International Refereed Journal of Engineering and Science (IRJES)
ISSN (Online) 2319-183X, (Print) 2319-1821
Volume 5, Issue 8(August 2016), PP.51-56

A Refined Neural Fuzzy Time Series Model for Forecasting

R.K.Srivastava[1], Atib Khan[2],
[1]*Head, Department of Computer Science* [*]*National Rehabilitation University, Lucknow*
[2]*Research Scholar, Shri Venkateshwara University, Amroha(U.P.)*

Abstract : Various methods have been proposed for forecasting based on soft computing techniques. In this paper we have proposed a neural fuzzy hybrid time series model. The developed model has been presented in simple form of computation in fuzzy time series and back propagation algorithm for training of neural network model for the output data set obtained from fuzzy time series. In present study we have used the historical data of production of mustard of Lucknow district (INDIA).
Keywords:- Artificial neural network ,Fuzzy time series, Fuzzy set, Linguistic value

I. INTRODUCTION

Prediction of future values play an important role in our daily life and for the future planning. Agricultural production of crops where neither a trend is appeared nor a pattern in variation of time series are analyzed with forecast data. Song and Chissom[1993]. Successfully presented the definition of fuzzy time series and applied the fuzzy time series model to forecast the enrollments of the university of Alabama. Chen[1996] proposed an alternative simplified method of defuzzification using arithmetical operation. In the present work the proposed model have been implemented based on the historical time series data of production of mustard crop in Lucknow district of U.P.(India). To forecast the production of mustard is very unpredictive, where the data in general contain imprecision and vagueness. The study is aimed to get reliable, improved and versatile method for forecasting using two methods of computation to form the fuzzy relation by using theory of fuzzy time series. It has also advantage of reducing the calculation time and simplifying the calculation process. We have also predicting the production pf mustard crop by using back propagation algorithm of artificial neural network using same data set. We compare the forecasting results of the proposed methods with existing Chen and Ching Hsu[2004] method.

II. FUZZY TIME SERIES MODELS

Let Y(t) (t = …, 0,1,2,..), is a subset of R1, be the universe of discourse on which fuzzy sets fi (t) (i = 1,2,…) are defined and F(t) is the collection of fi (i = 1,2,…). Then F(t) is called fuzzy time series on Y(t) (t = …, 0,1,2,…). Further F(t) can be understood as a linguistic variable and fi(t) (i = 1,2,…) as the possible linguistic values of F(t).

Definition 1: Suppose F(t) is caused by a F(t-1) only or by F(t-1) or (F(t-2)or…or F(t-m) (m>0).This relation can be expressed as the following fuzzy relational equation:

$$F(t) = F(t-1) \circ R(t, t-1) \quad (1)$$

or

$$F(t) = (F(t-1) \cup F(t-2) \cup \ldots \cup F(t-m)) \circ R_0 (t, t-m) \quad (2)$$

The equation is called the first order model of F(t).

Definition 2: Suppose F(t) is caused by a F(t-1), F(t-2),…, and F(t-m) (m>0) simultaneously. This relation can be expressed as the following fuzzy relational equation

$$F(t) = (F(t-1) \times F(t-2) \times \ldots \times F(t-m)) \circ R_a(t, t-m) \quad (3)$$

and is called the m^{th} order model of F(t).

Definition 3: If in (2) or (3) or (4), the fuzzy relation R(t, t-1) or R_a (t, t-m) or R_a (t, t-m) of F(t) is dependent of time t , that is to say for different times t_1 and t_2, $R(t_1, t_1 - 1) = R(t_2, t_2 - 1)$, or $R_a(t_1, t_1 - m) = R_a (t_2, t_2 - m)$ or $R_0(t_1, t_1 - m) = R_0(t_2, t_2 - m)$, then F(t) is called a time invariant fuzzy time series .Otherwise it is called a time variant fuzzy time series,

In the case of time invariant fuzzy time series ,

$R(t, t - 1) = R$,

$R_a (t , t - m) = R_a (m)$,

$R_0(t , t - m) = R_0(m)$

In general at different times t_1 and t_2, $R(t_1, t_1 - 1) \neq R(t_2, t_2 - 1)$, $R_a(t_1, t_1 - m) \neq R_a(t_2, t_2 - m)$ and $R_o(t_1, t_1 - m) \neq R_o(t_2, t_2 - m)$. There are two reasons for this: first, the universes of discourse on which the fuzzy sets are defined may be different at different times: second the value of F(t) at different times may be different.

Depending upon the complexity of the system, fuzzy time series modeling for a forecast process may use type of relations $R(t, t-1)$, $R_a(t, t-m)$, $R_o(t, t-m)$. Several methods Dubois and Parde[1991] , Wu[1986] and Mamdani[1977] are available to determine these relations.

III. PRODUCTION OF MUSTARD FORECASTING : PROPOSED MODEL VERSUS CHEN AND CHING HSU[2004] MODEL :

Fuzzy time series model deals with situation where the data are linguistic values, in contrast to the conventional time series approaches that typically manipulate numerical data. If data are available in crisp form, it is to be fuzzified before the fuzzy time series methodology can be applied. Fuzzification process starts with defining the universe of discourse U, which contains the historical data and upon which the fuzzy sets are defined. The study deals with the production of mustard for the Lucknow district(India) in various years starting from 1999-2000 to 2013-2014 with assumption that it includes some vagueness incurred due to statistical sampling.

The algorithm for application of fuzzy time series in production of mustard crop forecasting comprises of the following steps using proposed method and Chen and Ching Hsu[2004]:

Proposed method :

Step 1: Universe of Discourse U=[4-11]

Step 2: Partition of universe of Discourse U in the seven intervals

u_1=[4-5], u_2=[5-6], u_3=[6-7], u_4= [7-8], u_5= [8-9] , u_6= [9-10], u_7=[10-11]

Step 3: Define seven fuzzy sets $A_1, A_2, \ldots A_7$ as linguistic variables on the universe of discourse U. These fuzzy variables are being defined as

A_1: poor A_2: Average A_3: Good A_4: Very Good

A_5:Better A_6: Excellent A_7: Outstanding

And the membership grades to these fuzzy sets of linguistic values are defined as

$A_1 = \{u_1/1, u_2/0.5, u_3/0, u_4/0, u_5/0, u_6/0, u_7/0\}$

$A_2 = \{u_1/0.5, u_2/1, u_3/0.5, u_4/0, u_5/0, u_6/0, u_7/0\}$

$A_3 = \{u_1/0, u_2/0.5, u_3/1, u_4/0.5, u_5/0, u_6/0, u_7/0\}$

$A_4 = \{u_1/0, u_2/0, u_3/0.5, u_4/1, u_5/0.5, u_6/0, u_7/0\}$

$A_5 = \{u_1/0, u_2/0, u_3/0, u_4/0.5, u_5/1, u_6/0.5, u_7/0\}$

$A_6 = \{u_1/0, u_2/0, u_3/0, u_4/0.5, u_5/0.5, u_6/1, u_7/0.5\}$

$A_7 = \{u_1/0, u_2/0, u_3/0, u_4/0.5, u_5/0, u_6/0.5, u_7/1\}$

Step 4: Fuzzify the production of mustard data to find out the equivalent fuzzy set to each year's production using the step- 3 . The equivalent fuzzy set to each year's production are shown in table-1:

Table-1

Year	Production/Quintal	Fuzzified value
1999-00	5.7	A_2
2000-01	8.02	A_5
2001-02	4.6	A_1
2002-03	7.4	A_4
2003-04	6.54	A_3
2004-05	6.87	A_3
2005-06	9.41	A_6
2006-07	8.14	A_5
2007-08	8.02	A_5
2008-09	9.24	A_6
2009-10	8.37	A_5
2010-11	8.8	A_5
2011-12	5.21	A_2
2012-13	10.43	A_7
2013-14	8.42	A_5

Step 5: The fuzzified historical time series data of production of mustard are obtained and fuzzy logical relations are established in table-2

Table- 2

$A_1 \to A_4$	$A_5 \to A_1, A_5, A_6, A_5, A_2$
$A_2 \to A_5, A_7$	$A_6 \to A_5, A_5$
$A_3 \to A_3, A_6$	$A_7 \to A_5$
$A_4 \to A_3$	

Step 6: Using the proposed method the computations have been carried out and forecasted values obtained are placed in table-3.

Table-3:

Year	Production/Quintal	Forecasted value
1999-00	5.7	8.17
2000-01	8.02	7.5
2001-02	4.6	6
2002-03	7.4	7.0
2003-04	6.54	7.5
2004-05	6.87	7.5
2005-06	9.41	8.8
2006-07	8.14	7.5
2007-08	8.02	7.5
2008-09	9.24	8.8
2009-10	8.37	7.5
2010-11	8.8	7.5
2011-12	5.21	8.17
2012-13	10.43	9.5
2013-14	8.42	7.5

Chen and Ching Hsu Method :

zGet a statistics of the distribution of the historical production of mustard in each interval. Sort the intervals based on the number of historical production of mustard data in each interval from the highest to the lowest. Find the interval having the largest number of historical production of mustard and divide it into the three sub intervals of equal length. If there are no data distributed in an interval, then discard this interval. For example, the distributions of the historical production of mustard data in different intervals are summarized as shown in table.

Step 1: Universe of Discourse U=[4-11]

Step 2: Partition of universe of Discourse U in the seven intervals

u_1=[4-5], u_2=[5-6], u_3=[6-7], u_4= [7-8], $u_{5,1}$= [8.0-8.3] $u_{5,2}$=[8.3-8.6] $u_{5,3}$=[8.6-9.0] , u_6= [9-10], u_7=[10-11]

And the membership grades to these fuzzy sets of linguistic values are defined as

$A_1 = \{u_1/1, u_2/0.5, u_3/0, u_4/0, u_{5,1}/0, u_{5,2}/0, u_{5,3}/0, u_6/0, u_7/0\}$

$A_2 = \{u_1/0.5, u_2/1, u_3/0.5, u_4/0, u_{5,1}/0, u_{5,2}/0, u_{5,3}/0 ,u_6/0, u_7/0\}$

$A_3 = \{u_1/0, u_2/0.5, u_3/1, u_4/0.5, u_{5,1}/0, u_{5,2}/0, u_{5,3}/0, u_6/0, u_7/0\}$

$A_4 = \{u_1/0, u_2/0, u_3/0.5, u_4/1, u_{5,1}/0.5, u_{5,2}/0, u_{5,3}/0, u_6/0, u_7/0\}$

$A_{5,1} = \{u_1/0, u_2/0, u_3/0, u_4/0.5, u_{5,1}/1, u_{5,2}/0.5, u_{5,3}/0, u_6/0, u_7/0\}$

$A_{5,2} = \{u_1/0, u_2/0, u_3/0, u_4/0, u_{5,1}/0.5, u_{5,2}/1, u_{5,3}/0.5, u_6/0, u_7/0\}$

$A_{5,3} = \{u_1/0, u_2/0, u_3/0, u_4/0, u_{5,1}/0, u_{5,2}/0.5, u_{5,3}/1, u_6/0.5, u_7/0\}$

$A_6 = \{u_1/0, u_2/0, u_3/0, u_4/0, u_{5,1}/0, u_{5,2}/0, u_{5,3}/0.5, u_6/1, u_7/0.5\}$

$A_7 = \{u_1/0, u_2/0.5, u_3/1, u_4/0.5, u_{5,1}/0, u_{5,2}/0, u_{5,3}/0, u_6/0.5, u_7/1\}$

Step 3*:* Fuzzify the production of mustard data to find out the equivalent fuzzy set to each year's production using the step- 2 . The equivalent fuzzy set to each year's production are shown in table-4:

Table-4

Year	Production/Quintal	Fuzzified value
1999-00	5.7	A_2
2000-01	8.02	A_{51}
2001-02	4.6	A_1
2002-03	7.4	A_4
2003-04	6.54	A_3
2004-05	6.87	A_3
2005-06	9.41	A_6
2006-07	8.14	A_{53}
2007-08	8.02	A_{51}
2008-09	9.24	A_6
2009-10	8.37	A_{52}
2010-11	8.8	A_{53}
2011-12	5.21	A_2
2012-13	10.43	A_7
2013-14	8.42	A_{52}

Step 4: The fuzzified historical time series data of production of mustard are obtained and fuzzy logical relations are established in table-5

Table-5

$A_1 \rightarrow A_4$	$A_6 \rightarrow A_{53}, A_{52}$
$A_2 \rightarrow A_{51}, A_7$	$A_{51} \rightarrow A_1, A_6$
$A_3 \rightarrow A_3, A_6$	$A_{52} \rightarrow A_2, A_{53}$
$A_4 \rightarrow A_3$	$A_{53} \rightarrow A_2, A_{51}$
$A_7 \rightarrow A_{52}$	

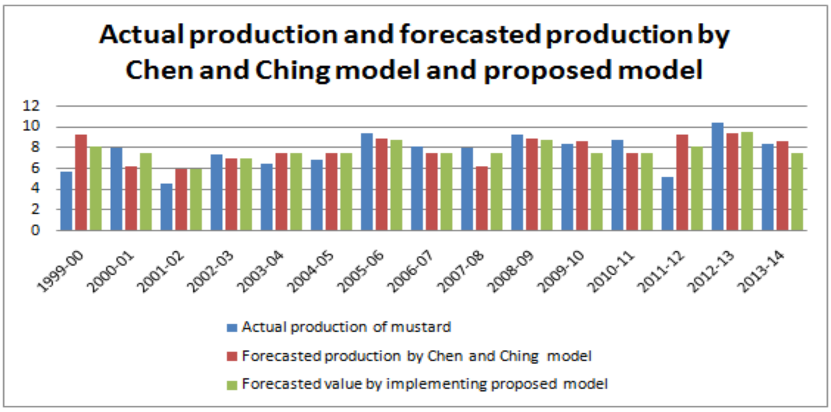
Step 5: Using the proposed method the computations have been carried out and forecasted values obtained are placed in table-6

Table-6

Year	Production/Quintal	Forecasted value
1999-00	5.7	9.32
2000-01	8.02	6.32
2001-02	4.6	6
2002-03	7.4	7.0
2003-04	6.54	7.5
2004-05	6.87	7.5
2005-06	9.41	8.91
2006-07	8.14	7.48
2007-08	8.02	6.32
2008-09	9.24	8.91
2009-10	8.37	8.62
2010-11	8.8	7.48
2011-12	5.21	9.32
2012-13	10.43	9.47
2013-14	8.42	8.62

Table-7: Comparison of forecasted result of proposed method, Chen's method with actual production

Year	Production/Quintal	Forecasted value by Chen and Ching Hsu method	Forecasted value by Proposed Method
1999-00	5.7	9.32	8.17
2000-01	8.02	6.32	7.5
2001-02	4.6	6	6
2002-03	7.4	7.0	7.0
2003-04	6.54	7.5	7.5
2004-05	6.87	7.5	7.5
2005-06	9.41	8.91	8.8
2006-07	8.14	7.48	7.5
2007-08	8.02	6.32	7.5
2008-09	9.24	8.91	8.8
2009-10	8.37	8.62	7.5
2010-11	8.8	7.48	7.5
2011-12	5.21	9.32	8.17
2012-13	10.43	9.47	9.5
2013-14	8.42	8.62	7.5



In time series forecasting, the forecasting accuracy of a model is commonly measured in terms of mean square error (MSE) or in terms of percentage error. Lower the mean square error or average percentage error, better the forecasting method. The mean square error and average forecasted error have been computed and compared with Chen and Ching Hsu method in Table-8.

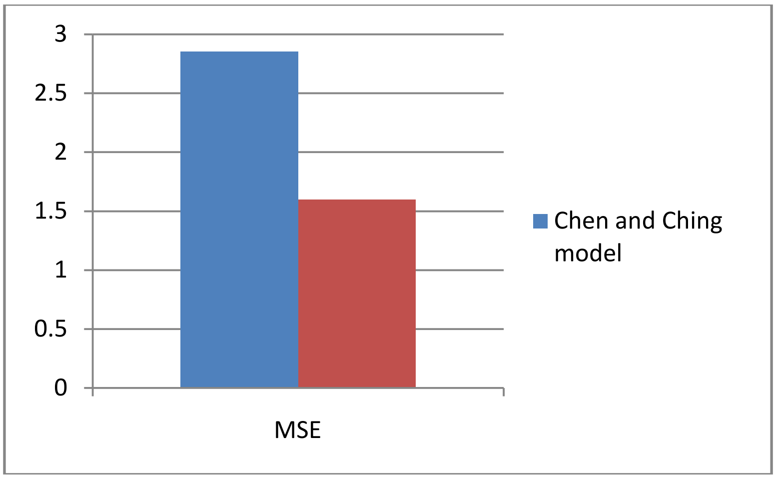
Mean square error = $\underline{\sum_{i=1}^{n}(\text{actual value}_i - \text{forecasted value}_i)^2}$
and forecasting error as

$$\text{Forecasting error (in percent)} = \frac{(\text{forecasted} - \text{actual value})}{\text{Actual value}} \times 100$$

$$\text{Average forecasting error (in percent)} = \frac{\text{sum of forecasting error}}{\text{Numbers of errors}}$$

Table-8

Model	Chen and Ching Model	Proposed Model
MSE	2.851733	1.598353
Average Forecasting Error	7.536333%	4.796095%



IV. ANN BASED METHOD

The production of mustard in different years of data have been used for the training in the development process of the ANN model and production of mustard of a lead year have also been used as desired output for validation of developed model outside the training set. The model have been developed for forecast based on different training sets and validation adopted as production of year 1999-2000 to 2012-2013 as input set and fuzzy output of 2013-2014 as desired output.

The algorithm has been implemented through C programming language, considering two hidden layers and computations have been made by various iterations levels like: 500 & 1000. Out of these, the best suitable forecasted values have been obtained by model with 1000 iterations. The results so obtained has been illustrated in Table-9.

Table-9

Year	Fuzzy Forecasted Production of Mustard	Forecasted Production by implementing ANN model
1999-00	8.17	
2000-01	7.5	
2001-02	6	
2002-03	7.0	
2003-04	7.5	
2004-05	7.5	
2005-06	8.8	
2006-07	7.5	
2007-08	7.5	
2008-09	8.8	
2009-10	7.5	
2010-11	7.5	
2011-12	8.17	
2012-13	9.5	
2013-14	7.5	7.8435

V. RESULT

In this paper, we have presented the proposed fuzzy time series method which has been implemented to have short term forecast for the production of mustard of Lucknow district(India). Prediction of production of mustard is neither governed by any mathematical formula nor any statistical or deterministic process because its production is highly nonlinear so almost unpredictive crop. It is also sensitive crop which is affected by disease and adverse whether we have considered the indirect relation of various parameters for time series data having imprecision. The motivation of the study is that the data collected through various sampling technique involving vagueness. Production forecast has also been obtained through ANN using back propagation algorithm. In the ANN study the target year for the forecast has been considered to be as year 2013-14.

Hence the fuzzy forecasted production of the target year 2013-14 have not been used in training set data of back propagation algorithm. The production forecast of the proposed method has been compared by Chen and Ching Hsu[2004] method and it is clear that the forecasted result of proposed method is quite impressive in terms of MSE and percentage error as shown in Table-8. The comparison of MSE and average percentage forecasting error in Table-8 shows the superiority of the proposed model over Chen and Ching Hsu[2004] model as it provides forecast of higher accuracy. Further to examine the suitability of forecasting result of proposed method the output values of the proposed fuzzy time series method has been used as data set for the training of back propagation algorithm and the result of ANN method for the forecasting of target year 2013-14 is also good, significantly close accordance to the actual production.

REFERENCES

[1]. Chen Shyi-Ming [1996],Forecasting enrollments based on fuzzy time series ,Fuzzy sets and Systems 81,311-319.

[2]. Chou-An Chen,(2007), Neuro –Fuzzy Technology as a Predictor of Parathyroid Harmonse Level in Hemodialysis Patients,Tohku J,Exp.Med., 211, pp. 81-87.

[3]. Duboi, S. D and H Parde [1991], Fuzzy sets in approximate reasoning Part I : interface with possibility distributions, Fuzzy Sets and Systems 40, 143-202.

[4]. Huarng Kunhuang [2001], Heuristic models of fuzzy time series for forecasting, Fuzzy Sets and Systems 123,369-386.

[5]. K.K.Shukla & Raghunath, (1999),An Efficient Global Algorithm for Supervised Training of Neural Networks,International Journal of Computers and EE,PERGAMON Press, 25, pp.195-218.

[6]. Mamdani, E.E.[1977], Application of Fuzzy logic to approximate reasoning using linguistic synthesis, IEEE Trans Comput 26, 1182-1191.

[7]. Shyi-Ming Chen[1996],Forecasting enrollments based on fuzzy time series ,Fuzzy

[8]. Sets and Systems 81,311-319.

[9]. Shyi-Ming Chen[2004],A New Method to Forecasting enrollments using fuzzy time series ,International Journal of Applied Science and Engineering, 2,3,234-244.

[10]. Song Q. and Chissom B.[1993], Fuzzy time series and its models, Fuzzy Sets and

a. Systems ,54,269-277.

[11]. Song Q. and Chissom B.[1993],Forecasting enrollments with fuzzy time series -Part I, Fuzzy sets and Systems 54,1-9.

[12]. Song Q. and Chissom B.[1994],Forecasting enrollments with fuzzy time series -Part II, Fuzzy Sets and Systems 64, 1-8.

[13]. S.R.Singh,(2006), A Robust Method of Forecasting Based on Fuzzy Time Series, Elsevier Applied Mathematics and Computation, 188, pp.472-484.

[14]. Tanu Verma R.K.Srivastava, Artificial Neural Network based Heart Disease Predictive Approach, Volume 4, Issue 3, March 2015, IJAITEM, 2319-4847.

[15]. Wu W [1986], Fuzzy reasoning and fuzzy relational equations, Fuzzy Sets and Systems 20, 67-78.

[16]. Werbos Paul J. [1990], Back propagation through time: what it does and how to do it,Proceeding of the IEEE,78,1550-1559.

[17]. Yang CC[1998],The use of back propagation neural network for the sim. & analysis of time seriesdata ,Transcations of the ASAE,41,1181-1187 .

[18]. Zadeh L. A., [1965], Fuzzy sets, Information & Control ,8,38-353 .