Skilful rainfall forecasts from artificial neural networks with long duration series and single-month optimization

John Abbot\textsuperscript{a,b,c,*}, Jennifer Marohasy\textsuperscript{a,b}

\textsuperscript{a} Climate Modelling Laboratory, Noosa, Queensland, Australia
\textsuperscript{b} Institute of Public Affairs, Melbourne, Victoria, Australia
\textsuperscript{c} James Cook University, Townsville, Queensland, Australia

A B S T R A C T

General circulation models, which forecast by first modelling actual conditions in the atmosphere and ocean, are used extensively for monthly rainfall forecasting. We show how more skilful monthly and seasonal rainfall forecasts can be achieved through the mining of historical climate data using artificial neural networks (ANNs). This technique is demonstrated for two agricultural regions of Australia: the wheat belt of Western Australia and the sugar growing region of coastal Queensland. The most skilful monthly rainfall forecasts measured in terms of Ideal Point Error (IPE), and a score relative to climatology, are consistently achieved through the use of ANNs optimized for each month individually, and also by choosing to input longer historical series of climate indices. Using the longer series restricts the number of climate indices that can be used.

1. Introduction

Two very different approaches exist for medium-term rainfall forecasting: general circulation modelling (GCMs) (Hawthorne et al., 2013; Scheppen and Wang, 2014) and statistical modelling (Fawcett and Stone, 2010) – including with artificial neural networks (ANNs). Statistical models typically use a set of lagged-input variables, and of particular importance are climate indices (Risbey et al., 2009). These are provided as input, together with historical temperature and rainfall data to compute a desired output - in this study monthly rainfall at 3 and 12 month lead-times for two important agricultural regions of Australia with very different rainfall patterns.

Wheat is a major crop in Western Australia, with 11–13 million ha planted annually, approximately 40% of this in south-west Western Australia. Aside from soil fertility and other agronomic considerations, the major constraints on production are meteorological and climatological (Anderson et al., 2005; Anderson, 2010; Sharma et al., 2008; Zhang et al., 2010). Soil moisture profile at the time of planting influences crop success (French and Schultz, 1984), which in turn is influenced by rainfall during the summer-autumn fallow period. Water availability during the growing period is the most critical factor affecting crop yields (French and Schultz, 1984). The timing of the arrival of adequate autumn rains for sowing is a key factor in the crop establishment phase (Pook et al., 2009). In particular, the timing of rainfall events in relation to crop requirements is regarded as more important than the total rainfall received during the crop life-cycle (Pook et al., 2012, 2009).

Sugarcane is an important agricultural crop on the east coast of Australia, cultivated in a coastal region extending from northern Queensland into northern New South Wales. The timing and amount of rainfall is critical in determining both the yield of sugar, and scheduling of harvesting operations (Skocaj and Everingham, 2014; Du et al., 2010; Valade et al., 2014). In Australia, water stress is estimated to cost the sugar industry AUS$260 million per annum in lost production (Inman-Bamber et al., 2012). Skocaj and Everingham (2014) investigated the impact of variables on sugar yields for the Tully region in the wet tropics of Queensland, reporting that spring and summer rainfall were the most important determinants. The optimal time of cane harvesting to achieve maximum yield is influenced by the rainfall profile during the growing season, and accurate prior knowledge of rainfall patterns would be beneficial in harvesting decisions (Jiao et al., 2005; Wood et al., 2005).

Prolonged rainfall over large areas of Queensland led to extreme flooding in December 2010, extending into January 2011 (Queensland Flood Commission of Enquiry, 2012). This not only affected the sugar industry, but also Brisbane, the state capital of Queensland which is in the south east of the state, not far from the sugar-town of Maryborough.

Currently seasonal rainfall forecasts issued by the Australian Bureau of Meteorology (BOM) use output from a GCM, specifically the Predictive Ocean Atmosphere Model for Australia, known as POAMA...
This model attempts to simulate actual atmospheric and oceanic conditions, and then forecast rainfall for specific grid areas in these agricultural regions. These BOM forecasts are issued in the form of probabilities relative to the median seasonal rainfall, and do not differentiate between anticipated rainfall slightly above the median and an extreme rainfall event as occurred in Queensland during the period December 2010 and January 2011 (Abbot and Marohasy, 2015a). Not only is the information content of probabilistic forecasts less than corresponding deterministic forecasts, where actual numerical values of predicted rainfall are provided, farmers often have difficulty interpreting the forecasts (Coventry and Dalgleish, 2014).

In the Western Australian wheat belt, the BOM’s seasonal forecasts are also considered too unreliable as a basis for major cropping decisions (Grains RDC, 2014; Petersen and Fraser, 2001). In regard to Australian sugarcane cultivation, Kingston (2011) concluded that currently available climate forecasting methods can identify high probability of above average wet season rainfall, but reliability is not good for spring months which are crucial for the sugar harvest.

Before adopting the GCM POAMA for forecasting, that is prior to May 2013, the BOM issued seasonal rainfall forecasts generated from relatively simple statistical models. These used two principle input values for the entire continent: two lagged sea surface temperatures, SST1 and SST2 (Fawcett and Stone, 2010). The predictions were made using the technique of linear discriminant analysis.

In reviewing the performance of this statistical model, Fawcett and Stone (2010) described the skill level demonstrated for seasonal rainfall forecasting as “only moderate, better than climatology or randomly guessed forecasts”. There is no evidence, however, to suggest that current forecast from the GCM POAMA are superior to the earlier forecast based on the sea surface temperatures.

Over the last two decades, artificial intelligence (AI) methods, particularly artificial neural networks (ANNs), have been applied to rainfall forecasting in many parts of the world (Darji et al., 2015; Wu and Chau, 2013; Nayak et al., 2013). Other AI methods applied to rainfall forecasting include adaptive neuro-fuzzy inference systems (ANFIS) (Akrami et al., 2014; Jeong et al., 2012; Akrami et al., 2013; Chaudhuri et al., 2016; Kajornrit et al., 2014; Mekanik et al., 2016; Hashim et al., 2016; Nhita and Adiwijaya, 2013; Srivastava et al., 2010; de Castro et al., 2011; Asklayn et al., 2011; Kajornrit et al., 2012; Nhita et al., 2015) and support vector machines (SVM) (Ortiz-Garcia et al., 2014; Feng et al., 2015; Hasan et al., 2015; Sanchez-Monedero et al., 2014).

ANNs have been applied to rainfall forecasting over a wide range of forecast periods from hourly (Wei, 2013), daily (Nastos et al., 2014; Altunkaynak and Nigussie, 2015; Partal et al., 2015; Namitha et al., 2015; Nagahamulla et al., 2012; Sang et al., 2013), weekly, monthly and seasonal (Table 1), through to annual (Philip and Joseph, 2003; Goswami and Srividya, 1996).

An historic survey of published rainfall forecasting using ANNs undertaken by Darji et al. (2015) identified 25 studies dating back to 1996. Most of the investigations were from India, but this survey also includes one study from Australia (Abbot and Marohasy, 2012). In Table 1 we list studies since the Darji et al. (2015) review. Our list (Table 1) includes 16 studies from Australia, 10 from China, and 8 from India. Studies from other parts of Asia include Indonesia, Japan, Thailand, and Malaysia. Studies have also been reported from the middle-east (Iran and Jordan) and Africa, as shown in Table 1. None of these studies apply the novel single-month optimization technique detailed in this study, and most use relatively short data sets.

ANNs have never been used to produce operational forecasts by the Australian BOM. Various studies from many parts of the world, however, suggest the technique has relevance if the objective is skillful monthly and seasonal rainfall forecasts (Acharya et al., 2012; Dahamshaha and Aksoy, 2009; Shukla et al., 2011) – including for Australia (Abbot and Marohasy, 2012, 2014, 2015a, 2016a, 2016b, 2017; Deo and Sahin, 2015; Mekanik et al., 2013).

There is a specialist climate science literature describing the climate indices and how they affect Australian rainfall (Cai et al., 2001). A strong concurrent relationship between rainfall and a particular input variable does not necessarily translate into a strong lagged relationship (Scheppen et al., 2012). It is lagged relationships that are essential for forecasting (Kirono et al., 2010). Chiow et al. (1998) examined linear correlations between rainfall and key climate indices for the Queensland region, including lagged relationships. The strongest linear correlations were found for spring rainfall – which is the period of most interest to the Queensland sugar industry. The geographical coverage and intensity of the strongest correlations declined as lags were progressively increased from 0 to 1, 2 and 3 months. Correlations with summer rainfall were less expansive geographically, and diminished with lag time.

In this study we used commercially-available artificial neural network software and conventional climate indices – indices that can be readily downloaded from research institute archives. Input data included 7 long-duration climate indices (based on temperature and pressure patterns) with monthly values available back to the late 1800s, and an additional 5 short-duration climate indices with monthly data available from at least 1960.

We used the 7-long duration series to generate: (i) forecast where all months in each year were optimized together, and also, (ii) a more elaborate time consuming technique that involved generating forecasts where each month was optimized individually. The single-month optimization technique gave superior results. We then used this technique on a shorter dataset that also included the 5 climate indices available from the 1960s.

2. Data and methods

Rainfall forecasts were made for three towns within the wheat belt of Western Australia and eight sites in coastal Queensland where sugar cane is cultivated; as shown by the maps in Figs. 1 and 2 respectively.

These towns and sites were chosen because they have the longest and most complete rainfall records – extending back at least 100 years. The locations and the first year of observations of each BOM rainfall record are shown in Table 2.

The Queensland locations have much higher annual rainfall that is more consistent across years. This is evident from a comparison of 10 years of rainfall as recorded at Victoria mill in Queensland relative to Southern Cross in Western Australia, as shown in Fig. 3.

The rainfall patterns for the two regions are distinct, as shown in Fig. 4. At coastal Queensland locations such as Plane Creek, there is heavy rainfall during the summer months, with low winter rainfall. While in the Western Australian wheat belt, illustrated by Southern Cross, the rainfall is more evenly distributed throughout the year with heaviest rainfall occurring during the winter months.

In addition to climate indices, local rainfall, and also maximum (MaxT) and minimum temperature (MinT) were input attributes. Each attribute considered was lagged up to 12 months. The data was divided into training (75%), evaluation (15%) and test sets (10%).

The seven long-duration climate indices used in this study listed are the Southern Oscillation Index (SOI), the Inter-Decadal Pacific Oscillation (IPO), four Nino indices (Nino 1.2, Nino 3, Nino 4, Nino 3.4), associated with the Pacific Ocean, and the Dipole Mode Index (DMI) associated with the Indian Ocean – as shown in Table 3.

In addition, five shorter-duration indices, also listed in Table 3, were inputted: south-eastern Indian Ocean index (SEIO), Western Indian Ocean Index (WIO), The Southern Annular Mode (SAM) and Quasi-Biannual Oscillation (QBO). Data were variously sourced from the Australian Bureau of Meteorology (BOM), Climate Explorer Website (CE) and the UK Met Office (UKMO), as shown in Table 3.

The above climate indices, together with local rainfall and local temperatures, were provided as input to Neurosolutions Infinity.
The GRNN has the following advantages:

- Single-pass learning so no backpropagation is required;

For the datasets used in this investigation, the optimal AI model automatically selected by Neurosolutions Infinity was a general regression neural network (GRNN) (Specht, 1991). The topology of this type of neural network is shown Fig. 5 as a feedforward network that can be used to estimate a vector Y from a measurement vector X. The network “learns” in one pass through the data and can generalise from examples as soon as they are stored. During training, the estimate converges to the conditional mean regression surfaces as more and more examples are observed. It forms very reasonable regression surfaces based on only a few samples, and the estimate is bounded by the minimum and maximum of the observations. The general regression neural network (GRNN) is similar in form to the probabilistic neural network (PNN). Whereas the PNN makes a single-pass learning process, the GRNN estimates values for continuous dependent variables.

The GRNN has the following advantages:

- Single-pass learning so no backpropagation is required;

For the datasets used in this investigation, the optimal AI model automatically selected by Neurosolutions Infinity was a general regression neural network (GRNN) (Specht, 1991). The topology of this type of neural network is shown Fig. 5 as a feedforward network that can be used to estimate a vector Y from a measurement vector X. The network "learns" in one pass through the data and can generalise from examples as soon as they are stored. During training, the estimate converges to the conditional mean regression surfaces as more and more examples are observed. It forms very reasonable regression surfaces based on only a few samples, and the estimate is bounded by the minimum and maximum of the observations. The general regression neural network (GRNN) is similar in form to the probabilistic neural network (PNN). Whereas the PNN makes a single-pass learning process, the GRNN estimates values for continuous dependent variables.

The GRNN has the following advantages:

- Single-pass learning so no backpropagation is required;
Each input attribute considered was lagged up to 12 months. The data were divided into training (75%), evaluation (15%) and test sets (10%). Two approaches were used for ANN optimization. With the first approach, designated as “all-month optimization”, data for all 12 months of the year was included as input and optimized together, as in our previous studies (Abbot and Marohasy, 2017). With the second approach, designated as “single-month optimization”, forecasts

### Table 2
Locations of weather stations used in this study.

<table>
<thead>
<tr>
<th>Location</th>
<th>State</th>
<th>Start Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bingera</td>
<td>QLD</td>
<td>1900</td>
</tr>
<tr>
<td>Fairymead</td>
<td>QLD</td>
<td>1881</td>
</tr>
<tr>
<td>Macknade</td>
<td>QLD</td>
<td>1890</td>
</tr>
<tr>
<td>Maryborough</td>
<td>QLD</td>
<td>1870</td>
</tr>
<tr>
<td>Mossman</td>
<td>QLD</td>
<td>1910</td>
</tr>
<tr>
<td>Plane Creek</td>
<td>QLD</td>
<td>1909</td>
</tr>
<tr>
<td>Victoria Mill</td>
<td>QLD</td>
<td>1909</td>
</tr>
<tr>
<td>Merredin</td>
<td>WA</td>
<td>1903</td>
</tr>
<tr>
<td>Narrogin</td>
<td>WA</td>
<td>1891</td>
</tr>
<tr>
<td>Southern Cross</td>
<td>WA</td>
<td>1889</td>
</tr>
</tbody>
</table>

### Fig. 1
Map of the Western Australian wheat belt showing the POAMA grid lines and the towns of Narrogin, Merredin and Southern Cross.

### Fig. 2
Locations of rainfall observation sites for coastal Queensland, and Brisbane, the capital of Queensland.

### Fig. 3
Monthly rainfall for Southern Cross and Victoria Mill over a 10-year period.

- high accuracy in the estimation since it uses Gaussian functions;
- it can handle noises in the inputs.
corresponding to each calendar month were performed individually, so that 12 optimizations were carried out to produce forecasts for the entire year.

Optimizations occur of both the neural network architecture and the associated inputs. In each treatment there were initially 120 lagged input variables available to relate to each required output rainfall value for long-duration time series (Table 2). In all cases, the fraction of these inputs retained by the Infinity software in the optimized model was a small subset of the initial input set made available, typically only 5% to 10%.

The concept of Ideal Point Error (IPE) has been applied to provide a statistical evaluation of the performance of the different models used in hydrological forecasting including river flows (Dawson et al., 2012) and rainfall (Malamos and Koutsoyiannis, 2016). The IPE is a dimensionless composite index that measures model performance with respect to an ideal point in an n-dimensional space (where n is the number of model attributes used).
performance evaluation metrics employed). IPE standardises a set of model performance evaluation statistics to an ideal point lying at \([0, 0, 0, \ldots, 0]\). The worst case is at \([1, 1, 1, \ldots, 1]\). The overall performance of a model in terms of IPE is measured as the Euclidian distance from that ideal point (i.e. smaller is better). If IPE is applied to a set of model outputs computed on the same data set, an IPE value of unity corresponds to the worst performing model; an IPE value of zero corresponds to a perfect (ideal) model. Eq. (1) was applied to the results generated

```
0.0
0.2
0.4
0.6
0.8
IPE

0.0
0.2
0.4
0.6
0.8
IPE

0.0
0.2
0.4
0.6
0.8
IPE
```

Fig. 7. Ideal point errors corresponding to optimizations for sites in coastal Queensland at 3 months lead time.

Fig. 8. Ideal point errors corresponding to optimizations for sites in coastal Queensland at 12 months lead time.

Fig. 9. Skill scores for monthly rainfall forecasts for Southern Cross at 12 months lead time for (i) all-month extended-time; (ii) single-month extended-time; and (iii) single-month extended-time.

Fig. 10. Skill scores for monthly rainfall forecasts for sites in the Western Australia wheat belt for single-month extended-time datasets.

Fig. 11. Observed and forecast monthly rainfall at 12 months lead time for Southern Cross corresponding to test period.

Fig. 12. Skill scores for monthly forecasts for Victoria Mill: (i) all-month extended-time; (ii) single-month extended-time; and (iii) single-month extended-time.
from the GRNN to produces IPE values after Malamos and Koutsoyiannis (2016).

\[
\text{IPE3} = 0.33((\text{RMSE max} - \text{RMSE}) + (\text{Min} R^2 - 1)^2
+ (\text{MAE max} - \text{MAE})^2)^{1/2}
\]  

(1)

Evaluations of forecasts were also made by application of skill scores calculated relative to climatology from Eq. (2). This is analogous to the method used by the BOM where a skill score is calculated for forecasts using POAMA (Hawthorne et al., 2013).

\[
\text{Skill score} = \frac{\text{RMSE (climatology)} - \text{RMSE (model)}}{\text{RMSE (climatology)}} \times 100%
\]  

(2)

Applying this equation, it follows that if the calculated values of 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 skill score worse than climatology.

3. Results

3.1. Statistical analysis using Ideal Point Error

Fig. 6 shows IPE values for rainfall forecasts at 12 month lead for sites in the wheat-belt of WA. Fig. 6 (upper) shows lower IPE values using single-month optimization when compared to all-month-optimization. Fig. 6 (lower) shows higher IPE values associated with the winter months rather that the summer months, and clear higher skill (better forecast) for single-month optimization.

Fig. 13. Skill scores for monthly rainfall forecasts for sites in coastal Queensland for single-month extended-time datasets at 12 months lead time and 3 months lead.

Fig. 7 shows IPE values for rainfall forecasts at 3 month lead for sites in coastal Queensland. Fig. 7 (upper) shows average IPE values for the 8 sites investigated in coastal Queensland progressing from the most northerly site (Mossman) to the most southerly site (Maryborough). No strong trend in IPE is evident, and the IPE values do not exhibit strong difference depending on whether single-month or all-month neural network optimization was used. Fig. 7 (lower) shows average IPE values for the sites in coastal Queensland corresponding to each individual

Fig. 14. Observed monthly rainfall for Bingera and forecast monthly rainfall at 3 months lead time for test period October 2004 to September 2014 for all-month extended-time, and single-month extended time.

Fig. 15. Observed monthly rainfall for Plane Creek and forecast monthly rainfall at 12 months lead time for test period October 2005 to September 2014 for all-month extended-time; and single-month extended-time.
forecast month. There is a clear annual trend evident, with lower values of IPE associated with the winter months and higher IPE values associated with summer months. Again, there is no consistent preference for single-month or all-month optimization in terms of lower IPE values.

Fig. 8 shows IPE values for rainfall forecasts at 12 month lead for sites in coastal Queensland. Fig. 8 (upper) shows average IPE values for the 8 sites investigated in coastal Queensland. Again there is no clear trend in IPE moving geographically from north to south along the Queensland coast. However, in contrast to 3 month lead shown in Fig. 8 (upper) there a clearer preference for single-month optimization when compared to all-month-optimization as IPE values are consistently lower. Fig. 8 (lower) shows lower values of IPE during the winter months, and clear preference for single-month optimization.

3.2. Comparisons based on skill scores

Comparisons of skill scores with respect to climatology were determined for the three methods of forecast determination. The single-month extended-time datasets produces superior results compared to either the all-month extended-time or single-month reduced-time input datasets. The extent of the difference is evident in Fig. 9, which compares skill scores for the three different treatments for Southern Cross in the Western Australia wheat belt.

Applying the single-month input dataset generates a positive skill score typically in the 40% to 60% range for the three wheat belt towns, as shown in Fig. 10. This relatively high level of skill is shown in Fig. 11, where observed and forecast rainfall is plotted over the test period for Southern Cross.

The single-month extended-time dataset also produced superior results for the sugarcane growing regions in Queensland. Fig. 12 shows rainfall forecast skill scores for Victoria Mill with a 12 month lead time for monthly rainfall for each month generated by the ANN corresponding to the three different input datasets. For all months the skill scores are positive and fall in the range 20% to 90%.

Fig. 13 shows results for all 8 Queensland sites, using lead times of 12 and 3 months. Figs. 14 and 15 show plots of observed and forecast rainfall over the test period for Bingera and Plane Creek, corresponding to all-month and single-month datasets at lead times of 12 and 3 months respectively. Clearly the single-month datasets give superior rainfall forecasts.

This ANN single-month optimization technique was also able to anticipate the extreme rainfall at Maryborough in December 2010. Fig. 16 shows the ANN forecasts for Maryborough, at 3-month lead time, for the months of December and January during the test period. It is evident that this forecast method is able to distinguish between the anticipated rainfall in December 2010 and January 2011 compared to the much lower rainfall experienced in the corresponding month the previous year.

4. Discussion

The ANNs, using single-month optimization and longer duration series, produce average skill scores between 35% and 65%, as shown in Fig. 17. These results, for monthly rainfall forecasting at lead times of 12 months, were consistent across the climatically diverse regions of the Western Australian wheat belt and sugarcane regions of Queensland. There was only a marginal improvement in the skill score with the reduction in lead time from 12 to 3 months.

Clarke et al. (2010) described a statistical method for seasonal rainfall forecasting at long lead times for the sugar cane growing region of north eastern Australia using the climate index Nino 3.4 as input. Results at 9 months lead time for Plane Creek and Tully for the three-month period September–October–November (SON) are shown expressed as skill scores relative to climatology in Table 5. For Plane Creek, this method (Clarke et al., 2010) gave a skill score of 2.6%, slightly above climatology. Our single-ANN method at 12 months lead time gave a skill score of 50% over the SON period. The skill score using Clarke's statistical model for Tully was 23.9% for the period SON. The closest site to Tully in the present study was Macknade, 90 km to the south, giving a skill score of 49% at 12 month lead, as shown in Table 5.

Hawthorne et al. (2013) used the output from the GCM 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, and far north Queensland. For lead times between 3, 4, 5, 6, 7 and 8 months, for both grid areas approximately 60% of the forecasts give skill levels relative to climatology below 0%. For south-east Queensland only about 20% of the forecasts had a skill level in the 15–20% range, with only about 25% of the monthly forecasts reaching this skill level for a grid location in far north Queensland (Hawthorne et al., 2013). Skill scores corresponding to the grid area to the WA wheat belt were on average about equivalent to climatology with most results in the range −10% to 10% at lead times of 6 to 8 months (Hawthorne et al., 2013).

Hawthorne et al. (2013) described the skill of their monthly rainfall forecasts as low, and concluded that forecasting with POAMA remained a challenge.

Comparing these skill scores with the results from our study suggests that ANNs can produce superior monthly and seasonal rainfall forecasts than other statistical techniques.

The skill of the ANN forecasts is improved with single-month optimization. This is not surprising given that different climate indices are known to exhibit strong lagged relationships with rainfall at different times during the year (Schepen et al., 2012). The most skilful forecasts are achieved when single-month optimization is combined with the longer duration series. This is evident when the results for individual locations are charted – with Victoria Mill in Queensland shown in Fig. 8, and Southern Cross in Western Australia in Fig. 5. This result suggests that there is more useful information for the ANNs embedded in the early historical climate data, than the additional climate indices.

GCMS are currently used for operational forecasting in Australia. The demonstrated skill of GCMS in predicting seasonal or monthly rainfall is low for many regions of the world including India, the United States and Australia (Zimmerman et al., 2016; Tiwari et al., 2014; Schepen and Wang, 2013) often with little enhancement in forecast skill above climatology. A number of different approaches have been investigated to improve the skill of GCMS in predicting rainfall, including using weighted multi-model ensembles (Acharya et al., 2014a, 2014b) and developing hybrid statistical models (Yang and Ke, 2012; Sun and Chen, 2012). Post-processing of the output of the GCM POAMA using statistical models has been reported in attempting to improve the skill of rainfall forecasts for Australia (Schepen and Wang, 2013).

An additional advantage in using ANNs – additional to the more skilful forecasts – is that it is practical to tailor medium-term forecasts to the requirements of particular end users. The BOM forecasts are currently only provided to the public in the form of probabilities relative to the median, and do not differentiate between anticipated rainfall slightly above the median and extreme rainfall events such as occurred in January 2011 (van den Honert and McAneney, 2011).

During the summer of 2010–11, major flooding occurred throughout most of the Brisbane River catchment, with an estimated 18,000 properties inundated in metropolitan Brisbane, and elsewhere in the Brisbane River Valley (van den Honert and McAneney, 2011). The event has been termed a “dam release flood”, suggesting that the sudden release of water from the Wivenhoe Dam was a principal cause of flooding.

The decision to suddenly release water resulted because the capacity of the Wivenhoe Dam did not allow for the extreme rainfall event, and the official medium-term rainfall forecasts available were inadequate, and did not provide sufficient warning of the impending heavy rainfall in December 2010 and January 2011. If the magnitude of the rainfall had been forecast at sufficient lead time, water stored in the Wivenhoe dam could have been released gradually over an extended period.

Results from this study, as demonstrated for the sugar regions of Maryborough, indicate that ANNs can forecast extreme flood events at a lead time of at least 3 months.

5. Conclusion

Until recently, official monthly and seasonal rainfall forecasts issued by the BOM were based on a simple statistical scheme using only two climate indices relating to sea surface temperatures in the Pacific and Indian oceans respectively. The BOM switched to the use of the General Circulation Model POAMA in June 2013. There is no evidence, however, that the GCMS can produce skill monthly rainfall forecasts. Despite substantial efforts to enhance skill over three decades, forecast issued by the BOM are of limited skill, and do not provide sufficiently skilful information at appropriate lead times.

There is, in fact, little evidence that GCMS produce results that offer much enhancement beyond climatology – the simple long term rainfall average – for anywhere in the world.

In contrast, this study builds on work summarized by Darji et al. (2015) that indicates, ANNs offer a more skilful alternative approach to monthly rainfall forecasting. We have demonstrated in the present investigation that monthly rainfall can be consistently forecast for different geographical regions in Australia, generating skill scores in the range 40% to 60%. The most skilful forecasts are achieved when the longest duration datasets are combined with single-month ANN optimization. Single-month optimization is a novel technique not applied in any of studies reviewed by Darji et al. (2015).

 Clearly this ANN technique has application for the two very different agricultural industries of sugar in Queensland, and wheat in Western Australia – operating in regions with very different rainfall profiles bordering the Pacific and Indian Oceans, respectively. Application of the technique is therefore likely to result in improved forecasts in other parts of the world. The technique is likely to have application for any location for which there are long historical temperature and rainfall series.

This single-month optimization technique using ANNs can not only potentially improve agricultural productivity, but also improved dam management, as it can potentially provide warning of extreme flooding.

Acknowledgements

This research was funded by the B. Macfie Family Foundation.

References


Feng, Q., Wen, X., Li, J., 2015. Wavelet analysis-support vector machine coupled models


benmahdjoub, K., Amour, Z., Boulifa, M., 2013. Forecasting of rainfall using time delay

du, J.H., et al., 2010. Analysis on correlation between sugar content and planting con-

Anderson, W.K., et al., 2005. The role of management in yield improvement of the wheat

Deo, R.C., Sahin, M., 2015. Application of the artificial neural network: a review with optional incorporation of an explanatory variable. Part 2: application to syn-


Dahamshba, A., Akosy, H., 2009. Artificial neural network models for forecasting in-

Darji, M.P., Dabhi, V., HarshaKumar, B.P., 2015. Rainfall forecasting using neural net-


Deo, R.C., Sahin, M., 2015. Application of the artificial neural network model for pred-

Diah, M., et al., 2010. Analysis on correlation between sugar content and planting con-


