A Genetic Algorithm based Feature Selection Approach for Rainfall Forecasting in Sugarcane Areas

Ali Haidar and Brijesh Verma
Center for Intelligent Systems
School of Engineering and Technology
Central Queensland University, Australia
a.haidar@cqu.edu.au, b.verma@cqu.edu.au

Abstract— **Rainfall is a vital phenomenon that contributes in the success of sugar industry season. The ability to determine the amount of precipitation in sugarcane areas enhances the profitability of the season. Different types of climate indices and attributes are usually applied to model rainfall forecasting systems. In this paper, we present a novel genetic algorithm based feature selection approach to determine which climate indices and attributes are most significant for rainfall forecasting in sugarcane areas. The most significant features are features that return the highest accuracy for rainfall forecasting through artificial neural networks. The approach is evaluated on real-world data that contain different weather forecasting features. A set that contains maximum temperature values and Southern Oscillation Index (SOI) has proven to be the best combination among the other models with a Root Mean Square Error (RMSE) of 0.027 in November. An Average RMSE of 0.0638 for the genetic algorithm based forecasts was recorded. The proposed model was compared to other models and the proposed model revealed higher accuracy in forecasting monthly rainfall.**

Keywords—Artificial Neural Networks; Genetic Algorithms; Rainfall Forecasting; Climate Indices; Feature Selection

I. INTRODUCTION

The information about rainfall over a specific area at a specific time can be beneficial for various types of industries. Accurate forecasts enhance decision making and management almost through all aspects in human life. Rainfall values are essential for different agricultural crops including sugarcane. Researchers have linked the success of sugarcane cropping season to the ability of forecasting rainfall precisely [1].

Different types of forecasting models have been developed and employed to predict precipitation for several durations. Various models have shown their applicability for rainfall forecasting across regions around the world and different climate features were used in weather forecasting problems. Appropriate feature data is essential for various forecasting systems to perform well. Different types of climate attributes and indices were applied in rainfall forecasting and sugarcane related studies. He et al. used Southern Oscillation Index (SOI), Pacific Decadal Oscillation (PDO), Southern Annular Mode (SAM) and Indian Ocean Dipole (IOD) to forecast South Australia monthly rainfall anomaly [2]. Deo and Sahin used 13 different attributes and a learning approach to predict monthly values for Effective Drought Index (EDI) in Eastern Australia. These attributes were categorized into site-specific and climate variables. The dataset composed of year, month, latitude, longitude elevation, monthly mean rainfall, monthly mean temperature, monthly maximum temperature, monthly mean air temperature, SOI, PDO, SAM and IOD [3]. Nasseri et al. used rainfall historical data to forecast hourly rainfall in Parramatta catchment, Sydney, Australia [4]. Climate attributes were applied not only for rainfall but also for sugarcane yields. Everingham et al. used SOI to forecast Sugarcane yields for Northern Australia [1]. Everingham et al. have also used SOI to generate a long lead rainfall prediction model that aids the sugarcane industry decisions in several locations across Eastern Australia [5]. The addition of these climate attributes into different studies was considered when setting up the dataset in this study.

Kishtawal et al. have utilized a genetic algorithm to forecast rainfall for India [6]. Past rainfall values of three months (June, July and August) were gathered to conduct the experiment. Authors aimed to find the optimal equation that represent the temporal variations of seasonal rainfall in India. An analytical expression was generated by the Genetic Algorithm (GA) and then used to perform forecasting. 132 years of rainfall were used in the study. 122 years were applied to find the equation that best fit data with GA and the remaining 10 years to validate. To measure the accuracy of the model, standard error and fitness strength were calculated. The fitness of the final equation was reported to be 0.644 for the testing set (1993-2003), while RMSE was 28.6 mm for the period 1943-2003.

Artificial neural networks were applied with genetic algorithms in weather forecasting problems. Meng conducted an experiment to optimize backpropagation neural network weights using a genetic algorithm [7]. Author aimed to combine the advantages of the two machine learning algorithms. Daily temperature values were targeted in the study. 246 samples were used for training, and 31 samples for testing. Results showed an error of 0.001 with the improved genetic neural network.

Nagahamulla et al. presented a genetic algorithm and k-means clustering algorithm to select the most suitable artificial

neural networks to form an ensemble in two locations in Srilanka: Colombo and Katugastota [8]. To generate the ensemble members, authors developed a pool that contains General Regression Neural Network (GRNN) topologies that are trained with varied data. These training datasets were achieved by different pre-processing and data sampling techniques. The use of GA was to find the best ANN models in the pool so that the total Mean Square Error (MSE) of the results is reduced. A GRNN was used to combine the selected ANN members of the ensemble for each method (k-means clustering and GA). To measure the accuracy, RMSE and Mean Absolute Error (MAE) were recorded. The collected data set contained 41 years of daily observed data for 26 features. 25 years were used for training, 8 for validation and remaining 8 for testing. The pool consisted of 1023 GRNN trained with different datasets. The two proposed methods were compared to bagging and boosting. Best ensembles in the two locations with each model were compared where GA based ensemble revealed the lowest RMSE in both Colombo and Katugastotata (7.33, 6.25) followed by k-means clustering method with 7.38 and 6.37 respectively.

In this paper, we investigate a novel model for forecasting monthly rainfall which is based on a genetic algorithm and an artificial neural network. A genetic algorithm is deployed to select the best input features for generating outlooks, while a neural network is created to assess the features and to produce the forecasts. The remainder of this paper consists of three sections. Section II contains a description of the developed model. Section III shows experimental setup, results and discussion. Section IV presents conclusion and future research directions.

II. Proposed Feature Selection Model

The aim of this study is to find the best subset for generating monthly rainfall forecasts for each month in sugarcane areas. The proposed model is shown in Fig. 1

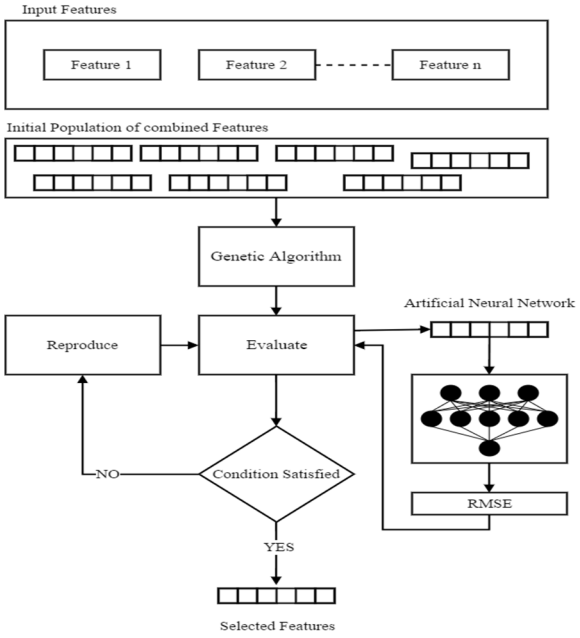


Fig. 1. Proposed Model

TABLE I. Weather Features

Number	*Feature*
1	Mean Maximum Temperature(MaxT)
2	Mean Minimum Temperature(MinT)
3	SOI
4	NINO 1.2
5	NINO 3.0
6	NINO 3.4
7	NINO 4.0
8	DMI
9	IPO
10	SUNSPOTS

A. Study Area

The study area was selected to be Bingera, a town located in Queensland, Australia. Bingera has an annual rainfall average of 1023 mms. Multiple weather stations are located in the area.

B. Data Collection (Input Features)

Local and global climatic attributes were collected to setup the dataset for Bingera. Mean Monthly Rainfall, Mean Minimum Temperature, Mean Maximum Temperature, and Southern Oscillation Index (SOI) values were gathered from the Australian Bureau of Meteorology (BOM). Nino 1.2, Nino 3.0, Nino 3.4, Nino 4.0, Dipole Mode Index (DMI) and sunspots were collected from KNMI climate explorer, a web application that contains various types of global weather attributes. Finally, Interdecadal Pacific Oscillation (IPO) was collected from Climate of 20th Century(C20C) project website. Rainfall was selected as target in the study. In total, around 115 years were used in the study. The 10 input features used to setup the study are shown in Table I.

C. Genetic Algorithm

A genetic algorithm is designed and incorporated in the proposed model to select the best features from a given dataset that would enhance the performance of the forecasting model. Feature selection is the mechanism of identifying a subset from the whole data set that generate the best optimal solution [9]. Feature selection is a mapping of a set of features into a smaller set that would reveal higher results for a specific problem. An initial population is supplied to the genetic algorithm, then population members are subjected to evolutionary processes. These processes are: selection, crossover and mutation. Each genetic algorithm has stopping criteria which is specified based on number of generations, no improvement has been recorded for a duration of time or desired solution obtained. Reproduction of off-springs continues until reaching the stopping criteria. Different types of genetic algorithms have been applied in different applications for classification and feature selection. The genetic algorithm for proposed model is represented in Fig. 2. It has been added into the study to investigate its ability in selecting the optimal combination of

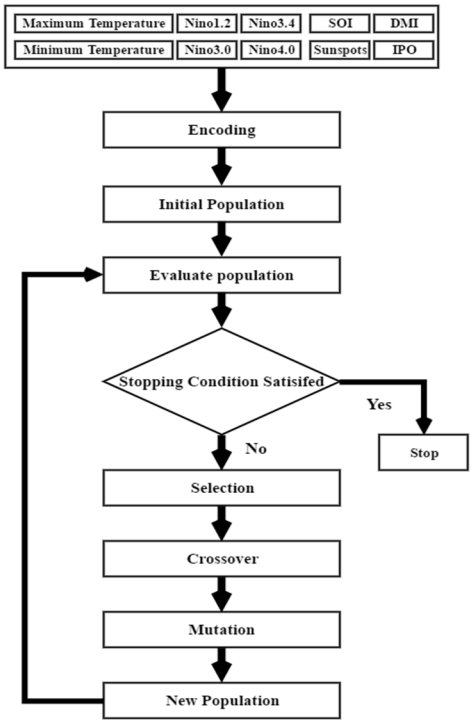


Fig. 2. Genetic Algorithm

climate attributes to ensure highest accuracy in forecasting monthly rainfall.

D. Artificial Neural Networks

A feed forward neural network is combined with genetic algorithm and incorporated in the proposed model. It is formed of interconnected processing elements called neurons that process information by learning. The interconnected nodes are organised into layers. Three types of layers are in the structure of ANN: input layer, hidden layer and output layer. The feed-forward architecture of a neural network is shown in Fig. 3. The input layer receives features, perform calculations and send results to hidden layer where information manipulation is held. The hidden layer processes the information and sends to output layer that returns the network results. ANN learns to generalize data through the training process, where weights and connections between layers are modified to obtain desired values [10]. With weather forecasting problems, ANN minimizes the error between actual and predicted values to increase the performance. Particular learning algorithms that hold different mechanisms are usually used with ANN. ANN has been widely applied to forecast different climate attributes as rainfall temperature and relative humidity. In addition, it has been formulated to forecast for various times: hourly, daily, weekly, monthly, etc. Neural network can be taught the dynamics of the system which would lead to improvements in overall approximation accuracy of outlooks [11]. Output of neural network can be completely different when changing a small part in its parameters [8]. Data is essential for several forecasting models including ANN. Discrete weather attributes are usually collected in an attempt to setup a model for

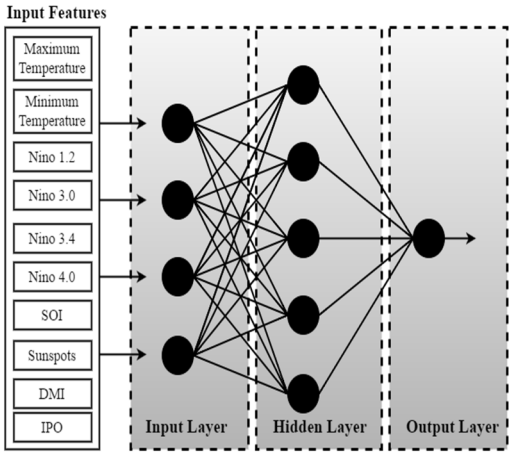


Fig. 3. Feedforward Neural Network Structure

prediction using artificial neural networks. Climate indices and attributes represent specific situation on land or in oceans. The formation of rainfall in a specific location can be related to climatic events over different parts of the globe. With the vast spread of technology through the previous century, the ability to record and save climate variables became much easier. This abundance of data revealed a new challenging task which is the selection of climate features that highly incorporate in generating accurate conjectures over a specific location. To specify the climate attributes to be added within a model, close geographic indices to targeted area can be chosen. But, a variation at one corner of the globe may produce tornado to another place geographically far away (butterfly effect)[11]. Considering this issue, there is still a need to look at global climate indices. Increasing the dataset size would definitely rise processing time. Typically, neural network performs better with larger dataset, but appending some features may return lower performance. The high dimensionality in data may affect the performance [9]. Hence, climate attributes should be chosen carefully to ensure accurate rainfall forecasts. To select the optimal input data (e.g. features) that would reveal highest accuracy, trial and error based manual method can be used. In trial and error, different input data are formed based on user preferences and added to the network. In addition to, diverse computational techniques could be utilized to select the optimal subset as Particle Swarm Optimization (PSO), K-mean clustering, Principle Component Analysis (PCA) and Genetic Algorithms (GA). We propose the genetic algorithm to search for the optimal subset.

E. Model Development

GA models natural evolution. Its population consists of a number of elements called chromosomes. A chromosome is composed of genes (climate features) that represent a possible solution for the problem GA is trying to solve. Each month has its own predicators, therefore for each month, the same genetic algorithms was deployed. The following steps were followed to setup the proposed GA:

1) Encoding

Encoding is a mechanism where intended features are mapped into chromosomes. 10 climate features shown in Table I were collected to setup the experiment. Chromosome type was selected as binary where each possible gene can has a value of 0 or 1 only. If the gene is included in the chromosome 1 will be found in its index, otherwise 0. Since there is 10 features, chromosomes consisted of 10 genes (each gene 1 index). The chromosome highly depends on the sequence of features being selected. The index of each climate feature is shown in Fig.4.

2) Initial Population

To ensure multiple combinations are formulated through generations, the initial population should contain some chromosomes that are formed of 2 or more genes. Hence, random initial population was generated.

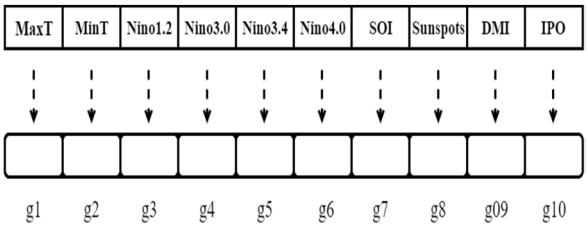


Fig. 4. Encoding of Climate Features into Binary Chromosome.

3) Evolution of Chromosomes

a) Selection

Selection is the process of choosing individuals that is used in reproduction. In this study, it has been set to 4. Four chromosomes out of the generated population are used for generating the next population.

b) Crossover

Crossover is an operator for the GA where performance highly depends on. Two chromosomes are combined to generate a new off spring that would reveal better results. Through crossover, diverse combinations can be created which will assist finding the best solution.

c) Mutation

Mutation includes random transforming of values of genes in the chromosome. The combination of different genes may generate better forecasts. Mutation probability refers to the chance where a gene in the chromosome is to be flipped. Through this technique, removing a feature from the dataset may enhance the performance (transforming 1 into 0). On the other hand, adding a new feature to a given dataset may generate a better chromosome (transforming 0 into 1).

d) Fitness Function

Fitness function is used in each generation by the genetic algorithm to evaluate the performance of the chromosomes. It returns a scalar value that determines the effectiveness of the chromosome (collection of features). A fitness function consisting of a Feed Forward Neural Network (FFNN) with three layers was proposed. Based on the size of the dataset, the number of hidden neurons in the hidden layers was selected. Number of neurons should be proportional to data size to avoid low accuracy. Gradient descent based backpropagation algorithm was considered and used as the training algorithm for the proposed network. Tansig transfer functions were used between input to hidden and hidden to output layers. To train and validate the model, a dataset that ranges between 01/1901 and 09/2004 was utilized. 85 % of this dataset were partitioned for training and 15 % for validation. A new dataset composing of 10 values for the latest 10 years of each month were used for testing.

To measure the fitness of each chromosome, Root Mean Square Error (RMSE) was calculated for the testing dataset. RMSE mathematical equation (equation 1) is shown below:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - x_i')^2}{N}} \quad (1)$$

Where i represents a value in the testing set, x_i is the actual value and x_i' is the predicted value. N is the total number of values in a dataset.

RMSE is a positive number, the closer the number to zero the better the performance. In other way, the closer the number to zero the better the forecasts are. Fitness function return value was selected as RMSE measurement of the developed neural network. The aim of the study was to get the best features for forecasting rainfall. Setting the initial weights randomly each time fitness function would have generated different results for the best chromosome. Therefore, neural network connection weights and bias connections were generated from same set of random values.

III. Experiments and Results

Climate features varied based on length. Each feature was manipulated to ensure the combined dataset comprises same duration for all the features. Missing values were replaced with alternatives from nearby weather stations. Monthly rainfall values were selected as target while the remaining 10 features were used as predicators. For sugarcane areas, monthly rainfalls vary between different seasons and months. Rainfall ranges fluctuate between those months. Therefore, each month may have different climate attributes that affect formation of rain. SOI which is a climate attribute (gene) may be paramount

TABLE II. Genetic Algorithm Parameters.

Parameter	***Value***
Population	50
Population Type	Bit String
Generations	20
Crossover	Crossover Two Points
Mutation	Uniform Mutation
Fitness function	FFNN
Mutation Probability	0.3
Selection	Selection tournament (size 4)

for forecasting January rainfall, but not July. Therefore, we divided the dataset into 12 months to find the optimal subset for each month. 10 climate features were used to predict rainfall. In total, 1024 distinct datasets can be generated for forecasting. To find up the best combination for each month, the proposed approach was deployed.

As discussed in Section II, various steps are followed to setup the GA. The initial population contained 50 chromosomes (4.88%) that were randomly created, ensuring that all genes are not zero. Number of iterations was set to 20. Hence, a maximum value of 1000 networks could be tested (not including the first 50 for fitness). We aimed to increase the initial population so that the algorithm can start identifying the best genes in earlier generations and to avoid local minimum. Trapping in local minimum could be encountered when the algorithm begin to check closely similar features. Two points crossover where two chromosomes are combined by selecting genes from the second parent and adding them to the same location (index) for the parents was used. Uniform Mutation was selected with a probability of 0.3. The stopping criterion was selected to be number of generations (20). The fitness function was selected to be a FFNN. 13 neurons were added to the hidden layer. The number of epochs for the network was 500. Table II summarizes the genetic algorithm specifications and Table III summarizes the neural network specifications.

Neural network can learn better when having normalised values. The range of the climate attributes was totally different. For each feature, the upper and lower bounds were calculated. Then the difference between each value and the minimum was divided by the difference between upper and lower boundaries. All the features were normalised to a range between 0 and 1.

Each month dataset was used by the GA to find the best subset of the whole dataset. The best selected subset for each month is shown in Table IV. Fitness value and mean value for each genetic algorithm optimization through generations are shown in Fig. 5

TABLE III. NEURAL NETWORK SPECIFICATION.

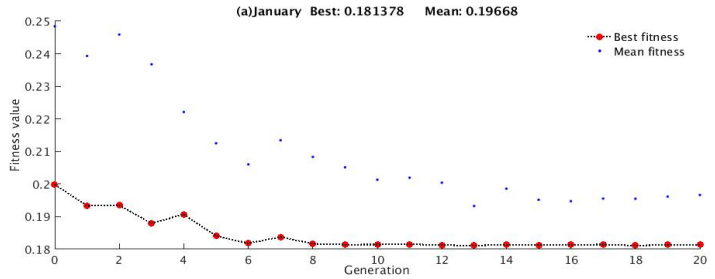
Attribute	*Value*
Type	Feed Forward Neural Network
Layers	Three layers
Neurons in hidden layer	13
Transfer functions	Tansig
Training Algorithm	RMSE
Epochs	500
Training ratio	85 %
Validation ratio	15 %

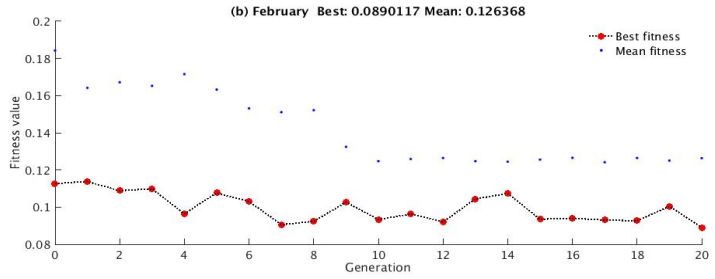
Based on the results, February used nine features which was recorded as the largest subset after optimization. The range of errors varied between 0.027 and 0.181. The average error for each month is shown in the table IV. Maximum temperature feature, Nino 3.0 and SOI were selected in 7 out of the 12 months (Table V). Nino 3.4 and IPO were used in 3 out of the 12 subsets. Some months need no time to capture the best solution while in other months, multiple iterations were required to enhance the optimization. These months hold low average monthly values compared to others. In February the GA fluctuated through most of the iterations while in July and August results were found from early iterations. Using 50 individuals in the initial population allowed the network to optimize earlier.

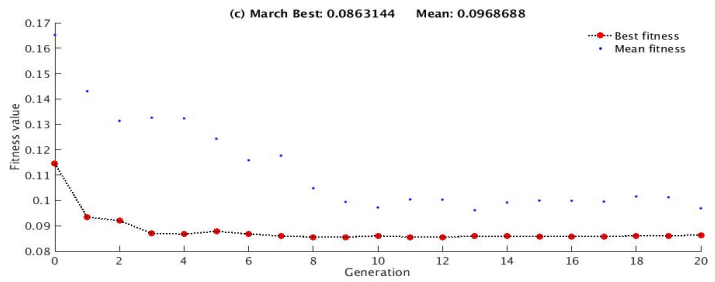
To compare the proposed GA based forecasting model, another 12 neural networks that hold the same specifications the network in fitness function held were developed. A FFNN with three layers having 13 neurons in the hidden layer was designed.

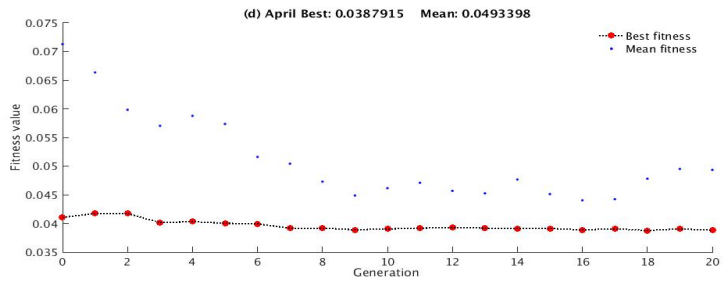
TABLE IV. RESULTS USING PROPOSED APPROACH

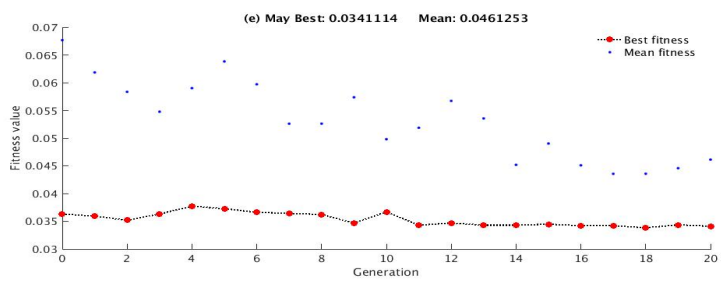
Month	*Selected subset*	*Decoded Subset*	*RMSE(mms)*	*Number of Attributes*
January	[4,7]	[Nino3.0, SOI]	0.1813	2
February	[1,2,3,4,5,6,8,9,10]	[MaxT, MinT, Nino1.2, Nino3.0, Nino3.4, Nino4.0, Sunspots, DMI, IPO]	0.0890	9
March	[3,4,8,9]	[Nino1.2, Nino3.0, Sunspots, DMI]	0.0863	4
April	[1,2,5]	[MaxT, MinT, Nino4.0]	0.0387	3
May	[1,4,8,9]	[MaxT, Nino3.0, Sunspots, DMI]	0.0341	4
June	[1,6,7]	[MaxT, Nino4.0, SOI]	0.0459	3
July	[1,4,6]	[MaxT, Nino3.0, Nino4.0]	0.0300	3
August	[2,3,5,6,7,8,9,10]	[MinT, Nino1.2, Nino3.4, Nino4, SOI, Sunspots, DMI, IPO]	0.0344	8
September	[6,7,8]	[Nino3.4, Nino4.0, SOI]	0.0346	3
October	[2,4,7,8,10]	[MinT, Nino3.0, SOI, Sunspots, IPO]	0.0342	5
November	[1,7]	[MaxT, SOI]	0.0270	2
December	[1,2,3,4,7]	[MaxT, Mint, Nino1.2, Nino3.0, SOI]	0.1311	5

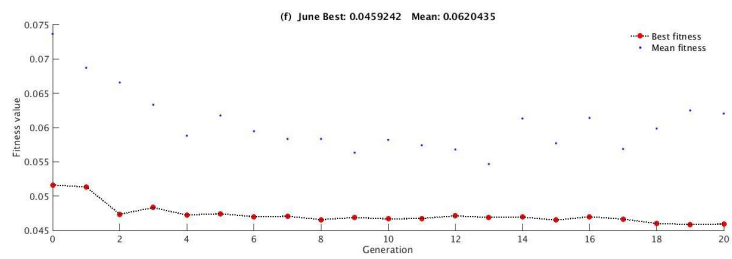

(a)January Best: 0.181378 Mean: 0.19668
Best fitness
Mean fitness
Fitness value
Generation


(b) February Best: 0.0890117 Mean: 0.126368
Best fitness
Mean fitness
Fitness value
Generation


(c) March Best: 0.0863144 Mean: 0.0968688
Best fitness
Mean fitness
Fitness value
Generation


(d) April Best: 0.0387915 Mean: 0.0493398
Best fitness
Mean fitness
Fitness value
Generation


(e) May Best: 0.0341114 Mean: 0.0461253
Best fitness
Mean fitness
Fitness value
Generation


(f) June Best: 0.0459242 Mean: 0.0620435
Best fitness
Mean fitness
Fitness value
Generation

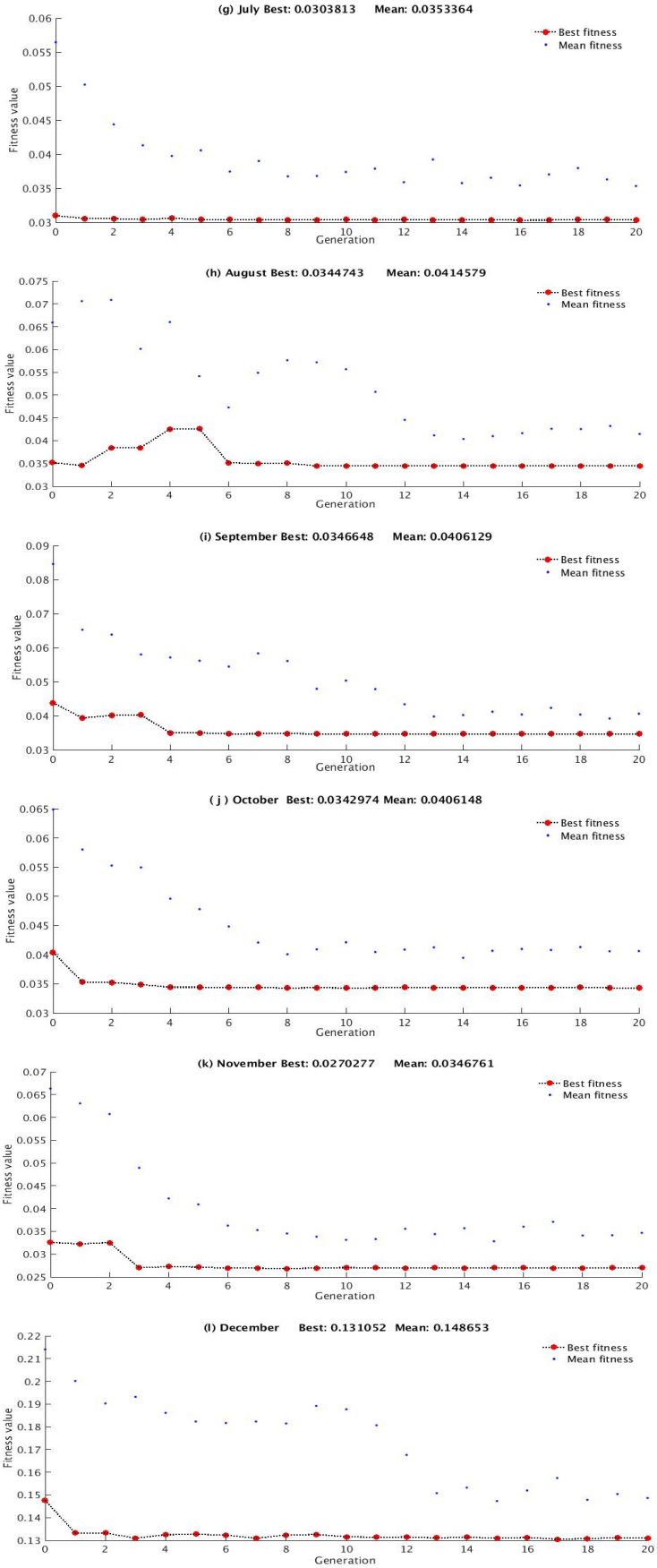


Fig. 5. Genetic algorithm for each of the twelve months: (a) January, (b) February, (c) March, (d) April, (e) May,(f) June, (g) July,(h) August, (i) September, (j) October, (k) November, (L) December.

TABLE V. NUMBER OF TIMES EACH FEATURE WAS IN BEST SUBSET.

Feature	*Gene index in chromosome*	*Number of times used in selected subsets*
MaxT	1	7
MinT	2	5
Nino 1.2	3	4
Nino 3.0	4	7
Nino 3.4	5	3
Nino 4.0	6	5
SOI	7	7
Sunspots	8	6
DMI	9	4
IPO	10	3

Gradient descent was used to train the networks. The only difference was the number of inputs. For each month, the whole dataset was trained and tested. The performance of the whole dataset networks was compared against proposed GA selected results. To measure the accuracy, RMSE over the testing set was calculated. Results of the two models are shown in Table VI. RMSE for all feature based networks varied between 0.039 in November and 0.2354 in February. The RMSE of the network that used the all features exceeded the genetic algorithm based approach in all the months. The highest difference in performance was recorded in both February and December with a difference of 133.58 and 73.56 mms respectively. The average error of GA based forecasts was reported 0.0638 (58.35 mms). This proves that using appropriate climate attributes to forecast could reveal higher performance than using all the available climate attributes.

TABLE VI. COMPARISON OF RESULTS IN TERMS OF RMSE.

Month	*RMSE (All features)*	*RMSE (Proposed Approach)*
January	0.2301	0.1813
February	0.2354	0.0890
March	0.1603	0.0863
April	0.0711	0.0387
May	0.0526	0.0341
June	0.0904	0.0459
July	0.0591	0.0300
August	0.0598	0.0344
September	0.0506	0.0346
October	0.0474	0.0342
November	0.0393	0.0270
December	0.2115	0.1311
Average	0.1089	0.0638

IV. CONCLUSION

In this paper, a genetic algorithm based approach was proposed to select the best features for monthly rainfall forecasting in Eastern Australia. Many experiments were conducted and improved results were obtained. The results revealed that for most of the months, several climate attributes such as maximum temperature values, Nino3.0 and SOI are essential. There is no need to combine all the features since some of them obtained lower performance when added to the classifier. Results showed that a set consisting of maximum temperature values and SOI achieves the highest accuracy for November forecasts. In future, this research will be extended by including data from more locations near sugarcane industry and new climate variables. In addition, new optimisation algorithms will be investigated and compared.

ACKNOWLEDGMENT

This research is funded by CQUniversity's research division. We gratefully acknowledge CQUniversity's eResearch support and the use of the high performance computing facility (www.cqu.edu.au/hpc) in the completion of this work.

REFERENCES

[1] Y. L. Everingham, R. C. Muchow, R. C. Stone, and D. H. Coomans, "Using Southern Oscillation Index Phases to Forecast Sugarcane Yields: A Case Study for Northeastern Australia," *International Journal of Climatology,* vol. 23, pp. 1211-1218, Aug. 2003.

[2] X. He, H. Guan, X. Zhang, and C. T. Simmons, "A Wavelet‐based Multiple Linear Regression Model for Forecasting Monthly Rainfall," *International Journal of Climatology,* vol. 34, pp. 1898-1912, Aug. 2013.

[3] R. C. Deo and M. Şahin, "Application of the Extreme Learning Machine Algorithm for the Prediction of Monthly Effective Drought Index in Eastern Australia," *Atmospheric Research,* vol. 153, pp. 512-525, Feb. 2015.

[4] M. Nasseri, K. Asghari, and M. J. Abedini, "Optimized Scenario for Rainfall Forecasting using Genetic Algorithm Coupled with Artificial Neural Network," *Expert Systems with Applications,* vol. 35, pp. 1415-1421, Oct. 2008.

[5] Y. L. Everingham, A. J. Clarke, and S. Van Gorder, "Long Lead Rainfall Forecasts for the Australian Sugar Industry," *International Journal of Climatology,* vol. 28, pp. 111-117, May 2007.

[6] C. M. Kishtawal, S. Basu, F. Patadia, and P. K. Thapliyal, "Forecasting Summer Rainfall over India using Genetic Algorithm," *Geophysical Research Letters,* vol. 30, no. 23, Dec. 2003.

[7] X. Meng, "Weather Forecast Based on Improved Genetic Algorithm and Neural Network," in *Proceedings of the International Conference on Information Engineering and Applications (IEA) 2012: Volume 4*, Z. Zhong, Ed., ed London: Springer London, 2013, pp. 833-838.

[8] H. Nagahamulla, U. Ratnayake, and A. Ratnaweera, "Selecting Most Suitable Members for Neural Network Ensemble Rainfall Forecasting Model," in *Recent Advances on Soft Computing and Data Mining: Proceedings of The First International Conference on Soft Computing and Data Mining (SCDM-2014) Universiti Tun Hussein Onn Malaysia, Johor, MalaysiaJune 16th-18th, 2014*, T. Herawan, R. Ghazali, and M. M. Deris, Eds., ed Cham: Springer International Publishing, 2014, pp. 591-601.

[9] O. H. Babatunde, E. Cowan, L. Armstrong, J. Leng, and D. Diepeveen, "A Genetic Algorithm-based Feature Selection," *International Journal of Electronics and Communication and Computer Engineering*, IJECCE, 2014.

[10] J. Zou, Y. Han, and S.-S. So, "Overview of Artificial Neural Networks," *Methods in molecular biology (Clifton, N.J.),* vol. 458, pp. 15-23, Dec. 2008.

[11] N. S. Philip and K. B. Joseph, "A Neural Network Tool for Analyzing Trends in Rainfall," *Computers & Geosciences,* vol. 29, pp. 215-223, Mar. 2003.