Proceedings of the Technical Sessions, 22 (2006) 1-16
Institute of Physics – Sri Lanka

Rainfall Forecasting: An Artificial Neural Network Approach

A.D. Kumarasiri and D.U.J. Sonnadara
Department of Physics, University of Colombo, Sri Lanka

ABSTRACT

An innovative technique is utilized for rainfall forecasting using Artificial Neural Networks based on feed-forward back-propagation architecture. Focus is set upon making successful predictions from the available data, not on incorporating the physical aspects of the atmosphere or the actual process of rainfall occurrence. Both short term and long term forecasting was attempted for ground level data collected by the meteorological station in Colombo, Sri Lanka (Lat: 79.87 E, Long: 6.90 N, Altitude: 7.3 m).

Three Neural Network models were developed; a one-day-ahead model for predicting the rainfall occurrence of the next day, which was able to make predictions with a 74.25% accuracy, and two long term forecasting models for monthly and yearly rainfall depth predictions with 58.33% and 76.67% accuracies within a 5% uncertainty level. Each of these models was extended to make predictions several time steps into the future, where accuracies were found to be decreasing with the number of time steps. The success rates and rainfall trends within the monsoon seasons were also studied and presented.

1. INTRODUCTION

Over the last few decades, several models have been developed, attempting the successful forecasting of rainfall in Sri Lanka. Though some of these models show notable accuracies in short term rainfall occurrence prediction [1, 2], long term prediction and rainfall depth prediction has proven to be somewhat difficult using traditional statistical methods. The reason being that the rainfall dynamics are dependant upon highly unpredictable physical parameters such as humidity, wind speed, wind direction, pressure, temperature and cloud amount.

Considering each of these physical parameters generate increasing degrees of sophistication in the statistical models. Importance of each parameter has to be taken into account, thus, making initial assumptions about the parameters on hand may become necessary. Fortunately, recent developments in artificial intelligence and pattern recognition provide an answer for this dilemma.

Using Artificial Neural Networks (ANNs) which are based upon the neural structure of the human brain, complex pattern recognition can be attempted without making any initial assumptions where the data set used is allowed to govern the process by itself. An ANN provides the user a model free tool, which can generate input-output mapping for any set of data, however complex. Training the network with the relevant data enables the network the ability of making predictions based on any input it encounters [3, 4].

1.1 Artificial Neural Networks

Recognition of the fact that the functioning of the human brain is completely different from a Von Neumann based traditional digital computer has inspired the development of Artificial Neural Networks which are capable of performing complex computations such as complex pattern recognizance that traditional computers lack the capacity to handle.

The fundamental processing element of an ANN is an artificial neuron. Just like the natural neuron in human brain, it can receive inputs, process them and produce the relevant output. A simple mathematical model can be used in explaining a neuron quantitatively.

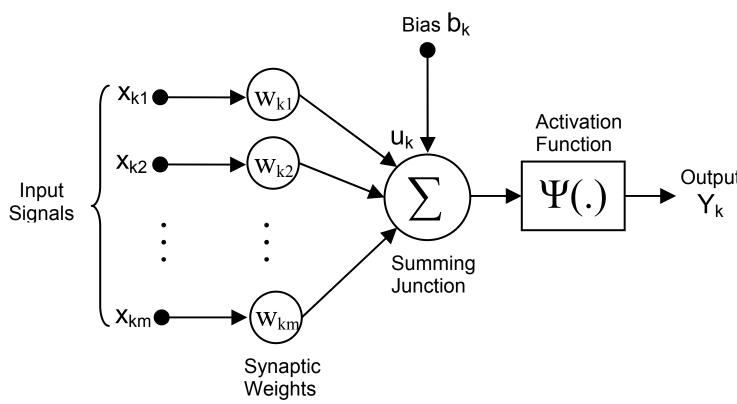


Figure 1: Modeling a Simple Neuron

The net input at the summing junction can be written as: $U_K = \sum_{j=1}^{m} W_{kj} X_{kj}$ [5]

The output U_k is worked upon an activation function, whose sole purpose is to limit the output of the neuron to a desired value, increasing the net performance of the network. Then the output from the k^{th} Neuron is:

$$Y_K = \Psi(U_K + b_K)$$

Linking a neuron with others via synapses result in a network of neurons which can be single layered or multi layered. A multi layer ANN contains an input layer of neurons, an output layer of neurons, and may also contain one or more hidden layers of neurons, so called because they are not directly visible to the outside.

According to the neuron positioning and connection types, various types of ANNs have been developed. The type utilized in this paper is feed-forward back-propagation Networks.

1.2 Feed-Forward Back-Propagation Architecture

The feed-forward, back-propagation architecture, developed in the early 1970s, is still the most popular and most effective model for complex, multi layered networks. The model is extensively used in weather and financial forecasting networks.

The term *feed-forward* refers to the inputs being swept forward through the network, getting multiplied by each synaptic weight and being summed at each node until the output node is met. Then, *back-propagation* occurs where the desired output is compared with the produced output and errors are backwardly propagated through the network where synaptic weights for each layer are adjusted in proportion to this error, limited by a factor defined by the user so that large adjustments may cease to occur for stray values in the data set.

The same cycle continues for multiple times, for the same input and output vectors as the weights are fine tuned to produce the best possible output; in this case, the closest value to the desired output. Global accuracy of the network is increased in each cycle as the error becomes smaller with the number of cycles.

The priority is given to determining the weights which makes the maximum contribution towards the error, and making adjustments before moving on to less significant nodes. Changing the weight of an inactive node may not improve the network performance at all.

The typical back-propagation network contains an input layer, an output layer, and at least one hidden layer. The number of neurons at each layer and the number of hidden layers determine the network's ability on producing accurate outputs for a particular data set. Unfortunately, there are no well defined methods for determining these characteristics. There is no quantifiable best answer to the layout of a network for any particular application. It is solely based on trial and error methods and the network designer's own experience, where finding the most effective network for a particular data set can be considered as an art.

1.3 Training and Testing Methods

Once a network had been structured for a particular application, inputs and the corresponding targets are used to train a network until it learns to associate a particular input with a reasonable output. A network is ideally trained until the change in weights in a training cycle approaches a minimum, improving its performance through the learning process.

After the network is sufficiently trained, it has to be tested for its ability to produce accurate outputs. Large multilayered networks with multiple nodes in each layer are capable of memorizing data due to the vast number of synaptic weights available on such a network. Thus, generating correct outputs for input vectors encountered within the training process does not justify the ability of a network to generate accurate outputs.

2. SHORT TERM FORCASTING

In developing a model for short term forecasting, an attempt is made at forecasting the rainfall occurrence of the next day, i.e. whether tomorrow would be a dry day or a wet day.

2.1 The One-Day-Ahead Model

A wet day was defined as a day in which the total rainfall depth of 24 hours from 12 AM exceeds or is equal to 0.1 mm, a value based upon the minimum measurement made by the meteorological department where depths under 0.1mm are not numerically recorded. Any day which had a rainfall depth under 0.1mm was considered a dry day. The basic parameters considered here upon which rainfall occurrence depends are: humidity (due to high variability, four values measured at 0600, 1200, 1800 and 2400 hours), Pressure (measured at 0900 hours), wind run (distance the wind had traveled within 24 hours), temperature (average for the last 24 hours) and cloud amount (cloud cover index).

When all other parameters were considered, temperature did not improve the results of the network, thus was omitted from the final design for the sake of improving the network efficiency. Including the past 3 days rainfall occurrence data in the input vector immensely improved its accuracy of the prediction.
Input value to the network for rainfall occurrence was defined as,

$X_t = 0$, if $R_t < 0.1$ mm $\quad X_t = 1$, if $R_t > 0.1$ mm

Where R_t is the rainfall depth of the t^{th} day and X_t is the corresponding input value to the network. Each of the parameters, except rainfall occurrences was normalized to limit the inputs between [-1,1] to improve the training efficiency. The 10-6-1 network architecture showed best potential within a short training and testing period, and was considered as the final network.

Figure 2 represents the logical interpretation of the network.

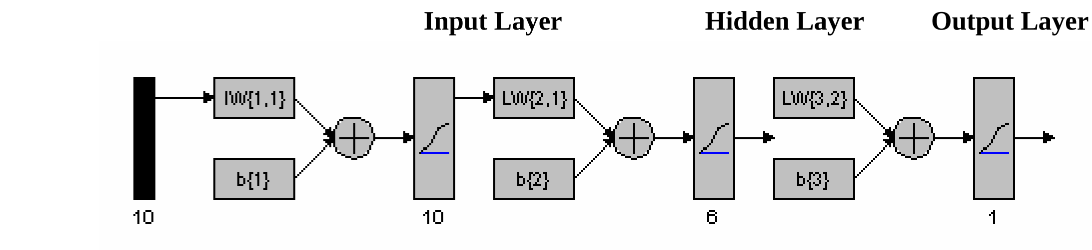


Figure 2: The one-day-ahead ANN model

IW{1,1} - Weights of the Input layer
LW{2,1} - Weights between Layer 1 and Layer 2
LW{3,2} - Weights between Layer 2 and Layer 3
b{1}, b{2}, b{3} - Biases for each layer

Synaptic weights and Biases are set randomly before the training begins. Output of each neuron was limited within [0,1] by using the Log-Sigmoid Activation function to increase the training efficiency. The output neuron uses a Log-Sigmoid function as well, since the output should be within the interval [0,1].

The network was trained for a period of 9 years from 1994 to 2002, a total data set of 3284 input/output vector pairs, using the error back-propagation algorithm. Batch training was carried out for batches of six months of data, which was determined as the best batch size. During the training period, a limit of 100 epochs and a training mean square error of 10^{-30} were set.

In testing the network performance, two sets of six-monthly testing data were used to collaborate the six-monthly batch training. First set consisted of the first six months of 2003 (181 Days), and the second set consisted of the last six months (184 Days). The total set contained 169 wet days and 196 dry days. The output of the network during the testing phase was interpreted as,

if $X_t < 0.5$, day t is dry   if $X_t \geq 0.5$, day t is wet

The prediction success rate of the network was calculated on the basis of above criteria using,

$$\text{Prediction Success rate} = \frac{x_c}{x_{tot}} \times 100$$

where X_c = Number of correct predictions, X_{tot} = Total number of predictions

In addition the RMS Error of the networks were used which is defined as,

$$\sqrt{\overline{X^2}} = \sqrt{\frac{\sum_i (X_{Pr} - X_{Exp})^2}{i}}$$

where X_{pr} = Predicted Output
X_{exp} = Expected Output

2.2 Results of One-Day-Ahead Forecasting

Out of 365 days in the total data set, 271 correct predictions were made, giving a success rate of 74.25%. This rate is higher than the 67% obtained in reference 1 using ANN approach, and the 69% obtained in reference 2 using the Markov model for Colombo.

Table 1: Prediction Success Rates of the one-day-ahead model

Type of Forecasting	Data Set 1	Data Set 2	Total Set
Rain Classification	72.50%	76.4%	74.55%
No Rain Classification	79.2%	68.42%	73.98%
Total Rain / No Rain Classification	76.24%	72.28%	74.25%

As illustrated in table 1, the success rate drops from 76.24% to 72.28% for the second six months. This behavior is expected as the network is only trained up to December 2002. The ability to make correct predictions drops with time even if the rainfall pattern changes very slightly, thus, immediate predictions have a higher tendency of being accurate. This can be avoided by retraining the network constantly with time, preferably each year.

However, when separate predictions for two data sets are analyzed, there is a notable difference. No Rain classification in data set 1 is significantly high, while the rain classification in data set 2 is higher than that of set 1. An explanation can be provided on the basis of the patterns of wet and dry days in each set. The first half of the year contained more dry days, 101 to be exact, where long strings of dry days ('0' values) were present during March-April inter-monsoon season, thus, making a correct dry day prediction becomes highly likely. The second set contained a high number of rainy days, 89 to be exact. During the late southwest monsoon season, strings of continuous wet days ('1' values) were present in the months of July and August, while dry days were scattered throughout the data set. This gives rise to a high rain classification rate and low no rain classification rate in set 2.

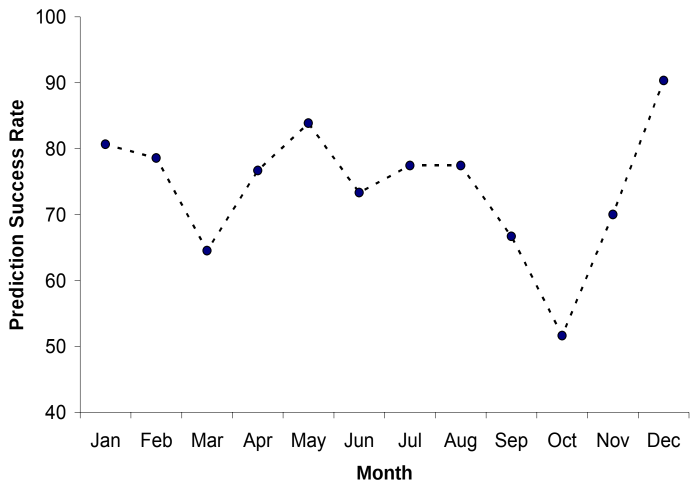


Figure 3: Prediction Success Rates by month for year 2003

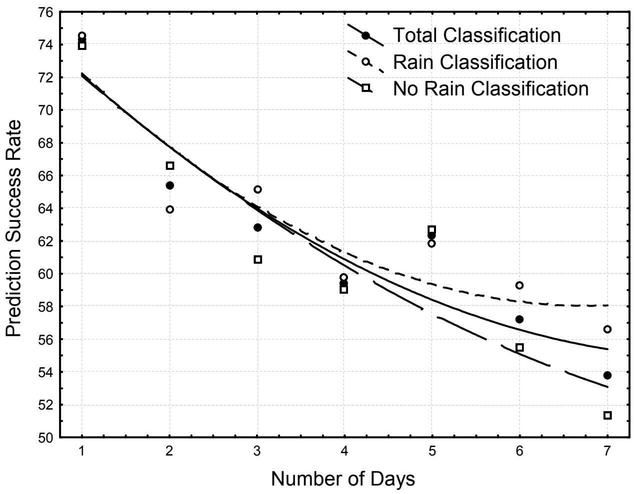
The prediction success rate for each month is illustrated in Figure 3. The highest rate of 90.32 was observed for December, while the lowest rate of 51.61 was observed for October. These unusually high and low values can be explained by the dry and wet day distributions in the data set. In December, only 3 wet days were present at the beginning of the month, and a long string of dry days thereafter. Thus the high rate occurs due to successful prediction of dry days within the month. October contained randomly scattered dry and wet days throughout the month and predictions were found relatively difficult to make.

2.3 Multiple Steps Ahead Forecasting

Using the same network, an attempt was made at predicting rainfall occurrences seven days into the future; namely for days t, (t+1)…up to the $(t+6)^{th}$ day. Here the input vector for each case consisted of predicted values for the rainfall occurrence for the past three days instead of actual values, and the $(t-1)^{th}$ day's remaining weather parameters.

The results closely followed the trend expected. The accuracy of each classification, along with the total classification should decrease with time. This is true for each case except for day 5 where a sudden increase in success rates can be seen.

An explanation for this unusual behavior can be formulated by observing the actual test data. The output of the network contained a number of values near the 0.5 benchmark. As a result, a few successful dry day forecasts were made based on values closer to 0.47, and a few wet day predictions were made on values of 0.51. This gives rise to a slightly high rate of success than expected, but the RMS error has still increased for day 5 beyond than that of day 4, indicating a continuous drop in network performance. Thus, the expected drop in accuracies with the number of days had prevailed.



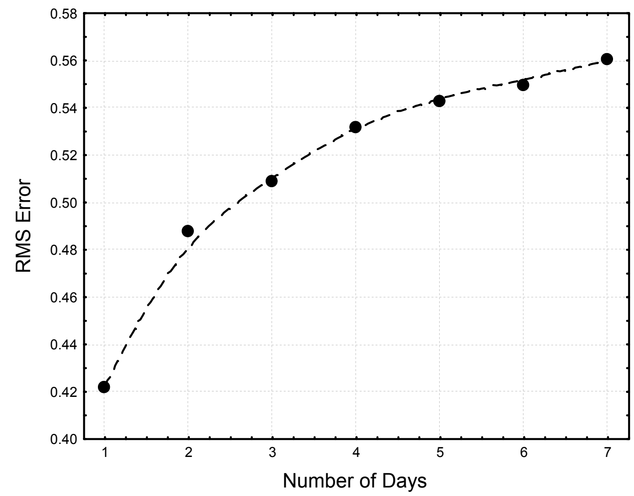


Figure 4: Variation of success rates with number of days

Figure 5: Variation of RMS Errors with number of days

Rain classification rate in set 2 is higher than that of set 1 and no rain classification in set 2 is higher than that of set 1, the same pattern as explained earlier. The success rate drops to almost 50% on the 7th day, thus, making forecasts beyond 7 days was found to be pointless.

Although the total number of dry days was higher than wet days, long strings of wet days could be seen in both northeast and southwest monsoon seasons. This explains the slow decrement rate of rain classification with time, compared with no rain classification.

Deciphered in figure 5, the RMS errors precisely follow the expected pattern, increasing steadily at first, and then tend to even off as the prediction success rate comes near the 50% mark. It can clearly be seen that the increase of success in day 5 observed earlier is purely by inconsistencies in the data set, as the RMS error on day 5 is consistent with the pattern. This RMS Error graph is a handy tool to analyze the performance of a network, and is unique for each network.
An analysis of predictions during the monsoon seasons were required to explain the high rain classification success rates in the above model. The network was tested for the southwest monsoon which occurs from May to September, and the northeast monsoon from December to February.

Table 2: Rain Classification Success Rates of Monsoon Seasons

Day	Southwest Monsoon	Northeast Monsoon
Day 1	80.23%	72.22%
Day 2	77.91%	77.77%
Day 3	76.24%	72.22%
Day 4	79.07%	72.22%
Day 5	77.91%	70.59%
Day 6	81.39%	68.75%
Day 7	77.91%	60.00%

Table 2 justifies the explanation given on high rain classification rates. During the southwest monsoon a total of 86 wet days were observed in 2003, and a universally high prediction rate is observed for rain classification, no matter the number of days ahead. Hence, high rain classification rates are observed in data set 2.

During inter-monsoon seasons, long stretches of dry days could be seen. This gives rise to high no rain classification rates in set 1, as opposed to monsoons.

3. LONG TERM FORECASTING

To make the model for rainfall forecasting comprehensive, developing ANNs for making long term predictions were a necessity. Two such networks were developed based on the previously discussed one-step-ahead approach to make monthly and yearly rainfall depth forecasts.

3.1 The One-Month-Ahead Model

Using the same approach as the one-day-ahead model, an ANN was developed to make monthly rainfall depth predictions. Making actual depth predictions for each month was quickly found to be unfeasible as the error was too large, thus, monthly rainfall totals were classified into 6 categories as follows.

Table 3: Classification of Monthly Rainfall Depths

Actual Rainfall Depth (mm)	Input Value	Interpretation of the Output Value
0< X ≤100	0.5	0< Y ≤1.0
100< X ≤200	1.5	1.0< Y ≤2.0
200< X ≤300	2.5	2.0< Y ≤3.0
300< X ≤400	3.5	3.0< Y ≤4.0
400< X ≤500	4.5	4.0< Y ≤5.0
X > 500	5.5	Y > 5.0

The last interval was kept open ended since very few values were encountered beyond 500mm in the data set.

The Input vector comprised of rainfall depths of twelve months who are the immediate predecessors of the month upon which the forecasting is done, and the output vector is obviously the target month's rainfall depth, as determined by table 3.

The output should not be limited as setting an upper boundary on monthly rainfall depths is not desirable. Thus, a pure linear transfer function was used for the output neuron. For all other neurons, the usual Log-Sigmoid function was utilized.

The network was trained for 50 years from 1949 to 1998 and testing was done for 60 inputs, from 1999 to 2003. The network performance was analyzed using the prediction success rate and RMS errors.

3.2 Results of the One-Month-Ahead Network

Out of the 60 months used for testing, the network was able to predict 35 correctly within the testing basis presented in table 3. Thus, the prediction success rate was 58.33 and the RMS error of the network was calculated as 0.755.

Table 4: Success Rates of Rainfall Depth Classification

Rainfall Depth	Prediction Success Rate
0< X ≤100	57.89%
100< X ≤200	66.67%
200< X ≤300	62.50%
300< X ≤400	25.00%
400< X ≤500	00.00%
X > 500	50.00%

According to table 4, the success rates are somewhat similar up to a 300mm depth. Rates beyond them cannot be viably interpreted as there were only 4, 1 and 2 months in each of the last data sets. For example, the network wasn't able to make a successful prediction for the single month in the 400<X≤500 range, producing a 0% rate. However, this is due to the limitations of the data set, and doesn't necessarily indicate no correct prediction for that range is possible.

Multiple steps ahead forecasting was attempted for several months into the future using the same network.

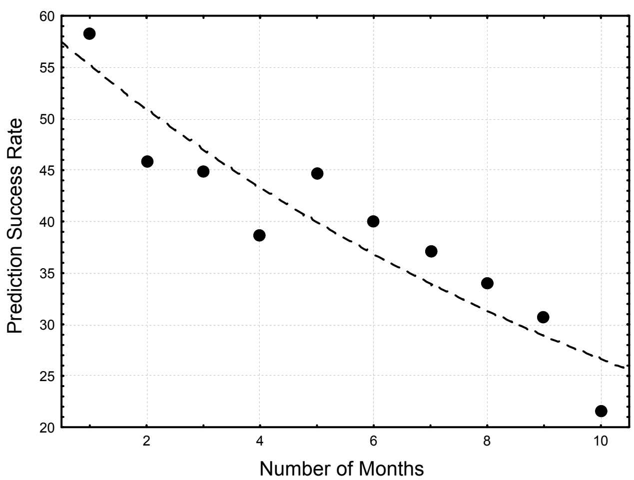


Figure 6: Variation of Success Rates with number of months

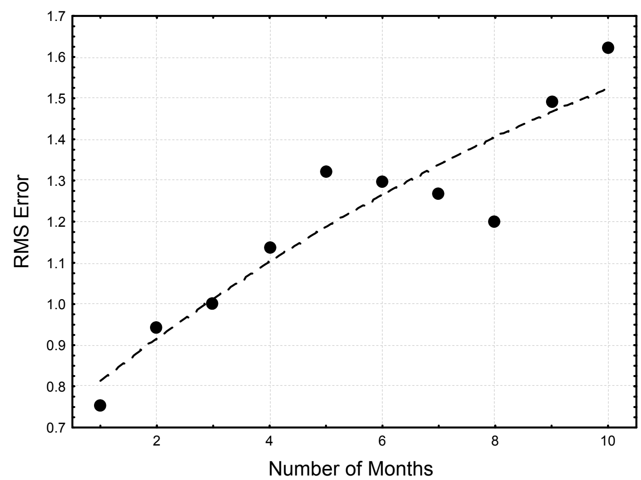


Figure 7: Variation of RMS Errors with number of months

By the 10th month, the network correctly predicts only one value for five inputs. Thus, further steps into the future were deemed unfeasible.

Expected behavior is met with only slight deviations. Up to the fourth month, the decrease in the success rates and the increase in RMS errors are as expected. But within the next few months, both parameters tend to show a slightly erratic behavior, which an effect probably brought on by the complications in the data sets used. However, the general trend is clear from the figures.

3.3 The One-Year-Ahead Model

Unlike the previous models, making actual depth predictions were found to be viable. The Input vector comprised of rainfall depths of ten years who are the immediate predecessors of the year upon which the forecasting is done, and the output vector is the target year's rainfall depth.

The network was trained for a period of 105 years, from 1869 to 1973, using the error back-propagation algorithm, and was tested for a separate data set of 30 years from 1974 to 2003. A new parameter, the percentage error was introduced in analyzing the one-year-ahead network

$$\text{Percentage Error} = \frac{(X_P - X_A)}{X_A} \times 100 \text{ , Where } X_P = \text{Prediction, } X_A = \text{Actual Value}$$

In the case of total prediction accuracy, absolute values of the percentage errors were used in order to prevent the positive and negative errors being compensated for each other for the whole network.

Out of the 30 annual depths tested for, the network successfully predicted 23 depths within an error of ±100, approximately 5% of the mean annual rainfall depth within the total period.

Thus, the prediction success rate for the network is (23/30)*100 = 76.67% and the RMS error was calculated as 0.267. The absolute percentage error was found to be 6.99%.

3.4 Results of the One-Year-Ahead Model

Figure 8 compares the actual annual rainfall depths with the predicted depths for the testing period from 1974 to 2003. Deviations from actual values tend to increase slightly when the year moves away from the last training year of 1973, for the same reasons explained in the one-day-ahead network. Two large positive deviations can bee seen for years 2001 and 1995.

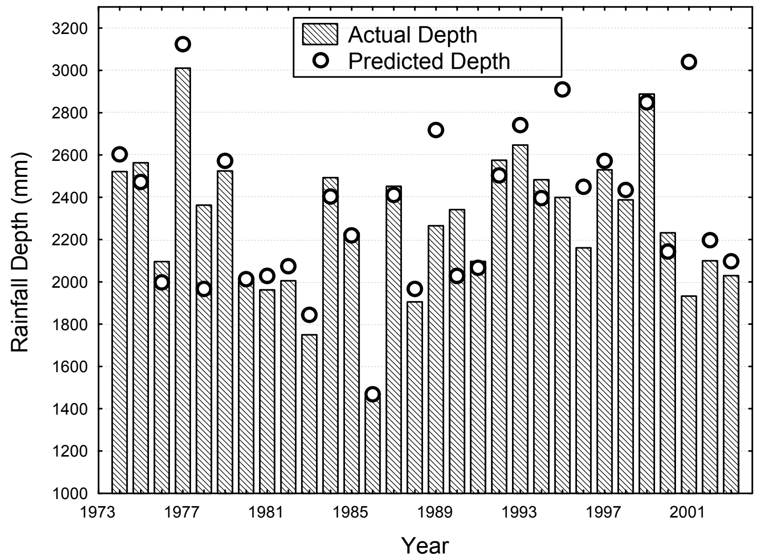


Figure 8: Actual and Predicted Rainfall Depths for 1974-2003

In both these years, severe droughts were experienced, limiting the rainfall amounts throughout the country [6]. The network is not equipped to handle such incidents and its predictions point out the annual rainfall depths for 2001 and 1995 if the natural pattern had prevailed.

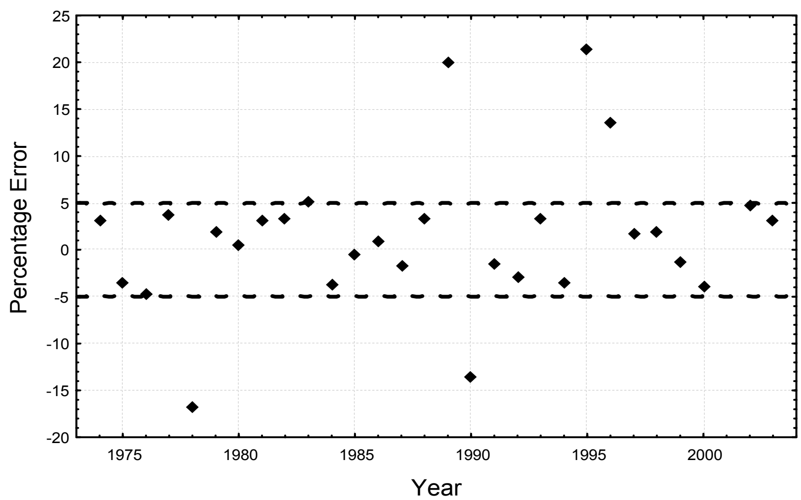


Figure 9: Percentage Error of each year

The percentage error plot clearly shows that the bulk of the predictions have percentage errors between ±5. The stray value for 2001 (62% error) was ignored for the clarity of figure 9 as it enlarged the Y axis scale and the graph became indistinct.

Table 5 indicates the success rates of the network for multiple time steps ahead forecasting.

Table 5: Success Rates for multiple steps ahead forecasting

Year	Prediction Success Rate	RMS Error	Absolute Percentage Error
Year 1	76.67%	0.267	6.99%
Year 2	31.03%	0.598	15.89%
Year 3	17.86%	0.889	26.93%
Year 4	18.52%	1.121	37.65%
Year 5	3.85%	2.254	68.16%

The success rate drops by more than a half by the second year. Though the success rate slightly increases for year 4, both the RMS error and the Absolute Percentage Error indicate that the network performance is still rapidly decreasing. The rapid drop of rates is owing to the fact that the actual depths are being predicted, rather than rainfall occurrence or depth categories like in previous models. The relatively large prediction errors for first year depth forecasts present in the network for years 1978, 1989-90, 1995-96 and 2001 highly affect the entire multiple steps ahead predictions. At least one of these large errors is present in most of the inputs vectors used in testing, thus higher order forecasting successes quickly perish. By year 5, only one depth was predicted correctly out of 26 predictions.

Thus, this network should not be extended for forecasting beyond 2 years ahead, as accuracies become exceedingly low.

4. DISCUSSION AND CONCLUSIONS

- The One-Day-Ahead network is successful in forecasting the rainfall occurrence of the next day, with a success rate of 74.25%. When the model was extended for multiple days ahead, the success rate of the network decreased and RMS error increased with the number of days. Seven days was the limit up to where reasonable predictions could be made.

- The One-Month-Ahead network was reasonably successful in forecasting the next month's rainfall depth category with a success rate of 58.33% within a ±50mm error limit. Extension for multiple months into the future produced the same result as above. The success rate dropped to 41% within 6 months into the future.

- The One-Year-Ahead network is highly successful in forecasting the annual rainfall depth of the next year, with a success rate of 76.67% within a ±100mm error limit. When the stray values for droughts in 2001 and 1995-96 are ignored, the rate increased to 83.33%. The network performance dropped rapidly when multiple years ahead predictions were made. Hence, this model ideally should not be used beyond the second year into the future.

- The ability for actual rainfall depth prediction using the feed-forward back-propagation ANN models used for this research showed a tendency to be significantly improved when forecasting over long time periods was done. Yearly depths could be accurately predicted while monthly depths could be predicted only when classified into different categories, and daily predictions could be made only for occurrence.

- Using ground level data undoubtedly reduced the accuracy of predictions. However, this problem is impenetrable since cloud level data is not available adequately to train such a network. An alternate approach would be to combine ANNs for several whether stations close by, making a multiple point forecasting network which would eliminate hectic variations in spatial coordinates [7].

Acknowledgements: Financial assistance by National Science Foundation, Sri Lanka (NSF grant number RG/2000/P/01) is acknowledged.

REFERENCES

1. Dharmaratne W.G.D. and Premarathna L.D., Development of a Rainfall forecasting model for Sri Lanka using Artificial Feed-Forward Neural Network, Proceedings of the 2nd Science Symposium-University of Ruhuna, Sri Lanka (2004) 29-36
2. Perera, H.K.W.I., Sonnadara, D.U.J. and Jayewardene, D.R., Forecasting the Occurrence of Rainfall in Selected Weather Stations in the Wet and Dry Zones of Sri Lanka, Sri Lankan J. Physics, 3 (2002) 39-53
3. Philip, N.S. and Joseph K.B., A Neural Network tool for analyzing trends in rainfall, Computers & Geosciences, 29 (2002) 215-223
4. Luk, K.C., Ball, J.E. and Sharma, A., An Application of Artificial Neural Networks for Rainfall Forecasting , Mathematical and Computer Modeling, 33 (2001) 683-693
5. Haykin, S. Neural Networks: A Comprehensive Foundation, Prentice-Hall Inc, (1999) 2nd Edition

6. Hettiarachchi N.D. , National Disaster management policy: Implications for Development, Proceedings of Conference on Natural Disaster Management, Sri Lanka (2004) 28-41
7. James, N.K.L. and Raymond S.T.L., Rainfall forecasting from Multiple Point Sources using Neural Networks, IEEE journal, (1999) 429-434