# Fuzzy Time Series Model to Forecast Rice Production

**Conference Paper** *in* IEEE International Conference on Fuzzy Systems · July 2013

DOI: 10.1109/FUZZ-IEEE.2013.6622509



Bindu Garg
Bharati Vidyapeeth(Deemed to be University)College of Engineering, Pune



Abdul Quaiyum Ansari
National Islamic University

# Fuzzy Time Series Model to Forecast Rice Production


Bindu Garg
*Department of Computer Eng.*
*Jamia Millia Islamia,*
*New Delhi-110025, India*
bindugarg80@gmail.com

M.M. Sufyan Beg
*Department of Computer Eng.*
*Jamia Millia Islamia,*
*New Delhi-110025, India*
mmsbeg@cs.berkley.edu

A.Q. Ansari
*Department of Electrical Eng.*
*Jamia Millia Islamia,*
*New Delhi-110025, India*
aqansari@ieee.org



***Abstract*- Crop production is considered as one of the real world complex problem due to its non-deterministic nature and uncertain behavior. Particularly, forecasting of rice production for a lead year is pre-eminent for crop planning, agro based resource utilization and overall management of rice production. As such, main challenge in rice production forecasting is to generate realistic method that must be capable for handling complex time series data and generating forecasting with almost negligible error. The objective of present work is to design & implement such a competent fuzzy time series model for forecasting of rice production.**

**We have proposed forecasting model based on fuzzy time series that highlights the impact of trend & seasonal components by yielding dynamic change of values from time t to t+1. The aim of using fuzzy time series is to deal with forecasting under the fuzzy environment that contains uncertainty, vagueness and imprecision. This method assigns importance to fuzzy intervals on the basis of frequency of number of time series data. Subsequently, computed fuzzy logical relations are used for analysis of time series rather than random and non-random functions as in case of usual time series analysis. Performance of the proposed model is demonstrated and compared with few pre-existing forecasting methods on rice production. To prove robustness and accuracy of the presented model, analysis is performed on forecasting of enrollment data of university of Alabama.**



***Keywords- Fuzzy Logic; Time Series; Accuracy; Forecasting;***


## I. INTRODUCTION

In view of the fact that the utmost challenge to the world in coming years is to supply food to burgeoning population, which is expected to rise 8,909 million by 2050. This situation would be more dreadful as detriment in world food production has been observed since past three decades. Although, for better results, advanced practices of cropping are acclimatized in agriculture production system. However, uncertainty exists in crop production because of some unrestrained parameters such as agro metrological variables, weather, etc. Moreover, crop production is being handled with the field data. So, accuracy of data is always matter of concern. It is studied that most of the work on time series has been done for dealing of forecasting problems like prediction in management information system, health care domain, economic forecasting, sales forecasting, budgetary analysis, stock market fluctuations, and business analysis etc. Nevertheless, there is tenacious need for precise forecasting technique to deal with non linear and intricate behavior of crop production system. Accuracy of such real time prediction has always been a challenging task. Selection of pertinent and efficient forecasting method is very crucial factor in such problem domains. Necessity of soft computing based scientifically sound and robust forecasting method is solicited.

In past, conventional statistical methods were used for time series forecasting but future values of time series data of complex system like crop production is neither exactly governed by a deterministic mathematical model nor by the probabilistic model. In the present scenario, soft computing techniques such as probabilistic reasoning, fuzzy logic, genetic algorithms, etc are being used as alternatives to the traditional statistical time series methods. Design of effectual and consistent forecasting model is required to attain accurate prediction. Due to uncertainty and some unknown parameters in crop production system, fuzzy time series prediction model becomes first choice for forecasting purpose. The proposed method gives a strong foundation for the development and application of fuzzy time series methods for short period crop production forecasting for a small area.

This paper is organized into nine sections. Section 1 is introduction. Section 2 highlights the significance and role of rice production in agriculture. Section 3 is discussion of related work on the evolution of various fuzzy time series models. Section 4 defines basic fuzzy concepts. Section 5 describes the new computational model for forecasting. Section 6 demonstrates usage of proposed model in application domain (predicting rice production). Section 7 evaluates and compares the result of proposed model with previous forecasting models and reveals its performance. Section 8 tested and verified the robustness and correctness of presented work on enrollment data. Section 9 has conclusion and future work.

## II. IMPORTANCE OF RICE PRODUCTION

Rice farming is acceded as one of the world's most productive and sustainable cropping system. Rice is the basic food for 2.7 billion people, approximately half of the world population and more than half of the world's farmers grow this crop. It is estimated that rice consumption soars by nearly 60 million tonnes almost in every eight year. Asia [1] alone consume and cultivate 90 percent of total rice production. Fundamentally, change in demand for rice is compelled by population growth, rank of per capital income, and changes in the price of rice with respect to other substitute crops [2]. It is believed that recent sharp growth of rice production in Asia will aid to mitigate the poverty since flat-out growth of rice leads to lower rice prices for indigent rural and urban consumers. As a result, comparatively lesser prices of rice will stimulate farmers to spread out into higher valued crops and therefore, provide plus income to farmers and enhanced nutrition for consumers.

A major challenge during the coming decade is to develop effective forecasting technology to increase the ability of farmers to manage the resources at their disposal more efficiently. Evaluating and predicting rice production is a complex field and challenging task. It is revealed from past work on rice that rice production varies; plus variations are not only unpredictable but also time consuming. Moreover, forecasting method and its results should be understandable not only to administrators in agriculture but also to those who use the results in decision making. So far limited research has been done in domain of crop resource utilization and crop planning using soft computing techniques. Existing methods of forecasting are confined to laboratory application and case studies. These cannot depict the complexity of crop production, making them impracticable for agriculture system.

## III. RELATED WORK

Forecasting using fuzzy time series has been emanated as an intelligent approach in the domains where information is vague and imprecise. Moreover, fuzzy time series can handle situations where neither viewing of trend is possible nor visulaization of patters in time series is handled. Substantial work has been done on forecasting problems using fuzzy time series.

Fuzzy time series definitions were proposed by Song and Chissom. Song and Chissom presented the concepts of time variant and time invariant time series [3-4]. It was applied on the time series data of university of Alabama to forecast enrollments. Song and Chissom [5] also proposed an average auto correlation function as a measure of dependency. Chen [6-7] presented simplified arithmetic operations in place of max-min composition operations which were used by Song & Chissom and then, designed high order fuzzy time series forecasted model. Hunrag [8-9] Hwang and Chen [10], Lee Wang and Chen [11], Li and Kozma [12], all developed number of fuzzy forecasting methods with some variations. Singh [13-14] developed forecasting models using computational algorithm. Lee et al. proposed a fuzzy candlestick pattern to improve forecasting results [15]. Then, a multivariate heuristic model was proposed to achieve highly complex matrix computations [16]. Work on determination on length of interval of fuzzy time series was done [17]. Qiu et al. [18] did generalization of forecasting model. Jilani et al. [19] proposed multivariate high order fuzzy time series based forecasting method for car road accidents. Fuzzy inductive model was proposed to predict dynamic student performance in e-learning course [20]. Yu [21] used concept of recurrent relationship to generate forecasting model. The expectation and the grade-selection method based on transitional weight were developed for calculating weights [22]. However this model requires large number of historical data for training process that was too complex to maintain. Forecasting models based on event discretization function was presented [23] and used for forecasting of average length of stay of patient [24]. Garg [25-26] developed forecasting model by using the concept OWA weights. This model was success to a certain extent to reduce forecasting error. Subsequently, Garg [27-28] also proposed optimized genetic-fuzzy-OWA based forecasting models. Garg [29] developed fuzzy based model to forecast number of outpatient visits in hospital. Then, accuracy of ordered weighted averaging based fuzzy predictor is enhanced using genetic algorithm [30].

However, most of the models were implemented for forecasting of other problem domains and not rice production. In this direction, we designed model to forecast rice production for lead year on basis of historical time series rice data. The proposed forecasting method controls the forecasting error by determining percentage change of historical time series data for universe of discourse. Percentage change of values from time t to t+1 helps to find out the impact of trend & seasonal components. Formation of fuzzy relations is easy as compared to min-max composition operators used by previous approaches. The presented method has time complexity in linear order. It is applied to forecast rice production on real time data of Patnanagar farm, G.B. Pant University of Agriculture & Technology, India and results have been compared with models on same rice data to prove its superiority. Further, robustness and efficiency of proposed model is tested on historical enrollment data of university of Alabama.

## IV. FUZZY TIME SERIES

Some basic fuzzy concepts are presented in own discourse to make present method self contained. For this, [1-26] are viewed.

### *1) Fuzzy Set :*

Fuzzy set is a pair (A, m), where A is a set and m: A→ |0,1|. For a finite set $A=\{x_1, ..., x_n\}$, fuzzy set (A, m) is often denoted by $\{(m(x_1)/x_1), ... , (m(x_n) / x_n)\}$. For each, x Є A, m(x) is called the grade of membership of x in (A, m). Let x Є A, then x is not included in the fuzzy set (A, m) if m(x) = 0, x is fully included if m(x) = 1 and x is called fuzzy member if $0 < m(x) < 1$. The set x Є A | m(x)>0 is called the support of (A, m) and the set x Є A | m(x) =0 is called its kernel.

### *2) Time Series*

A series of observations made sequentially in time. In time domain analysis, a time series is represented by a mathematical model G (t) = O(t) + R(t), where O(t) represents a systematic or ordered part and R(t) represents a random part. The fact is that the two components cannot be observed separately and may involve several parameters.

### *3) Fuzzy Time Series*

Let Y(t) (t = . . ., 0,1,2, . . .), subset of real numbers, be the Universe of Discourse on which fuzzy sets $f_j(t)$ are defined. If

F(t) is collection of $f_1(t)$, $f_2(t)$, . . ., then F(t) is called a fuzzy time-series defined on Y(t).

### *4) Fuzzy Relationship*

Let F(t - 1) = $A_i$ and F(t) = $A_j$. Relationship between two consecutive observations, F(t) and F(t - 1), referred to as a fuzzy logical relationship (FLR), can be denoted by $A_i \rightarrow A_j$, where $A_i$ is called the left-hand side (LHS) and $A_j$ is the right-hand side (RHS) of the FLR.

## V. PROPOSED METHOD FOR FORCASTING

In this section, computation based fuzzy time series model is proposed. We present stepwise procedure for forecasting on basis of historical time series data.

*Step 1*: Convert time series data in percentage change from time t to time t+1. That is determined as PerCh (t+1) = (X(t+1) - X(t)) \ X(t), where X(t+1) is value at time t+1 index and X(t) is actual value at time t index. PerCh is the percentage change of value from time t to t+1.

This first step makes the universe of discourse suitable for numerical evaluation by associating data at different times and eliminates subtlety of the universe of discourse. Various forecasting methods have used differences of time series data as the universe of discourse. However, increasing and decreasing rate of time series cannot be captured from difference alone.

*Step 2*: Define the universe of discourse on PerCh as U= [$D_{min}$ - $D_1$, $D_{max}$+$D_2$], where $D_{max}$ and $D_{min}$ are maximum value and minimum value of PerCh respectively. $D_1$ and $D_2$ are positive real values. Partition it into equal intervals say; $u_1$, $u_2$, $u_3$...$u_m$ of equal lengths. Number of intervals will be in accordance with the number of primary linguistic variables $B_1, B_2, \ldots, B_m$ to be considered. Here, we have m primary linguistic variables

B1: poor
$B_2$: below average
. . . .
$B_{m-2}$: very good
$B_{m-1}$: excellent
$B_m$: extraordinary

*Step 3*: Since, forecasting accuracy varies with length of intervals [8], intervals get further sub divided using frequency distribution based approach i.e. interval having maximum PerCh gets further divided into smaller intervals. It is done as:

a) Count number of PerCh fall in each interval.

b) Check for the interval that is having maximum frequency of PerCh. Break it into four intervals.

c) Do the repetition for next two intervals having higher frequency. Thereafter, break these intervals into three and two respectively. Let remaining intervals remain unchanged.

*Step 4*: After frequency distribution, number of interval will be increased by six. Now, define fuzzy sets $F_1, F_2, F_3, \ldots F_{q-1}$, Fq as linguistic variables on the re-divided intervals. Here, q=m+6. Apply triangular membership function, due to its simplicity [31], to define the fuzzy sets $F_i$. Subsequently, fuzzify time series data where fuzzy set $F_i$ denotes a fuzzy value of the PerCh represented by a fuzzy set.

*Step 5:* Fuzzy logical relationships (FLR) are then defined in such a way that influence of previous fuzzy and next fuzzy observation can also be accounted on current fuzzy observation $F_j$; besides the impact of fuzzy intervals $F_{j-1}$ and $F_{j+1}$ on $F_j$. Order of the model is three. Since it utlizes fuzzy production at time t, time t-1 and time t-2. Establish FLR using following rule:

***Rule***: If $F_j$ is the fuzzy production at time period t, $F_i$ is the fuzzify production at time period t-1 and $F_k$ is the fuzzify production at time period t+1 then the fuzzy logical relation is denoted as $F_j \rightarrow F_i$ and $F_j \rightarrow F_k$. Here, $F_j$ is called current state, $F_i$ is the previous state and $F_k$ is next state.

*Step 6*: Defuzzified value D of $F_j$ is obtained using defuzzification formula in eq. (1) [19]:

$$D=1.5/((1/f_{j-1})+(0.5/f_j)) \quad \text{if } j=1$$
$$D=2/((0.5/f_{j-1})+(1/f_j)+(0.5/f_{j+1})), \quad \text{where } 1<j<n \qquad (1)$$
$$D=1.5/((0.5/f_{j-1})+(1/f_j)) \quad \text{if } j=n$$

In eq. (1) $f_{j-1}$, $f_j$, $f_{j+1}$ are the mid points of the fuzzy intervals $F_{j-1}$, $F_j$, $F_{j+1}$ respectively. eq. (1) fulfills the axioms of fuzzy set like monotonicity, boundary condition, continuity and idempotency.

*Step 7:* Forecasting production in terms of FPerCh at time t+1 is computed by first determining TV as time variant parameter. TV is difference of PerCh at time t-1, t and t+1. Forecasting is done by performing computations in form of arithmetic operations on defuzzified value D by adjusting time variant parameter TV. These arithmetic operations generate the relation between time series data for time t, t-1 and t+1. Finally, forecasted PerCh for the time t+1 is obtained. For each time series data t=2 to n, forecasting for time t+1 is done as:

$TV_i$, time variant parameter is calculated using eq. (2):

$$TV_i = \big|\, |(PerCh_j - PerCh_i)| - |(PerCh_k - PerCh_j)| \,\big| \qquad (2)$$

Here, $PerCh_i$, $PerCh_j$ and $PerCh_k$ are percentage change of $F_i$, $F_j$ and $F_k$ respectively, such that $F_i$, $F_j$ and $F_k$ are fuzzy values at time t-1, t, t+1 and logical relation between them is $F_j \rightarrow F_i$ and $F_j \rightarrow F_k$.

*Defining of Variables*

$TV_i$, $TV_i$ /2, $TV_i$ /4, $TV_i$ /8 are added fractions values in D.

E is total defuzzified value at any iteration,

I is intermittent fuzzy value & count is iteration count

[$*F_j$] is corresponding interval $u_j$ for which membership in $F_j$ is Supremum (i.e. 1),

L[$*F_j$] is the lower bound of interval $u_j$ and U[$*F_j$] is the upper bound of interval $u_j$

*Computations*

Calculate $TV_i$ /2, $TV_i$ /4 and $TV_i$ /8 from $TV_i$.

Let, X(1)= $TV_i$, X(2)= $TV_i$ /2, X(3)= $TV_i$ /4, X(4)=D/8.

Initialize count as 1 and E=D

For i=1 to 4 do

I=D + X(i)

If U [$*F_j$] ≤ I ≤ L [$*F_j$] than E=E+I and count=count+1

I=D - X(i)

If U [$*F_j$] ≤ I≤ L [$*F_j$] than E=E+I and count=count+1

$FPerCh_t$ = E/count ($FPerCh_t$ is optimized $F_j$ at time t)

Return $FPerCh_t$

*Step 8:* For time series data t=1 to n-1, calculate forecasted value as $Fore.val_{t+1}=(X(t)*FPerCh_{t+1})+X(t)$. Here, X(t) is actual value at time t and $FPerCh_t$ is corresponding value obtained in step 7.

## VI. FORECASTING OF RICE PRODUCTION

Proposed method is being applied to forecast rice production. The historical time series data of rice production [13] are of the farm of G.B. Pant University,Pantnagar, India. The historical time series data of rice production is in terms of productivity in kg per hectare. Step wise computations are as:

*Step 1:* Histological Data is given year wise. Calculate $PerCh_t$ of year 1981 and onwards is shown in Table I.

TABLE I Calculation of PerCh

| Year | Production (kg) | PerCh (%) | Year | Production (kg) | PerCh (%) |
|---|---|---|---|---|---|
| 1981 | 3552 | - | 1991 | 3851 | 2.69 |
| 1982 | 4177 | 17.60 | 1992 | 3231 | -16.10 |
| 1983 | 3372 | -19.27 | 1993 | 4170 | 29.06 |
| 1984 | 3455 | 2.46 | 1994 | 4554 | 9.21 |
| 1985 | 3702 | 7.15 | 1995 | 3872 | -14.98 |
| 1986 | 3670 | -0.86 | 1996 | 4439 | 14.64 |
| 1987 | 3865 | 5.31 | 1997 | 4266 | -3.90 |
| 1988 | 3592 | -7.06 | 1998 | 3219 | -24.54 |
| 1989 | 3222 | -10.30 | 1999 | 4305 | 33.74 |
| 1990 | 3750 | 16.39 | 2000 | 3928 | -8.76 |

*Step 2*: Obtain $D_{min}$= -24.54, $D_{max}$=33.74 from Table I. Define the universe of discourse *U* and partition it into intervals $u_1$, $u_2 \dots u_m$ of equal length. Thus universe of discourse will be U = [-30, 40]. Where $D_1$=5.46 and $D_2$ = 6.26. Partitioning U into seven equal intervals as $u_1$=[-30 -20], $u_2$=[-20 -10], $u_3$=[-10 0], $u_4$=[0 10], $u_5$=[ 10 20], $u_6$=[20 10], $u_7$=[30 40]

*Step 3*: The frequency density based distribution and partitioning of PerCh is given in Table II and Table III.

TABLE II Frequency of PerCh

| Intervals | [−30,−20] | [−20,-10] | [-10, 0] | [0,10] | [10,20] | [20,30] | [30,40] |
|---|---|---|---|---|---|---|---|
| Number of Data | 1 | 4 | 4 | 5 | 3 | 1 | 1 |

TABLE III Frequency Distribution and Fuzzy set

| Fuzzy | Intervals | Fuzzy | Intervals |
|---|---|---|---|
| $F_1$ | [−30,−20] | $F_8$ | [2.5, 5] |
| $F_2$ | [-20,−16.67] | $F_9$ | [5, 7.5] |
| $F_3$ | [−16.67, -13.33] | $F_{10}$ | [7.5, 10] |
| $F_4$ | [-13.33, -10] | $F_{11}$ | [10 20] |
| $F_5$ | [-10,-5] | $F_{12}$ | [20 30] |
| $F_6$ | [-5,0] | $F_{13}$ | [30 40] |
| $F_7$ | [0, 2.5] | | |

Define thirteen fuzzy sets as linguistic variables on redivided intervals. These fuzzy variables are defind as:

TABLE IV Frequency Distribution and Fuzzy set

| Fuzzy Set | Linguistic Variable |
|---|---|
| $F_1$ | *very very poor rice production* |
| $F_2$ | *very poor Production* |
| $F_3$ | *poor rice production* |
| $F_4$ | *above poor rice production* |
| $F_5$ | *below average rice production* |
| $F_6$ | *average rice production* |
| $F_7$ | *above average rice production* |
| $F_8$ | *good rice production* |
| $F_9$ | *very good rice production* |
| $F_{10}$ | *very very good rice production* |
| $F_{11}$ | *Excellent rice production* |
| $F_{12}$ | *Above Excellent rice production* |
| $F_{13}$ | *Extraordinary rice production* |

The membership grades to these fuzzy sets which are assign to linguistic values are defined as:

$F_1 = 1/ u_1 +0.5/ u_2 +0/ u_3+ 0/ u_4 + 0/ u_5 +0/ u_6 +0/ u_7 +0/ u_8+0/ u_9 + 0/ u_{10} +0/ u_{11} +0/ u_{12}+0/ u_{13}$

$F_2= 0.5/ u_1 +1/ u_2 +0.5/ u_3+ 0/ u_4 + 0/ u_5 +0/ u_6 +0/ u_7 +0/ u_8+0/ u_9 + 0/ u_{10} +0/ u_{11} +0/ u_{12}+0/ u_{13}$

$F_3= 0/ u_1 +0.5/ u_2 +1/ u_3+ 0.5/ u_4 + 0/ u_5 +0/ u_6 +0/ u_7 +0/ u_8+0/ u_9 + 0/ u_{10} +0/ u_{11} +0/ u_{12}+0/ u_{13}$

$F_4 = 0/ u_1 +0 / u_2 +0.5/ u_3+ 1/ u_4 + 0.5/ u_5 +0/ u_6 +0/ u_7 +0/ u_8+0/ u_9 + 0/ u_{10} +0/ u_{11} +0/ u_{12}+0/ u_{13}$

$F_5 = 0/ u_1 +0/ u_2 +0/ u_3+ 0.5/ u_4 + 1/ u_5 +0.5/ u_6 +0/ u_7 +0/ u_8+0/ u_9 + 0/ u_{10} +0/ u_{11} +0/ u_{12}+0/ u_{13}$

$F_6 = 0/ u_1 +0/ u_2 +0/ u_3+ 0/ u_4 + 0.5/ u_5 +1/ u_6 +0.5/ u_7 +0/ u_8+0/ u_9 + 0/ u_{10} +0/ u_{11} +0/ u_{12}+0/ u_{13}$

$F_7 = 0/ u_1 +0.5/ u_2 +0/ u_3+ 0/ u_4 + 0/ u_5 +0.5/ u_6 +1/ u_7 +0.5/ u_8+0/ u_9 + 0/ u_{10} +0/ u_{11} +0/ u_{12}+0/ u_{13}$

$F_8= 0/ u_1 +0.5/ u_2 +0/ u_3+ 0/ u_4 + 0/ u_5 +0/ u_6 +0.5/ u_7 +1/ u_8+0.5/ u_9 + 0/ u_{10} +0/ u_{11} +0/ u_{12}+0/ u_{13}$

$F_9 = 0/ u_1 +0.5/ u_2 +0/ u_3+ 0/ u_4 + 0/ u_5 +0/ u_6 +0/ u_7 +0.5/ u_8+1/ u_9 + 0.5/ u_{10} +0/ u_{11} +0/ u_{12}+0/ u_{13}$

$F_{10}= 0/ u_1 +0.5/ u_2 +0/ u_3+ 0/ u_4 + 0/ u_5 +0/ u_6 +0/ u_7 +0/ u_8+0.5/ u_9 + 1/ u_{10} +0.5/ u_{11} +0/ u_{12}+0/ u_{13}$

$F_{11} = 0/ u_1 +0.5/ u_2 +0/ u_3+ 0/ u_4 + 0/ u_5 +0/ u_6 +0/ u_7 +0/ u_8+0/ u_9 + 0.5/ u_{10} +1/ u_{11} +0.5/ u_{12}+0/ u_{13}$

$F_{12} = 0/ u_1 +0.5/ u_2 +0/ u_3+ 0/ u_4 + 0/ u_5 +0/ u_6 +0/ u_7 +0/ u_8+0/ u_9 + 0/ u_{10} +0.5/ u_{11} +1/ u_{12}+0.5/ u_{13}$

$F_{13} = 0/ u_1 +0.5/ u_2 +0/ u_3+ 0/ u_4 + 0/ u_5 +0/ u_6 +0/ u_7 +0/ u_8+0/ u_9 + 0/ u_{10} +0/ u_{11} +0.5/ u_{12}+1/ u_{13}$

*Step 4*: Do fuzzification of data as shown in Table V.

*Step 5*: Create FLR as shown in Table V.

TABLE V Fuzzy Sets & Fuzzy Relationships

| Year | Fuzzy | FLR | Year | Fuzzy | FLR |
|---|---|---|---|---|---|
| 1981 | | | 1991 | $F_8$ | $F_8 \rightarrow F_{11},\ F_8 \rightarrow F_3$ |
| 1982 | $F_{11}$ | $F_{11} \rightarrow F_2$ | 1992 | $F_3$ | $F_3 \rightarrow F_8,\ F_3 \rightarrow F_{12}$ |
| 1983 | $F_2$ | $F_2 \rightarrow F_{11},\ F_2 \rightarrow F_7$ | 1993 | $F_{12}$ | $F_{12} \rightarrow F_3,\ F_{12} \rightarrow F_{10}$ |
| 1984 | $F_7$ | $F_7 \rightarrow F_2,\ F_7 \rightarrow F_2$ | 1994 | $F_{10}$ | $F_{10} \rightarrow F_{12},\ F_{10} \rightarrow F_3$ |
| 1985 | $F_9$ | $F_9 \rightarrow F_7,\ F_9 \rightarrow F_7$ | 1995 | $F_3$ | $F_3 \rightarrow F_{10},\ F_3 \rightarrow F_{11}$ |
| 1986 | $F_6$ | $F_6 \rightarrow F_9,\ F_6 \rightarrow F_9$ | 1996 | $F_{11}$ | $F_{11} \rightarrow F_3,\ F_{11} \rightarrow F_3$ |
| 1987 | $F_9$ | $F_9 \rightarrow F_6,\ F_9 \rightarrow F_5$ | 1997 | $F_6$ | $F_6 \rightarrow F_{11},\ F_6 \rightarrow F_1$ |
| 1988 | $F_5$ | $F_5 \rightarrow F_9,\ F_5 \rightarrow F_4$ | 1998 | $F_1$ | $F_1 \rightarrow F_6,\ F_1 \rightarrow F_{13}$ |
| 1989 | $F_4$ | $F_4 \rightarrow F_5,\ F_4 \rightarrow F_{11}$ | 1999 | $F_{13}$ | $F_{13} \rightarrow F_1,\ F_{13} \rightarrow F_5$ |
| 1990 | $F_{11}$ | $F_{11} \rightarrow F_4,\ F_{11} \rightarrow F_8$ | 2000 | $F_5$ | $F_5 \rightarrow F_{13},$ |

*Step 6*: Defuzzified value D of PerCh is presented in Table VI.

TABLE VI Defuzzification of PerCh

| Year | Fuzzy | D (%) | Year | Fuzzy | D (%) |
|---|---|---|---|---|---|
| 1981 | | | 1991 | $F_8$ | 7.22 |
| 1982 | $F_{11}$ | | 1992 | $F_1$ | 18.64 |
| 1983 | $F_1$ | -18.98 | 1993 | $F_{12}$ | 26.98 |
| 1984 | $F_8$ | 2.35 | 1994 | $F_9$ | 17.98 |
| 1985 | $F_9$ | 5.56 | 1995 | $F_2$ | 45.01 |
| 1986 | $F_7$ | -8.33 | 1996 | $F_{10}$ | -12.79 |
| 1987 | $F_9$ | 5.88 | 1997 | $F_6$ | -5.17 |
| 1988 | $F_5$ | -14.79 | 1998 | $F_1$ | -8.96 |
| 1989 | $F_3$ | -16.80 | 1999 | $F_{12}$ | -37.50 |
| 1990 | $F_{11}$ | 12.73 | 2000 | $F_4$ | -12.86 |

*Step7 and Step 8*: Optimized value of D is calculated as FPerCh in Table VII. Final forecasted value (Fore. Rice) is in column 5 of Table VII.

TABLE VII FPerCh, Fore. Rice, AFER &MSE

| Year | Production(kg) | D(%) | FPerCh | Fore. Rice (kg/ha) | $(A_i - F_i)^2$ | $\lvert A_i - F_i\rvert/A_i$ |
|---|---|---|---|---|---|---|
| 1981 | 3552 | | | | | |
| 1982 | 4177 | | | | | |
| 1983 | 3372 | -18.98 | -0.19 | 3384.21 | 148.97 | 0.00362 |
| 1984 | 3455 | 2.35 | 0.02 | 3451.24 | 14.12 | 0.001088 |
| 1985 | 3702 | 5.56 | 0.06 | 3671.97 | 901.56 | 0.008111 |
| 1986 | 3670 | -8.33 | -0.01 | 3671.27 | 1.62 | 0.000347 |
| 1987 | 3865 | 5.88 | 0.07 | 3910.02 | 2026.62 | 0.011648 |
| 1988 | 3592 | -14.79 | -0.10 | 3468.84 | 15169.00 | 0.034288 |
| 1989 | 3222 | -16.80 | -0.13 | 3129.71 | 8517.52 | 0.028644 |
| 1990 | 3750 | 12.73 | 0.15 | 3694.99 | 3026.14 | 0.014669 |
| 1991 | 3851 | 7.22 | 0.04 | 3898.13 | 2220.77 | 0.012237 |
| 1992 | 3231 | 18.64 | -0.17 | 3196.72 | 1175.45 | 0.010611 |
| 1993 | 4170 | 26.98 | 0.28 | 4129.86 | 1610.88 | 0.009625 |
| 1994 | 4554 | 17.98 | 0.10 | 4586.17 | 1034.65 | 0.007063 |
| 1995 | 3872 | 45.01 | -0.16 | 3829.46 | 1809.77 | 0.010987 |
| 1996 | 4439 | -12.79 | 0.16 | 4474.48 | 1259.06 | 0.007994 |
| 1997 | 4266 | -5.17 | -0.04 | 4244.57 | 459.17 | 0.005023 |
| 1998 | 3219 | -8.96 | -0.25 | 3184.14 | 1215.05 | 0.010829 |
| 1999 | 4305 | -37.50 | 0.36 | 4362.71 | 3330.52 | 0.013406 |
| 2000 | 3928 | -12.86 | -0.11 | 3842.21 | 7359.50 | 0.02184 |
| | | | | | **MSE= 2848.91** | **AFER= 1.177934** |

Comparison between proposed model, existing models and actual value for rice production is shown in Table VIII.

TABLE VIII Actual Vs Forecasted rice production

| Year | Actual (kg/ha) | Proposed | Singh [13] | Chen [6] | Song&[3] Chissom |
|---|---|---|---|---|---|
| 1981 | 3552 | | | | |
| 1982 | 4177 | | | | |
| 1983 | 3372 | 3384.21 | | | |
| 1984 | 3455 | 3451.24 | 3490 | 3900 | 3900 |
| 1985 | 3702 | 3671.97 | 3690.25 | 3700 | 3700 |
| 1986 | 3670 | 3671.27 | 3700 | 3800 | 3800 |
| 1987 | 3865 | 3910.02 | 3874.16 | 3800 | 3800 |
| 1988 | 3592 | 3468.84 | 3537 | 3767 | 3767 |
| 1989 | 3222 | 3129.71 | 3322 | 3700 | 3700 |
| 1990 | 3750 | 3694.99 | 3700 | 3900 | 3900 |
| 1991 | 3851 | 3898.13 | 3879 | 3800 | 3800 |
| 1992 | 3231 | 3196.72 | 3300 | 3767 | 3767 |
| 1993 | 4170 | 4129.86 | 4100 | 3900 | 3900 |
| 1994 | 4554 | 4586.17 | 4459.5 | 3900 | 3900 |
| 1995 | 3872 | 3829.46 | 3922.5 | 4000 | 4100 |
| 1996 | 4439 | 4474.48 | 4498 | 3767 | 3767 |
| 1997 | 4266 | 4244.57 | 4318.33 | 4000 | 4100 |
| 1998 | 3219 | 3184.14 | 3300 | 3600 | 3600 |
| 1999 | 4305 | 4362.71 | 4300 | 3900 | 3900 |
| 2000 | 3928 | 3842.21 | 3900 | 3600 | 3600 |

Fig. 1 shows the existing trends of forecasted rice production obtained by proposed method, Singh [13], Chen [6] and Song & Chissom [3]. It can be observed from Fig. 1 that the forecasted values computed by the proposed model are significantly in closed accordance to the actual productions.

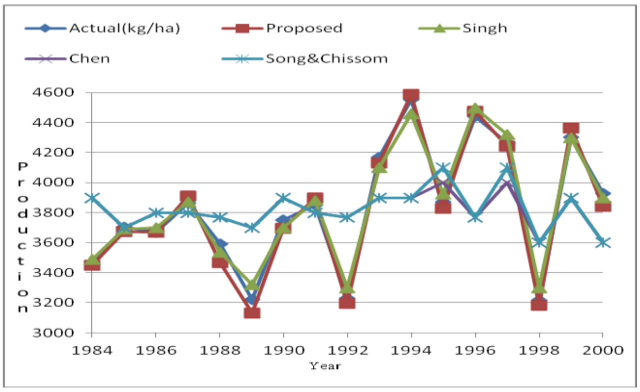


Fig 1 Actual Rice Production Vs Forecast rice Production of various models

## VII. Performance Evaluation and Comparison

In this section, we assessed the forecasting efficacy of our proposed model for rice production and juxtaposed the results with previous selective models [3,6,13]. These models used same data set. Forecasting accuracy is measured in terms of mean square error (MSE) and average forecasting error rate (AFER). Lower value of MSE and AFER are measure of higher forecasting precisioness.

*1)* Average Forecasting Error Rate

AFER can be defined as

$$AFER = \left( \sum_{n}^{t=1} (\lvert A_t - F_t\rvert / A_t) \right)/n \ * 100\% \qquad (3)$$

*2)* Mean Square Error

MSE can be defined as

$$MSE=\sum_{n}^{t=1}(A_t - F_t)^2/n \quad (4)$$

where, $A_t$ is actual value and $F_t$ is forecasted value of time series data at time t and n is total number of time series data.

TABLE IX MSE & AFER of various forecasting models

| Method | MSE | AFER |
|---|---|---|
| Proposed Method | **2848.91** | **1.177934%** |
| Singh[13] | 3140.4 | 1.304412% |
| Chen[6] | 132,162.9 | 7.934613% |
| Song &Chissom[3] | 131,715.9 | 7.948644% |

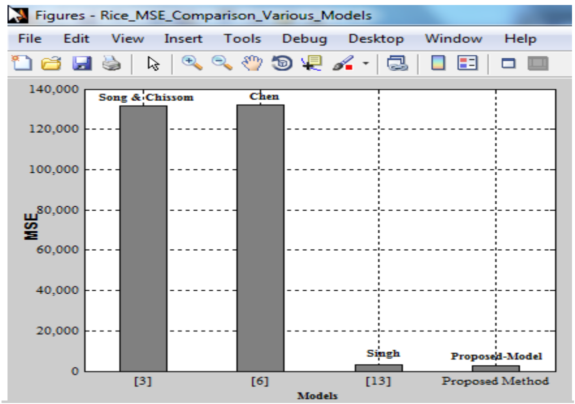


Fig 2 MSE Comparison for Rice Production

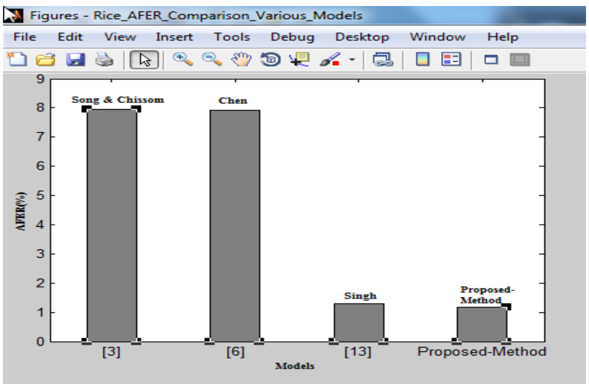


Fig 3 AFER Comparison for Rice Production

*Discussion:* From Table IX, Fig. 2 and Fig. 3, it can be seen that MSE and AFER are lowest in case of proposed method, clearly indicating its transcendence over existing like Song & Chissom, Chen and Singh forecasting models for rice production. Proposed method determined number of intervals by using frequency distribution. This eliminates the inadequacies by way of defining universe of discourse on static historical data so that intervals get weight-age on the basis of maximum rate of change along with frequency of PerCh. Re-division of intervals into effective length enhanced forecasted accuracy to a certain factor. Subsequently, FLR are defined on basis of impact of past and future fuzzy observation on current fuzzy observation. It used 3rd order time series model, sufficient to remove ambiguities while preserving efficiency. Dominant reason for accurateness of proposed method is that first it determined defuzzified value for time t+1. Then it performed effective computations by calculating a time variant factor TV. TV captured time varying difference of current observation from past and future observation. Then, time variant difference is attuned by performing arithmetic operation on defuzzified value.

## VIII. Accuracy & Robustness: Forecasting of Enrollment Data

Correctness and potency of proposed model is tested and verified on time series enrollment data of university of Alabama that is used as benchmark data by most of the fuzzy time series forecasting models. The following experiment on enrollment time series data can be viewed as general applicability of proposed method on historical time series data also. Computations of proposed method on the time series data of enrollments [4, 7, 9, 10, 13, 14, 19] are shown in Table X.

TABLE X FPerCh, Forecasted, AFER &MSE(Enroll)

| Year | Enroll. | D | FPerCh | Fore.Enroll | $(A_i - F_i)^2$ | $\lvert A_i - F_i \rvert / A_i$ |
|---|---|---|---|---|---|---|
| 1971 | 13055 | | | | | |
| 1972 | 13563 | 3.33% | 3.33% | 13489.732 | 5368.4929 | 0.054021 |
| 1973 | 13867 | 2.67% | 2.49% | 1390.719 | 1136.95073 | -0.0024316 |
| 1974 | 14696 | 5.76% | 5.49% | 14628.298 | 4583.52018 | 0.006068 |
| 1975 | 15460 | 4.96% | 4.96% | 15424.922 | 1230.49415 | 0.002269 |
| 1976 | 15311 | 2.38% | 0.65% | 15560.49 | 62245.2601 | -0.0162948 |
| 1977 | 15603 | 1.71% | 1.38% | 15522.292 | 6513.81355 | 0.0051726 |
| 1978 | 15861 | 1.71% | 1.20% | 15790.236 | 5007.5437 | 0.0044615 |
| 1979 | 16807 | 5.76% | 5.67% | 16760.319 | 2179.14377 | 0.0027775 |
| 1980 | 16919 | 0.54% | 0.82% | 16944.817 | 666.538143 | -0.0015259 |
| 1981 | 16388 | -2.14% | -2.66% | 16468.955 | 6553.64726 | -0.0049399 |
| 1982 | 15433 | -4.10% | -4.68% | 15621.042 | 35359.6433 | -0.0121844 |
| 1983 | 15497 | 0.48% | 0.30% | 15479.299 | 313.325401 | 0.0011422 |
| 1984 | 15145 | -2.14% | -2.36% | 15131.271 | 188.490933 | 0.0009065 |
| 1985 | 15163 | 0.48% | 0.12% | 15163.174 | 0.030276 | -1.148E-05 |
| 1986 | 15984 | 5.76% | 5.76% | 16036.389 | 2744.58637 | -0.0032776 |
| 1987 | 16859 | 5.76% | 5.49% | 16861.522 | 6.35846656 | -0.0001496 |
| 1988 | 18150 | 6.49% | 6.16% | 17897.514 | 63748.9782 | 0.0139111 |
| 1989 | 18970 | 4.22% | 4.48% | 18963.12 | 47.3344 | 0.0003627 |
| 1990 | 19328 | 1.71% | 1.71% | 19294.387 | 1129.83377 | 0.0017391 |
| 1991 | 19337 | 0.48% | -0.45% | 19241.024 | 9211.39258 | 0.0049633 |
| 1992 | 18876 | -2.41% | -2.41% | 18870.978 | 25.2174709 | 0.000266 |
| | | | | | **MSE= 9917.17122** | **AFER= 0.34%** |

Graphical comparison among actual enrollment data and forecasted enrollment data by various models can be seen in Fig. 4. It is observed that curve of predicted enrollment data by proposed method is closest to curve of actual enrollment data. Fig. 4 also displays the current trends among forecasted values attained by various forecasting models.

TABLE XI Actual Vs Forecasting Enrollment, MSE & AFER

| Year | Actual Enroll. | Song Chis-som [4] | Chen [7] | Hwa-ng [10] | Hur-ang [9] | Singh [13] | Jilani, [28] | Pro-posed Method |
|---|---|---|---|---|---|---|---|---|
| 1971 | 13055 | – | – | – | – | | 13579 | |
| 1972 | 13563 | 14000 | 14000 | – | 14000 | | 13798 | 13489.732 |
| 1973 | 13867 | 14000 | 14000 | – | 14000 | | 13798 | 13900.719 |
| 1974 | 14696 | 14000 | 14000 | – | 14000 | 14286 | 14452 | 14628.298 |
| 1975 | 15460 | 15500 | 15500 | – | 15500 | 15361 | 15373 | 15424.922 |
| 1976 | 15311 | 16000 | 16000 | 16260 | 15500 | 15468 | 15373 | 15560.49 |
| 1977 | 15603 | 16000 | 16000 | 15511 | 16000 | 15512 | 15623 | 15522.292 |
| 1978 | 15861 | 16000 | 16000 | 16003 | 16000 | 15582 | 15883 | 15790.236 |
| 1979 | 16807 | 16000 | 16000 | 16261 | 16000 | 16500 | 17079 | 16760.319 |
| 1980 | 16919 | 16813 | 16833 | 17407 | 17500 | 16361 | 17079 | 16944.817 |
| 1981 | 16388 | 16813 | 16833 | 17119 | 16000 | 16362 | 16497 | 16468.955 |
| 1982 | 15433 | 16789 | 16833 | 16188 | 16000 | 15735 | 15737 | 15621.042 |
| 1983 | 15497 | 16000 | 16000 | 14833 | 16000 | 15446 | 15737 | 15479.299 |
| 1984 | 15145 | 16000 | 16000 | 15497 | 15500 | 15498 | 15024 | 15131.271 |
| 1985 | 15163 | 16000 | 16000 | 14745 | 16000 | 15306 | 15024 | 15163.174 |
| 1986 | 15984 | 16000 | 16000 | 15163 | 16000 | 15442 | 15883 | 16036.389 |
| 1987 | 16859 | 16000 | 16000 | 16384 | 16000 | 16558 | 17079 | 16861.522 |
| 1988 | 18150 | 16813 | 16833 | 17659 | 17500 | 18500 | 17991 | 17897.514 |
| 1989 | 18970 | 19000 | 19000 | 19150 | 19000 | 18475 | 18802 | 18963.12 |
| 1990 | 19328 | 19000 | 19000 | 19770 | 19000 | 19382 | 18994 | 19294.387 |
| 1991 | 19337 | 19000 | 19000 | 19928 | 19500 | 19487 | 18994 | 19241.024 |
| 1992 | 18876 | – | 19000 | 15837 | 19000 | 18744 | 18916 | 18870.978 |
| **MSE** | **-** | **775687** | **321418** | **226611** | **86694** | **90997** | **41426** | **9917.16** |
| **AFER** | **-** | **4.38%** | **3.12%** | **2.45%** | **1.53%** | **1.53%** | **1.02%** | **0.34%** |

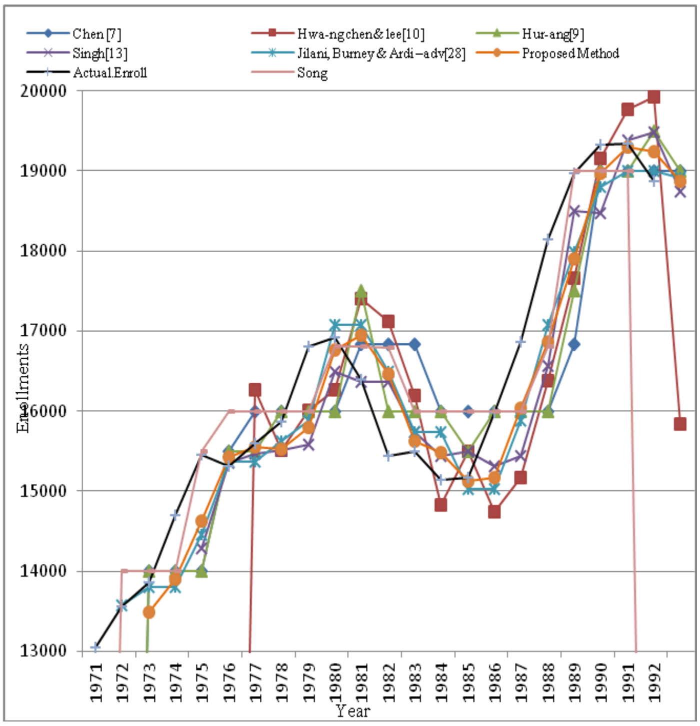


Fig 4 Actual Vs Forecasting of Enroll data

From Fig. 5 and Fig. 6 we can see that least value of MSE & AFER of enrollment ascertains correctness and robustness of proposed method than ones presented [4, 7, 10, 9, 13, and 19]. Proposed method started designing of model by capturing trend in form of increasing and decreasing rate of time series data. As discussed, frequency distribution on intervals and computations using time variant parameter are major factors to enhance accuracy by minimizing the disparity between actual and forecasting enrollments.

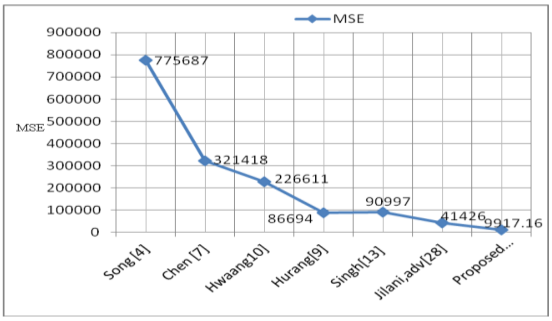


Fig 5 MSE Comparison of Various forecasting models on Enrollment Data

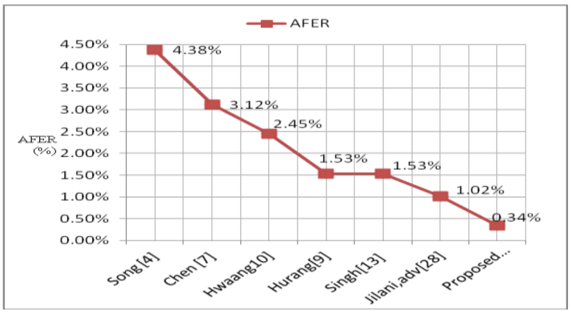


Fig 6 AFER Comparison of Various forecasting models on Enrollment Data

## IX. CONCLUSION

In this paper, a new fuzzy time series approach based upon percentage change, determination of effective length of intervals and emphatic computations on time series data is presented to improve prediction accuracy. First, method is implemented for forecasting of rice production. Subsequently, we proved correctness and robustness of the proposed approach by testing and verifying it on historical time series enrollment data. It is observed that new technique introduced instates the highest accuracy having smallest mean square error and average forecasting error rate than ones presented forecasting model in this domain. Thus, pioneered computational method can be employed as an accurate and reliable means for estimation and prediction of crop production. Proposed concept can also be considered as a strong standard methodology for better resource allocation, planning and management in other domains. In future work, the proposed model can be extended to deal with multidimensional time series data and optimized with genetic algorithm.

## REFERENCES


[1] L.T. Evans, Feeding the Ten Billion: plants and population growth. University Press, Cambridge, UK, Pg: 247,(1998).

[2] D.J. Greenland, The Sustainability of Rice Farming. CAB International, Wallingford,UK, Pg:273, (1997).

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

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

[5] Q. Song, A Note on Fuzzy Time Series Model Selection with Sample Autocorrelation Functions, An International Journal of Cybernetics and Systems, Vol. 34, Pg: 93-107,( 2003).

[6] S.M. Chen., Forecasting Enrollments based on Fuzzy Time Series, Fuzzy Sets and Systems, Vol. 81, Pg: 311-319, (1996).

[7] S. M. Chen, Forecasting Enrollments based on High Order Fuzzy Time Series, Intl Journal of Cybernetics and Systems, Vol. 33, Pg:1-16, (2002).

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

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

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

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

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

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

[14] S.R Singh, A Computational Method of Forecasting based on Fuzzy Time Series, International Journal of Mathematics and Computers in Simulation , Vol. 79, Pg:539-554, (2008).

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

[16] K.H. Huanrg, T.H.K. Yu and Y.W. Hsu, A Multivariate Heuristic Model for Forecasting, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 37, Pg: 836–846, (2007).

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

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

[19] T. A. Jilani, S. M. A. Burney, C. Ardil, Multivariate High Order Fuzzy Time Series Forecasting for Car Road Accidents, International Journal of Computational Intelligence, Vol. 4, Pg:15-20, (2007).

[20] Nebol. A, Fuzzy Predictive Models to Help Teachers in E-Learning Courses,2010 International Joint Conference on Neural Networks (IJCNN), Pg:1-7, 18-23 July 2010.

[21] H.K. Yu, Weighted Fuzzy Time Series Models for TAIEX Forecasting, Physica A, Vol. 349, Pg: 609–624, (2005).

[22] Ching-Hsue Cheng, Jia-Wen Wang, Chen-Hsun Li, Forecasting the Number of Outpatient Visits using a New Fuzzy Time Series based on Weighted-Transitional Matrix, Expert Systems with Applications, Vol. 34, Pg: 2568–2575, (2008).

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

[24] B. Garg, M. M. S. Beg, A.Q. Ansari and B. M. Imran, Soft Computing Model to Predict Average Length of Stay of Patient,, Communications in Computer and Information Science, Springer Verlag Berlin Heidelberg ,ISBN978-3-642-19423-8, Vol. 141, Pg: 221–232, (2011).

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

[26] B. Garg, M. M. S. Beg, and A. Q. Ansari, OWA based Fuzzy Time Series Forecasting Model, World Conference on Soft Computing, Berkeley, San Francisco, CA, Pg: 141-177, May 23-26, (2011).

[27] B. Garg, M. M. S. Beg and A. Q. Ansari, Enhanced Accuracy of Fuzzy Time Series Predictor using Genetic Algorithm, Third IEEE World Congress on Nature and Biologically Inspired Computing (NaBIC2011), Pg:273-278, Spain, (2011 ).

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

[29] B. Garg, M. M. S. Beg and A. Q. Ansari, 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 (NAFIPS 2012), University of California at Berkeley, USA, Pg:1-6, August 6-8, (2012).

[30] B. Garg, M. M. S. Beg and A. Q. Ansari, Enhanced Accuracy of Ordered Weighted Averaging based Fuzzy Predictor using Genetic Algorithm, IEEE Transactions on Systema, Man and Cybernetics, Part B, Submitted.

[31] W. Pedrycz, Why Triangular Membership Function, Fuzzy Set And Systems, Elsevier, Vol. 64, Pg: 21-30, (1994).