ABSTRACT

Local Monsoonal Precipitation forecast (LMP) based on temperature, relative humidity and Pan Evaporation is very crucial factor not only in agricultural activities but also in many other areas such as airplane authority of India (AAI), large scale organization of an event like common wealth game of a country or to decide the date of large scale play like cricket and other matches, even to decide the date of examination, educational institutions have to take into consideration the monsoonal rainfall activities.

In this study, the main emphasis is given on Agricultural weather forecasting so that the farmers of India can plan their farming activities such as irrigation, sowing and cutting the ripe crops etc. in anticipation and can increase the productivity and hence make an increase in national GDP of India, which is the main significance of this thesis.

Soft computing is an innovative approach to formulate artificially intelligent systems which can think and processes like human experience within a given range of parameters. Neural networks work as human neuron and have the capability to think and learn so that they can perform better in rapidly changing environments.

Adaptive Neural networks have the capability to adapt themselves according to the input data range and learn themselves by adjusting the weights in such a way that they can predict the output variable in some another input data irrespective of location and year of prediction. This can be utilized in two ways. First, if input data is changed to the next year of the current area under study i.e. Hisar, then the proposed model will predict the monsoonal rainfall of next year of the same location i.e. Hisar. Secondly, if we change the area under study let us say, Istanbul, Pakistan then the proposed model is so much flexible that it will adapt itself to the different input data set and can predict the monsoonal rainfall of Istanbul.

Moreover, the proposed model based on Adaptive Neuro Fuzzy is easy to implement and produces desirable forecasting result by training on the given dataset in comparative less time and space complexity.

The technique of soft computing such as Fuzzy Theory, Neural Networks and Adaptive Neuro Fuzzy Inference System which is the hybridization of neuro and fuzzy
techniques are used in this study to test the performance of which technique for Local Monsoonal Precipitation forecasting is most effective and have minimum MSE and RMSE.

In this study 40 years (1970-2010) of daily data of CCS Haryana Agriculture University (HAU) observatory, Hisar, Haryana is taken as input data, observed daily at 7:00 am IST and 2:30 p.m. ISI by expert observers. The whole data is collected, processed to make it normalized in such away that all values comes under the range of closed interval [0,1] and then it is divided into three parts as follows:

- 60% of Primary Data as Training Data Set (23 years)
- 20% of Primary Data as Checking Data Set to avoid model over fitting (8 years)
- 20% of Primary Data as Testing data Set (8 years)

First of all, this data is applied on multidimensional response surface tool rstool of MatLab and predicted value is compared with observed value in terms of output deviation \( \Delta \). It has been concluded that linear model shows the lesser output deviation 0.10491 as compared to other Models Pure Quadratic, Full Quadratic and interaction model having output deviation as 0.24476, 0.34646 and 0.20094 respectively. The best curve fit of successful linear model is given by equation (i)

\[
y = 0.80021 - 0.059745x_1 - 0.15918x_2 - 0.27851x_3 - 0.50302x_4 \quad \ldots \quad (i)
\]

Where \( X_1 \) is Temperature-Temp (measured in °C), \( X_2 \) is Average vapor Pressure- AVP (measured in mm), \( X_3 \) is Relative Humidity-RH (measured in %), \( wX_4 \) is Pan Evaporation-Pan Evap (measured in mm) and \( Y_1 \) is the output variable Local Monsoonal Precipitation – LMP (measured in mm).

However, **Root Mean Squared Error (RMSE)** in this plot is \( 3.3939e-02 \).

* The data of year 1987 is not included in training data set as the year 1987 was declared as draught year by Indian Meteorological Department (IMD, New Delhi).
Then the same data is uploaded in MatLab Anfis Editor as input data and a lot of Fuzzy inference systems are created in ANFIS editor to test which architecture of Anfis suits best to predict Local Monsoonal Precipitation (mm).

A lot of Experiments has been done on various data size, types and on various ANFIS architecture with different number of epochs and various error tolerance values. Various experiments have been done to verify which data set to train the neurons gives maximum reliability. The essence of all these experiments results in a successful and comparatively low cost Model in terms of space, time and cost as compared to statistical models.

A Total of Six experiment got succeeded in Anfis editor and result in fruitful prediction of output value. First two experiments, Experiment 2a and 2b, are done to check whether grid partitioning method is best or the subtractive clustering algorithm is best. This has been observed that with 40 epochs and 1.0e-05 error tolerance value grid partitioning algorithm on (3, -3, -3, -3, -1) Anfis Architecture Predict output with RMSE 1.7624e-02 while in subtractive clustering algorithm, with same 40 epochs and 1.0e-05 error tolerance value and Membership function are taken automatically depending upon the range of clusters and other parameters, predict output with RMSE 2.6226e-02. Since the value of RMSE is much less in grid partitioning algorithm Therefore in further experiment, grid partitioning algorithm is used to generate Fuzzy Inference Structure with hybridization of SD (Steepest Deviation) and LSE (Least Square Estimator).

Then to check the effect of error tolerance two experiments (experiment 2c(i) with two sub parts) are done. The same data is applied on same Number of membership function and other values as that of first two experiments. 40 epochs are taken in both experiments but with different error tolerance value 1.0e-10 and 1.0e-02 resp. and it is observed that the data shows higher efficiency when error tolerance is decreased. The Value of root mean squared error (RMSE) between predicted output and observed output got decreased significantly with decreasing value of error tolerance.

In experiment 2c (i) part I, at 1.0e-10 error tolerance, the value of RMSE is 9.1077e-02 while in experiment 2c (i) part II, at 1.0e-02 error tolerance, the value of RMSE is 9.9372-02.
This is visible clearly from these experiments 2c(i)- part I and part II, that decreasing the error tolerance, RMSE decreases significantly, however time and space complexity increases which may be accepted in case of such a sensitive prediction of monsoonal rainfall, that may sometime save thousands of lives if predicted correctly in advance.

Next, the effect of no. of epochs has been checked. Again two experiments are done as experiment 2c (ii) - part I and part II. Data, Membership function, its type and all other value are taken same as that of experiment 2c(i). Value of error tolerance is taken as same 1.0e-02 in both the experiments but with different number of epochs.

In experiment 2c (ii) part I, with 40 numbers of epochs, the value of RMSE is 9.1077e-02 while in experiment 2c (ii) part II, with 10 numbers of epochs, the value of RMSE is 1.2238e-01.

It is concluded that if we increase no. of epochs then RMSE decreases slightly, but with more space and time complexity. Now, it depends upon the level of predicted area and the amount spent on research program whether it is required to increase the slight performance of prediction at the cost of high space, time and hence monetary complexity. If more finance is provided to the researcher then it will be beneficial to increase number of epochs.

Last three experiments are the main experiment of this research. They are the true models if applied on data of Hisar then one can predict, at a given level of significance, the coming monsoonal seasonal rainfall whether it will be a heavy flood or a draught which may further help the farmers in doing their farming activities in advance.

Experiment 2d, 2e, 2f are done to finalize the architecture of ANFIS Model.

In experiment 2d, ANFIS Model (3,-3,-3,-5,-1) with 100 epochs the value of RMSE is 4.3633e-04, while in experiment 2e, ANFIS Model (3,-5,-5,-3,-1) with 25 epochs and 1.0e-03 error tolerance, the value of RMSE is 9.7795e-04 and in experiment 2f, ANFIS with model (5-3-3-5,-1) with 40 epochs, and 1.0e-03 error tolerance, the value of RMSE is 6.245e-04.
While comparing the RMSE of the three ANFIS models in Experiment 2d, 2e and 2f, it is concluded that the minimum RMSE is 4.3633e-04 in ANFIS Model (3,-3,-3,-5,-1) which is trained for 100 epochs and nearly zero error tolerance.

Finally, the result of ANFIS model and Multidimensional Response Surface Tool are compared, as shown in table 4.3.3 and it is clear that RMSE 4.3633e-04 is much lesser in (3,-3,-3,-5,-1) Architecture of ANFIS Model than RMSE 3.3939e-02 in linear model.

<table>
<thead>
<tr>
<th>Model</th>
<th>Successful curve fit and Architecture</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical Model</td>
<td>Linear, $y = 0.80021 - 0.059745x_1 - 0.15918x_2 - 0.27851x_3 - 0.50302x_4$</td>
<td>3.3939e-02</td>
</tr>
<tr>
<td>ANFIS Model</td>
<td>(3,-3,-3,-5,-1) Architecture</td>
<td>4.3633e-04</td>
</tr>
</tbody>
</table>

Table 4.3.3: Comparison of Statistical and ANFIS Models for minimum error rate

Therefore, ANFIS is a better predictor of monsoonal precipitation as compared to other statistical and neuro fuzzy models.