the previous section.

Not only is the information content of probabilistic forecasts less than corresponding deterministic forecasts, where actual numerical values of predicted rainfall are provided, studies have demonstrated that the general public and specific classes of end-users, such as farmers from south-east Queensland, often have difficulty interpreting the meaning of the probabilistic forecasts [33, 34]. The distinction between median and average was often not understood. In addition, the statement that there was a 30% probability of above the median rainfall was often misinterpreted to mean a forecast of a specific quantity of rainfall 30% higher than the median. There is confusion between a probabilistic and deterministic forecast that can be understood by the public as the expectation of higher rather than lower rainfall as intended by the forecaster.

The Bureau bases its probabilistic forecasts on the simulation of actual physical climatic processes through general circulation models, specifically POAMA [36–38]. For the period from 2002 through until 2011, seasonal rainfall predictions were made using POAMA 1.5.

In July 2011, the Bureau provided us with monthly forecasts for the period to March 2011 for 17 individual sites as simple bilinear interpolations of surrounding grid points which were calculated from the ensemble mean which in turn had been calculated from many runs of this

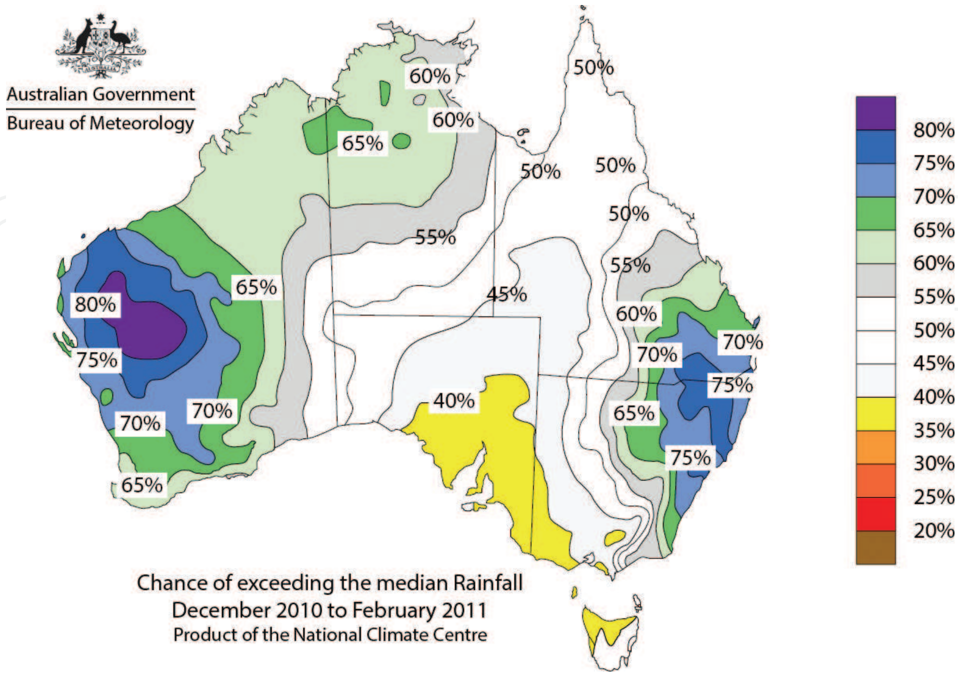

Figure 6. Seasonal rainfall forecast for Australia by the BoM issued in November 2010. http://www.bom.gov.au/climate/ahead/archive/rainfall/20101123.national.hrweb.gif [35].

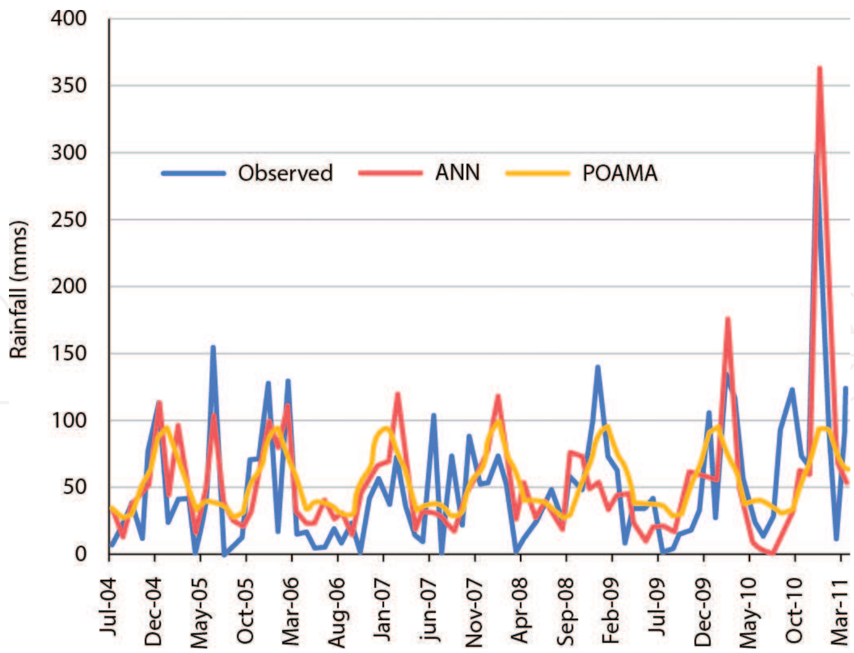

Figure 7. Forecasts for Miles from POAMA 1.5 (yellow line), and the ANN (red line) benchmarked against actual observed rainfall (blue line) for a 3-month lead for the period July 2004 to March 2011.

	RMSE (mm)	MAE (mm)	r
	Lead 3 months		
ANN	32.6	25.2	0.74
POAMA	45.2	32.8	0.44
Climatology	45.5	33.0	0.43
	Lead 6 months		
ANN	33.8	26.2	0.74
Climatology	47.6	33.6	0.33
POAMA	48.0	34.7	0.33
	Lead 9 months		
ANN	33.2	25.6	0.74
Climatology	47.3	33.9	0.34
POAMA	47.9	34.2	0.33

Table 3. Comparison of rainfall forecast skill for Miles using different forecasting methods at leads of 3, 6 and 9 months for the period July 2004 to March 2011.

general circulation model, POAMA 1.5. **Figure 7** shows the difference between the observed monthly rainfall, together with forecasts generated by POAMA 1.5 for the site of Miles. The forecasts generated by POAMA show little variation between the 6-year period 2004–2009 and the occurrence of heavy rainfall in 2010: each showing a maximum of approximately 100 mm. In contrast, the forecast generated by the ANN clearly show much higher rainfall of about 350 mm, closely corresponding to the observed value, **Figure 7** (**Table 3**).

Evaluations of forecast skill were also made by comparisons of skill scores calculated relative to climatology from Eq. (1). This is analogous to the method used by the Bureau where a skill score is calculated for forecasts using POAMA [39].

$$\text{Skill score} = \left[\text{RMSE (climatology)} - \text{RMSE (model)}/\text{RMSE (climatology)}\right]\ 100\% \qquad (2)$$

When the calculated value of the RMSE from climatology and for a particular model are equal, the forecast skill score will be zero. For a perfect forecast, the RMSE for the model will be zero, and the calculated skill score 100%. Negative values calculated from Eq. (2) indicate a forecast

Month	ANN (lead 12 months)	POAMA (lead 8 months)
January	30.2	6.0
February	22.2	-1.7
March	49.3	-6.3
April	39.3	-40.3
May	33.1	-5.9
June	43.3	-1.2
July	62.1	-3.4
August	30.6	0.0
September	69.0	-2.0
October	43.0	-8.0
November	22.0	8.1
December	48.6	-16.1
Average	41.0	-5.9

Table 4. Forecast skill as a percentage for monthly rainfall forecasts for Miles with reference to climatology for the composite ANN model and POAMA.

skill score worse than climatology. **Table 4** shows skill scores for each individual month using an ANN with a lead time of 12 months and POAMA with a lead time of 8 months. In all cases the ANN skill scores are positive and lie in the range 23–69%. In contrast, the skills cores for POAMA forecasts are negative, that is worse than climatology.

Hawthorne et al. [39] used the output from POAMA to produce monthly rainfall forecasts with up to 8 months lead time for 250 km × 250 km grid areas over continental Australia. The skill of their monthly rainfall forecasts was described as being generally low. Skill scores fell between −20% and 20%, including grid locations in coastal south-east Queensland. For lead times between 3, 4, 5, 6, 7 and 8 months, with approximately 60% of the forecasts give skill levels relative to climatology below 0%. For southeast Queensland only about 20% of the forecasts had a skill level in the 15–20% range [39] Hawthorne et al. [39] described the skill of their monthly rainfall forecasts as low, and concluded that monthly rainfall forecasting with POAMA remained a challenge.

7. Other ANN studies

The comparative data we have shown for specific locations, and also the isohyet for the sub-region, indicate that ANNs can give more skillful forecasts than the general circulation models currently relied upon by the Australian Bureau of Meteorology.

We have also published a series of articles describing the application of ANNs to monthly rainfall forecasting in other parts of Australia including northern and central Queensland [11–13, 15–17] the Murray Darling Basin [14] and Western Australia [19]. The forecasts for these regions also use a combination of large-scale climate indices and the local variables temperature and rainfall with lead times between 1 and 18 months.

The island state of Tasmania lies to the south of mainland Australia, and has an area similar to that of the region of south-eastern Queensland considered above. Tasmania is primarily reliant on hydro-electric power, and therefore the state is highly dependent on rainfall for electricity generation. Until 2005, Tasmania was self-reliant for electricity when the Basslink interconnector to mainland Victoria was completed. This link was intended to both provide energy security, in case of drought; and provide renewable energy to Victoria, which relies heavily on coal for power generation. Hydro Tasmania water storage levels decreased from over 60% capacity in 2013 to below 15% in early 2016. The state experienced an energy crisis in 2016 following 2 years of high volumes of energy export to Victoria, followed by low rainfall and a fault in the link to the mainland.

This situation emphasized the need to incorporate more accurate rainfall forecasts in planning decisions relating to hydro-electricity generation. The skill of monthly forecasts for this region of Australia at 2–3 months lead time generated by the Australian Bureau of Meteorology for this region are in the range −10 to 10%, averaging close to climatology [39]. Tasmania has at least 30 sites with long rainfall records extending back about a century or longer, distributed over the island. It should therefore be possible to construct regional or catchment forecasts for medium-term rainfall. One example is shown for Hobart, the capital of Tasmania in **Figure 8**. This shows forecasts of monthly rainfall at a 3 month lead time for a test period of

http://dx.doi.org/10.5772/intechopen.72619

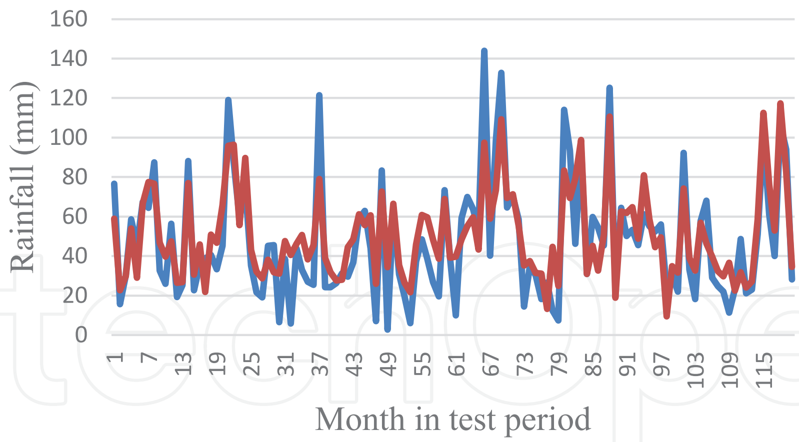

Figure 8. Observed (blue) and forecast (red) monthly rainfall for Hobart, Tasmania with a lead time of 3 months.

Month	ANN (lead 3 months)
January	35.1
February	50.9
March	50.6
April	77.1
May	44.9
June	43.9
July	39.0
August	34.0
September	53.3
October	40.8
November	67.7
December	61.8
Average	49.9

Table 5. Forecast skill as a percentage for monthly rainfall forecasts for Hobart with reference to climatology for the composite ANN model and POAMA.

10 years from January 2004 to December 2013. **Table 5** gives the corresponding skill scores for each month, which lie in the range 34–71%. This is significant improvement on that currently achievable using general circulation models.

Other studies, including by Bagirov et al. [40] predicted monthly rainfall for 8 sites in Victoria with 5 input meteorological variables, using data over the period 1889–2014. Montazerolghaem et al. [41] made monthly rainfall forecasts from 136 weather stations in south-eastern Australia in the states of Victoria, NSW, South Australia and southern Queensland with 1 month lead time. Using large scale climate indices, He et al. [42] made forecasts of monthly rainfall at 255 stations across Australia with training data between 1959 and 1889 test data between 1999 and 2008 with large-scale climate indices as inputs.

ANNs have been applied to rainfall forecasting in other parts of the world [25, 26, 43], with **Table 6** showing a survey of 36 articles published between 2014 and 2017 where ANNs have been used to forecast monthly or seasonal rainfall. The majority of these investigations were from Asia, particularly India and China.

ANNs have also been applied to rainfall forecasting at both shorter forecast periods, such as hourly [44] and daily [45] as well as longer forecast periods including annual [46]. ANNs have also been used to forecasting other climatic and meteorological variables including temperatures [47–50] and wind speed [51, 52]. Other machine learning methods have also been applied to rainfall forecasting; including support vector machines (SVM) [53] and adaptive neuro-fuzzy inference systems (ANFIS) [54–56].

The survey results detailed in **Table 6** indicate that there have been very few studies of ANNs and their application to rainfall forecasting in North America and Europe. These are regions that are likely to have the longest rainfall and temperature series and therefore are potentially ideally suited to such studies.

Country	Region	Forecast period	Reference
AUSTRALIA			
Australia	Queensland and Western Australia	Monthly	[20]
Australia	Queensland	Monthly	[21]
Australia	Victoria	Monthly	[40]
Australia	Bowen Basin, Queensland	Monthly	[18]
Australia	Western Australia	Monthly	[19]
Australia	south-eastern and eastern Australia	Monthly	[41]
Australia	Miles Murray Darling Basin,	Monthly	[14]
Australia	Brisbane River catchment, Queensland	Monthly	[15]
Australia	Continental Australia	Monthly	[42]
Australia	Bowen Basin, Queensland	Monthly	[16]
Australia	Queensland	Monthly	[13]
ASIA			
China	Aksu-Tarim River basin	Seasonal	[57]
China	Liaoyuan city	Monthly	[58]
China		Monthly	[59]

Country	Region	Forecast period	Reference
China	Yellow River	Seasonal	[60]
China		Monthly	[61]
China	western Jilin Province,	Monthly	[62]
China	Arid region north-west China	Monthly	[63]
India	Indian subcontinent monsoon	Seasonal	[64]
India	All India monsoon	Seasonal	[65]
India	Coonoor region	Monthly	[66]
India	Thanjavur district, province of Tamil Nadu	Monthly	[67]
India	Andhra Pradesh state	Monthly	[68]
India	Assam	Monthly	[69]
India	South peninsula	Seasonal	[70]
Malaysia	Johor state	Monthly	[71]
Indonesia	Ampel, Boyolali	Monthly	[72]
Indonesia	Tenggarong Station, East Kalimantan	Monthly	[73]
Thailand	Ping Basin	Seasonal	[74]
Thailand	North-east region	Monthly	[75]
MIDDLE EAST			
Iran	Mashhad	Monthly	[76]
Iran	Qara-Qum catchment	Monthly	[77]
Jordan	Arid region	Monthly	[27]
AFRICA			
Ethiopia	Upper Blue Nile Basin	Seasonal	[78]
West Africa	Sahel region	Seasonal	[79]
NORTH AMERICA			
USA	California	Seasonal	[80]

Table 6. Survey of publications on rainfall forecasting 2014-2017.

8. Conclusion

While general circulation models are used by meteorological agencies around the world for rainfall forecasting, they do not generally perform well at forecasting medium-term rainfall, despite substantial efforts to enhance performance over many years [36, 81–83]. These are the same models used by the Intergovernmental Panel on Climate Change (IPCC) to forecast climate change over decades. Though recent studies suggest ANNs have considerable application here, including to evaluate natural versus climate change over millennia, and also to better understand equilibrium climate sensitivity [84].

While machine learning is now a well-established discipline, and artificial neural networks a well understood subcomponent, this technology is only beginning to be applied to rainfall forecasting, so far with most of this effort concentrated in China, India and Australia as shown in **Table 6**. The technology is likely to have application to many more regions, particularly North America and Europe where longer time series of rainfall and temperature is likely to exist.

In this chapter, we demonstrated the application of ANNs for rainfall forecasting for specific locations and also a sub-region within south eastern Queensland, Australia. We demonstrated that the ANN could forecast the extreme flood event of December 2010 that resulted in the catastrophic flooding of Brisbane in January 2011. We show that the skill of the ANN forecast can be improved through single-month optimization of the ANN model, and also by using datasets that extend further back in time. We compare forecasts from the ANN with forecasts from a general circulation model, POAMA, and show that both the deterministic nature of the forecasts, and also their actual skill, indicates that ANNs have much to offer by way of improved monthly rainfall forecasts at least 1 year in advance.

Acknowledgements

This work was supported by the B. Macfie Family Foundation.

Author details

John Abbot[1,2,3]* and Jennifer Marohasy[2,3]

*Address all correspondence to: johnwabbot@gmail.com

1 University of Tasmania, Tasmania, Australia

2 Institute of Public Affairs, Melbourne, Victoria, Australia

3 Climate Modelling Laboratory, Noosa, Queensland, Australia

References

[1] Drosdowsky W, Chambers LE. Near-global sea surface temperature anomalies as predictors of Australian seasonal rainfall. Journal of Climate. 2001;**14**:1677-1687

[2] Fawcett RJB, Stone RC. A comparison of two seasonal rainfall forecasting systems for Australia. Australian Meteorological and Oceanographic Journal. 2010;**60**:15-24

[3] Mitchell T. Machine Learning. Maidenhead, UK: McGraw Hill; 1997:2

[4] Simon P. Too Big to Ignore: The Business Case for Big Data. Somerset, New Jersey: John Wiley & Sons; 2013:89

[5] Bishop CM. Pattern Recognition and Machine Learning. New York, NY: Springer; 2006

[6] Kirono DGC, Chiew FHS, Kent DM. Identification of best predictors for forecasting seasonal rainfall and runoff in Australia. Hydrological Processes. 2010;**24**:1237-1247

[7] Singh P, Borah B. Indian summer monsoon rainfall prediction using artificial neural network. Stochastic Environmental Research and Risk Assessment. 2013;**27**:1585-1599

[8] van den Honert RC, McAneney J. The 2011 Brisbane floods: Causes, impacts and implications. Water. 2011;**3**:1149-1173

[9] Queensland Flood Commission of Enquiry, Final Report 2012. http://www.floodcommission.qld.gov.au/__data/assets/pdf_file/0007/11698/QFCI-Final-Report-March-2012.pdf

[10] Maurice Blackburn Lawyers. Queensland Floods Class Action. https://www.mauriceblackburn.com.au/current-class-actions/queensland-floods-class-action/

[11] Abbot J, Marohasy J. Application of artificial neural networks to rainfall forecasting in Queensland, Australia. Advances in Atmospheric Sciences. 2012;**29**(4):717-730

[12] Abbot J, Marohasy J. The potential benefits of using artificial intelligence for monthly rainfall forecasting for the Bowen Basin, Queensland, Australia. Water Resources Management VII. WIT Transactions on Ecology and the Environment. 2013;**171**:287-297

[13] Abbot J, Marohasy J. Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks. Atmospheric Research. 2014;**138**:166-178

[14] Abbot J, Marohasy J. Forecasting of monthly rainfall in the Murray Darling Basin, Australia: Miles as a case study. WIT Transactions on Ecology and the Environment. 2015;**197**:149-159

[15] Abbot J, Marohasy J. Improving monthly rainfall forecasts using artificial neural networks and single-month optimisation: A case study of the Brisbane catchment, Queensland, Australia. WIT Transactions on Ecology and the Environment. 2015;**196**:3-13

[16] Abbot J, Marohasy J. Using artificial intelligence to forecast monthly rainfall under present and future climates for the Bowen Basin, Queensland, Australia. International Journal of Sustainable Development and Planning. 2015;**10**(1):66-75

[17] Abbot J, Marohasy J. Using lagged and forecast climate indices with artificial intelligence to predict monthly rainfall in the Brisbane Catchment, Queensland, Australia. International Journal of Sustainable Development and Planning. 2015;**10**(1):29-41

[18] Abbot J, Marohasy J. Forecasting monthly rainfall in the Bowen Basin of Queensland, Australia, using neural networks with Niño Indices for El Niño-Southern Oscillation. Lecture Notes in Computer Science. 2016;**9992**:88-100

[19] Abbot J, Marohasy J. Forecasting monthly rainfall in the Western Australian wheat-belt up to 18 months in advance using artificial neural networks. Lecture Notes in Computer Science. 2016;**9992**:71-87

[20] Abbot J, Marohasy J. Skilful rainfall forecasts from artificial neural networks with long duration series and single-month optimization. Atmospheric Research. 2017;**197**:289-299

[21] Abbot J, Marohasy J. Forecasting extreme monthly rainfall events in regions of Queensland, Australia using artificial neural networks. International Journal of Sustainable Development and Planning. 2017;**12**(7):1117-1131

[22] Schepen A, Wang QJ, Robertson D. Evidence for using lagged climate indices to forecast Australian seasonal rainfall. Journal of Climate. 2012;**25**:1230-1246

[23] Murphy BF, Ribbe J. Variability of southeastern Queensland rainfall and climate indices. International Journal of Climatology. 2004;**24**:703-721

[24] Risbey JS, Pook MJ, Mcintosh PC, Wheeler MC, Hendon HH. On the remote drivers of rainfall variability in Australia. Monthly Weather Review. 2009;**137**:3233-3253

[25] Nayak DR et al. A survey on rainfall prediction using artificial neural network. International Journal of Computers and Applications. 2013;**72**(16):32-40

[26] Darji MP, Dabhi V, Harshadkumar BP. Rainfall forecasting using neural network: A survey. In: International Conference on Advances in Computer Engineering and Applications (ICACEA); 2015

[27] Dahamsheh A, Aksoy H. Markov chain-incorporated artificial neural network models for forecasting monthly precipitation in arid regions. Arabian Journal for Science and Engineering. 2014;**39**(4):2513-2524

[28] Cai W, van Rensch P. The 2011 southeast Queensland extreme summer rain: A confirmation of a negative Pacific Decadal Oscillation phase? Geophysical Research Letters. 2012;**39**:L08702

[29] Duan W, Zhang R. Is model parameter error related to a significant spring predictability barrier for El Nino events? Results from a theoretical model. Advances in Atmospheric Sciences. 2010;**27**(5):1003-1013

[30] Duan W, Wei C. The 'spring predictability barrier' for ENSO predictions and its possible mechanism: Results from a fully coupled model. International Journal of Climatology. 2013;**33**:1280-1292

[31] Izumo T, Vialard J, Lengaigne M, et al. Influence of the state of the Indian Ocean Dipole on the following year's El Niño. Nature Geoscience. 2010;**3**:168-172

[32] Marohasy J, Abbot J. Assessing the quality of eight different maximum temperature time series as input when using artificial neural networks to forecast monthly rainfall at Cape Otway, Australia. Atmospheric Research. 2015;**166**:141-149

[33] Coventry WL, Dalgleish LI. Farmers' accuracy interpreting seasonal climate forecast probability. International Journal of Climatology. 2014;**34**(6):2097-2107

[34] Coventry WL, Dalgleish LI. Differences between the single-event and frequency formats of seasonal-climate-forecast probability. International Journal of Climatology. 2015;**35**:444-451

[35] Bureau of Meteorology. Archive of Rainfall Forecasts. http://www.bom.gov.au/climate/ahead/archive/rainfall/20101123.national.hrweb.gif

[36] Marshall AG, Hudson D, Hendon HH, et al. Simulation and prediction of blocking in the Australian region and its influence on intra-seasonal rainfall in POAMA-2. Climate Dynamics. 2014;**42**:3271-3288

[37] Hudson D, Alves O, Hendon HH, et al. Bridging the gap between weather and seasonal forecasting: Intraseasonal forecasting for Australia. Quarterly Journal of the Royal Meteorological Society. 2011;**137**:673-689

[38] Schepen A, Wang QJ, Robertson DE. Seasonal forecasts of Australian rainfall through calibration and bridging of coupled GCM outputs. Monthly Weather Review. 2014; **142**(5):1758-1770

[39] Hawthorne S, Wang QJ, Schepen A, Robertson D. Effective use of general circulation model outputs for forecasting monthly rainfalls to long lead times. Water Resources Research. 2013;**49**:5427-5436

[40] Bagirov AM, Mahmood A, Barton A. Prediction of monthly rainfall in Victoria, Australia: Clusterwise linear regression approach. Atmospheric Research. 2017;**88**:20-29

[41] Montazerolghaem M, Vervoort W, Minasny B, et al. Spatiotemporal monthly rainfall forecasts for south-eastern and eastern Australia using climatic indices. Theoretical and Applied Climatology. 2016;**124**(3-4):1045-1063

[42] He X, Guan H, Qin J. A hybrid wavelet neural network model with mutual information and particle swarm optimization for forecasting monthly rainfall. Journal of Hydrology. 2015;**527**:88-100

[43] Wu CL, Chau KW. Prediction of rainfall time series using modular soft computing methods. Engineering Applications of Artificial Intelligence. 2013;**26**(3):997-1007

[44] Wei C-C. Soft computing techniques in ensemble precipitation nowcast. Applied Soft Computing. 2013;**13**(2):793-805

[45] Nastos PT, Paliatsos AG, Koukouletsos KV, et al. Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens, Greece. Atmospheric Research. 2014;**144**:141-150

[46] Philip NS, Joseph KB. A neural network tool for analyzing trends in rainfall. Computational Geosciences. 2003;**29**(2):215-223

[47] Cakir S, Kadioglu M, Cubukcu N. Multischeme ensemble forecasting of surface temperature using neural network over Turkey. Theoretical and Applied Climatology. 2013;**111**(3-4):703-711

[48] De SS, Chattopadhyay G, Bandyopadhyay B, Suman P. A neurocomputing approach to the forecasting of monthly maximum temperature over Kolkata, India using total ozone concentration as predictor. Comptes Rendus Geoscience. 2011;**343**(10):664-676

[49] Al-Shamisi MH, Assi AH, Hejase HAN. Artificial neural networks for predicting global solar radiation in Al Ain City – UAE. International Journal of Green Energy. 2013;**10**(5):443-456

[50] Chen SX, Gooi HB, Wang MQ. Solar radiation forecast based on fuzzy logic and neural networks. Renewable Energy. 2013;**60**:95-201

[51] Cardenas-Barrera JL, Meng J, Castillo-Guerra E, Chang L. A neural network approach to multi-step-ahead, short-term wind speed forecasting. In: 12th International Conference on Machine Learning and Applications (ICMLA). Vol. 2; 2013. pp. 243-248

[52] Fazelpour F, Tarashkar N, Rosen MA. Short-term wind speed forecasting using artificial neural networks for Tehran, Iran. International Journal of Energy and Environmental Engineering. 2016;**7**(4):377-390

[53] Ortiz-Garcia EG, Salcedo-Sanz S, Casanova-Mateo C. Accurate precipitation prediction with support vector classifiers. A study including novel predictive variables and observational data. Atmospheric Research. 2014;**139**:128-136

[54] Akrami SA, El-Shafie A, Jaafar O. Improving rainfall forecasting efficiency using modified adaptive neuro-fuzzy inference system (MANFIS). Water Resources Management. 2013;**27**(9):3507-3523

[55] Jeong C, Shin J-Y, Kim T, et al. Monthly precipitation forecasting with a neuro-fuzzy model. Water Resources Management. 2012;**26**(15):4467-4483

[56] Mekanik F, Imteaz MA, Talei A. Seasonal rainfall forecasting by adaptive network-based fuzzy inference system (ANFIS) using large scale climate signals. Climate Dynamics. 2016;**46**(9-10):3097-3111

[57] Hartmann H, Snow JA, Su B, et al. Seasonal predictions of precipitation in the Aksu-Tarim River basin for improved water resources management. Global and Planetary Change. 2016;**147**:86-96

[58] Luo J, Lu W, Ji Y, et al. A comparison of three prediction models for predicting monthly precipitation in Liaoyuan city, China. Water Science and Technology: Water Supply. 2016;**16**(3):845-854

[59] Ouyang Q, Lu W, Xin X, et al. Monthly rainfall forecasting using EEMD-SVR based on phase-space reconstruction. Water Resources Management. 2016;**30**(7):2311-2325

[60] Yuan F, Berndtsson R, Uvo CB, et al. Summer precipitation prediction in the source region of the Yellow River using climate indices. Hydrology Research. 2016;**47**(4):847-856

[61] Jim L, Zhu J, Huang Y, et al. A nonlinear statistical ensemble model for short-range rainfall prediction. Theoretical and Applied Climatology. 2015;**119**(3-4):791-807

[62] Lu W, Chu H, Zhang Z. Application of generalized regression neural network and support vector regression for monthly rainfall forecasting in western Jilin Province, China. Journal of Water Supply: Research and Technology - AQUA. 2015;**64**(1):95-104

[63] Yang Y, Luo Y. Using the back propagation neural network approach to bias correct TMPA data in the arid region of Northwest China. Journal of Hydrometeorology. 2014;**15**(1):459-473

[64] Kurian N, Venugopal T, Singh J, et al. A soft-computing ensemble approach (SEA) to forecast Indian summer monsoon rainfall. Meteorological Applications. 2017;**24**(2):308-314

[65] Chaudhuri S, Das D, Goswami S, et al. Long-range forecast of all India summer monsoon rainfall using adaptive neuro-fuzzy inference system: Skill comparison with CFSv2 model simulation and real-time forecast for the year 2015. Climate Dynamics. 2016;**47**(9-10):3319-3333

[66] Tyagi N, Kumar A. Comparative analysis of Backpropagation and RBF Neural Network on Monthly Rainfall Prediction 2016 International Conference on Inventive Computation Technologies (ICICT), Vol. 1; 2016. pp. 314-319

[67] Manek AH, Singh PK. Comparative study of neural network architectures for rainfall prediction. In: 2016 IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development (TIAR); 2016. pp. 171-174

[68] Rani BK, Srinivas K, Govardhan A. Rainfall prediction with TLBO optimized ANN. Journal of Scientific and Industrial Research. 2014;**73**(10):643-647

[69] Goyal MK. Monthly rainfall prediction using wavelet regression and neural network: An analysis of 1901-2002 data, Assam, India. Theoretical and Applied Climatology. 2014; **118**(1-2):25-34

[70] Acharya N, Shrivastava NA, Panigrahi BK, et al. Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: An application of extreme learning machine. Climate Dynamics. 2014;**43**(5-6):1303-1310

[71] Beheshti Z, Firouzi M, Shamsuddin SM, et al. A new rainfall forecasting model using the CAPSO algorithm and an artificial neural network. Neural Computing and Applications. 2016;**27**(8):2551-2565

[72] Purnomo HD, Hartomo KD, Prasetyo SYJ. Artificial neural network for monthly rainfall rate prediction. IOP Conference Series-Materials Science and Engineering. 2017;**180**(1):012057

[73] Mislan H, Hardwinarto S, et al. Rainfall monthly prediction based on artificial neural network: A case study in Tenggarong Station, East Kalimantan – Indonesia. International Conference on Computer Science and Computational Intelligence (ICCSCI 2015). Procedia Computer Science. 2015;**59**:142-151

[74] Babel MS, Sirisena TAJG, Singhrattna N. Incorporating large-scale atmospheric variables in long-term seasonal rainfall forecasting using artificial neural networks: An application to the Ping Basin in Thailand. Hydrology Research. 2017;**48**(3):867-882

[75] Kajornrit J, Wong KW, Fung CC, et al. An integrated intelligent technique for monthly rainfall time series prediction. In: 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). Book Series: IEEE International Fuzzy Systems Conference Proceedings; 2014. pp. 1632-1639

[76] Khalili N, Khodashenas SR, Davary K, et al. Prediction of rainfall using artificial neural networks for synoptic station of Mashhad: A case study. Arabian Journal of Geosciences. 2016;**9**(13):624

[77] Amiri MA, Amerian Y, Mesgari MS. Spatial and temporal monthly precipitation forecasting using wavelet transform and neural networks, Qara-Qum catchment, Iran. Arabian Journal of Geosciences. 2016;**9**(5):421

[78] Elsanabary MH, Gan TY. Wavelet analysis of seasonal rainfall variability of the upper Blue Nile basin, its teleconnection to global sea surface temperature, and its forecasting by an artificial neural network. Monthly Weather Review. 2014;**142**(5):1771-1791

[79] Badr HS, Zaitchik BF, Guikema SD. Application of statistical models to the prediction of seasonal rainfall anomalies over the Sahel. Journal of Applied Meteorology and Climatology. 2014;**53**(3):614-636

[80] Le JA, El-Askary HM, Allali M, et al. Application of recurrent neural networks for drought projections in California. Atmospheric Research. 2017;**188**:100-106

[81] Shao Q, Li M. An improved statistical analogue downscaling procedure for seasonal precipitation forecast. Stochastic Environmental Research and Risk Assessment. 2013; **27**:819-830

[82] Vaze J, Teng J, Chiew FHS. Assessment of GCM simulations of annual and seasonal rainfall and daily rainfall distribution across south-east Australia. Hydrological Processes. 2011;**25**:1486-1497

[83] Langford S, Hendon HH. Improving reliability of coupled model forecasts of Australian seasonal rainfall. Monthly Weather Review. 2013;**141**:728-741

[84] Abbot J, Marohasy J. The application of machine learning for evaluating anthropogenic versus natural climate change. GeoResJ. 2017;**14**:36-46