Ratio Mathematica Volume 45, 2023

Forecasting of Annual Rainfall using Fuzzy Logic Interval Based Partitioning in Different Intervals

Rajan D*
Sugunthakunthalambigai R†

Abstract

Fuzzy time series models have been proposed by many researchers around the world for rainfall forecasting, but the forecasting has not been as accurate as existing methods. Frequency density or ratio-based segmentation methods have been used to represent discourse segmentation. In this paper, to make such predictions, we used interval-based segmentation as the discourse segmentation and the urban mean rainfall in the Trichy district as the discourse universe. Fuzzy models are used for forecasting in many fields such as admissions prediction, stock price analysis, agricultural production, horticultural production, marine production, weather forecasting, and more.

Keywords. Mean Square Error; Fuzzy time series; Average Forecast Error Rate.

AMS Subject Classification: 05C78‡.

---

*Associate Professor of Mathematics, (TBML College, Affiliated to Bharathidasan University Porayar-609 307, Mayiladuthurai Dist.); dan_rajan@rediffmail.com.

†Assistant Professor of Mathematics, HC&RI (W), Trichy (TNAU, Tamilnadu, India); suguntha@tnau.ac.in.

‡ Received on July 10, 2022. Accepted on October 15, 2022. Published on January 30, 2023. doi: 10.23755/rm. v45i0.1006. ISSN: 1592-7415. eISSN: 2282-8214.

1. Introduction

A forecasting process issued to predict future outcomes. Relevant data and figures are carefully analyzed to make accurate forecasts and make optimal choices regarding the future. There are two main reasons for choosing time series forecasting. First, most of the data that exists in the real world, such as economics, business, and finance, are time series. Second, evaluation of time series data is easy, and many techniques are available for evaluation of time series forecasts. If the future is in doubt, the forecasting process is mandatory. Implement using the fuzzy time series method to forecast precipitation and compare results with other existing techniques. Accurate rainfall information is essential for planning and managing water resources. Moreover, in urban areas, rainfall has a strong impact on transportation, sewerage, N. Q. Hung (nguyenquang.hung@ait.ac.th) systems, and other human activities. Nevertheless, precipitation is best understood and modelled for hydrological cycles because of the complexity of the atmospheric processes that generate precipitation and the tremendous range of variability over a wide range of scales in both space and time. It is one of the complex and difficult factors (French et al., 1992). Accurate rainfall forecasting is thus one of the greatest challenges in operational hydrology, despite many advances in weather forecasting in recent decades (Gwangseob and Ana, 2001). The data for prediction has all the uncertainties.

For this purpose, we generate rainfall simulations that reproduce in a distributional sense the set of key rainfall statistics obtained from the observational dataset (Benoit and Mariethoz, 2017). The practical interest of probabilistic rainfall models is, among other things, to complement numerical weather models for simulating rainfall heterogeneity at fine scales and to quantify uncertainties associated with rainfall reconstructions. . Indeed, numerical weather models face the challenge of reproducing spatial and temporal rainfall heterogeneity, especially at fine scales (Bauer et al., 2015; Bony et al., 2015). Some of the main applications of probabilistic rainfall models are therefore for local impact studies, e.g. related to hydrology (Paschalis et al., 2014; Caseri et al., 2016).

Fuzzy time series forecasting is a smart method in areas where information is explicit, imprecise, and approximate. Fuzzy time series can also tackle situations that do not provide trend investigation and analysis, or visualization of time series patterns. In-depth research work has been done on forecasting problems using this concept. Vikas [1] proposed various techniques for predicting crop yields and used artificial neural networks to predict wheat yields. Did. Adesh [2] conducted a comparative study of

Various techniques, including neural networks and fuzzy models. Askar [3] also tried to predict yield using time series models. Sachin [4-5] specifically worked on rice yield prediction using fuzzy time series models. Narendra [6] attempted to predict wheat yields. Pankaj [7] used an adaptive neuro-fuzzy system for predicting wheat yield. W. Qiu, X. Liu, and H. Li proposed a generalized method of forecasting based on fuzzy time series models [30].

The concept and definition of fuzzy time series was devised and published by Song & Chissom. They also delineated concepts and notions of variant and invariant time series [8-9]. First, the time-series data for the University of Alabama were obtained, the admission prediction was performed, and after a few years [10] average auto, then Chen [11-12] was the max-min configuration operation previously used by Song &Chissom. We depicted simplified arithmetic operations instead of using, and used higher-order fuzzy time series to organize our forecasting models. Huarng [13-14], Hwang and Chen [15], Lee Wang and Chen [16], Li and Kozma [17] all produced a number of fuzzy prediction methods with slight variations in each. Lee et al. Subsequently, a multivariate heuristic model was designed and implemented to obtain highly complex and complex matrix computations [20]. Research work has been done to ascertain the interval length of fuzzy time series [21].

Wong et al. (2003) used SOM and back propagation neural networks to build a fuzzy rule base and used the rule base to develop a forecast model for Swiss rainfall using spatial interpolation. Bardossy et al. (1995) implemented fuzzy logic for classification of atmospheric circulation patterns. Ozerkan et al. (1996) compared the performance of regression analysis and fuzzy logic in studying the relationship between monthly atmospheric circulation patterns and precipitation. Pesti et al. (1996) implemented fuzzy logic for drought assessment. Baum et al. (1997) developed a cloud classification model using fuzzy logic. Fujibe (1989) classified precipitation patterns in Honshu using the fuzzy C-means method. Garanboshi et al. (1999) Using fuzzy logic, we investigated the effects of ENSO and macro circulation patterns on precipitation over Arizona. Vivekanandan et al. (1999) developed and implemented a fuzzy logic algorithm for water meteor particle identification that is simple and efficient enough to run in real time for operational use. Wuwardi et al. (2006) use a neural fuzzy system to model tropical rainfall during the rainy season. The model results in lower RMSE values, indicating that the forecast model is reliable in representing recent inter annual variability in tropical rainfall during the wet season.

2. Basic Concepts of Fuzzy time Series Modeling

Definition 1.1. (Song and Chissom, 1993a, 1994; Liaw, 1997). A fuzzy number on the real line $\Re$ is a fuzzy subset of $\Re$ that is normal and convex.

Definition 1.2. (Song and Chissom, 1993a, 1994). Suppose that $R_1 = \cup_{ij} R_1 ij(t, t-1)$ and $R_2(t, t-1) = \cup_{ij} R_2 ij(t, t-1)$ are two fuzzy relations between $F(t)$and$F(t-1)$. If, for any$f_j(t) \in F(t)$, where$j \in J$, there exists$f_i(t-1) \in F(t-1)$, where$i \in I$, and fuzzy relations $R_1\ ij(t, t-1)$ and $R_2\ ij(t, t-1)$ such that $f_i(t) = f_i(t-1) \circ R_1\ ij(t, t-1)$ and$f_i(t) = f_i(t-1) \circ R_2\ ij(t, t-1)$, then define$R_1(t, t-1) = R_2(t, t-1)$.

Definition 1.3. (Song and Chissom, 1993a, 1994). Suppose that $F(t)$ is only caused by $F(t-1), F(t-2), \ldots,$or$F(t-m)$ $(m > 0)$. This relation can be expressed as the following fuzzy relational equation: $F(t) = F(t-1) \circ R_0(t, t-m)$, which is called the first-order model of $F(t)$.

Definition 1.4 (Song and Chissom, 1993a, 1994). Suppose that $F(t)$ is simultaneously caused by $F(t-1), F(t-2), \ldots,$ and $F(t-m)$ $(m > 0)$. This relation can be expressed as the following fuzzy relational equation: $F(t) = (F(t-1) \times F(t-2) \times \ldots \times F(t-m)) \circ Ra(t, t-m))$, which is called the m th-order model of $F(t)$.

Definition 1.5 (Chen, 1996). $F(t)$ is fuzzy time series if $F(t)$ is a fuzzy set. The transition is denoted as $F(t-1) \longrightarrow F(t)$.

Definition 1.6 (Chou, 2011). Let $d(t)$ be a set of real numbers $d(t) \subseteq R$. A lower interval $d(t)$ is a number b such that $x \geq b$ for all $x \in d(t)$. The set $d(t)$ is said to be an interval below if $d(t)$ has a lower interval. A number, min, is the minimum of $d(t)$ if min is a lower interval $d(t)$ and min $\in d(t)$.

Definition 1.7 (Chou, 2011). Let $d\ (t)$ be a set of real numbers $d(t) \subseteq R$. An upper interval $d(t)$ is a number b such that $x \leq b$ for all $x \in d\ (t)$. The set $d\ (t)$ is said to be an interval higher if $d\ (t)$ has an upper interval. A number, max, is the maximum of $d\ (t)$ if max is an upper interval $d\ (t)$ and max $\in d(t)$.

3. Proposed Method

In this section, we use real-world rainfall as the universe of discourse and propose a method for forecasting using interval-based segmentation. Relevant concepts and definitions regarding this can be found by referring to a previously published paper [29]. Another method for predicting the values provided in this paper is clearly explained in the following lines. The forecasting process follows these steps:

Step 1:

First, obviously analysis descriptive statistics. It helps in facilitating data visualization. Next, we describe the discourse universe U and the parcel U in intervals of equal length. Here, according to the data, 379.79 is the minimum value and 1163.59 is the maximum value. We need to specify the discourse universe, the intervals in which all given values of rainfall exist. So in this case the discourse universe would be [300, 1200]. Descriptive statistics and block-by-block rainfall data are shown in Tables I and II.

Step 2:

Fuzzy partitioning is a methodology for generating fuzzy sets that represent the underlying data. The techniques can be classified into three categories: grid partitioning, tree partitioning, and distributed partitioning. Among the various fuzzy partitioning methods, grid partitioning is the most commonly used in practice, especially in system control applications. Grid partitioning forms partitions by dividing the input space into several fuzzy slices.

Next, divide Universe of discourse in 6, 9 and 18 equal intervals these are as following. The discourse universe can be defined by $U = [300, 1200]$. U is then divided into 6 equal length intervals and the midpoint of the 6th interval is calculated as shown below.

Table 3.1. Annual rainfall in Trichy district (From2004 to 2010)

Blocks	**Rainfall(mm)**
Andanallur	1085.23
Lalgudi	1136.79
Manachanallur	569.44
Manapparai	837.43
Manikandam	379.79
Marungapuri	942.16
Musiri	794.31
Pullambadi	1163.59
T.pet	710.16
Thiruverumbur	960.9
Thottiam	866.73
Thuraiur	942.11
Uppiliyapuram	473.35
Vaiyampatty	922.27

Table 3.2. Descriptive statistics

Minimum	= 379.79
Maximum	= 1163.59
Range	= 783.8
Count	= 14
Sum	= 11784.26
Mean	= 841.733
Median	= 894.5
Mode	= No mode
Standard Deviation	= 237.65
Variance	= 56479.54

Table 3.3. a) 6 equal intervals

u_1	$[300 - 450]$	375
u_2	$[450 - 600]$	525
u_3	$[600 - 750]$	675
u_4	$[750 - 900]$	825
u_5	$[900 - 1050]$	975
u_6	$[1050 - 1200]$	375

Here, U is partitioned into 9 equal length intervals and calculated mid points of 9th intervals given below:

Table 3.4. b) 9 equal Intervals with Midpoints

v_1	[300 – 400]	350
v_2	[400 –500]	450
v_3	[500 – 600]	550
v_4	[600 –700]	650
v_5	[700 – 800]	750
v_6	[800 – 900]	850
v_7	[900 –1000]	950
v_8	[1000 – 1100]	1050
v_9	[1100 –1200]	1150

Here, U is partitioned into 18 equal length intervals and calculated mid points of 18th intervals given below.

Table 3.5. c) 18 equal intervals with Midpoints

w_1	$[300 - 350]$	325
w_2	$[350 - 400]$	375
w_3	$[400 - 450]$	425
w_4	$[450 - 500]$	475
w_5	$[500 - 550]$	525
w_6	$[550 - 600]$	575
w_7	$[600 - 650]$	625
w_8	$[650 - 700]$	675
w_9	$[700 - 750]$	725
w_{10}	$[750 - 800]$	775
w_{11}	$[800 - 850]$	825
w_{12}	$[850 - 900]$	875
w_{13}	$[900 - 950]$	925
w_{14}	$[950 - 1000]$	975

w_{15}	$[1000-1050]$	1025
w_{16}	$[1050-1100]$	1075
w_{17}	$[1100-1150]$	1125
w_{18}	$[1150-1200]$	1175

Step 3. Define a fuzzy set based on 6, 9, and 18 intervals to fuzz the historical data.

a. 6 Equal intervals

Let $U = \{u_1, u_2, u_3, u_4, u_5, u_6\}$ be the world of discourse. The number of intervals depends on the number of considered linguistic variables (fuzzy sets) $A_1, A_2, A_3, A_4, A_5, A_6$.Define 6 fuzzy sets $A_1, A_2, \ldots, A_6$ as linguistic variables in the discourse world U. These fuzzy variables are defined as:

Table 3.6. Label Linguistic value of enrolments

Fuzzified	Linguistic Value
A_1	very few
A_2	very very few
A_3	Moderate
A_4	High
A_5	very High
A_6	very very High

b. 9 equal Intervals

Let U be the universe of discourse, where$U = \{v_1, v_2, v_3, v_4, v_5, v_6 \ldots v_9\}$. The number of intervals will be in accordance with the number of linguistic variables (fuzzy sets) $B_1, B_2, B_3, B_4, \ldots B_9$, to be considered.
Define 9fuzzy sets $B_1, B_2, B_3, B_4, \ldots B_9$, as linguistic variables on the universe of discourse U. These fuzzy variables are being defined as:

Table 3.7. Label Linguistic value of enrolments

Fuzzified	Linguistic Value
B_1	very ^{3} few
B_2	very ^{2} few
B_3	very^1 few
B_4	Few
B_5	Moderate
B_6	High

B_7	Very^1 High
B_8	very ^{2} High
B_9	very ^{3} High

c. 18 Equal intervals

Let U be the universe of discourse, and let $U = \{w_1, w_2, w_3, \dots w_{18}\}$. The number of intervals depends on the number of considered linguistic variables (fuzzy sets) $C_1, C_2, C_3, \dots C_{18}$. Define 18 fuzzy sets $C_1, C_2, C_3, \dots C_{18}$ as linguistic variables in the universe of discourse U. These fuzzy variables are defined as:

Table 3.8. Label Linguistic value of enrolments

Fuzzified	Linguistic Value
C_1	very ^{8} few
C_2	very ^{7} few
C_3	very ^{6} few
C_4	very ^{5} few
C_5	very ^{4} few
C_6	very ^{3} few
C_7	very ^{2} few
C_8	very^1 few
C_9	Few
C_{10}	High
C_{11}	Very^1 High
C_{12}	very ^{2} High
C_{13}	very ^{3} High
C_{14}	very ^{4} High
C_{15}	very ^{5} High
C_{16}	very ^{6} High
C_{17}	very ^{7} High
C_{18}	very ^{8} High

Step 4:

Fuzzy set defined by U (all intervals).

A fuzzy set A_i is represented as

$$A_1 = {}^{1}/_{u_1} + {}^{0.5}/_{u_2} + {}^{0}/_{u_3} + {}^{0}/_{u_4} + {}^{0}/_{u_5} + {}^{0}/_{u_6}$$

$$A_2 = {}^{0.5}/_{u_1} + {}^{1}/_{u_2} + {}^{0.5}/_{u_3} + {}^{0}/_{u_4} + {}^{0}/_{u_5} + {}^{0}/_{u_6}$$

………………………………………………………

………………………………………………………

………………………………………………………

$$A_{10} = {}^{0}/_{u_1} + {}^{0}/_{u_2} + {}^{0}/_{u_3} + {}^{0}/_{u_4} + {}^{0.5}/_{u_5} + {}^{1}/_{u_6}$$

A fuzzy set B_i are expressed as follows:

$$B_1 = {}^{1}/_{u_1} + {}^{0.5}/_{u_2} + {}^{0}/_{u_3} + \cdots + {}^{0}/_{u_9}$$
$$B_2 = {}^{0.5}/_{u_1} + {}^{1}/_{u_2} + \cdots + {}^{0}/_{u_9}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$B_9 = {}^{0}/_{u_1} + {}^{0}/_{u_2} + \cdots + {}^{0.5}/_{u_5} + {}^{1}/_{u_9}$$

A fuzzy set C_i are expressed as follows:

$$C_1 = {}^{1}/_{u_1} + {}^{0.5}/_{u_2} + {}^{0}/_{u_3} + {}^{0}/_{u_4} + {}^{0}/_{u_5} + {}^{0}/_{u_6}$$
$$C_2 = {}^{0.5}/_{u_1} + {}^{1}/_{u_2} + {}^{0.5}/_{u_3} + {}^{0}/_{u_4} + {}^{0}/_{u_5} + {}^{0}/_{u_6}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$C_{18} = {}^{0}/_{u_1} + {}^{0}/_{u_2} + {}^{0}/_{u_3} + \cdots + {}^{0.5}/_{u_{17}} + {}^{1}/_{u_{18}}$$

Step 5: Fuzzify historical data.

Table 3.9. Linguistic values for the enrolments from 2004 to 2010

Block	Rainfall (mm)	Linguistic value 6th Interval	Linguistic value 9th Interval	Linguistic value 18th Interval
Andanallur	1085.23	C_6	D_8	E_{16}
Lalgudi	1136.79	C_6	D_9	E_{17}
Manachanallur	569.44	C_2	D_3	E_6
Manapparai	837.43	C_4	D_6	E_{12}
Manikandam	379.79	C_1	D_1	E_2
Marungapuri	942.16	C_5	D_7	E_{13}
Musiri	794.31	C_4	D_5	E_{10}
Pullambadi	1163.59	C_6	D_9	E_{18}
T.pet	710.16	C_4	D_5	E_9
Thiruverumbur	960.9	C_5	D_7	E_{14}
Thottiam	866.73	C_4	D_6	E_{12}
Thuraiur	942.11	C_5	D_7	E_{13}
Uppiliyapuram	473.35	C_2	D_2	E_4
Vaiyampatty	922.27	C_5	D_7	E_{13}

Step 6:
Calculate predicted registrations for 6, 9, and 18 intervals given below:

Table 3.10. Forecasted value for all intervals

Blocks	**Rainfall (mm)**	**Forecasted value (6 intervals)**	**Forecasted value (9 intervals)**	**Forecasted value (18 intervals)**
Andanallur	1085.23	-	-	-
Lalgudi	1136.79	-	-	-
Manachanallur	569.44	-	-	-
Manapparai	837.43	900	900	912.5
Manikandam	379.79	712.5	725	737.5
Marungapuri	942.16	675	675	687.5
Musiri	794.31	750	725	737.5
Pullambadi	1163.59	825	800	812.5
T.pet	710.16	937.5	900	900
Thiruverumbur	960.9	937.5	900	912.5
Thottiam	866.73	937.5	925	937.5
Thuraiur	942.11	900	875	875
Uppiliyapuram	473.35	825	800	812.5
Vaiyampatty	922.27	825	800	800

Step 7:
Calculate MSE and AFER values for 6, 9, and 18 intervals given below:
Mean Squared Error (MSE) measures the amount of error in a statistical model. Evaluate the mean squared difference between observed and predicted values. If the model has no errors, MSE is equal to zero.

$$\text{MSE formula} = \left({}^{1}\!/_{n}\right) * \Sigma(Actual - Forecast)2$$

Mean Absolute Percentage Error (MAPE), also known as Mean Absolute Percentage Deviation (MAPD), is a measure of the predictive accuracy of a forecasting method in statistics. Accuracy is usually expressed as a ratio defined by the following formula:
$\frac{|Actual - predicted|}{Actual} \times 100\%$ = Average Forecasting Error Rate (**AFER**)

4. Performance Evaluation and Comparative Studies

A. Performance rating:

Two parameters are used to compare the results of the proposed method with existing methods. These are MSE &AFER. MSE & AFER are the calculated values for intervals 6, 9 and 18 as shown in Tables XI, XII and XIII. Interval-based partitioning is calculated in Table XI. MSE indicates the deviation error from the actual value to the predicted value. The deviation is shown in Figure 4.1 in the form of a graphical representation for better visualization. As we can see the proposed algorithm gives values very close to what is the actual rainfall value. The same is done for the 9th and 18th intervals, as shown in Tables XII & XIII and Fig. 4.2 and 4.3.

Table 3.11. MSE and AFER Values (6 Intervals)

Block	A_i **(Rainfall mm)**	F_i	**MSE** $(A_i - F_i)^2$	**AFER** $\lvert A_i - F_i \rvert / A_i$
Andanallur	1085.23	-	-	-
Lalgudi	1136.79	-	-	-
Manachanallur	569.44	-	-	-
Manapparai	837.43	900	3915.005	0.074717
Manikandam	379.79	712.5	110695.9	0.876037
Marungapuri	942.16	675	71374.47	0.283561
Musiri	794.31	750	1963.376	0.055784
Pullambadi	1163.59	825	114643.2	0.290987
T.pet	710.16	937.5	51683.48	0.320125
Thiruverumbur	960.9	937.5	547.56	0.024352
Thottiam	866.73	937.5	5008.393	0.081652
Thuraiur	942.11	900	1773.252	0.044698
Uppiliyapuram	473.35	825	123657.7	0.742896
Vaiyampatty	922.27	825	9461.453	0.105468
			MSE = 35337.42	AFER = 20.72%

Table 3.12. MSE AND AFER VALUES (9 INTERVALS)

Block	$\boldsymbol{A_i}$ **(Rainfall mm)**	$\boldsymbol{F_i}$	**MSE** $(\boldsymbol{A_i - F_i})^2$	**AFER** $\lvert\boldsymbol{A_i - F_i}\rvert/\boldsymbol{A_i}$
Andanallur	1085.23			
Lalgudi	1136.79			
Manachanallur	569.44			
Manapparai	837.43	912.5	5635.505	0.089643
Manikandam	379.79	737.5	127956.4	0.941863
Marungapuri	942.16	687.5	64851.72	0.270294
Musiri	794.31	737.5	3227.376	0.071521
Pullambadi	1163.59	812.5	123264.2	0.30173
T.pet	710.16	900	36039.23	0.26732
Thiruverumbur	960.9	912.5	2342.56	0.050369
Thottiam	866.73	937.5	5008.393	0.081652
Thuraiur	942.11	875	4503.752	0.071234
Uppiliyapuram	473.35	812.5	115022.7	0.716489
Vaiyampatty	922.27	800	14949.95	0.132575
			MSE=35914.42	AFER=21.39%

Table 3.13. MSE AND AFER values (18 intervals)

Block	A_i (Rainfall mm)	F_i	MSE $(A_i - F_i)^2$	AFER $\|A_i - F_i\|/A_i$
Andanallur	1085.23	-	-	-
Lalgudi	1136.79	-	-	-
Manachanallur	569.44	-	-	-
Manapparai	837.43	900	3915.005	0.074717
Manikandam	379.79	725	119169.9	0.90895
Marungapuri	942.16	675	71374.47	0.283561
Musiri	794.31	725	4803.876	0.087258
Pullambadi	1163.59	800	132197.7	0.312473
T.pet	710.16	900	36039.23	0.26732
Thiruverumbur	960.9	900	3708.81	0.063378
Thottiam	866.73	925	3395.393	0.06723
Thuraiur	942.11	875	4503.752	0.071234
Uppiliyapuram	473.35	800	106700.2	0.690081
Vaiyampatty	922.27	800	14949.95	0.132575
			MSE=35768.45	AFER=21.13%

The following figures (Fig.4.1, Fig 4.2 and Fig 4.3) are compared in Forecasted and Actual rainfall for the corresponding intervals respectively 6,9&18.And also Fig. 4.4 compares the MSE for all intervals.

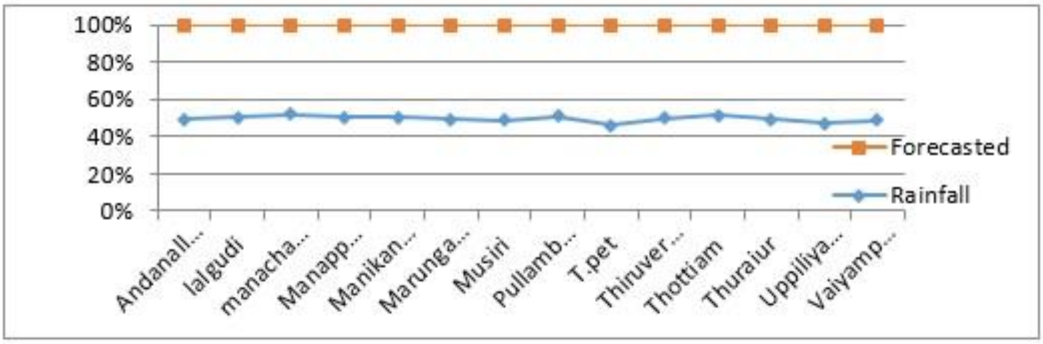

Fig. 4.1. Forecasted vs. Rainfall (6 intervals)

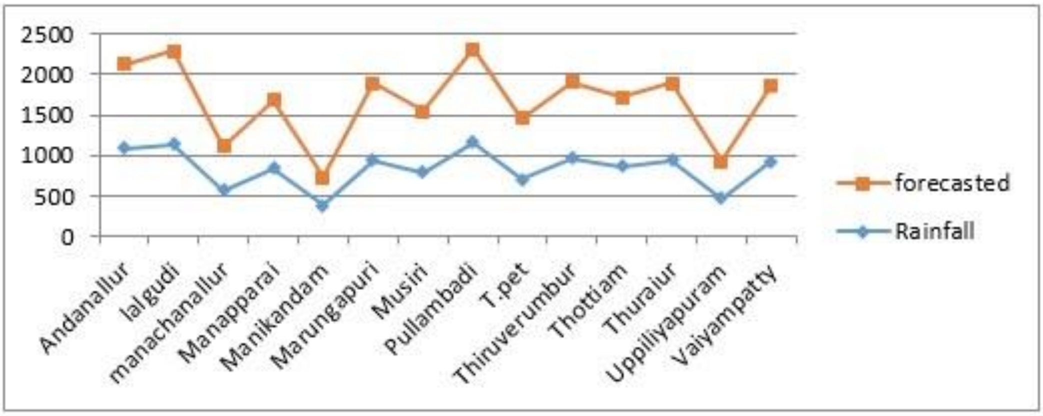

Fig. 4.2. Forecasted vs. Rainfall (9 intervals)

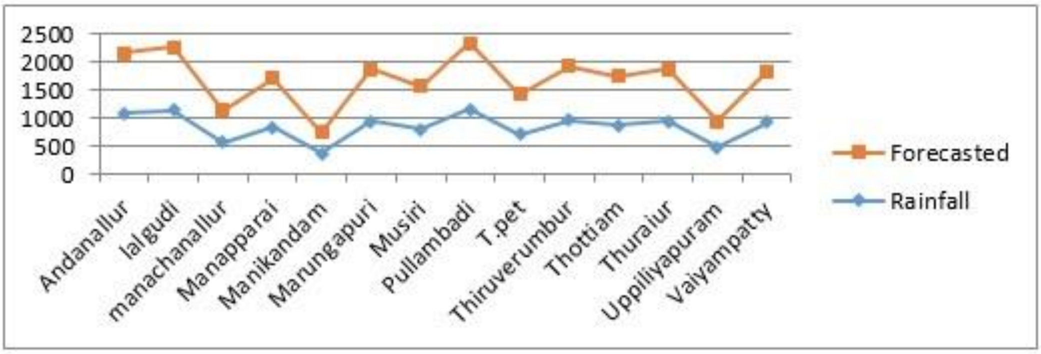

Fig. 4.3. Forecasted vs. Rainfall (18 intervals)

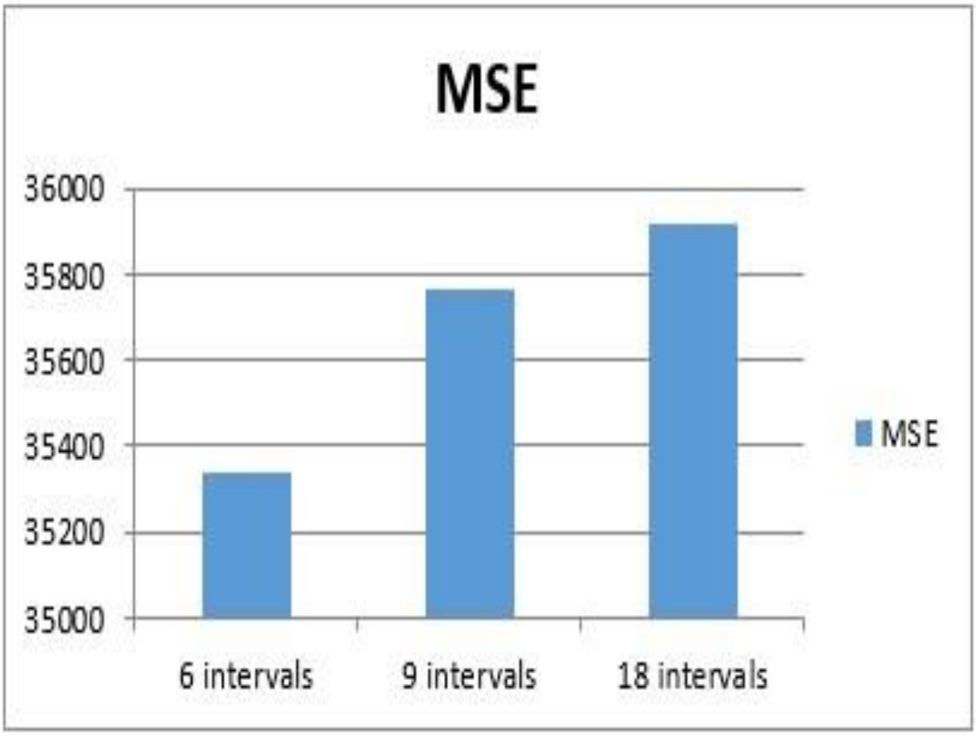

Fig. 4.4. Graph of MSE 6, 9 and 18 intervals

B. Results and Discussion.

The MSE and AFER calculated in Tables 3.11, 3.12 and 3.13 above were analyzed. The paper shows working with different intervals such as 6 , 9 and 18 . The majority of recently published papers work in one of these intervals. The focus of this paper was to propose a new algorithm and check its predicted variability over all these intervals. The results show that prediction works best on the 6th interval among all other intervals. All results are presented in an easy-to-understand bar chart format to reduce the complexity of this study and present it in a more understandable manner. To prove that this algorithm is efficient, a comparison is made with other existing methods proposed by Heuristic and FTS first order in Table XIV. As can be seen from Fig. 4.5, the proposed algorithm was able to achieve significantly lower MSE compared to other methods. This model not only provides a lower MSE, but also explains why researchers making fuzzy logic predictions choose the 6th interval for their work. All other intervals do not give better results than the 6th interval partitioning. A possible reason for increasing the number of intervals is that the data is overcrowded. As such, the relevant data between intervals is not included in the forecasting algorithm and affects forecasting results. Keeping the interval lower than the 6th interval will spread the data too much. Therefore, the 6th interval partitioning seems to be the best overall for fuzzy logic-based predictive models.

Table 4.1. Comparison to prove efficiency

Method	MSE	AFER
Heuristic	109728.7	44.87%
Proposed Method	35337.42	20.72%
First order	151459.9	46.90%

5. Conclusion

First, divide the data set into 6, 9, and 18 intervals and compute the values for each block. Use these midpoint values to compute rainfall forecasts for all blocks. The results were then validated by comparing with other existing methods using accuracy, precision and robustness of the proposed model. We found the new method to be the best advanced algorithm. The proposed model can be extended by working with more intervals. You can also apply frequency-based partitioning to intervals to fine-tune the distribution.

References

[1] Askar Choudhury, James Jones, Crop Yield Prediction Using Time Series Models, Journal of Economic and Econo Education Research, Volume15, Number 3, (2014).

[2] Bindu Garg, Rohit Garg, Enhanced Accuracy of Fuzzy Time Series Model using Ordered Weighted Aggregation, Applied Soft computing, Elsevier, Vol. 8, pp. 265-280, (2016).

[3] Chen S.M, Forecasting Enrolments based on High Order Fuzzy Time Series, Intl. Journal of Cybernetics and Systems, Vol. 33, pp. 1-16, (2010).

[4] Chen S.M., Forecasting Enrolments based on Fuzzy Time Series, Fuzzy Sets and Systems, Vol. 81, pp. 311-319, (1996).

[5] Garg B., M. Beg M. S., A. Q. Ansari, Employing OWA to Optimize Fuzzy Predicator, World Conference on Soft Computing (WConSC2011), San Francisco State University, USA, pp. 205-211, (2011).

[6] Garg B., M. Beg M. S., and. Ansari A. Q, OWA based Fuzzy Time Series Forecasting Model, World Conference on Soft Computing, Berkeley, San Francisco, CA, and pp. 141-177, May23-26, (2011).

[7] Garg B., M. Beg M. S., Ansari A.Q. and Imran B.M., Fuzzy Time Series Prediction Model, Communications in Computer and Information Science, Springer–Verlag Berlin Heidelberg, ISBN978-3-642-19423-8, Vol. 141, pp. 126-137, (2011).

[8] Garg B., M. Beg M. S and Ansari A. Q, A New Computational Fuzzy Time Series Model to Forecast Number of Outpatient Visits, ―Proc. 31st Annual Conference of the North American Fuzzy Information Processing Society (NAFIPS2012), University of California at Berkeley, USA, pp. 1-6, August 6-8, (2012).

[9] Ansari A. Q, Enhanced Accuracy of Fuzzy Time Series Predictor using Genetic Algorithm, Third IEEE World Congress on Nature and Biologically Inspired Computing (NaBIC2011), pp. 273-278, Spain, (2011).

[10]Garg B., M. Beg M. S, Ansari A. Q. and Imran B. M, Soft Computing Model to Predict Average Length of Stay of Patient, Communications in Computer and Information Science, Springer Verlag Berlin Heidelberg, ISBN 978-3-642-19423-8, Vol. 141, pp. 221–232, (2011).

[11]Garg B., M. Beg M. S and Ansari A. Q, Employing Genetic Algorithm to Optimize OWA-Fuzzy Forecasting Model, Third IEEE World Congress on Nature and Biologically Inspired Computing (NaBIC2011), pp. 285-290, Spain, (2011).

[12]Huarng K, Heuristic Models of Fuzzy Time Series for Forecasting; Fuzzy Sets and Systems, Vol. 123, pp. 369-386, (2002).

[13]Huarng K. H, Yu T. H. K and Hsu Y.W, a Multivariate Heuristic Model for Forecasting, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 37, pp. 836–846, (2007).

[14]Huarng K, Effective Lengths of Intervals to Improve Forecasting in Fuzzy Time Series, Fuzzy Sets and Systems, Vol. 12, pp. 387-394, (2001).

[15]Hwang J. R., Chen S. M. and Lee C. H., Handling Forecasting Problems using Fuzzy Time Series, Fuzzy Sets and Systems, Vol. 100, Pg.: 217-228,(1998).

[16]Lee C. H. L, Lin A. and Chen W.S., Pattern Discovery of Fuzzy Time Series for Financial Prediction, IEEE Transaction on Knowledge Data Engineering, Vol.18 pp. 613–625, (2006).

[17]Lee L. W., Wang L. W. Chen S. M, Handling Forecasting Problems based on Two-Factors High-Order Time Series, IEEE Transactions on Fuzzy Systems, Vol. 14, pp. 468-477, (2006).

[18]Li H. and Kozma R., A Dynamic Neural Network Method For Time Series Prediction using the KIII Model, Proceedings of the 2003 International Joint Conference on Neural Networks, pp. 347-352, (2003).

[19]Narendra kumar, Sachin Ahuja, Vipin Kumar and Amit Kumar, Fuzzy time series forecasting of wheat production, International Journal on Computer Science and Engineering, Vol. 02, pp. 635-640, (2010).

[20]Pankaj Kumar, Crop Yield Forecasting by Adaptive Neuro Fuzzy Inference System, Mathematical Theory and Modelling, ISSN2225-0522, Vol.1(3), (2011).

[21]Qiu W., Liu X. and Li H., A generalized method for forecasting based on fuzzy time series, Expert Systems with Applications, Vol. 38, pp. 10446-1045, (2011).

[22]Sachin Kumar and Dr. Narendra Kumar, A Novel Method for Rice Production Forecasting Using Fuzzy Time Series, International Journal of Computer Science Issues, Vol. 9, Issue 6, No. 2, (2012).

[23]Sachin Kumar and Dr. Narendra Kumar, Two Factor Fuzzy Time Series Model for Rice Forecasting, International Journal of Computer & Mathematical Sciences, ISSN2347–8527, Vol. 4, Issue 1, (2015).

[24]Singh S. R, A Robust Method of Forecasting based on Fuzzy Time Series, International Journal of Applied Mathematics and Computations, Vol. 188, pp. 472-484, (2007).

[25]Song Q. and Chissom B. S., Fuzzy Time Series and its Models, Fuzzy Sets and Systems, Vol. 54, pp. 269-277, (1993).

[26]Song Q, A Note on Fuzzy Time Series Model Selection with Sample Auto correlation Functions, an International Journal of Cybernetics and Systems, Vol. 34, and pp. 93-107, (2003).

[27]Song Q, Chissom B. S., Forecasting Enrolments with Fuzzy Time Series: Part II, Fuzzy Sets and Systems, Vol. 62, pp. 1-8, (1994).

[28]Vikas Lamba and Dhaka V.S., Wheat Yield Prediction Using Artificial Neural Network and Crop Prediction Techniques, International Journal for Research in Applied Science and Engineering Technology, Vol.2Issue IX,ISSN:2321-9653,(2014).

[29]Yolcu U., Egrioglu E., Vedide R. Uslu R., Basaran M. A., Aladag C. H.,A new approach for determining the length of intervals for fuzzy time series, Applied Soft Computing, Vol.9, pp. 647-651, (2009).