Forecasting Rainfall Using Adaptive Neuro-Fuzzy Inference System (ANFIS)

Jignesh Patel¹, Dr. Falguni Parekh²

¹PG Student, Water Resources Engineering and Management Institute (WREMI), Faculty of Technology and Engineering, The Maharaja Sayajirao University of Baroda, Samiala-391410, Gujarat, India

²Offg. Director & Associate Professor, Water Resources Engineering and Management Institute (WREMI), Faculty of Technology and Engineering, The Maharaja Sayajirao University of Baroda, Samiala-391410, Gujarat, India

Abstract
Rainfall is one of the most complex and difficult elements of the hydrologic cycle. The trend and forecasting of rainfall is very difficult to understand and to model due to the complexity of the atmospheric processes that generate rainfall. This paper investigates the development of an efficient model to forecast monthly monsoon rainfall for Gandhinagar station using Adaptive Neuro Fuzzy Inference System (ANFIS). Eight models are developed using various membership functions and climatic parameters as inputs. In this study, the generalized bell-shaped built-in membership function (gbell) has been used as a membership function in both Hybrid and Back propagation method for ANFIS. The four evaluation parameters Root mean square error (RMSE), Correlation Coefficient (r), Coefficient of Determination (R²) and Discrepancy ratio (D) are used to evaluate the developed model. The study reveals that hybrid Model with seven membership functions and using three inputs, temperature, relative humidity and wind speed gives best result to forecast rainfall for study area.

Keywords: Adaptive Neuro Fuzzy Inference System (ANFIS), Back Propagation Method, Hybrid Method, Rainfall forecasting.

1. INTRODUCTION
Rainfall is essentially random in nature and it is a natural phenomenon resulting from atmospheric and oceanic circulation. Rainfall forecasting is extremely important in water resource engineering like proper mitigation and management of floods, droughts, environmental flows, water demand by different sectors, maintaining reservoir levels, and disasters. Rainfall is main source for water supply, which is basic requirement for the crop production. Therefore forecasting of rainfall is useful for proper planning of cultivation. Rainfall is a stochastic process. Rainfall depends on weather parameters which are mean temperature, surface pressure, relative humidity and wind speed.

Banik, S. et al. (2009) have developed rainfall forecasting model using ANN, ANFIS and GA processes and results obtained by these models are also compared to the statistical. The ANFIS forecasting model and the GA forecasting model can be used to forecast monthly monsoon rainfall more accurately than the ANN model and the statistical model, forecasting method, namely linear multiple regression model.

Nayak, D. R et al. (2013) have done survey on rainfall predictions using different neural network architectures over twenty-five years. From the survey it has been found that most of the researchers used back propagation network for rainfall prediction and got significant results. The survey also gives a conclusion that the forecasting techniques that use MLP, BPN, RBFN, SOM and SVM are suitable to predict rainfall than other forecasting techniques such as statistical and numerical methods.

El-Shafie et al (2011) have developed an adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN) model to forecast the rainfall for Klang River in Malaysia on monthly basis. The result shows performance of ANFIS method is better than ANN method and concludes that ANFIS method is superior to the ANN method in forecasting monthly rainfall.

True quantitative rainfall forecasting is generally difficult and also a challenging task for anyone because of our complex atmospheric processes. Thus, rainfall is treated as one of the most complex and difficult events among other hydrological events. Forecasting techniques, namely the statistical methods (ARIMA, regression model, hidden Markov model, exponential smoothing etc.) and the AI methods (neural networks (artificial neural network (ANN), adaptive network based fuzzy inference system (ANFIS) etc.), fuzzy inference system (FIS), genetic algorithm (GA) and others) are proposed. In recent years, Adaptive Neuro-Fuzzy Inference System (ANFIS) have become extremely popular for forecasting in number of areas. Rainfall data are multi-dimensional, non-linear and dynamic, therefore to search for an appropriate model, the powerful techniques ANFIS have been chosen.

2. OVER VIEW ANFIS
This architecture is proposed by Jang (1993) and is developed based on the theory of fuzzy set and fuzzy logic. It is a combination of two intelligence systems, namely ANN system and FIS system in such a way that the ANN learning
algorithm is used to determine the parameters of the FIS. ANN is a non-linear statistical data-modelling tool, which can capture and model any input-output relationship (or can learn detect complex patterns in data). FIS (involves membership function (mf), fuzzy logic operator and if-then-rules) is the process of formulating the mapping from a given input to an output using fuzzy logic.

3. ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM

ANFIS is the fuzzy-logic based paradigm that grasps the learning abilities of ANN to enhance the intelligent system’s performance using knowledge gained after learning. Using a given input-output data set, ANFIS constructs a fuzzy inference system whose membership function parameters are tuned or adjusted using hybrid type of neural algorithms.

Adaptive neuro-fuzzy inference system can be classified into three categories:
- A fuzzy rule-based model constructed using a supervised NN learning technique.
- A fuzzy rule-based model constructed using reinforcement-based learning.
- A fuzzy rule-based model constructed using NN to construct its fuzzy partition of the input space.

Each fuzzy system contains three main parts: fuzzification, inference, and defuzzification. The fuzzy inference system that we have considered is a model that maps,
1. Input characteristics to input membership functions,
2. Input membership function to rules,
3. Rules to a set of output characteristics,
4. Output characteristics to output membership functions, and
5. The output membership function to a single valued output, or
6. A decision associated with the output.

The architecture of an ANFIS model with two input variable is shown in Figure 1. Suppose the rule base of ANFIS contains two fuzzy IF-THEN rules of sugeno type as follow:

If - then rules is defined as

Rule 1: If x1 is A1 and x2 is B1, then f1 = a1x1 + b1x2 + c1  (1)
Rule 2: If x1 is A2 and x2 is B2, then f2 = a2x1 +b2x2 + c2  (2)

Where a1, a2 and b1, b2 are membership; of input variable x and y which are parameters of the output function f1 and f2, respectively. A basic sugeno inference system that produce an output function f from input variables x and y by applying triangular membership function is illustrated schematically in Figure 1, and also the corresponding equivalent ANFIS architecture is shown in Figure 2.

![Two-input first-order Sugeno fuzzy model with two rules](image)
The functions of each layer are described as follows:

Layer 1: In the first layer, all the nodes are adaptive nodes. The output of layer 1 is the fuzzy membership grade of the inputs, which are given by,

\[
\Omega_1^i = \mu_{A_i}(x) \quad \text{for } i = 1, 2
\]

or

\[
\Omega_1^i = \mu_{B_i}(y) \quad \text{for } i = 3, 4
\]

where \(x\) (or \(y\)) is the input to the node; \(A_i\) (or \(B_i\)) is a fuzzy set associated with this node, characterized by the shape of the MFs in this node and can be any appropriate functions that are continuous and piecewise differentiable such as Gaussian, generalized bell shaped, trapezoidal shaped and triangular shaped functions. Assuming a generalized bell function as the MF, the output \(\Omega_1^i\) can be computed as,

\[
\Omega_1^i = \Phi_{\text{bell}} = \frac{1}{1 + \left(\frac{x - c_i}{a_i}\right)^2}
\]

Layer 2: The nodes in this layer are fixed nodes and they perform a simple multiplier. The output of this layer is represented as:

\[
\Omega_2^i = w_i = \mu_{A_i}(x) \mu_{B_i}(y), \quad i = 1, 2
\]

Layer 3: The nodes in this layer are also fixed nodes and they play normalization role in this network. The output of this layer can be represented as:

\[
\Omega_3^i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2
\]

Layer 4: All nodes in this layer are adaptive; whose output is simply the product of the normalized firing strength and a first order polynomial (for a first order sugeno model). Thus, the output of this layer is:
\( O_5 = \sum_j w_j f_i = w(p_i x + q_j y + r_i) \) (8)

Where \( O_5 \) is the output of layer 3 and \( \{p_i, q_j, r_i\} \) is the parameter set.

Layer 5: This single node in this layer is a fixed node and computes the overall output \( O_F \) as the summation of all incoming signal, that is,
\[
O_F = \text{Overall output} = \sum_j w_j f_i = \frac{\text{Error}}{\sum_j w_j} \quad (9)
\]

The task for learning algorithm for this architecture is to tune all the modifiable parameters to make ANFIS output match the training data.

4. STUDY AREA AND DATA COLLECTION

The study was carried out for Gandhinagar district which is located between 23.22° N Latitude and 72.68° E Longitude. Gandhinagar has a monsoon climate with three main seasons: summer, monsoon and winter. The climate is generally dry and hot outside of the monsoon season. The weather is hot to severely hot from March to June when the maximum temperature stays in the range of 36 to 42 °C (97 to 108 °F), and the minimum in the range of 19 to 27 °C (66 to 81 °F). It is warm from December to February, the average maximum temperature is around 29 °C (84 °F), the average minimum is 14 °C (57 °F), and the climate is extremely dry. The southwest monsoon brings a humid climate from mid-June to mid-September.

For this study, monthly climatic data are collected of weather station Ghandhingar. Monthly metrological data such as Mean Air Temperature, Relative Humidity, Wind Speed and Rainfall of ten years are collected.

5. METHODOLOGY

The Rainfall forecasting is nonlinear system so ANFIS model has been developed with a view to predict the rainfall based on previous year data. Model has been designed, trained and tested with different membership functions and different number of members. After Defining ANFIS Model, it is run with various FIS Algorithm, error tolerance and number of epochs to analyse the effect of all these parameters on RMSE and predict rainfall. The parameter optimization is done in such a way during training session that the error between the target and the actual output is minimized.

In this study, mean temperature (Tmean), relative humidity (RH), wind speed (U) and rainfall data are used to develop ANFIS models for rainfall forecasting. The data are divided into sets of 70-30% ratio. For example 70% data for training period and 30% data for validation period to develop ANFIS model.

ANFIS models are developed using different method, membership function and different alternative of inputs as shown in Table 1.

<table>
<thead>
<tr>
<th>Particular</th>
<th>Method</th>
<th>Input</th>
<th>Membership Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model-1</td>
<td>Hybrid</td>
<td>Tmean + RH + U</td>
<td>7 MF</td>
</tr>
<tr>
<td>Model-2</td>
<td>Hybrid</td>
<td>Tmean + RH + U</td>
<td>6 MF</td>
</tr>
<tr>
<td>Model-3</td>
<td>Back</td>
<td>Tmean + RH + U</td>
<td>7 MF</td>
</tr>
<tr>
<td>Model-4</td>
<td>Back</td>
<td>Tmean + RH + U</td>
<td>6 MF</td>
</tr>
<tr>
<td>Model-5</td>
<td>Hybrid</td>
<td>Tmean +RH</td>
<td>9 MF</td>
</tr>
<tr>
<td>Model-6</td>
<td>Hybrid</td>
<td>Tmean +RH</td>
<td>8 MF</td>
</tr>
<tr>
<td>Model-7</td>
<td>Back</td>
<td>Tmean +RH</td>
<td>6 MF</td>
</tr>
<tr>
<td>Model-8</td>
<td>Back</td>
<td>Tmean +RH</td>
<td>5 MF</td>
</tr>
</tbody>
</table>

In this study eight models are developed. Model 1 is developed using three inputs mean temperature, relative humidity and wind speed which are model-1 hybrid method with 7 MF, model-2 hybrid method with 6 MF, model-3 back...
propagation method with 7 MF and model-4 back propagation method with 6 MF. Model 5-8 developed using two inputs temperature and relative humidity which are model-5 hybrid method with 9 MF, model-6 hybrid method with 8 MF, model-7 back propagation with 6 MF, model-8 back propagation method with 5 MF.

The result obtains from Gandhinagar station evaluated by following evaluation parameters:

Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q(i) - \bar{Q})^2}$$  \hspace{1cm} (10)

Correlation coefficient (r):

$$r = \frac{\sum_{i=1}^{n}(Q(i) - \bar{Q})(Q(i) - \bar{Q})}{\sqrt{\sum_{i=1}^{n}(Q(i) - \bar{Q})^2 \sum_{i=1}^{n}(Q(i) - \bar{Q})^2}}$$  \hspace{1cm} (11)

Discrepancy ratio (D):

$$D = \frac{\sum_{i=1}^{n}(Q(i) - Q(i))}{\sum_{i=1}^{n}Q(i) - \bar{Q}}$$  \hspace{1cm} (12)

Where Q(i) is the n predicted rainfall; Q(i) is the n observed rainfall ; Q is the mean of the observed rainfall and $\bar{Q}$ is the mean of the predicted rainfall.

6. RESULTS AND ANALYSIS

ANFIS models are introduced into rainfall science as a powerful, flexible, and statistical modelling technique to address complex pattern recognition problems. The forecasting performances of the considered models are evaluated using four evaluation parameters namely; root mean square error (RMSE) correlation coefficient (r) and coefficient of determination (R²) and discrepancy ratio (D) are given in Table 2.

<table>
<thead>
<tr>
<th>Station-Gandhinagar</th>
<th>Model-1 Hybrid method - 7 MF</th>
<th>Model-2 Hybrid method - 6 MF</th>
<th>Model-3 Back propagation -7 MF</th>
<th>Model-4 Back propagation 6 - MF</th>
<th>Model-5 Hybrid method 9 - MF</th>
<th>Model-6 Hybrid method - 8 MF</th>
<th>Model-7 Back propagation - 6 MF</th>
<th>Model-8 Back propagation -5 MF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training 0.99</td>
<td>Training 0.99</td>
<td>Training 0.76</td>
<td>Training 0.88</td>
<td>Training 0.99</td>
<td>Training 0.99</td>
<td>Training 0.81</td>
<td>Training 0.68</td>
</tr>
<tr>
<td></td>
<td>Validation 0.92</td>
<td>Validation 0.92</td>
<td>Validation 0.58.31</td>
<td>Validation 0.20</td>
<td>Validation 0.99</td>
<td>Validation 0.72</td>
<td>Validation 0.42</td>
<td>Validation 0.32</td>
</tr>
<tr>
<td></td>
<td>R² 0.99</td>
<td>R² 0.99</td>
<td>R² 0.99</td>
<td>R² 0.78</td>
<td>R² 0.99</td>
<td>R² 0.51</td>
<td>R² 0.67</td>
<td>R² 0.47</td>
</tr>
<tr>
<td></td>
<td>D 0.99</td>
<td>D 1.33</td>
<td>D 1.30</td>
<td>D 1.26</td>
<td>D 1.00</td>
<td>D 0.60</td>
<td>D 1.48</td>
<td>D 1.40</td>
</tr>
<tr>
<td></td>
<td>RMSE 0.95</td>
<td>RMSE 76.27</td>
<td>RMSE 131.24</td>
<td>RMSE 98.13</td>
<td>RMSE 125.79</td>
<td>RMSE 258.10</td>
<td>RMSE 140.54</td>
<td>RMSE 141.04</td>
</tr>
</tbody>
</table>

Models 1 to 4 have three inputs and Models 5 to 8 have two inputs. Model 1 is developed using hybrid method and 7 generalized bell membership functions; having RMSE, r, R² and D; 0.95, 0.99, 0.99, 0.99 for training and 76.27, 0.92, 0.85, 1.33 respectively for validation. Model 2 is developed using 6 generalized bell membership function and have RMSE, r, R² and D; 3.02, 0.99, 0.99, 0.99 for training and 162.71, 0.80, 0.64, 0.81 respectively for validation. i.e. Model having seven MF gives better result with same input and method. Model 3 developed using back propagation method and 7 generalized bell membership functions with three inputs having RMSE, r, R² and D; 131.24, 0.76, 0.58, 1.3 for training and 121.47, 0.46, 0.21, 1.08 respectively for validation. Model 4 developed using 6 generalized bell membership functions have RMSE, r, R² and D; 98.13, 0.88, 0.78, 1.26 for training and 170.07, 0.20, 0.04, 2.06 for
validation respectively. Model 3 and 4 have less value of r, R² and D compared to Model 1 and 2. i.e. Hybrid method gives better results compared to Back propagation method using three inputs.

Model 5 is developed using hybrid method and 9 generalized bell membership functions which have RMSE, r, R² and D; 1.87, 0.99, 0.99, 1.0 for training and 125.79, 0.77, 0.60, 0.80 for validation respectively. Model 6 developed using 8 generalized bell membership functions, have RMSE, r, R² and D; 0.94, 0.99, 0.99, 0.99 for training and 258.10, 0.72, 0.51, 0.6 for validation respectively. Model 5 and 6 are developed using hybrid method but with two inputs and MF 9 and MF 8 respectively. But RMSE is more than Model 1 and 2. Model 7 is developed using back propagation method and 6 generalized bell membership functions with two inputs having RMSE, r, R² and D; 140.54, 0.81, 0.67, 1.48 for training and 136.13, 0.42, 0.17, 1.27 for validation respectively. Model 8 developed using 6 generalized bell membership functions have RMSE, r, R² and D; 179.86, 0.68, 0.47, 1.9 for training and 141.04, 0.32, 0.10, 1.4 for validation respectively. For Model 6 and 7 values of r, R² and D are very less hence these models are not performing well. i.e. Back propagation method is not giving good results.

From the above discussion, model 1 is having minimum RMSE for training and validation, maximum r and R² for training and validation, D is near to 1 for training and validation.

Figure 3 and 4 shows the comparison of observed and predicted rainfall for training and validation using grid partition algorithm and hybrid optimal method (Model-1). Figure 5 and 6 shows scatter plot of observed and predicted rainfall for training and validation for Model-1.
Figure 6 Scatter plot of observed and predicted rainfall for validation for Model-1

Model-1 has the highest correlation coefficient 0.99 in training period as well as 0.92 for validation, Root mean square error for training period is 0.95 and for validation it is 76.27. The value of discrepancy ratio is 0.99 for training and 1.33 for validation period.

7. CONCLUSION

- Rainfall forecasting is important for many areas of human activities such as agriculture, water resources, hydro-electric power projects, happening of droughts or floods and others.
- This paper modelled the complex multi-dimensional behaviors of monthly monsoon rainfall for station namely Gandhinagar using ANFIS.
- For this study, different algorithm, different number of membership functions and different inputs are used to develop ANFIS model. Model-1 is developed using hybrid method with 7 MF and generalized bell type membership function, which has the highest correlation coefficient 0.99 in training and 0.92 for validation and $R^2$ is 0.99 for training and 0.85 for validation. The root mean square error of model-1 is 0.95 for training and 76.27 for validation and discrepancy ratio of model-1 is near one.
- Thus the performance of model-1 is best compared to all other models and is suitable for Gandhinagar station for forecasting monsoon rainfall using ANFIS tool.
- It can be concluded that Hybrid method gives better results than Back Propagation Method for Gandhinagar station.

REFERENCES

AUTHOR’S BIOGRAPHY

Dr. Falguni Parekh has completed B.E. (Civil-IWM), M.E. (Civil) in Irrigation Water Management and Ph.D. in Civil Engineering from The M.S. University of Baroda, Vadodara, Gujarat, India. She is serving as Offg. Director and Associate Professor in Water Resources Engineering and Management Institute, Faculty of Technology and Engineering, The M. S. University of Baroda. She has 17 years of research experience and 16 years of teaching experience. Her areas of research include Reservoir Operation, Soft computing techniques, Micro Irrigation, Benchmarking of Irrigation Projects, Climate Change and its Impact on Water Resources, Rain Water Harvesting, and Low cost Micro Irrigation Systems. Dr. Parekh is Life Member of various professional bodies like Indian Society of Hydraulics, Indian Water Resources Society, Association of Hydrologists of India and Association of Agrometeorologists and Indian Society of Geomatics. She is chairman of Board of studies of WREMI and member of Faculty Board of Tech. & Engg. She is also joint secretary of Gujarat Chapter of Association of Hydrologists of India. She worked as Principal Investigator for Research Project funded by Gujarat Council on Science and Technology and completed five institutional consultancy projects. She is recipient of “Prof. S.C. Puranik Young Scientist Award” for the award winning paper in 2004 by Association of Hydrologists of India. She has published 32 Research Papers in various International/ National Journals/ Conferences.

Jignesh Patel completed B.E. (Civil) from Lalbhai Dalpatbhai College of Engineering, Ahmedabad. He is pursuing M.E. with specialization in Water Resource Engineering from Water Resource Engineering and Management Institute (WREMI), The Maharaja Sayajirao University, Samiala- Baroda. He is a life member of Indian Society of Geomatics.