RASHI 3 (1) : 46-51 (2018)

Prediction of rice production using Fuzzy based forecasting model

Surjeet Kumar, Manas Kumar Sanyal
[1]Satyendra Nath Mandal and [2]Asit Naskar

Dept. of Business Administration, University of Kalyani, Nadia- 741235,W.B)
[1]Dept. of I.T, Kalyani Govt. Engg College, Kalyani ,Nadia(W.B),India,
[2]Dept. of CSE, Kalyani Govt. Engg College, Kalyani , Nadia(W.B),India

Received : 03-10-2018 ; Revised : 15-12-2018 ; Accepted : 30-12-2018

ABSTRACT

This Paper represents a new approach depends on fuzzy logical relationships in the fuzzy time series forecasting method. The main objective of the paper is to develop a model which reduces the average forecasting error of the existing fuzzy time series forecasting method and increases the accuracy of prediction value. The historical secondary published data relating to production of rice is used. A better forecasting accuracy and minimum average error over the traditional time series model is observed.

Keywords : Average errors, fuzzy time series, fuzzy logical relationships, fuzzy logical relationship group, MSE

1. INTRODUCTION

Time series forecasting is an important problem dealing with variety of application in our everyday life to make future decisions such as weather forecasting, university enrollments, production of grain etc. The concept of fuzzy time series, capable of dealing with vague and imprecise data presented in terms of linguistic variables was developed by Song and Chissom [1] by using the theory of sets and linguistic variable given by Zadeh [2,3].Song and Chissom [4,5] further extended his theory of fuzzy time series to be capable of dealing with numerical data by introducing the concept of fuzzification and defuzzification and applied it to forecast the student enrollments of university of Alabama. Chen [6] resolved the problem of large computational requirements of Song and Chissom method of computing fuzzy relations using max-min composition by replacing it with the simplified arithmetic operations and applied the method on the student enrollments of University of Alabama. During the last few decades, several methods for time series forecasting have been presented by Tsai and Wu [7], Chen and Hwang [8], Huarng [9], Sullivan and Woodall [10] made a comparative study of fuzzy time series forecasting and Markov modeling. Kim and Lee [11] proposed a fuzzy time series prediction method based on consecutive values. Another direction for the improvement of fuzzy time series forecasting emerged as the time variant models by application of the high-order methods in the fuzzy time series forecasting. The major problem associated with Song and Chissom [5] is the need of large computational requirements. To forecast enrollments of year t, the number of past years of the enrollments data used was called the window basis and obtained the enrollments forecast for various window basis and succeeded in improvement in the forecasting errors. The objective of the present work is to develop a computational method of forecasting based on high-order (order 4) fuzzy logical relations with a motivation to reduces the average forecasting error of the existing fuzzy time series forecasting method and increases the accuracy of prediction value in agricultural production system to help and support the farmers, producers or decision makers. The new approach is applied on the time series data of rice(paddy) crop production of Pantnagar farm G.B.Pant University of Agriculture and Technology, Pantnagar (INDIA). Here, the rice production has been recorded in terms of quintal per hectare. The study comprise of model development, and testing on rice production forecast to compare its suitability in forecasting over the other forecasting.

2. BASICS OF FUZZY TIME SERIES

In view of making our exposition self contained, some basic definitions of fuzzy time series models presented in literature are summarized and is reproduced as [1,4,5].

Definition 2.1. A fuzzy set is a class of objects with a continuum of grade of membership. Let U be the Universe of discourse with U={ u1, u2, u3, …un },where U_i are possible linguistic values of U, then a fuzzy set of linguistic variables Ai of U is defined by -

Ai = $\mu_{Ai}(u_1)/(u_1) + \mu_{Ai}(u_2)/(u_2) + \mu_{Ai}(u_3)/(u_3) + \ldots\ldots + \mu_{Ai}(u_n)/(u_n)$ here, μ_{Ai} is the membership function of the fuzzy set Ai , such that $\mu_{Ai} : U \rightarrow [0,1]$. If uj is the member of Ai , then μ_{Ai} (uj) is the degree of belonging of uj to Ai.

Definition 2.2. Let's Y(t) (t = …, 0, 1, 2, 3, …), is a subset of R , be the universe of discourse on which fuzzy sets fi(t) , (i = 1, 2, 3, …) are defined and F(t) is the collection of fi, then F(t) is defined as fuzzy time series on Y(t).

E-mail: surjeetk05kgec@gmail.com

Definition 2.3. Let's F(t) is caused only by F(t-1) and is denoted by F(t-1) → F(t); then there is a fuzzy relationship between F(t) and F(t-1) and can be expressed as the fuzzy relational equation:-

F(t) =F(t-1) °R(t, t-1)

here, '°' is Max –Min composition operator. The relation R is called first order model of F(t). Further if Fuzzy relation R(t, t-1) of F(t) is independent of time t, that is to say for different times t1 and t2, R(t1, t1-1) = R(t2, t2-1), then F(t) is called a time invariant fuzzy time series.

Definition 2.4. If F(t) is caused by more fuzzy sets, F(t-n), F(t-n+1) , …, F(t-1), then the fuzzy relationship is represented by A_{i1} , A_{i2}, …., A_{in} → A_j here, F(t-n) = A_{i1}, F(t-n+1) = A_{i2}, … , F(t-1) = A_{in}. This relationship is called nth order fuzzy time series model.

Definition 2.5. Let's F(t) is caused by a F(t-1), F(t-2),…, and F(t-m) (m>0) simultaneously and the relations are time variant. The F(t) is said to be time variant fuzzy time series and the relation can be expressed as the fuzzy relational equation :- F(t) = F(t-1)° Rw (t, t-1) here, w>1 is a time (number of years) parameter by which the forecast F(t) is being affected. Various complicated computational methods are available to for the computations of the Relation Rw (t, t-1).

3. FORECASTING RICE PRODUCTION COMPUTATIONAL PROCEDURES

The implementation of the above algorithm have been carried out by the four models : Chen's arithmetic model (model-1),refined arithmetic model (model-2),Rajaram's modified approach model (model-3) and a combined approach of Rajaram's and Chen's arithmetic model (model-4). Forecasting of the total rice production is based on the 23 years (1981 to 2003) time series production data of India.

Step 1

Define the universe of discourse U as [D min -D1, D max + D2] to accommodate the time series data, where D min and D max are the minimum and maximum historical production respectively. From table 1, we get D min = 3219 and D max = 4554. The variables D1 and D2 are just two positive numbers, properly chosen by the user. If we let D1 =19 and D2 =46, we get U= [3200-4600].

Step 2

Partition the universes of discourse into 7 equal length intervals U1 ,U2 ,……..,U7 such that

U1 = [3200-3400], U2 = [3400-3600], U3 = [3600-3800], U4 = [3800-4000], U5 = [4000-4200],
U6 = [4200-4400], U7 = [4400-4600].

Step 3

Define 7 fuzzy sets A1 , A2 , ..., A7 having some linguistic values on the universe of discourse U. The linguistic values to these fuzzy variables are as follows :

A1 : poor production,
A2 : below average production
A3 : average production
A4 : good production
A5 : very good production
A6 : excellent production
A7 : Outstanding production

The membership grades to these fuzzy sets of linguistic variables are defined as :

A1= 1/ u1 + 0.5/ u 2 + 0/ u 3 + 0/ u4 + 0/ u5 + 0/ u6 + 0/ u7
A2= 0.5/ u1 + 1/ u 2 + 0.5/ u 3 + 0/ u4 + 0/ u5 + 0/ u6 + 0/ u7
A3= 0/ u1 + 0.5/ u 2 + 1/ u 3 + 0.5/ u4 + 0/ u5 + 0/ u6 + 0/ u7
A4= 0/ u1 + 0/ u 2 + 0.5/ u 3 + 1/ u4 + 0.5/ u5 + 0/ u6 + 0/ u7
A5= 0/ u1 + 0/ u 2 + 0/ u 3 + 0.5/ u4 + 1/ u5 + 0.5/ u6 + 0/ u7
A6= 0/ u1 + 0/ u 2 + 0/ u 3 + 0/ u4 + 0.5/ u5 + 1/ u6 + 0.5/ u7
A7= 0/ u1 + 0/ u 2 + 0/ u 3 + 0/ u4 + 0/ u5 + 0.5/ u6 + 1/ u7

Step. 4

In this context, fuzzification is the process of identifying associations between the historical values in the dataset and the fuzzy sets defined in the previous step. Each historical value is fuzzified according to its highest degree of membership. If the highest degree of belongingness of a certain historical time variable, say F(t-1), occurs at fuzzy set Ak , then F (t-1) is fuzzified as Ak . A complete overview of fuzzification of historical total rice production for different fuzzy time series models are shown in the table 2. To exemplify this, let us fuzzify year 1984 in model-1 and model-4. According to table 1, the rice production in 1984 was 3455 (000' tonnes) which lies within the boundaries of interval U2 . Since the highest membership degree of U2 occurs at A2 , the historical time variable F (1984) is fuzzified as A2.

The historical time series data[13] are fuzzified and are placed in table 1.

Table 1: The actual and fuzzified data of rice production

Year	Production Kg ha^{-1}**	Fuzzified production
1981	3552	A2
1982	4177	A5
1983	3372	A1
1984	3455	A2
1985	3702	A3
1986	3670	A3
1987	3865	A4
1988	3592	A2
1989	3222	A1
1990	3750	A3
1991	3851	A4
1992	3231	A1
1993	4170	A5
1994	4554	A7
1995	3872	A4
1996	4439	A7
1997	4266	A6
1998	3219	A1
1999	4305	A6
2000	3928	A4
2001	3978	A4
2002	3870	A4
2003	3727	A3

With these fuzzified value, we can create the fuzzy logical relations of various orders (order1,2, 3, 4, and so on).

**Rice crop production of Pantnagar farm G.B.Pant University of Agriculture and Technology, Pantnagar (INDIA)[13].

4. FOURTH ORDER FUZZY LOGICAL RELATIONSHIPS FOR THE RICE PRODUCTION

A2, A5, A1, A2 $\rightarrow$ A3
A5, A1, A2, A3 $\rightarrow$ A3
A1, A2, A3, A3 $\rightarrow$ A4
A2, A3, A3, A4 $\rightarrow$ A2
A3, A3, A4, A2 $\rightarrow$ A1
A3, A4, A2, A1 $\rightarrow$ A3
A4, A2, A1, A3 $\rightarrow$ A4
A2, A1, A3, A4 $\rightarrow$ A1
A1, A3, A4, A1 $\rightarrow$ A5
A3, A4, A1, A5 $\rightarrow$ A7
A4, A1, A5, A7 $\rightarrow$ A4
A1, A5, A7, A4 $\rightarrow$ A7
A5, A7, A4, A7 $\rightarrow$ A6
A7, A4, A7, A6 $\rightarrow$ A1

A4, A7, A6, A1 $\rightarrow$ A6
A7, A6, A1, A6 $\rightarrow$ A4
A6, A1, A6, A4 $\rightarrow$ A4
A1, A6, A4, A4 $\rightarrow$ A4
A6, A4, A4, A4 $\rightarrow$ A3

5. FUZZY LOGICAL RELATIONSHIP(FLR) WITHOUT REPETITION

A2 $\rightarrow$ A5 A5 $\rightarrow$ A1 A1 $\rightarrow$ A2 A2 $\rightarrow$ A3 A3 $\rightarrow$ A3 A3 $\rightarrow$ A4 A4 $\rightarrow$ A2 A2 $\rightarrow$ A1
A1 $\rightarrow$ A3 A4 $\rightarrow$ A1
A1 $\rightarrow$ A5 A5 $\rightarrow$ A7 A7 $\rightarrow$ A4 A4 $\rightarrow$ A7 A7 $\rightarrow$ A6 A6 $\rightarrow$ A1 A1 $\rightarrow$ A6 A6 $\rightarrow$ A4
A4 $\rightarrow$ A4 A4 $\rightarrow$ A3

6. FUZZY RELATIONSHIP GROUP (FLRG)

A1 $\rightarrow$ A2, A3, A5, A6 A2 $\rightarrow$ A1, A3, A5
A3 $\rightarrow$ A3, A4 A4 $\rightarrow$ A1, A2, A3, A4, A7
A5 $\rightarrow$ A1, A7 A6 $\rightarrow$ A1, A4
A7 $\rightarrow$ A4, A6

Table 2 : The available data are fuzzified based on Gaussian function given below

Actual value	Fuzzy set	A1	A2	A3	A4	A5	A6	A7
3552	A2	0.24	1	0	0	0	0	0
4177	A5	0	0	0	0.115	1	0.885	0
3372	A1	1	0.86	0	0	0	0	0
3455	A2	0.725	1	0.275	0	0	0	0
3702	A3	0	0.49	1	0.51	0	0	0
3670	A3	0	0.65	1	0.35	0	0	0
3865	A4	0	0	0.675	1	0.325	0	0
3592	A2	0.04	1	0.96	0	0	0	0
3222	A1	1	0.89	0	0	0	0	0
3750	A3	0	0.25	1	0.75	0	0	0
3851	A4	0	0	0.745	1	0.255	0	0
3231	A1	1	0.845	0	0	0	0	0
4170	A5	0	0	0	0.15	1	0.85	0
4554	A7	0	0	0	0	0	0.23	1
3872	A4	0	0	0.64	1	0.36	0	0
4439	A7	0	0	0	0	0	0.805	1
4266	A6	0	0	0	0	0.67	1	0.33
3219	A1	1	0.905	0	0	0	0	0
4305	A6	0	0	0	0	0.475	1	0.525
3928	A4	0	0	0.36	1	0.64	0	0
3978	A4	0	0	0.11	1	0.89	0	0
3870	A4	0	0	0.65	1	0.35	0	0
3727	A3	0	0.365	1	0.635	0	0	0

Step. 5: Prediction Value = The Value of fuzzy order(Ai)-A_{ij} where $A_j \rightarrow A_i$.
The table above gives the relational Matrix of A_i and A_j

7. FORECAST VALUE AND ERROR CALCULATION

The suitability of the proposed model in forecasting the rice production has been studied on the basis of mean square error (MSE) and average error of the forecast.

The MSE is defined as

$$\text{Mean Square Error} = \frac{\sum_{I=1}^{n}\left(\text{actual value}_{i\_} - \text{forecast value}_i\right)^2}{n}$$

The forecasted error and average forecasting error are calculated using the formula :-

Forecasting error = |(forecasted value - actual value)| / (actual value) * 100 %
Average forecasting error = (sum of forecasting errors) / (total no of errors).

Table 3 : Actual rice production vs forecasted outputs by fuzzy first, fourth order and proposed models

Year	Actual value	First order	Proposed model on first order	Fourth order	Proposed model on forth order
1981	3552				
1982	4177	4100	4100		
1983	3372	3300	3300		
1984	3455	3500	3499.275		
1985	3702	3700	3699.51	3700	3699.51
1986	3670	3700	3699	3700	3699
1987	3865	3900	3899.325	3900	3899.325
1988	3592	3500	3500	3500	3500
1989	3222	3300	3299.11	3300	3299.11
1990	3750	3700	3700	3700	3700
1991	3851	3900	3899.255	3900	3899.255
1992	3231	3300	3300	3300	3300
1993	4170	4100	4100	4100	4100
1994	4554	4500	4500	4500	4500
1995	3872	3900	3900	3900	3900
1996	4439	4500	4500	4500	4500
1997	4266	4300	4299.67	4300	4299.67
1998	3219	3300	3300	3300	3300
1999	4305	4300	4300	4300	4300
2000	3928	3900	3900	3900	3900
2001	3978	3900	3901	3900	3901
2002	3870	3900	3901	3900	3901
2003	3727	3700	3699.365	3700	3699.365
Average Error		1.3363	1.359	1.2423	1.2983
MSE		3092.59	3071.60	2888.85	2868.52

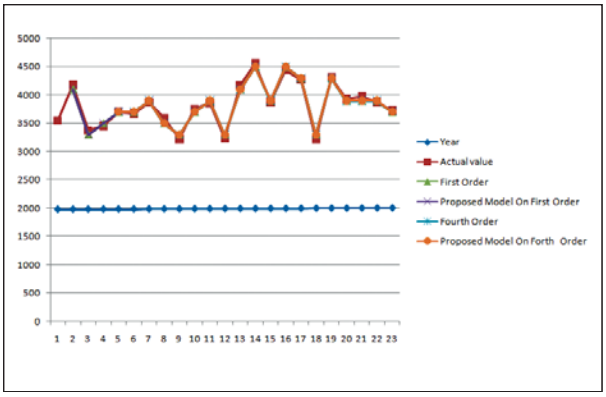

Fig. 1A : Graph of Actual rice production vs forecasted rice production:

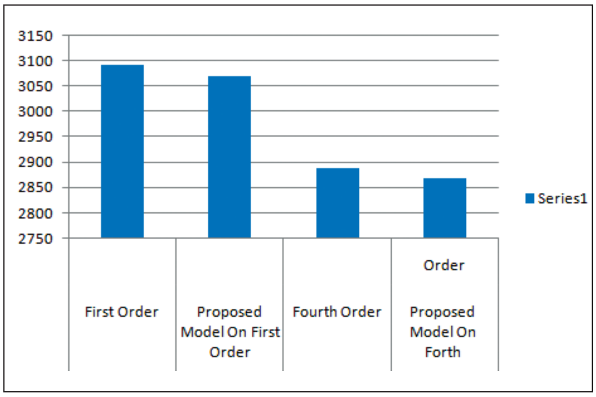

Fig. 1B : Comparison of mean square error of rice production forecasted values

8. CONCLUSION AND FUTURE SCOPE

The new methodology is used to find the forecast value using fuzzy time series technique based on the given historical data set. It is observed that the accuracy and calculated MSE is comparatively better than the traditional time series model. In future, the artificial neural network (ANN) based on the fuzzy input and genetic algorithm can be used to calculate error and based on the minimum error giving method can be selected to forecast rice production more accurately.

REFERENCES

Song,Q. and Chissom, B.S. 1993. Fuzzy time series and its models, *Fuzzy Sets Systems,* **54:** 269-77.

Zadeh, L.A. 1965. Fuzzy Set, *Information Control,* **8** : 338-53.

Zadeh, L.A. 1975. The concept of a linguistic variable and its application to approximate reasoning-Part I, *Information Sciences,* **8**:199–249.

Song,Q. and Chissom, B.S. 1993. Forecasting enrollments with fuzzy time series–Part I,*Fuzzy Sets Systems*, **54:** 1-9.

Song,Q. and Chissom, B.S. 1994 . Forecasting enrollments with fuzzy time series–Part II,*Fuzzy Sets Systems*, **64** : 1-8.

Chen,S.M. 1996. Forecasting enrollments based on fuzzy time series. *Fuzzy Sets Systems*, **81:**311-19.

Tsai, C.C. and Wu, S.J. 1999. A study for second-order modeling of fuzzy time series, *IEEE International Fuzzy Systems Conference Proceedings* Aug 22-25,1999 Seoul, Korea,IEEE:II 719-725.

Chen, S.M. and Hwang, J.R. 2000. Temperature prediction using fuzzy time series, *IEEE Transactions on Systems, man, and Cybernetics-Part B: Cybernetics* **30**:263-75.

Huarng, K. 2001. Heuristic models of fuzzy time series for forecasting, *Fuzzy Sets Systems*, **123** : 369-386.

Sullivan, J. and Woodall W.H. 1994. A comparison of fuzzy forecasting and Markov modeling, *Fuzzy Sets Systems*, **64** :279-93.

Kim, I. and Lee, S.R. 1999. Fuzzy time series prediction method based on consecutivevalues, *IEEE International Fuzzy Systems Conference, August 22-25, Seol Korea,Proceedings* II : 703-707.

Mondal, S. N, Palchowdhury. J, De, D. Bhadra Chowdhuri SR 2008. Roll of membership functions in fuzzy logic for prediction of shoot length of mustard plant based on residual analysis, Journal World Academy of Science, *Engineering and Technology*, 38 : 378-384.

Singh, A 2017. A Computational Method for Rice Production Forecasting Based on High Order Fuzzy Time Series, Banaras Hindu University.