



Weather Forecasting using Soft Computing and Statistical Techniques

¹Monika Sharma, ²Lini Mathew, ³S. Chatterji

¹Senior lecturer, Department of Electrical & Electronics Engineering, RKGIT, Ghaziabad, UP, India

² Associate Professor, Department of Electrical Engineering, NITTTR, Chandigarh, India

³ Professor, Department of Electrical Engineering, NITTTR, Chandigarh, India

ABSTRACT: Weather forecasting is the application of science and technology to predict the state of the atmosphere for a future time at a given location. It is carried out by collecting quantitative data about the current state of the atmosphere and past and/or present experiences. In this study Adaptive Neuro-Fuzzy Inference System (ANFIS) and Multiple linear regression model were used to analyze metrological data sets obtained from the metrological station. The Multiple linear regression model is simple due to the fact that it uses simple mathematical equation using Multiple Linear Regression (MLR) equations that can be easily understood by a medium educated farmer. Adaptive Neuro-Fuzzy Inference Systems (ANFIS) combines the capabilities of Artificial Neural Networks (ANN) and Fuzzy Inference Systems (FIS) to solve different kinds of problems. The data covers a five year period (2008-2012) were for the monthly means of minimum and maximum temperature, wind speed, and relative humidity and mean sea level pressure (MSLP). The results showed that both models could be applied to weather prediction problems. The performance evaluation of the two models that was carried out on the basis of root mean square error (RMSE) showed that the ANFIS model yielded better results than the multiple linear regression (MLR) model with a lower prediction error.

KEYWORDS: ANFIS, Regression, fuzzy inference system, RMSE, weather forecasting.

I. INTRODUCTION

Weather Forecasting is an important and necessary area of investigation in human life. Weather for future is one of the most important attributes to forecast because agriculture sectors as well as many industries are largely dependent on the weather conditions. Weather conditions are required to be predicted not only for future planning in agriculture and industries but also in many other fields like defence, mountaineering, shipping and aerospace navigation etc. It is often used to warn about natural disasters are caused by abrupt change in climatic conditions.

At macro level, weather forecasting is usually done using the data gathered by remote sensing satellites. Weather parameters like maximum temperature, minimum temperature, extent of rainfall, cloud, conditions, wind streams and their directions, are projected using images and data taken by these meteorological satellites to access future trends. In this paper Multiple Linear Regression (MLR) and adaptive neuro fuzzy inference system (ANFIS) are used to develop models for forecasting weather parameters. The proposed models are capable of forecasting the weather conditions for a particular station using the data collected locally.

In the data set, there are five parameters: Maximum Temperature, Minimum Temperature, Humidity and Wind speed, mean sea level pressure(MSLP) and this data are only on Delhi's weather condition. For analysis and forecast, we applied ANFIS and multiple linear regression techniques on this data and finally the performance of these two models is compared on the basis of root mean square error (RMSE).

In this paper, ANFIS MATLAB Fuzzy Logic Toolbox is used to design ANFIS model. These tools apply fuzzy inference techniques to data modeling. The ANFIS toolbox function constructs a fuzzy inference system (FIS) whose membership function parameters are tuned (adjusted) using back propagation algorithm. This allows our fuzzy systems to learn from the data they are modeling.



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2014

II. LITERATURE SURVEY

Oyediran et al. in [1] compared Adaptive Neuro-Fuzzy Inference System (ANFIS) and Multi-Layer Perceptron (MLP) Artificial Neural Network (ANN) models to analyze metrological data sets obtain from the metrological station. The data covers a ten year period (2002-2012) were for the monthly means of minimum and maximum temperature, rainfall, wind run, and relative humidity. The performance evaluation of the two models that was carried out showed that the ANFIS model yielded better results than the MLP ANN model with a lower prediction error.

Gholam et al. in [2] revealed that soft computing techniques are promising and efficient. The root mean square error by using Fuzzy inference system model was obtained 52 mm. Further it is also stated that unlike conventional artificial intelligence techniques the guiding principal of soft computing is to exploit tolerance for imprecision, uncertainty, robustness, partial truth to achieve tractability and better rapport with reality.

Zhou et al. in [3] revealed that Forecast of wind speed is very important for making out dispatch scheme of power system and operation in higher reliability, according to the forming mechanism of wind, its influencing factors and its inherent variation rule, one of pattern recognition called as adaptive neuron-fuzzy inference system (ANFIS) is used in wind speed forecast. The hybrid algorithm is used to train the parameter of fuzzy interference system, with the train samples, the model is constructed, and anticipated wind speed is easily gotten. The forecast result shows that applying ANFIS to the practice would be valid.

Tektaş in [4] presented a comparative study of statistical and neuro-fuzzy network models for forecasting the weather of Goztepe, Istanbul, Turkey. For developing the models, nine year data (2000-2008) comprising daily average temperature (dry-wet), air pressure, and wind-speed have been used. Adaptive Network Based Fuzzy Inference System (ANFIS) and Auto Regressive Moving Average (ARIMA) models have been applied. The criteria of performance evaluation are calculated for estimating and comparing the performances of ARIMA and ANFIS models. The performance comparisons of ANFIS and ARIMA models due to MAE (Moving Average Error), RMSE (Root-Mean-Square error) criteria, indicate that ANFIS yields better results.

Paras et al. in [5] developed a Multiple Linear Regression (MLR) which is used to develop a model for forecasting weather parameters. The proposed model is capable of forecasting the weather conditions for a particular station using the data collected locally. The data is processed to obtain some statistical indicators to extract the hidden information present in the time series. On the basis of correlation, features are chosen as inputs to the models. Regression equations are obtained for the parameter to be forecast, which is termed as target. The whole data set is divided in two parts, the first is used to obtain MLR equations and remaining is used for testing the model. The power of MS Excel has been used to process the data and present the results in simply understandable form.

III. MULTIPLE LINEAR REGRESSION (MLR) MODEL

Regression explains the nature of relationship, i.e., the average probable change in one variable given by a certain amount of change in the other variable. The general equation of regression of Y on X is

$$y = \alpha + \beta X + e \tag{1}$$

This equation is known as the mathematical model for linear regression. As the special case the form $Y = \alpha + \beta X$ is called the deterministic model.

Multiple Linear Regression equation, which describes the linear relationship with set of dependent variable Y, and k sets of independent variables $X_1, X_2, X_3, \dots, X_k$ is,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + e \tag{2}$$

where Y is predictant and $X_1, X_2, X_3, \dots, X_k$ are the predictors. 'e' is the error which is distributed normally with zero mean and variance s^2 , i.e. $e \sim N(0, s^2)$.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2014

To develop the proposed model, the parameters $\alpha, \beta_1, \beta_2, \dots, \beta_k$ are estimated using the training sample sets. $X_1, X_2, X_3, \dots, X_k$ are among the extracted features mentioned earlier.

The coefficient of determination, r^2 , explains the extent by which the variation of dependent variable Y is being expressed by the independent variable X. It is obtained as

$$r^2 = \frac{\text{variance explained}}{\text{total variance}} \quad (3)$$

A high r^2 shows that there exists a linear relationship between the two variables. If $r^2=1$, it indicates the perfect relationship between the two variables.

IV. ADAPTIVE NEURO FUZZY INFERENCE SYSTEM (ANFIS)

An adaptive network is a network of nodes and directional links. Associated with the network is a learning rule: for example back propagation. It's called adaptive because some, or all, of the nodes have parameters which affect the output of the node. The networks learn the relationship between the inputs and outputs. The ANFIS approach learns the rules and membership functions from data. The ANFIS architecture is presented in figure 2.

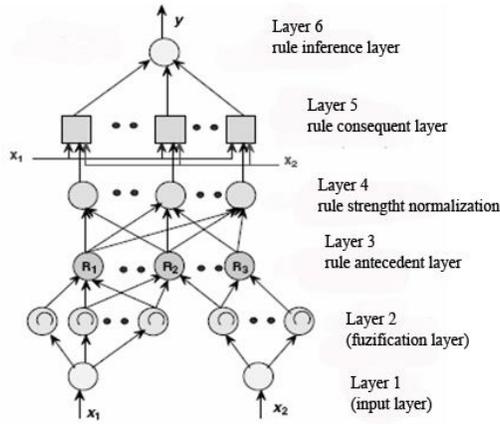


Fig.1. An ANFIS Architecture for a Two Rule Sugeno System

A Two Rule Sugeno ANFIS has rules of the form

$$\text{If } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ THEN } f_1 = p_1x + q_1y + r_1 \quad (4)$$

$$\text{If } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ THEN } f_2 = p_2x + q_2y + r_2 \quad (5)$$

When training the network there is a forward pass and a backward pass. The forward pass propagates the input vector through the network layer by layer. In the backward pass, the error is sent back through the network in a way similar to back propagation.

- **Layer 1**

In Layer 1, the output of each node is

$$O_{1,i} = \mu_{B_{i-2}}(y) \quad \text{for } i = 3, 4 \quad (6)$$

And $O_{1,i}(x)$ is membership functions grade for x and y . The membership functions could be any shape. Using the Gaussian membership function given by

$$\mu_A(x) = \frac{1}{1+x^2} \quad (7)$$

- **Layer 2**



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2014

Every node in this layer is a fixed node labeled Π , representing the firing strength of each rule and is calculated by the fuzzy AND connective of product of the incoming signals by

$$O_{2,i} = w_i = \mu_{A_i}(x) \mu_{B_i}(y) \quad \text{for } i=1, 2 \quad (8)$$

Where, $\mu_{A_i}(x)$ and $\mu_{B_i}(x)$ are membership grades of fuzzy sets A, B and also w_i is firing strength of each rule.

- **Layer 3**

Every node in this layer is a fixed node labeled N, representing the normalized firing strength of each rule. The i_{th} node calculates the ratio of the i_{th} rule's firing strength to the sum of two rule's firing strengths by using eq.(6)

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i=1, 2 \quad (9)$$

where, \bar{w}_i is normalized firing strength that is the ratio of the i_{th} rule's firing strength (w_i) to the sum of the first and second rule's firing strengths (w_1, w_2).

- **Layer 4**

Every node in this layer is an adaptive node with a node function in eq.(10), indicating the contribution of i_{th} rule toward the overall output.

$$O_{4,i} = \bar{w}_i z_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (10)$$

Where, z_i is equal to $(p_i x + q_i y + r_i)$ and also p_i, q_i and r_i are consequent parameters.

- **Layer 5**

The single node in this layer is a fixed node labeled Σ , indicating the overall output as the summation of all incoming signals calculated by

$$Z = \sum_i \bar{w}_i z_i = \frac{\sum_i w_i z_i}{\sum_i w_i} \quad (11)$$

Where, Z is the summation of all incoming signals.

V. MATERIALS AND METHODS

A. Data Preprocessing

The case data for the study was obtained from the weather website and it covers a period of 5 years from January 2008 to December 2012. The preprocessing of the data was first carried. Missing values were replaced with zeros. The meteorological dataset, their type and description are presented in Table 1.1.

Table1.1. List of Different Input Variables

S.No.	Input Variables	Units
1.	Max. Temperature	Deg.C
2.	Min. Temperature	Deg.C
3.	Relative Humidity	%
4.	Wind speed	Kmph
5.	MSLP	Hpa

B. The Performance Criteria used for Evaluating the Models

Root Mean-Squared Error-- This is simply the square root of the mean squared error. The mean-squared error gives the error value the same dimensionality as the actual and predicted values. It is given by

$$RMSE = \sqrt{\left(\frac{1}{P} \sum_{i=1}^P [(P_M) - (P_{es})]^2\right)} \quad (12)$$



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2014

VI. RESULTS AND DISCUSSION

In this study, MLR and ANFIS models are designed to forecast various weather parameters. Weather parameters to be predicted are maximum temperature, minimum temperature, relative humidity, wind speed and MSLP. For both models yearly data is divided into four seasons as shown in table1.2. and models are trained using this seasonal data as input.

Table1.2. Yearly Data Division into Four Seasons

S. No.	Season	Duration
1.	Winter	1 December – 28 February
2.	Spring	1 March – 31 May
3.	Summer	1 June – 31 August
4.	Autumn	1 September – 30 November

The best model suitable to forecast weather is evaluated by computing the root mean square error between the exact and predicted values. Theoretically, a prediction model is accepted as ideal when RMSE is small. Results for all weather parameters in terms of RMSE are shown in following table1.3.

Table.1. 3. Performance Comparison of ANFIS and Regression Models

Date	Max.temp.		Min.temp.		Relative humidity		Wind speed		MSLP	
	RMSE		RMSE		RMSE		RMSE		RMSE	
	Regression	ANFIS	Regression	ANFIS	Regression	ANFIS	Regression	ANFIS	Regression	ANFIS
01-01-2013	3.16	3.11	12.60	2.10	10.2	10.83	4.06	148.76	5.11	0.15
01-02-2013	2.72	2.60	14.60	2.64	15.93	16.57	2.97	76.74	6.26	1.89
01-03-2013	2.07	1.52	13.53	1.52	14.10	15.91	2.79	2.72	6.55	8.57
01-04-2013	2.54	2.20	14.69	2.43	13.10	13.51	3.11	2.23	6.81	6.81
01-05-2013	1.70	1.72	14.64	2.18	12.87	12.89	3.85	4.70	5.15	4.31
01-06-2013	3.15	2.81	10.15	2.53	12.22	12.69	4.29	4.79	3.57	30.07
01-07-2013	2.48	2.26	7.94	1.02	12.90	13.41	3.26	3.64	5.33	7.66
01-08-2013	2.09	1.93	7.48	0.99	14.67	14.69	4.61	4.26	6.34	0.11
01-09-2013	1.73	1.55	9.06	2.01	13.27	12.85	5.07	6.17	2.38	8.64
01-10-2013	1.15	1.32	12.79	2.31	6.28	8.57	2.66	2.77	7.62	25.86
01-11-2013	1.30	1.30	12.55	1.78	10.70	11.21	2.17	1.33	6.74	11.99
01-12-2013	2.25	2.20	14.58	2.49	10.12	10.40	1.83	144.20	4.53	1.00

VII.CONCLUSION

Based on the obtained results, it can be concluded that both of these models were able to capture the dynamic behavior of the weather data, resulting in a more compact and natural internal representation of the Temperature, MSLP, Wind speed and Relative humidity information contained in the weather profile.

However, regarding prediction accuracy, the ANFIS is highly appreciated for temperature forecasting, relative humidity, wind speed. For MSLP only, Regression model performed better than ANFIS model.



ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2014

REFERENCES

- [1] Oyediran O. F., Adeyemo A. B., “Performance Evaluation of Neural Network MLP and ANFIS models for Weather Forecasting Studies”, African Journal of Computing & ICT, IEEE, Vol. 6, No. 1, March 2013.
- [2] Fallah-Ghalhary, Gholam Abbas; Mousavi-Baygi, Mohammad, Nokhandan, Majid Habibi, “Annual Rainfall Forecasting by Using Mamdani Fuzzy Inference System”, Research Journal of Environmental Sciences, Vol. 3, No. 4, pp 400-407, August 2009.
- [3] Hui Zhou, Beijing Jiaotong, Mei Huang, Xinfhua Wu, “Forecast of Wind speed and power of wind generator based on pattern recognition”, Proceedings of the IEEE International Conference on Industrial Mechatronics and Automation, Chengdu, pp 504-508, May, 2009.
- [4] Mehmet Tektaş, “Weather Forecasting Using ANFIS and ARIMA MODELS: A Case Study for Istanbul”, Environmental Research, Engineering and Management, No. 1(51), pp 5-10, march, 2010.
- [5] Paras, Sanjay Mathur, “Simple Weather Forecasting Model Using Mathematical Regression”, Indian Research Journal of Extension Education Special Issue, Vol.1, January, 2012.
- [6] Mathur S., Shukla A.K., Pant R.P., “A Comparative Study of Neural Network and Regression Models for Estimating Stock Prices”, Proceedings of the IEEE International Conference on Optimization Techniques and its Applications in Engineering and Technology, September, 2001.
- [7] Rahman M., Islam A.H.M.S., Nadvi S.Y.M., Rahman, R.M., “Comparative study of ANFIS and ARIMA model for weather forecasting in Dhaka”, Proceedings of the IEEE International Conference on Informatics, Electronics & Vision (ICIEV), Dhaka, pp 1-6, May, 2013. .
- [8] Jiansheng Wu, A Novel Nonlinear Ensemble Rainfall Forecasting Model Incorporating Linear and Nonlinear Regression”, Proceedings of the IEEE Fourth International Conference on Natural Computation (ICNC '08), Jinan, Vol.3, pp 34-38, October, 2008.
- [9] Khadangi E., Madvar H.R., Ebadzadeh M.M., “Comparison of ANFIS and RBF models in daily stream flow forecasting”, Proceedings of the IEEE 2nd International Conference on Computer, Control and Communication, Karachi, pp 1-6, February, 2009.
- [10] Awan J.A., Deg-hyo Bae, “Application of Adaptive Neuro-Fuzzy Inference System for dam inflow prediction using long-range weather forecast”, Proceedings of the IEEE Eighth International Conference on Digital Information Management (ICDIM), Islamabad, pp 247-251, September, 2013.
- [11] Deg-Hyo Bae, Dae Myung Jeong, Gwangseob Kim, “Monthly dam inflow forecasts using weather forecasting information and neuro-fuzzy technique”, Hydrological Sciences, vol.52, No.1, pp 99-113, January, 2010.