Prediction of Indian Monsoon Rainfall by Interval based Simplified High Order Fuzzy Time Series Model

Amit Kumar Rana

***Abstract:** Rain is of uttermost importance for agriculture based economies. Most of the Asian countries, India in particular largely depend on a good rainfall. The prediction of rainfall will not only help government to make better future policies but also farmers and agro based companies can make better future management. Rainfall forecasting involves high degree of uncertainty and for such conditions fuzzy time series and other soft computing techniques are best to deal with. The utility of a forecasting method lies with the accuracy with the predicted values. In this paper rainfall prediction by fuzzy time series model is proposed in which two difference values of the interval corresponding to the fuzzified forecasted value is proposed. This model is tested on real time data of average monsoon rainfall in India. The predicted values are compared with Chen model. The results show that the proposed model have less error compared to Chen's model.*

***Keywords:** Difference intervals, Fuzzy relations (FR), Fuzzy sets (FS), Fuzzy time series model (FTSM)*

I. INTRODUCTION

Zadeh [5] was first to consider the uncertainty in mathematical formulations and presented FS theory. Since then FS become one of the best techniques in dealing with the daily life problems having vagueness and uncertainty in the information. Song and Chissom [8], [9] successfully implemented the idea of verbal variables for approximate reasoning using FTSM and applied it on Alabama University's enrollment data. Huarng [4] model improved the forecast of university enrollments using heuristic increasing and decreasing relation and tested this model on Taiwan Futures Exchange forecasting. Chen [10] used arithmetic operations instead of completed max-min operators and presented high order FTSM and got better results than Song and Chissom [8], [9]. Chen and Hsu [11] developed a new improved FTS method for forecasting enrollments. Optimal length of interval is the base for a good forecasting model. Singh [7] presented a review in which FTS based modeling techniques are discussed. Rana [1] studied on the rice production FTS Forecasting model. Panigrahi and Behera [12] in his study proposed an efficient computational model for forecasting using high order FTS. Rana [2] worked on time invariant models and presented a comparative study for forecasting crop production using time invariant FTS models. Chen [3] et al. used an important common problem of many countries and proposed a model for real time flood forecasting by FTSM. Bose and Mali [6] presented a survey on designing FTSM which provides a base for studying this soft computing technique. The proposed model uses a high order FTSM for forecasting a Indian monsoon rainfall using a 15 years historical data of rainfall from meteorological department of government of India.

Revised Manuscript Received on April 04, 2020.
Amit Kumar Rana, Assistant Professor, Department of Mathematics, Swami Vivekanand Subharti University, Meerut, Uttar Pradesh, India, ***Email:** akrana77@gmail.com*

II. PRELIMINARIES OF FS

Definition 1. FS A_i with membership function μ_{A_i} on $U = \{u_1, u_2, u_3 \ldots\ldots u_n\}$ is defined by

$$A_i = \frac{\mu_{A_1}(u_1)}{u_1} + \frac{\mu_{A_2}(u_2)}{u_2} + \frac{\mu_{A_2}(u_2)}{u_2} + \cdots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots + \frac{\mu_{A_n}(u_n)}{u_n}$$

Definition 2. Let FS $f_i(t), (i = 1,2,3 \ldots.)$ are defined on $Y(t)$ and $F(t)$ is the collection of all f_i then $F(t)$ is called FTS on $Y(t)$.

Definition 3. Suppose $F(t)$ is caused only by $F(t-1)$ and is denoted by $F(t-1) \to F(t) \Rightarrow$ a fuzzy relationship between $F(t)$ and $F(t-1)$ and $F(t)$ is caused only by $F(t-1)$. This can be expressed by the fuzzy relational equation: $F(t) = F(t-1) o R(t, t-1)$ where, '' o '' is Max–Min composition operator. Relation R is called model first-order of $F(t)$.Also if $R(t_1, t_1 - 1) = R(t_2, t_2 - 1)$, $\forall\, t_1 \neq t_2$ then $F(t)$ is called a time invariant FTS.

Definition 4. A n^{th}-order FTS is one in which $F(t)$ is caused by n fuzzy sets $F(t-n), F(t-n+1), \ldots F(t-1)$ and the fuzzy relationship is given by $A_{i_1}, A_{i_2}, A_{i_3} \ldots \ldots \ldots .. A_{i_n} \to A_j$

Here, $F(t-n) = A_{i_1}, , \ldots \ldots \ldots \ldots \ldots \ldots \ldots F(t-1) = A_{i_n}$

III. PROPOSED METHODOLOGY

Step 1. Define U as $U = [U_{min} - U_1, U_{max} - U_2]$ where U_1 and U_2 are two proper positive numbers and U_{min}, U_{max} are min. and max. values of real time rainfall data respectively.
Step 2. Construct equal length sub intervals $u_1, u_2, \ldots \ldots .. u_m$ from.
Step 3. Construct the FS A_i
Step 4. Fuzzify the data and establish FLR using If A_c , A_n are the fuzzy production for n^{th}, $(n+1)^{th}$ year then FLR is

$A_c \rightarrow A_n$ where A_c , A_n are called current and next state.
Step 5. Using FLR, obtain fuzzified output.
Step 6. defuzzified and obtained the forecasted value using

$[S_n{}^*]$ is corresponding interval u_n for which A_n has max. membership 1

$L[S_n{}^*]$ lower bound of u_n , $U[S_n{}^*]$ upper bound of u_n

$l[S_n{}^*]$ length of u_n , $M[S_n{}^*]$ mid value of u_n

For FLR $A_c \rightarrow A_n$

A_c : fuzzified data of n^{th} year,

A_n : fuzzified data of $(n+1)^{th}$ year

E_c : actual data of n^{th} year

E_{c-1} : actual data of $(n-1)^{th}$ year

E_{c-2} : actual data of $(n-2)^{th}$ year

F_n : crisp forecasted value of the $(n+1)^{th}$ year

Algorithm for forecasting rainfall
FLR for k^{th} to $(k+1)^{th}$ year is $A_c \rightarrow A_n$; $k = 3$ ...upto end of time series data
Calculating the differences as

$D_c = \left| |E_c - E_{c-1}| - |E_{c-1} - E_{c-2}| \right|$

$\alpha_c = E_c + \frac{D_c}{3}$ $\alpha'_c = E_c - \frac{D_c}{3}$, $\beta_c = E_c + \frac{2D_c}{3}$

$\beta'_c = E_c - \frac{2D_c}{3}$

$\gamma_c = E_c + D_c$ $\gamma'_c = E_c - D_c$

For $I = 1$ to 6
If $L[S_n{}^*] \leq \alpha_c \leq U[S_n{}^*]$ then $F_1 = \alpha_c$, $n_1 = 1$, Else $F_1 = 0$, $n_1 = 0$
Next : If $L[S_n{}^*] \leq \alpha'_c \leq U[S_n{}^*]$ then $F_2 = \alpha'_c$, $n_2 = 1$, Else $F_2 = 0$, $n_2 = 0$
Next : If $L[S_n{}^*] \leq \beta_c \leq U[S_n{}^*]$ then $F_3 = \beta_c$, $n_3 = 1$, Else $F_3 = 0$, $n_3 = 0$
Next : If $L[S_n{}^*] \leq \beta'_c \leq U[S_n{}^*]$ then $F_4 = \beta'_c$, $n_4 = 1$, Else $F_4 = 0$, $n_4 = 0$
Next : If $L[S_n{}^*] \leq \gamma_c \leq U[S_n{}^*]$ then $F_5 = \gamma_c$, $n_5 = 1$, Else $F_5 = 0$, $n_5 = 0$
Next : If $L[S_n{}^*] \leq \gamma'_c \leq U[S_n{}^*]$ then $F_6 = \gamma'_c$, $n_6 = 1$, Else $F_6 = 0$, $n_6 = 0$
Now $\Delta = F_1 + F_2 + F_3 + F_4 + F_5 + F_6$
If $\Delta = 0$ then $F_n = M[S_n{}^*]$, Else $F_n = (\Delta + M[S_n{}^*])/(n_1 + n_2 + n_3 + n_3 + n_4 + n_5 + n_6 + 1)$
Next k

IV. STEPWISE APPLICATION OF THE PROPOSED MODEL

Step 1. Universe of discourse U = [0940, 1240]
Step 2. Partitioning U into equal length of sub intervals with mid values as

$u_1 = [0940, 0960, 0980]$ $u_2 = [0980, 1000, 1020]$
$u_3 = [1020, 1040, 1060]$ $u_4 = [1060, 1080, 1100]$
$u_5 = [1100, 1120, 1140]$ $u_6 = [1140, 1160, 1180]$
$u_7 = [1180, 1200, 1220]$ $u_8 = [1200, 1220, 1240]$

Step 3. Defining FS A_n, $n = 1 \ldots 8$ as

A_1: Drought situation, A_2 : Very low rainfall,
A_3 : Low rainfall, A_4 : Average rainfall,
A_5 : Good rainfall, A_6 : Very good rainfall,
A_7 : Heavy rainfall, A_8 : Flood situation

These FS with membership grades are as

$$A_1 = \frac{1}{u_1} + \frac{.5}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8}$$

$$A_2 = \frac{.5}{u_1} + \frac{1}{u_2} + \frac{.5}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8}$$

$$A_3 = \frac{0}{u_1} + \frac{.5}{u_2} + \frac{1}{u_3} + \frac{.5}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8}$$

$$A_4 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{.5}{u_3} + \frac{1}{u_4} + \frac{.5}{u_5} + \frac{0}{u_6} + \frac{0}{u_7} + \frac{0}{u_8}$$

$$A_5 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{.5}{u_4} + \frac{1}{u_5} + \frac{.5}{u_6} + \frac{0}{u_7} + \frac{0}{u_8}$$

$$A_6 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{.5}{u_3} + \frac{0}{u_4} + \frac{.5}{u_5} + \frac{1}{u_6} + \frac{0}{u_7} + \frac{0}{u_8}$$

$$A_7 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{.5}{u_6} + \frac{1}{u_7} + \frac{0}{u_8},$$

$$A_8 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5} + \frac{0}{u_6} + \frac{.5}{u_7} + \frac{1}{u_8}$$

Step 4. After getting the fuzzified historical time series rainfall data, table 1 gives obtained FLR

Table 1: Fuzzified Rainfall

Year	Actual Rainfall	Fuzzified Rainfall
2000	1035.4	A_3
2001	1105.2	A_5
2002	0981.9	A_2
2003	1233.6	A_8
2004	1080.5	A_4
2005	1208.3	A_7
2006	1161.6	A_6
2007	1179.3	A_6
2008	1118.0	A_5
2009	0953.7	A_1
2010	1215.5	A_7
2011	1116.3	A_5
2012	1054.7	A_3
2013	1092.5	A_4
2014	1045.2	A_3

Step 5. The forecasted defuzzified rainfall is calculated by using the proposed algorithms and put in the table 2
Step 6. Calculating mean square error (MSE), forecasting error (FE) and average forecasting error (AFE) as

$$\text{MSE} = \frac{\sum_{i=1}^{n}((\text{act. val.})_i - (\text{fore. val.})_i)^2}{n} \quad (1)$$

$$\text{FE} = \frac{|\text{act. val} - \text{fore. val.}|}{\text{actual value}} \times 100 \quad (2)$$

$$AFE = \frac{\text{sum of FE}}{\text{numbers of errors}} \times 100 \quad (3)$$

Table 2: Forecasted rainfall

Year	Actual Rainfall	Forecasted rainfall Proposed Model	Forecasted rainfall Chen Model
2000	1035.4	_	_
2001	1105.2	_	1100
2002	981.9	_	1000
2003	1233.6	1240.00	1240
2004	1080.5	1080.00	1080
2005	1208.3	1200.00	1120
2006	1161.6	1166.50	1140
2007	1179.3	1140.75	1140
2008	1118	1120.00	1140
2009	953.7	0960.00	1000
2010	1215.5	1200.00	1200
2011	1116.3	1119.00	1140
2012	1054.7	1037.50	1000
2013	1092.5	1078.60	1000
2014	1045.2	1040.00	1120
MSE	_	199.961	2196.344
% FE	_	10.47055	46.17992
AFE	_	0.872546	3.298566

Figure 1 shows the year wise comparison between actual rainfall and forecasted rainfall by proposed and Chen's model

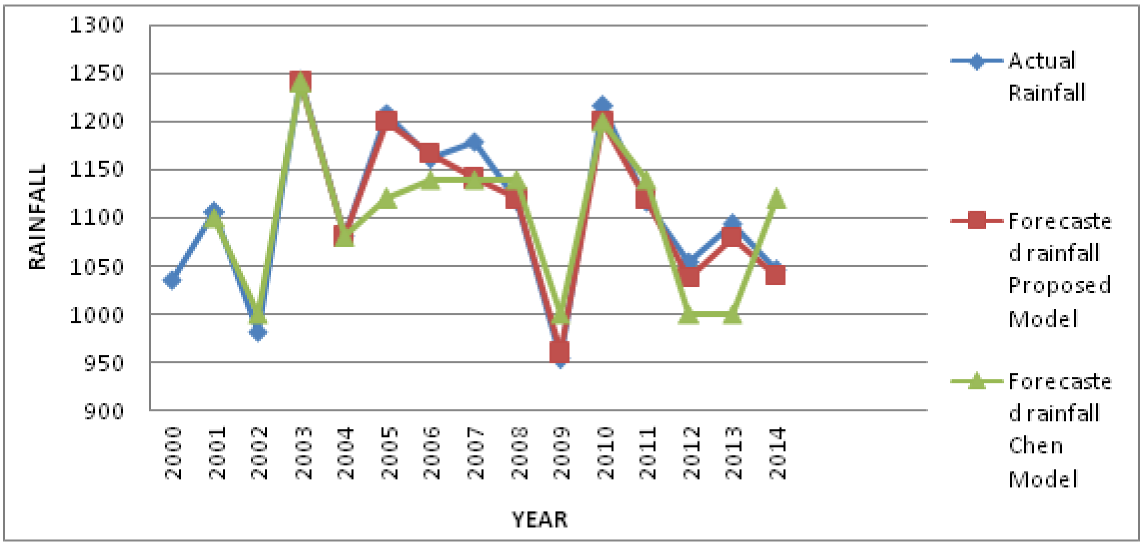

Figure 1: Forecasted and Actual Rainfall

V. CONCLUSION

In this paper a new method to predict monsoon rainfall in India is proposed. Rainfall prediction is done by taking values of two differences of the interval corresponding to the fuzzified forecasted value. The proposed method shows significant reduction in the errors of predicted values. Also in the present study the computational procedure is much easier than complicated min-max operation. The robustness of the proposed model is tested on real rainfall data and accuracy of the presented model is verified by comparing the results with Chen's model. The results show betterment in the forecasted values over the compared model.

ACKNOWLEDGEMENT

The rainfall data taken from website of Indian meteorological department is dully acknowledged.

REFERENCES

1. A.K. Rana, "Rice Production Forecasting Through Fuzzy Time Series", American International Journal of Research in Science, Technology, Engineering & Mathematics, vol. 23(1), pp. 158-162, 2018.
2. A.K. Rana, "Study on Fuzzy Time Invariant Series Models for Crop Production Forecasting", International Journal of Scientific Research and Reviews, vol. 8(2), pp. 3729-3741, 2019.
3. C.S. Chen, Y.D. Jhong, W.Z. Wu and S.T. Chen, "Fuzzy Time Series for Real Time Flood Forecasting", Stochastic Environmental Research and Risk Assessment, vol. 33, pp. 645-656, 2019.
4. K. Huarng, "Heuristic Models of Fuzzy Time Series for Forecasting", Fuzzy Sets and Systems, vol. 123, pp. 369–386, 2001.
5. L.A. Zadeh, "Fuzzy sets", Information and Control, vol. 8, pp. 338–353, 1965.
6. M. Bose and K. Mali, "Designing Fuzzy Time Series Forecasting Models: A Survey", International Journal of Approximate Reasoning, vol. 111, pp. 78-99, 2019.

7. P. Singh, “A Brief Review on Modeling Approaches Based on Fuzzy Time Series”, International Journal of Machine Learning and Cybernetics, vol. 8(2), pp. 397-420, 2017
8. Q. Song and B. S. Chissom, “Forecasting Enrollments with Fuzzy Time Series – Part I”, Fuzzy Sets and Systems, vol. 54, pp. 1–9, 1993
9. Q. Song and B. S. Chissom, “Fuzzy Time Series and Its Models”, Fuzzy Sets and Systems, vol.54, pp. 269–277, 1993
10. S.M. Chen, “Forecasting Enrollments Based on High Order Fuzzy Time Series”, Cybernetics and Systems: An International Journal, vol. 31, pp. 1-16, 2002
11. S.M. Chen and C.C. Hsu, “A New Method to Forecast Enrollments Using Fuzzy Time Series”, International Journal of Applied Sciences and Engineering, vol. 2(3), pp. 234-244, 2004
12. S. Panigrahi and H.S. Behera, “A Computationally Efficient Method for High Order Fuzzy Time Series Forecasting”, Journal of Theoretical and Applied Information Technology, vol. 96(21), pp. 7215-7226, 2018

AUTHOR PROFILE

Amit Kumar Rana received his M.Sc., Ph.D. (Mathematics) degree from the G. B. Pant University of Agriculture & Technology, Pantnagar, India in 2000 and 2005 respectively. During Ph.D. author was a recipient of Junior Research Fellowship provided by Council of Scientific and Industrial Research, Govt. of India. At present he is an Assistant Professor in Department of Mathematics, Swami Vivekanand Meerut, Uttar Pradesh, India. His research areas are mainly Fuzzy Time Series forecasting, MCDM and optimization techniques.