REGULAR PAPER

An effective neural network and fuzzy time series-based hybridized model to handle forecasting problems of two factors

Pritpal Singh · Bhogeswar Borah

Abstract   Fuzzy time series forecasting method has been applied in several domains, such as stock market price, temperature, sales, crop production and academic enrollments. In this paper, we introduce a model to deal with forecasting problems of two factors. The proposed model is designed using fuzzy time series and artificial neural network. In a fuzzy time series forecasting model, the length of intervals in the universe of discourse always affects the results of forecasting. Therefore, an artificial neural network- based technique is employed for determining the intervals of the historical time series data sets by clustering them into different groups. The historical time series data sets are then fuzzified, and the high-order fuzzy logical relationships are established among fuzzified values based on fuzzy time series method. The paper also introduces some rules for interval weighing to defuzzify the fuzzified time series data sets. From experimental results, it is observed that the proposed model exhibits higher accuracy than those of existing two-factors fuzzy time series models.

Keywords   Fuzzy time series · Two factors · Temperature · Fuzzy logical relation · Artificial neural network

1 Introduction

Fuzzy time series forecasting method has been applied in several domains, such as stock market price, temperature, sales, crop production and academic enrollments. Application of fuzzy time series theory in forecasting problems was first introduced by Song and Chissom [28–30]. They presented the fuzzy time series model by means of fuzzy relational equations involving max–min composition operation, and applied the model to forecast the enrollments in University of Alabama. In 1996, Chen [4] used simplified arithmetic operations avoiding

P. Singh (✉) · B. Borah
Department of Computer Science and Engineering, Tezpur University, Tezpur-28, India
e-mail: pritpal@tezu.ernet.in

B. Borah
e-mail: bgb@tezu.ernet.in

the complicated max–min operations, and their method produced better results. Later, many studies provided some improvements to the existing methods in terms of effective lengths of intervals, fuzzy logical relation and defuzzification techniques.

Hwang et al. in [12] used the differences of the available historical data as fuzzy time series instead of direct usage of raw numeric values. Unlike Song–Chissom and Chen approaches, Sah and Degtiarev's proposed model [24] utilizes variations of the available historical data as fuzzy time series. Huarng tried to improve the forecasting accuracy based on the determination of effective length of intervals [10] and heuristic approaches [11]. Lee and Chou [17] forecasted the university enrollments by defining the supports of the fuzzy numbers that represent the linguistic values of the linguistic variables more appropriately.

Cheng et al. in [6] used entropy minimization to create the intervals. They also used trapezoidal membership functions in the fuzzification process. Chang [2] presented cardinality-based fuzzy time series forecasting model which builds weighted fuzzy rules according to calculating the cardinality of fuzzy relations. To obtain less number of intervals, Cheng [7] proposed a model using fuzzy clustering technique to partition the data effectively. Kai et al. [13] applied the K-means clustering algorithm to partition the universe of discourse into different groups. Singh and Borah [26] forecasted the university enrollments with the help of new proposed algorithm by dividing the universe of discourse of the historical time series data into different length of intervals.

Chen and Hwang [5] forecasted the daily average temperature of Taipei based on two-factors fuzzy time series. In this model, first factor is daily temperature, whereas the second factor is daily cloud density. They proposed two algorithms—Algorithm-B and Algorithm-B*. Their experimental results show that the accuracy rate of Algorithm-B* is better than Algorithm-B. Lee et al. [20] proposed a new method to forecast the daily average temperature of Taipei and the Taiwan Futures Exchange (TAIFEX). In this model, high-order fuzzy logical relationship is constructed to increase the forecasting accuracy. Chang and Chen [3] forecasted the daily temperature using fuzzy C-means and fuzzy rules interpolation techniques. In this model, rules are constructed based on fuzzy C-means clustering algorithm. Then, this model performs fuzzy inference based on the multiple fuzzy rules interpolation scheme. Based on two-factors high-order fuzzy time series and automatic clustering techniques, Wang and Chen [32] proposed a new method to predict the daily average temperature and TAIFEX. Lee et al. [18,19] presented a new method for temperature prediction and the TAIFEX forecasting based on two-factors high-order fuzzy logical relationships by hybridizing genetic algorithms with fuzzy time series method.

In this paper, we present a new model to deal with the forecasting problems of two factors. The proposed model is designed using fuzzy time series and artificial neural network (ANN). In this study, high-order fuzzy logical relationships are also employed to design the model. Hence, we have entitled this model as "Two-factors high-order neuro-fuzzy hybridized model." The main purpose of designing such a hybridized model is explained next.

For fuzzification of time series data sets, the determination of length of intervals is very important. In case of most of the above discussed models [4,11,12,28,30], the lengths of the intervals were kept same. No any specific reason is mentioned for using the fixed lengths of intervals. Huarng [10] shows that effective lengths of intervals always affect the results of forecasting. Therefore, for the creation of effective length of intervals of the historical time series data sets, an ANN-based technique is adopted in this model.

Song and Chissom [28] adopted the following method to forecast enrollments of the University of Alabama:

$$Y(t) = Y(t - 1) \circ R \tag{1}$$

where $Y(t-1)$ is the fuzzified enrollment of year $(t-1)$, $Y(t)$ is the forecasted enrollment of year "t" represented by fuzzy set, "$\circ$" is the max–min composition operator and "R" is the union of fuzzy relations. This method takes a lot of time to compute the union of fuzzy relations [5]. Therefore, to improve the efficiency of the proposed model, some rules for intervals weighing are proposed to defuzzify the fuzzified time series data sets. The proposed model exhibits higher accuracy than those of existing models [3,5,18–20,32].

The rest of the paper is organized as follows: In Sect. 2, the basic concepts of fuzzy time series are briefly explained. Section 3 presents the application of ANN for creating intervals of historical time series data sets. In Sect. 4, new forecasting model based on hybridization of ANN with fuzzy time series is proposed. The performance of the model is assessed and presented in Sect. 5. Conclusions and directions for future work are discussed in Sect. 6.

2 Fuzzy sets and fuzzy time series-A brief overview

In 1965, Zadeh [35] introduced the theory of fuzzy sets. According to Zadeh, "A fuzzy set is a class of objects with continuum of grades of membership. Such a set is characterized by a membership function which assigns to each object a grade of membership ranging between zero and one." He also presented fuzzy arithmetic theory and its application [36–38]. Based on fuzzy sets theory, Song and Chissom [28–30] introduced the fuzzy time series concept. Here, we briefly reviewed some concepts of fuzzy time series from [28–30].

Definition 1 (*Fuzzy Set*) A fuzzy set is a class with varying degrees of membership in the set. Let U be the universe of discourse, which is discrete and finite, then fuzzy set A can be defined as follows:

$$A = \{\mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \cdots\} = \Sigma_i \mu_A(x_i)/x_i \qquad (2)$$

where μ_A is the membership function of A, $\mu_A : U \rightarrow [0, 1]$, and $\mu_A(x_i)$ is the degree of membership of the element x_i in the fuzzy set A. Here, the symbol "$+$" indicates the operation of union and the symbol "/" indicates the separator rather than the commonly used summation and division in algebra, respectively.

When U is continuous and infinite, then the fuzzy set A of U can be defined as:

$$A = \left\{ \int \mu_A(x_i)/x_i \right\}, \forall x_i \in U \qquad (3)$$

where the integral sign stands for the union of the fuzzy singletons, $\mu_A(x_i)/x_i$.

Fuzzy time series concept was proposed in [29], and the main difference between the traditional time series and the fuzzy time series is that the values of the former are crisp numerical values while the values of the latter are fuzzy sets. The crisp numerical values can be represented by real numbers, whereas in fuzzy sets, the values of observations are represented by linguistic values. The definitions of fuzzy time series are briefly reviewed as follows:

Definition 2 (*Fuzzy time series*) Let $Y(t)(t = 0, 1, 2, \ldots)$ be a subset of real numbers "R"L and the universe of discourse on which fuzzy sets $\mu_i(t)(i = 1, 2, \ldots)$ are defined, and let $F(t)$ be a collection of $\mu_i(t)(i = 1, 2, \ldots)$. Then, $F(t)$ is called a fuzzy time series on $Y(t)(t = 0, 1, 2, \ldots)$.

From Definition 2, we can see that $F(t)$ is a function of time t and $\mu_i(t)$ are the linguistic values of $F(t)$, where $\mu_i(t)(i = 1, 2, \ldots)$ are represented by fuzzy sets and the values of

$F(t)$ can be different at different times because the universe of discourse can be different at different times. Fuzzy time series can be divided into two categories which are the time-invariant fuzzy time series and the time-variant fuzzy time series.

If $F(t)$ is caused by $F(t - 1)$, that is, $F(t - 1) \rightarrow F(t)$, then this relationship can be represented as follows:

$$F(t) = F(t - 1) \circ R(t, t - 1) \tag{4}$$

where $R(t, t-1)$ is the fuzzy relationship between $F(t)$ and $F(t-1)$. Here, "R" is the union of fuzzy relations and "$\circ$" is max–min composition operator. It is also called the first-order model of $F(t)$.

Definition 3 (*Fuzzy time-variant and time-invariant series*) Let $F(t)$ be a fuzzy time series, and $R(t, t - 1)$ be a first–order model of $F(t)$. If $R(t, t - 1) = R(t - 1, t - 2)$ for any time t, and $F(t)$ only has finite elements, then $F(t)$ is referred as a time-invariant fuzzy time series. Otherwise, it is referred as a time-variant fuzzy time series.

3 ANN and its application for creation of intervals

ANN is a computational model that is inspired by the human brain [1,27]. ANN is composed of large number of interconnected *nodes* or *neurons*, which usually operate in parallel, and are configured in regular architectures. Researchers employ ANN in various forecasting problems (like electric load forecasting [31], short-term precipitation forecasting [16], long-rage summer monsoon rainfall forecasting [25], etc.), due to its capability to extract relationships between the input and output data.

Data clustering is a popular approach for automatically finding classes, concepts, or groups of patterns [9]. Time series data are pervasive across all human endeavors, and their clustering is one of the most fundamental applications of data mining [14,23]. In literature, many data clustering algorithms [8,22,33] have been proposed, but their applications are limited to the extraction of patterns that represent points in multi-dimensional spaces of fixed dimensionality [34]. In our proposed model, a distance-based clustering algorithm, that is, the self-organizing feature maps (SOFM) are employed for determining the intervals of the historical time series data sets by clustering them into different groups. SOFM is developed by Kohonen [15], which is a class of neural networks with neurons arranged in a low-dimensional (often two-dimensional) structure, and trained by an iterative unsupervised or self-organizing procedure [21]. SOFM converts the patterns of arbitrary dimensionality into response of one-dimensional or two-dimensional arrays of neurons, that is, it converts a wide pattern space into a feature space. The neural network performing such a mapping is called feature map. The training process of SOFM consists of the following steps [27]:

Step 1 *Initialize the weights (W_{uv}) and learning rate (α).*
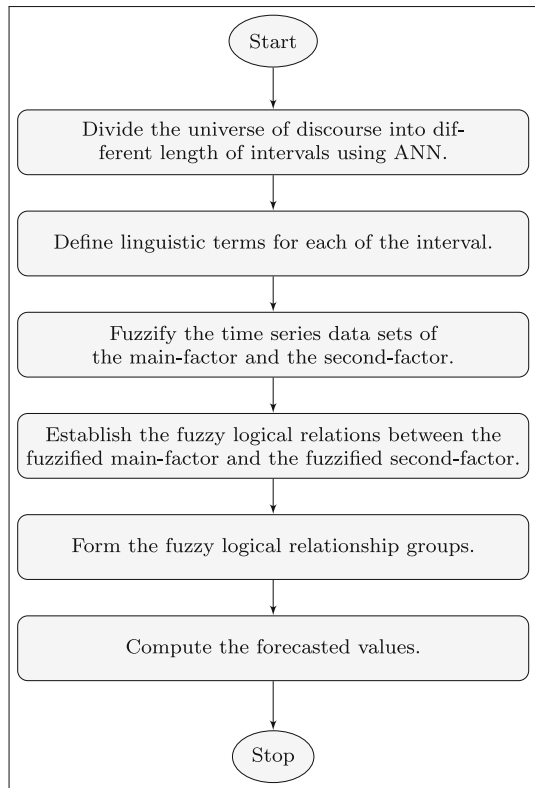
Step 2 *When stopping condition is false, then perform Steps 2–8.*

Step 3 *For each input vector (X), perform Steps 3–5.*

Step 4 *For each $v = 1$ to m, compute the square of the Euclidean distance as:*

$$D(v) = \sum_{u=1}^{n} (X_u - W_{uv})^2 \tag{5}$$

Fig. 1 Two-factors high-order
neuro-fuzzy hybridized model

Step 5 *Obtain winning unit index (J), so that $D(J) = minimum$.*

Step 6 *Calculate weights of winning unit as:*

$$W_{uv}(new) = W_{uv}(old) + \alpha[X_u - W_{uv}(old)] \qquad (6)$$

Step 7 *Reduce the learning rate (α) by using the following formula:*

$$\alpha(t + 1) = 0.5\alpha(t) \qquad (7)$$

Step 8 *Reduce radius of topological neighborhood network.*

Step 9 *Test for stopping condition of the network.*

Based on the above-mentioned algorithm, the historical time series data sets are partitioned into different length of intervals. These intervals are presented in Sect. 4.

4 Proposed ANN and fuzzy time series hybridized model

In this section, we introduce a new forecasting model based on hybridization of ANN with fuzzy time series. The architecture of the proposed model consists of six phases as shown in Fig. 1. For verification of model, the historical data sets of the daily average temperature and the daily cloud density from June 1996 to September 1996 in Taipei, Taiwan [5] are

Day	Month			
	June	July	August	September
1	26.1	29.9	27.1	27.5
2	27.6	28.4	28.9	26.8
3	29.0	29.2	28.9	26.4
4	30.5	29.4	29.3	27.5
5	30.0	29.9	28.8	26.6
6	29.5	29.6	28.7	28.2
7	29.7	30.1	29.0	29.2
8	29.4	29.3	28.2	29.0
9	28.8	28.1	27.0	30.3
10	29.4	28.9	28.3	29.9
11	29.3	28.4	28.9	29.9
12	28.5	29.6	28.1	30.5
13	28.7	27.8	29.9	30.2
14	27.5	29.1	27.6	30.3
15	29.5	27.7	26.8	29.5
16	28.8	28.1	27.6	28.3
17	29.0	28.7	27.9	28.6
18	30.3	29.9	29.0	28.1
19	30.2	30.8	29.2	28.4
20	30.9	31.6	29.8	28.3
21	30.8	31.4	29.6	26.4
22	28.7	31.3	29.3	25.7
23	27.8	31.3	28.0	25.0
24	27.4	31.3	28.3	27.0
25	27.7	28.9	28.6	25.8
26	27.1	28.0	28.7	26.4
27	28.4	28.6	29.0	25.6
28	27.8	28.0	27.7	24.2
29	29.0	29.3	26.2	23.3
30	30.2	27.9	26.0	23.5
31	–	26.9	27.7	–

Table 1 Historical data of the daily average temperature from June 1996 to September 1996 in Taipei (Unit: °C)

used, which are shown in Tables 1 and 2, respectively. In these data sets, the daily average temperature is called the main factor, and the daily average cloud density is called the second factor.

In the following, we apply the proposed model to predict the daily temperature of Taipei from June 1996 to September 1996. To explain the functionality of each phase of the model, the daily average temperature and the daily cloud density data sets from June 1, 1996 to June 30, 1996, are considered as an example. Each phase of the model is explained as follows:

Phase 1 *Divide the universe of discourse into different length of intervals.*

Define the universe of discourse "A" of the main factor and the universe of discourse "B" of the second factor of the historical time series data sets. Let $A = [M_{min}, M_{max}]$, where

Table 2 Historical data of the daily average cloud density from June 1996 to September 1996 in Taipei (Unit: %)

| Day | Month | | | |
---	June	July	August	September
1	36	15	100	29
2	23	31	78	53
3	23	26	68	66
4	10	34	44	50
5	13	24	56	53
6	30	28	89	63
7	45	50	71	36
8	35	34	28	76
9	26	15	70	55
10	21	8	44	31
11	43	36	48	31
12	40	13	76	25
13	30	26	50	14
14	29	44	84	45
15	30	25	69	38
16	46	24	78	24
17	55	26	39	19
18	19	25	20	39
19	15	21	24	14
20	56	35	25	3
21	60	29	19	38
22	96	48	46	70
23	63	53	41	71
24	28	44	34	70
25	14	100	29	40
26	25	100	31	30
27	29	91	41	34
28	55	84	14	59
29	29	38	28	83
30	19	46	33	38
31	–	95	26	–

M_{min} and M_{max} are the minimum and maximum values of the main factor, respectively. Let $B = [N_{min}, N_{max}]$, where N_{min} and N_{max} are the minimum and maximum values of the second factor, respectively.

Based on Tables 1 and 2, we have the universe of discourse of the daily average temperature $A = [26.1, 30.9]$, and the universe of discourse of the cloud density $B = [10, 96]$. By applying the SOFM algorithm, divide the universe of discourse "A" into different lengths of intervals as $a_1, a_2, \ldots,$ and a_n. Similarly, divide the universe of discourse "B" into different lengths of intervals as $b_1, b_2, \ldots,$ and b_n. For each interval, the centroid is calculated by taking the mean of the upper bound and lower bound of the interval. Each interval bears a weight equal to the frequency of the interval. The resulting intervals, centroids and weights for the considered data sets are shown in Tables 3 and 4.

Table 3 Intervals, elements of intervals, centroids and weights for the daily temperature for June 1996

Intervals	Elements of intervals	Centroids	Weights
$a_1 = [26.1,\ 26.1)$	(26.1)	26.1	1
$a_2 = [27.1,\ 27.1)$	(27.1)	27.1	1
$a_3 = [27.4,\ 27.4)$	(27.4)	27.4	1
$a_4 = [27.5,\ 27.6)$	(27.5, 27.6)	27.55	2
$a_5 = [27.7,\ 27.7)$	(27.7)	27.7	1
$a_6 = [27.8,\ 27.8)$	(27.8, 27.8)	27.8	2
$a_7 = [28.4,\ 28.5)$	(28.4, 28.5)	28.45	2
$a_8 = [28.7,\ 28.8)$	(28.7, 28.7, 28.8, 28.8)	28.75	4
$a_9 = [29.0,\ 29.0)$	(29, 29, 29)	29	3
$a_{10} = [29.3,\ 29.4)$	(29.3, 29.4, 29.4)	29.37	3
$a_{11} = [29.5,\ 29.7)$	(29.5, 29.5, 29.7)	29.57	3
$a_{12} = [30.0,\ 30.3)$	(30, 30.2, 30.2, 30.3)	30.18	4
$a_{13} = [30.5,\ 30.9)$	(30.5, 30.8, 30.9)	30.73	3

Table 4 Intervals, elements of intervals, centroids and weights for the daily cloud density for June 1996

Intervals	Elements of intervals	Centroids	Weights
$b_1 = [10,\ 10)$	(10)	10	1
$b_2 = [13,\ 14)$	(13, 14)	13.5	2
$b_3 = [15,\ 15)$	(15)	15	1
$b_4 = [19,\ 19)$	(19, 19)	19	2
$b_5 = [21,\ 21)$	(21)	21	1
$b_6 = [23,\ 23)$	(23, 23)	23	2
$b_7 = [25,\ 26)$	(25, 26)	25.5	2
$b_8 = [28,\ 29)$	(28, 29, 29, 29)	28.75	4
$b_9 = [30,\ 30)$	(30, 30, 30)	30	3
$b_{10} = [35,\ 35)$	(35)	35	1
$b_{11} = [36,\ 40)$	(36, 40)	38	2
$b_{12} = [43,\ 45)$	(44, 45)	44	2
$b_{13} = [46,\ 46)$	(46)	46	1
$b_{14} = [55,\ 56)$	(55, 55, 56, 60)	56.5	4
$b_{15} = [63,\ 63)$	(63)	63	1
$b_{16} = [96,\ 96)$	(96)	96	1

Phase 2 *Define linguistic terms for each of the interval.*

The universe of discourse "A" of the main factor is divided into n intervals (i.e., $a_1, a_2, \ldots,$ and a_n). Assume that there are n linguistic variables (i.e., $U_1, U_2, \ldots, U_n$) represented by fuzzy sets, where $1 \leq i \leq n$, shown as follows:

$$U_1 = 1/a_1 + 0.5/a_2 + 0/a_3 + 0/a_4 + 0/a_5$$
$$+ \cdots + 0/a_{n-2} + 0/a_{n-1} + 0/a_n,$$
$$U_2 = 0.5/a_1 + 1/a_2 + 0.5/a_3 + 0/a_4 + 0/a_5$$
$$+ \cdots + 0/a_{n-2} + 0/a_{n-1} + 0/a_n,$$

$$U_3 = 0/a_1 + 0.5/a_2 + 1/a_3 + 0.5/a_4 + 0/a_5$$
$$+ \cdots + 0/a_{n-2} + 0/a_{n-1} + 0/a_n,$$
$$\vdots$$
$$U_n = 0/a_1 + 0/a_2 + 0/a_3 + 0/a_4 + 0/a_5$$
$$+ \cdots + 0/a_{n-2} + 0.5/a_{n-1} + 1/a_n.$$

Similarly, the universe of discourse "B" of the second factor is divided into m intervals (i.e., $b_1, b_2, \ldots,$ and b_m). Assume that there are m linguistic variables (i.e., $V_1, V_2, \ldots, V_m$) represented by fuzzy sets, where $1 \leq i \leq m$, shown as follows:

$$V_1 = 1/b_1 + 0.5/b_2 + 0/b_3 + 0/b_4 + 0/b_5$$
$$+ \cdots + 0/b_{m-2} + 0/b_{m-1} + 0/b_m,$$
$$V_2 = 0.5/b_1 + 1/b_2 + 0.5/b_3 + 0/b_4 + 0/b_5$$
$$+ \cdots + 0/b_{m-2} + 0/b_{m-1} + 0/b_m,$$
$$V_3 = 0/b_1 + 0.5/b_2 + 1/b_3 + 0.5/b_4 + 0/b_5$$
$$+ \cdots + 0/b_{m-2} + 0/b_{m-1} + 0/b_m,$$
$$\vdots$$
$$V_m = 0/b_1 + 0/b_2 + 0/b_3 + 0/b_4 + 0/b_5$$
$$+ \cdots + 0/b_{m-2} + 0.5/b_{m-1} + 1/b_m.$$

The maximum membership values of both U_i and V_i occur at intervals a_i and b_i, respectively.

Phase 3 *Fuzzify the historical time series data sets of the main factor and the second factor.*

If the time series data of the main factor belong to the interval a_i, where $1 \leq i \leq n$, then fuzzify the time series data of the main factor into fuzzy set U_i. Similarly, if the time series data of the second factor belong to the interval b_i, where $1 \leq i \leq m$, then fuzzify the time series data of the second factor into fuzzy set V_i.

The fuzzified values of the main factor and second factor for June 1996 time series data sets are shown in Table 5. The fourth and fifth columns of Table 5 represent the centroids and weights of the corresponding intervals for the main factor, respectively. In Table 5, only fuzzified values of the second factor are shown (last column), because for forecasting the main factor, only fuzzified values of the second factor are required.

Phase 4 *Establish the fuzzy logical relationships between the fuzzified main factor and the fuzzified second factor.*

We can establish the nth-order fuzzy logical relationship based on the fuzzified main factor and the fuzzified second factor. If there exists a fuzzy logical relationship between U_i and V_i, where U_i and V_i denote the fuzzified main factor and second factor of day "i," respectively, then the two-factors nth-order fuzzy logical relationship can be represented as follows:

$$((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1})) \rightarrow U_i \qquad (8)$$

Here, $(U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1})$ represent fuzzified values of day $n - i, \ldots,$ day $n - 2$, day $n - 1$ and day i, respectively, where $2 \leq i \leq n$. The left-hand side and right-hand side of fuzzy logical relationship (8) are called the *previous state* and the *current state*,

Table 5 Fuzzified daily temperature (with their corresponding centroids and weights) and cloud density for June 1996

Day	Temperature	Fuzzified temperature	Centroid	Weight	Cloud density	Fuzzified cloud density
1	26.1	U_1	26.1	1	36	V_{11}
2	27.6	U_4	27.55	2	23	V_6
3	29.0	U_9	29.0	3	23	V_6
4	30.5	U_{13}	30.73	3	10	V_1
5	30.0	U_{12}	30.18	4	13	V_2
6	29.5	U_{11}	29.57	3	30	V_9
7	29.7	U_{11}	29.57	3	45	V_{12}
8	29.4	U_{10}	29.37	3	35	V_{10}
9	28.8	U_8	28.75	4	26	V_7
10	29.4	U_{10}	29.37	3	21	V_5
11	29.3	U_{10}	29.37	3	43	V_{12}
12	28.5	U_7	28.45	2	40	V_{11}
13	28.7	U_8	28.75	4	30	V_9
14	27.5	U_4	27.55	2	29	V_8
15	29.5	U_{11}	29.57	3	30	V_9
16	28.8	U_8	28.75	4	46	V_{13}
17	29.0	U_9	29.0	3	55	V_{14}
18	30.3	U_{12}	30.18	4	19	V_4
19	30.2	U_{12}	30.18	4	15	V_3
20	30.9	U_{13}	30.73	3	56	V_{14}
21	30.8	U_{13}	30.73	3	60	V_{14}
22	28.7	U_8	28.75	4	96	V_{16}
23	27.8	U_6	27.8	2	63	V_{15}
24	27.4	U_3	27.4	1	28	V_8
25	27.7	U_5	27.7	1	14	V_2
26	27.1	U_2	27.1	1	25	V_7
27	28.4	U_7	28.45	2	29	V_8
28	27.8	U_6	27.8	2	55	V_{14}
29	29.0	U_9	29.0	3	29	V_8
30	30.2	U_{12}	30.18	4	19	V_4

respectively. Here, $U_{ni}, \ldots, U_{n2}$ and U_{n1} represent the fuzzified values of the main factor of days $n - i, \ldots, n - 2$, and $n - 1$, respectively. Similarly, $V_{ni}, \ldots, V_{n2}$, and V_{n1} represent the fuzzified values of the second factor of days $n - i, \ldots, n - 2$, and $n - 1$, respectively.

Based on fuzzy logical relationship (8) and Table 5, the first-order and the second-order fuzzy logical relationships of two factors are formed, which are shown in Tables 6 and 7, respectively. In Tables 6 and 7, the symbol "?" represents an unknown value.

Phase 5 *Form the fuzzy logical relationship groups.*

If the nth-order fuzzy logical relationships have the same previous state, then, the nth-order fuzzy logical relationships can be divided into a nth-order fuzzy logical relationship group. Consider the following nth-order fuzzy logical relationships given as follows:

Table 6 Two-factors first-order fuzzy logical relationships between the fuzzified temperature and cloud density data of June 1996

Two-factors first-order fuzzy logical relationships
$(U_1, V_{11}) \rightarrow U_4$
$(U_4, V_6) \rightarrow U_9$
$(U_9, V_6) \rightarrow U_{13}$
$(U_{13}, V_1) \rightarrow U_{12}$
$(U_{12}, V_2) \rightarrow U_{11}$
$(U_{11}, V_9) \rightarrow U_{11}$
$(U_{11}, V_{12}) \rightarrow U_{10}$
$(U_{10}, V_{10}) \rightarrow U_8$
$(U_8, V_7) \rightarrow U_{10}$
$(U_{10}, V_5) \rightarrow U_{10}$
$(U_{10}, V_{12}) \rightarrow U_7$
$(U_7, V_{11}) \rightarrow U_8$
$(U_8, V_9) \rightarrow U_4$
$(U_4, V_8) \rightarrow U_{11}$
$(U_{11}, V_9) \rightarrow U_8$
$(U_8, V_{13}) \rightarrow U_9$
$(U_9, V_{14}) \rightarrow U_{12}$
$(U_{12}, V_4) \rightarrow U_{12}$
$(U_{12}, V_3) \rightarrow U_{13}$
$(U_{13}, V_{14}) \rightarrow U_{13}$
$(U_{13}, V_{14}) \rightarrow U_8$
$(U_8, V_{16}) \rightarrow U_6$
$(U_6, V_{15}) \rightarrow U_3$
$(U_3, V_8) \rightarrow U_5$
$(U_5, V_2) \rightarrow U_2$
$(U_2, V_7) \rightarrow U_7$
$(U_7, V_8) \rightarrow U_6$
$(U_6, V_{14}) \rightarrow U_9$
$(U_9, V_8) \rightarrow U_{12}$
$(U_{12}, V_4) \rightarrow ?$

$$((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1})) \rightarrow U_k$$
$$((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1})) \rightarrow U_s$$
$$\vdots$$
$$((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1})) \rightarrow U_n$$

Then, the nth-order fuzzy logical relationship group can be formed as follows:

$$((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1})) \rightarrow U_k, U_s, \ldots, U_n \tag{9}$$

The first-order fuzzy logical relationship groups are formed based on Table 6, which are shown in Table 8; and the second-order fuzzy logical relationship groups are formed based on Table 7, which are shown in Table 9. If the same fuzzy logical relationship appears more than once, it is included only once in the group

Table 7 Two-factors second-order fuzzy logical relationships between the fuzzified temperature and cloud density data of June 1996	Two-factors second-order fuzzy logical relationships
	$((U_1, V_{11}), (U_4, V_6)) \to U_9$
	$((U_4, V_6), (U_9, V_6)) \to U_{13}$
	$((U_9, V_6), (U_{13}, V_1)) \to U_{12}$
	$((U_{13}, V_1), (U_{12}, V_2)) \to U_{11}$
	$((U_{12}, V_2), (U_{11}, V_9)) \to U_{11}$
	$((U_{11}, V_9), (U_{11}, V_{12})) \to U_{10}$
	$((U_{11}, V_{12}), (U_{10}, V_{10})) \to U_8$
	$((U_{10}, V_{10}), (U_8, V_7)) \to U_{10}$
	$((U_8, V_7), (U_{10}, V_5)) \to U_{10}$
	$((U_{10}, V_5), (U_{10}, V_{12})) \to U_7$
	$((U_{10}, V_{12}), (U_7, V_{11})) \to U_8$
	$((U_7, V_{11}), (U_8, V_9)) \to U_4$
	$((U_8, V_9), (U_4, V_8)) \to U_{11}$
	$((U_4, V_8), (U_{11}, V_9)) \to U_8$
	$((U_{11}, V_9), (U_8, V_{13})) \to U_9$
	$((U_8, V_{13}), (U_9, V_{14})) \to U_{12}$
	$((U_9, V_{14}), (U_{12}, V_4)) \to U_{12}$
	$((U_{12}, V_4), (U_{12}, V_3)) \to U_{13}$
	$((U_{12}, V_3), (U_{13}, V_{14})) \to U_{13}$
	$((U_{13}, V_{14}), (U_{13}, V_{14})) \to U_8$
	$((U_{13}, V_{14}), (U_8, V_{16})) \to U_6$
	$((U_8, V_{16}), (U_6, V_{15})) \to U_3$
	$((U_6, V_{15}), (U_3, V_8)) \to U_5$
	$((U_3, V_8), (U_5, V_2)) \to U_2$
	$((U_5, V_2), (U_2, V_7)) \to U_7$
	$((U_2, V_7), (U_7, V_8)) \to U_6$
	$((U_7, V_8), (U_6, V_{14})) \to U_9$
	$((U_6, V_{14}), (U_9, V_8)) \to U_{12}$
	$((U_9, V_8), (U_{12}, V_4)) \to ?$

Phase 6 *Compute the forecasted values.*

To compute the forecasted values, the rules for interval weighing are proposed. These rules are presented as follows:

Rule 1. For forecasting day, $D(t)$, the previous state's fuzzified values of the main factor and the second factor are considered from days, $D(t-n), \ldots, D(t-2)$ to $D(t-1)$; where "t" is the current day which we want to forecast and "n" is the order of fuzzy logical relationships. The **Rule 1** is applicable only if there is only one fuzzified value in the current state. The steps under **Rule 1** are given as follows:

Step 1. For forecasting day, $D(t)$, obtain the previous state's fuzzified values of the main factor and the second factor from days $D(t-n)$ to $D(t-1)$ as $(U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2})$ and (U_{n1}, V_{n1}).

Step 2. Find the fuzzy logical relationship group whose previous state is $((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1}))$, and the current state is U_k, that is, the fuzzy logical rela-

Table 8 Two-factors first-order fuzzy logical relationship groups of the fuzzified temperature and cloud density data of June 1996

Two-factors first-order fuzzy logical relationship groups
Group 1: $(U_1, V_{11}) \rightarrow U_4$
Group 2: $(U_4, V_6) \rightarrow U_9$
Group 3: $(U_9, V_6) \rightarrow U_{13}$
Group 4: $(U_{13}, V_1) \rightarrow U_{12}$
Group 5: $(U_{12}, V_2) \rightarrow U_{11}$
Group 6: $(U_{11}, V_9) \rightarrow U_{11}, U_8$
Group 7: $(U_{11}, V_{12}) \rightarrow U_{10}$
Group 8: $(U_{10}, V_{10}) \rightarrow U_8$
Group 9: $(U_8, V_7) \rightarrow U_{10}$
Group 10: $(U_{10}, V_5) \rightarrow U_{10}$
Group 11: $(U_{10}, V_{12}) \rightarrow U_7$
Group 12: $(U_7, V_{11}) \rightarrow U_8$
Group 13: $(U_8, V_9) \rightarrow U_4$
Group 14: $(U_4, V_8) \rightarrow U_{11}$
Group 15: $(U_8, V_{13}) \rightarrow U_9$
Group 16: $(U_9, V_{14}) \rightarrow U_{12}$
Group 17: $(U_{12}, V_4) \rightarrow U_{12}$
Group 18: $(U_{12}, V_3) \rightarrow U_{13}$
Group 19: $(U_{13}, V_{14}) \rightarrow U_{13}, U_8$
Group 20: $(U_8, V_{16}) \rightarrow U_6$
Group 21: $(U_6, V_{15}) \rightarrow U_3$
Group 22: $(U_3, V_8) \rightarrow U_5$
Group 23: $(U_5, V_2) \rightarrow U_2$
Group 24: $(U_2, V_7) \rightarrow U_7$
Group 25: $(U_7, V_8) \rightarrow U_6$
Group 26: $(U_6, V_{14}) \rightarrow U_9$
Group 27: $(U_6, V_{14}) \rightarrow U_9$
Group 28: $(U_9, V_8) \rightarrow U_{12}$
Group 29: $(U_{12}, V_4) \rightarrow ?$

tionship group is in the form of $((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1})) \rightarrow U_k$, then, the forecasted value is calculated based on the following step.

Step 3. Find the interval where the maximum membership value of U_k occurs. Let this interval be a_k. This interval a_k has the corresponding centroid C_k. This centroid C_k is the forecasted value for day, $D(t)$.

Rule 2. This rule is applicable if there are more than one fuzzified values in the current state. The steps under **Rule 2** are given as follows:

Step 1. For forecasting day, $D(t)$, obtain the previous state's fuzzified values of the main factor and the second factor from days $D(t-n)$ to $D(t-1)$ as $(U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2})$ and (U_{n1}, V_{n1}).

Step 2. Find the fuzzy logical relationship group whose previous state is $((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1}))$, and the current state is $U_k, U_s, \ldots, U_n$, that is, the fuzzy logical relationship group is in the form of $((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1})) \rightarrow U_k, U_s, \ldots, U_n$, then, the forecasted value is calculated based on the following step.

Table 9 Two-factors second-order fuzzy logical relationship groups of the fuzzified temperature and cloud density data of June 1996

Two-factors second-order fuzzy logical relationship groups
Group 1: $((U_1, V_{11}), (U_4, V_6)) \rightarrow U_9$
Group 2: $((U_4, V_6), (U_9, V_6)) \rightarrow U_{13}$
Group 3: $((U_9, V_6), (U_{13}, V_1)) \rightarrow U_{12}$
Group 4: $((U_{13}, V_1), (U_{12}, V_2)) \rightarrow U_{11}$
Group 5: $((U_{12}, V_2), (U_{11}, V_9)) \rightarrow U_{11}$
Group 6: $((U_{11}, V_9), (U_{11}, V_{12})) \rightarrow U_{10}$
Group 7: $((U_{11}, V_{12}), (U_{10}, V_{10})) \rightarrow U_8$
Group 8: $((U_{10}, V_{10}), (U_8, V_7)) \rightarrow U_{10}$
Group 9: $((U_8, V_7), (U_{10}, V_5)) \rightarrow U_{10}$
Group 10: $((U_{10}, V_5), (U_{10}, V_{12})) \rightarrow U_7$
Group 11: $((U_{10}, V_{12}), (U_7, V_{11})) \rightarrow U_8$
Group 12: $((U_7, V_{11}), (U_8, V_9)) \rightarrow U_4$
Group 13: $((U_8, V_9), (U_4, V_8)) \rightarrow U_{11}$
Group 14: $((U_4, V_8), (U_{11}, V_9)) \rightarrow U_8$
Group 15: $((U_{11}, V_9), (U_8, V_{13})) \rightarrow U_9$
Group 16: $((U_8, V_{13}), (U_9, V_{14})) \rightarrow U_{12}$
Group 17: $((U_9, V_{14}), (U_{12}, V_4)) \rightarrow U_{12}$
Group 18: $((U_{12}, V_4), (U_{12}, V_3)) \rightarrow U_{13}$
Group 19: $((U_{12}, V_3), (U_{13}, V_{14})) \rightarrow U_{13}$
Group 20: $((U_{13}, V_{14}), (U_{13}, V_{14})) \rightarrow U_8$
Group 21: $((U_{13}, V_{14}), (U_8, V_{16})) \rightarrow U_6$
Group 22: $((U_8, V_{16}), (U_6, V_{15})) \rightarrow U_3$
Group 23: $((U_6, V_{15}), (U_3, V_8)) \rightarrow U_5$
Group 24: $((U_3, V_8), (U_5, V_2)) \rightarrow U_2$
Group 25: $((U_5, V_2), (U_2, V_7)) \rightarrow U_7$
Group 26: $((U_2, V_7), (U_7, V_8)) \rightarrow U_6$
Group 27: $((U_7, V_8), (U_6, V_{14})) \rightarrow U_9$
Group 28: $((U_6, V_{14}), (U_9, V_8)) \rightarrow U_{12}$
Group 29: $((U_9, V_8), (U_{12}, V_4)) \rightarrow ?$

Step 3. Find the intervals where the maximum membership values of $U_k, U_s, \ldots, U_n$ occur, and let these intervals be $a_k, a_s, \ldots, a_n$, respectively. These intervals have the corresponding centroids $C_k, C_s, \ldots, C_n$ and weights $W_k, W_s, \ldots, W_n$, respectively.

Step 4. The forecasted value for day, $D(t)$ is calculated as follows:

$$Forecast\,(t) = \frac{\sum_{i=1}^n C_k W_k + C_s W_s + \cdots + C_n W_n}{\sum_{i=1}^n W_k + W_s + \cdots + W_n} \tag{10}$$

Rule 3. This rule is applicable only if there is an unknown value in the current state. The steps under **Rule 3** are given as follows:

Step 1. For forecasting day, $D(t)$, obtain the previous state's fuzzified values of the main factor and the second factor from days $D(t-n)$ to $D(t-1)$ as $(U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2})$ and (U_{n1}, V_{n1}).

Step 2. Find the fuzzy logical relationship group whose previous state is $((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1}))$, and the current state is "?" (the symbol "?" represents

an unknown value), that is, the fuzzy logical relationship group is in the form of $((U_{ni}, V_{ni}), \ldots, (U_{n2}, V_{n2}), (U_{n1}, V_{n1})) \rightarrow ?$, then, the forecasted value is calculated based on the following step.

Step 3. Find the intervals where the maximum membership values of $U_{ni}, \ldots, U_{n2}, U_{n1}$ occur, and let these intervals be $a_{n-i}, \ldots, a_{n-2}, a_{n-1}$, respectively. These intervals have the corresponding centroids $C_{n-i}, \ldots, C_{n-2}, C_{n-1}$ and weights $W_{n-i}, \ldots, W_{n-2}, W_{n-1}$, respectively.

Step 4. The forecasted value for day, $D(t)$ is calculated as follows:

$$Forecast\,(t) = \frac{\sum_{i=1}^{n} C_{n-i} W_{n-i} + \cdots + C_{n-2} W_{n-2} + C_{n-1} W_{n-1}}{\sum_{i=1}^{n} W_{n-i} + \cdots + W_{n-2} + W_{n-1}} \tag{11}$$

Based on the proposed method, we have presented here two examples to compute forecasted values of daily average temperature as follows:

Ex 1. Based on two-factors first-order fuzzy logical time series, an example is presented here to forecast the temperature on day, $D(t)$. Suppose, we want to forecast the temperature on June 7, 1996, in Taipei. To compute this value, the fuzzified temperature and cloud density values of the previous state are required. For forecasting day, D(June 7), the fuzzified temperature and cloud density values for day, D(June 6) are obtained from Table 5, which are U_{11} and V_9, respectively. Then, obtain the fuzzy logical relationship group whose previous state is (U_{11}, V_9) from Table 8. In this case, the fuzzy logical relationship group is $(U_{11}, V_9) \rightarrow U_{11}, U_8$ (i.e., Group 6). Therefore, **Rule 2** is applicable here, because the current state has two fuzzified values. Now, find the intervals where the maximum membership values of U_{11} and U_8 occur from Table 3, which are a_{11} and a_8, respectively. The corresponding centroid and weight for the interval a_{11} are 29.57 and 3, respectively. The corresponding centroid and weight for the interval a_8 are 28.75 and 4, respectively. Now, based on Eq. 10, the forecasted temperature for day, D(June 7) can be computed as:

$$\frac{(29.57 \times 3 + 28.75 \times 4)}{3 + 4} = 29.10$$

Ex 2. Based on two-factors second-order fuzzy logical time series, an example is presented here to forecast the temperature on day, $D(t)$. Suppose, we want to forecast the temperature on June 4, 1996, in Taipei. To compute this value, the fuzzified temperature and cloud density values of the previous state are required. For forecasting day, D(June 4), the fuzzified temperature and cloud density values for days, D(June 2) and D(June 3) are obtained from Table 5, which are (U_4, V_6) and (U_9, V_6), respectively. Then, obtain the fuzzy logical relationship group whose previous state is $((U_4, V_6), (U_9, V_6))$ from Table 9. In this case, the fuzzy logical relationship group is $((U_4, V_6), (U_9, V_6)) \rightarrow U_{13}$ (i.e., Group 2). Therefore, **Rule 1** is applicable here, because in the current state, only one fuzzified value is available. Now, find the interval where the maximum membership value for fuzzy set U_{13} occurs from Table 3, which is a_{13}. The interval a_{13} has the centroid 30.73, which is the forecasted temperature for day, D(June 4).

The daily average temperature of June 1996 is forecasted based on the two-factors second-order fuzzy logical time series, which is shown in Table 10.

Table 10 Forecasted daily average temperature of June 1996 based on the two-factors second-order fuzzy logical time series (Unit: °C)

Day	Actual temperature	Actual cloud density	Forecasted temperature
1	26.1	36	–
2	27.6	23	–
3	29.0	23	29.00
4	30.5	10	30.73
5	30.0	13	30.18
6	29.5	30	29.57
7	29.7	45	29.10
8	29.4	35	29.37
9	28.8	26	28.75
10	29.4	21	29.37
11	29.3	43	29.37
12	28.5	40	28.45
13	28.7	30	28.75
14	27.5	29	27.55
15	29.5	30	29.57
16	28.8	46	29.10
17	29.0	55	29.00
18	30.3	19	30.18
19	30.2	15	30.18
20	30.9	56	30.73
21	30.8	60	30.73
22	28.7	96	28.75
23	27.8	63	27.80
24	27.4	28	27.40
25	27.7	14	27.70
26	27.1	25	27.10
27	28.4	29	28.45
28	27.8	55	27.80
29	29.0	29	29.00
30	30.2	19	30.18

5 Experimental results

The proposed model computes the forecasted values with the help of hybridization of ANN (SOFM neural network) with the fuzzy time series. For training process, the daily temperature and the daily cloud density data sets from June 1, 1996 to June 30, 1996, are employed. In the testing process, the data sets of the daily temperature and the daily cloud density from July 1996 to September 1996 are used. During the learning process of neural network, different experiments were made to set additional parameters like learning rate, epochs, initial weight, learning radius, etc. to obtain optimal results, and we have chosen the ones that exhibit the best behavior in terms of accuracy. The determined optimal values of all these parameters are listed in Table 11.

Table 11 Additional parameters and their values during the learning process of SOFM neural network

Serial Number	List of additional parameter	Value
1	Learning Rate	0.5
2	Epochs	100
3	Initial Weight	0.3
4	Learning Radius	3
5	Optimum number of intervals for main factor (June)	13
	Optimum number of intervals for second factor (June)	15
6	Optimum number of intervals for main factor (July)	12
	Optimum number of intervals for second factor (July)	15
7	Optimum number of intervals for main factor (August)	16
	Optimum number of intervals for second factor (August)	15
8	Optimum number of intervals for main factor (September)	15
	Optimum number of intervals for second factor (September)	15

The main downside of fuzzy time series forecasting model is that increase in the number of intervals increases accuracy rate of forecasting, and decreases the fuzziness of time series data sets. Therefore, in this study, the parameter called "optimum number of intervals" for the main-factor and second-factor time series data sets are decided using a heuristic approach. We have tried different values for this parameter, and calculate the average forecasting error rate (AFER) for different orders for the months – June, July, August and September. The equation for AFER is presented next.

$$AFER = \frac{|F_i - A_i|/A_i}{N} \times 100\%$$ (12)

Here, A_i and F_i denote the actual and forecasted temperature for day i, and N denotes the total number of days to be forecasted.

All these experimental results are plotted in graphs for different orders and intervals as shown in Fig. 2, and we have chosen the "optimum number of intervals" (shown in Table 11) for the main factor and second factor that exhibit the best behavior in terms of AFER. The experimental results of our proposed model are presented in Table 12 in terms of AFER. The forecasting results of the proposed model are then compared with existing models proposed by Chen and Hwang [5], Lee et al. [20], Lee et al. [18], Lee et al. [19], Chang and Chen [3], and Wang and Chen [32]. The comparative analyses in Tables 12, 13, 14, 15, 16, 17 and 18 signify that our proposed model exhibits higher accuracy than those of considered competing models [5,3,18–20,32].

6 Conclusions and directions for future work

In this paper, a new model is proposed for handling two-factors forecasting problems based on the hybridization of ANN with fuzzy time series. For generation of intervals of time series data sets, SOFM neural network is used. Then, some proposed rules of interval weighing are used to compute the forecasted values. From empirical analyses of experimental results, it is evident that our model is superior compared to the considered competing models in terms of accuracy.

Fig. 2 AFER *curves* for June, July, August and September (*top to bottom*) with different orders and intervals
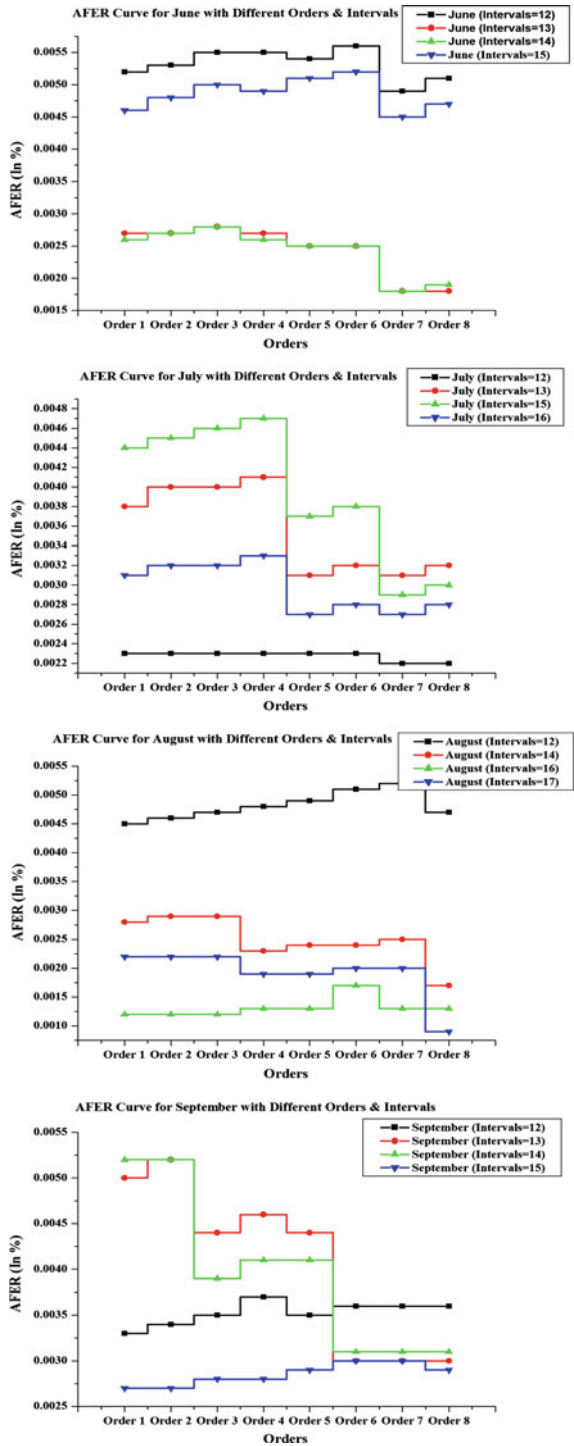
Table 12 The average forecasting error rates to forecast the temperature from June 1996 to September 1996 in Taipei for different orders based on proposed model

Month	Order							
	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)	Sixth (%)	Seventh (%)	Eighth (%)
June	0.27	0.27	0.28	0.27	0.25	0.25	0.18	0.18
July	0.23	0.23	0.23	0.23	0.23	0.23	0.22	0.22
August	0.12	0.12	0.12	0.13	0.13	0.17	0.13	0.13
September	0.27	0.27	0.28	0.28	0.29	0.30	0.30	0.29

Table 13 The average forecasting error rates to forecast the temperature from June 1996 to September 1996 in Taipei with different window bases based on Chen and Hwang [5] model

Month	Window bases						
	w = 2 (%)	w = 3 (%)	w = 4 (%)	w = 5 (%)	w = 6 (%)	w = 7 (%)	w = 8 (%)
June	2.88	3.16	3.24	3.33	3.39	3.53	3.67
July	3.04	3.76	4.08	4.17	4.35	4.38	4.56
August	2.75	2.77	3.30	3.40	3.18	3.15	3.19
September	3.29	3.10	3.19	3.22	3.39	3.38	3.29

Table 14 The average forecasting error rates to forecast the temperature from June 1996 to September 1996 in Taipei for different orders based on Lee et al. [20] model

Month	Order							
	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)	Sixth (%)	Seventh (%)	Eighth (%)
June	1.44	0.80	0.76	0.79	0.76	0.79	0.79	0.81
July	1.59	0.96	0.96	0.98	0.97	1.00	0.98	0.99
August	1.26	1.07	1.06	1.08	1.08	1.09	1.07	1.07
September	1.89	1.01	0.9	0.94	0.96	0.95	0.95	0.92

Table 15 The average forecasting error rates to forecast the temperature from June 1996 to September 1996 in Taipei for different orders based on Lee et al. [18] model

Month	Order							
	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)	Sixth (%)	Seventh (%)	Eighth (%)
June	1.24	0.74	0.64	0.72	0.65	0.66	0.64	0.65
July	1.23	0.78	0.73	0.83	0.70	0.71	0.68	0.69
August	1.09	0.92	0.88	1.07	0.75	0.76	0.75	0.73
September	1.28	0.91	0.86	1.03	0.87	0.97	0.84	0.82

Table 16 The average forecasting error rates to forecast the temperature from June 1996 to September 1996 in Taipei for different orders based on Lee et al. [19] model

Annealing constant α	Month	Order							
		First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)	Sixth (%)	Seventh (%)	Eighth (%)
0.25	June	0.79	0.44	0.42	0.42	0.42	0.44	0.40	0.40
	July	0.66	0.45	0.42	0.41	0.41	0.40	0.41	0.40
	August	0.64	0.43	0.47	0.40	0.41	0.38	0.40	0.45
	September	0.69	0.58	0.59	0.57	0.56	0.57	0.58	0.47
0.5	June	0.84	0.50	0.45	0.42	0.38	0.43	0.39	0.46
	July	0.66	0.50	0.47	0.44	0.40	0.38	0.44	0.42
	August	0.69	0.40	0.38	0.37	0.37	0.39	0.42	0.45
	September	0.66	0.62	0.59	0.59	0.56	0.54	0.56	0.53
0.9	June	0.79	0.46	0.42	0.44	0.42	0.41	0.46	0.39
	July	0.62	0.46	0.45	0.44	0.44	0.41	0.40	0.40
	August	0.66	0.40	0.40	0.40	0.36	0.41	0.39	0.44
	September	0.62	0.59	0.61	0.57	0.54	0.59	0.57	0.50

Table 17 The average forecasting error rates to forecast the temperature from June 1996 to September 1996 in Taipei with different window bases (fifteen generated fuzzy rules) based on Chang and Chen [3] model

Month	Window bases						
	w = 2 (%)	w = 3 (%)	w = 4 (%)	w = 5 (%)	w = 6 (%)	w = 7 (%)	w = 8 (%)
June	1.70	1.50	1.38	1.37	1.28	1.13	0.97
July	1.62	1.77	1.74	1.68	1.77	1.72	3.04
August	1.60	1.48	1.24	1.30	1.28	2.41	2.97
September	1.44	1.51	1.35	1.20	2.02	2.49	2.42

Table 18 The average forecasting error rates to forecast the temperature from June 1996 to September 1996 in Taipei for different orders based on Wang and Chen [32] model

Month	Order							
	First (%)	Second (%)	Third (%)	Fourth (%)	Fifth (%)	Sixth (%)	Seventh (%)	Eighth (%)
June	0.53	0.28	0.29	0.30	0.29	0.29	0.28	0.29
July	0.71	0.34	0.35	0.34	0.34	0.35	0.33	0.32
August	0.32	0.23	0.22	0.22	0.22	0.23	0.23	0.22
September	0.74	0.51	0.49	0.51	0.51	0.53	0.5	0.51

Still, there are scopes to apply the model in some other domains in a flexible way as follows:

1. To check the accuracy and performance of the model by forecasting the temperature for different regions,

2. To test the performance of the model for different types of financial, stocks and marketing data sets, and
3. To enhance this model so that it can deal with multi-dimensional time series data set.

References

1. Bose NK, Liang P (1998) Neural network fundamentals with graphs, algorithms, and applications. Tata McGraw-Hill, New Delhi
2. Chang J, Lee Y, Liao S, Cheng C (2007) Cardinality-Based Fuzzy Time Series for Forecasting Enrollments. In: new trends in applied artificial intelligence, vol 4570. Japan. pp 735–744
3. Chang YC, Chen SM (2009) Temperature prediction based on fuzzy clustering and fuzzy rules interpolation techniques. In: Proceedings of the 2009 IEEE international conference on systems, man, and cybernetics. San Antonio, TX, USA, pp 3444–3449
4. Chen SM (1996) Forecasting enrollments based on fuzzy time series. Fuzzy Sets Syst 81:311–319
5. Chen SM, Hwang JR (2000) Temperature prediction using fuzzy time series. IEEE Trans Syst Man Cybern 30:263–275
6. Cheng C, Chang J, Yeh C (2006) Entropy-based and trapezoid fuzzification-based fuzzy time series approaches for forecasting IT project cost. Technol Forecast Soc Chang 73:524–542
7. Cheng CH, Cheng GW, Wang JW (2008) Multi-attribute fuzzy time series method based on fuzzy clustering. Expert Syst Appl 34:1235–1242
8. Estivill-Castro V (2002) Why so many clustering algorithms: a position paper. ACM SIGKDD Explor Newsl 4(1):65–75
9. Gondek D, Hofmann T (2007) Non-redundant data clustering. Knowl Inf Syst 12:1–24
10. Huarng K (2001) Effective lengths of intervals to improve forecasting in fuzzy time series. Fuzzy Sets Syst 123:387–394
11. Huarng K (2001) Heuristic models of fuzzy time series for forecasting. Fuzzy Sets Syst 123:369–386
12. Hwang JR, Chen SM, Lee CH (1998) Handling forecasting problems using fuzzy time series. Fuzzy Sets Syst 100:217–228
13. Kai C, Ping FF, Gang CW (2010) A novel forecasting model of fuzzy time series based on k-means clustering. In: 2010 second international workshop on education technology and computer science, China. pp 223–225
14. Keogh E, Lin J (2005) Clustering of time-series subsequences is meaningless: implications for previous and future research. Knowl Inf Syst 8(2):154–177
15. Kohonen T (1990) The self organizing maps. In: Proceedings of IEEE vol 78. pp 1464–1480
16. Kuligowski RJ, Barros AP (1998) Experiments in short-term precipitation forecasting using artificial neural networks. Mon Weather Rev 126:470–482
17. Lee HS, Chou MT (2004) Fuzzy forecasting based on fuzzy time series. Int J Comput Math 81(7):781–789
18. Lee LW, Wang LH, Chen SM (2007) Temperature prediction and TAIFEX forecasting based on fuzzy logical relationships and genetic algorithms. Expert Syst Appl 33(3):539–550
19. Lee LW, Wang LH, Chen SM (2008) Temperature prediction and TAIFEX forecasting based on high-order fuzzy logical relationships and genetic simulated annealing techniques. Expert Syst Appl 34(1):328–336
20. Lee LW, Wang LH, Chen SM, Leu YH (2006) Handling forecasting problems based on two-factors high-order fuzzy time series. IEEE Trans Fuzzy Syst 14:468–477
21. Liao TW (2005) Clustering of time series data-a survey. Pattern Recognit 38(11):1857–1874
22. Ordonez C (2003) Clustering binary data streams with K-means. Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery. ACM Press, New York, USA. pp 12–19
23. Rakthanmanon T, Keogh E, Lonardi S, Evans S (2012) MDL-based time series clustering. Knowl Inf Syst. pp 1–29: doi:10.1007/s10115-012-0508-7
24. Sah M, Degtiarev K (2005) Forecasting enrollment model based on first-order fuzzy time series. In: Proceedings of world academy of sciences, engineering and technology vol 1, pp 132–135
25. Sahai AK, Soman MK, Satyan V (2000) All India summer monsoon rainfall prediction using an artificial neural network. Clim Dyn 16:291–302

26. Singh P, Borah B (2011) An efficient method for forecasting using fuzzy time series. In: Sharma U, Nath B, Bhattacharya DK (eds) Machine intelligence. Tezpur University, Assam, pp 67–75
27. Sivanandam SN, Deepa SN (2007) Principles of soft computing. Wiley India (P) Ltd, New Delhi
28. Song Q, Chissom BS (1993) Forecasting enrollments with fuzzy time series—part I. Fuzzy Sets Syst 54(1):1–9
29. Song Q, Chissom BS (1993) Fuzzy time series and its models. Fuzzy Sets Syst 54(1):1–9
30. Song Q, Chissom BS (1994) Forecasting enrollments with fuzzy time series—part II. Fuzzy Sets Syst 62(1):1–8
31. Taylor JW, Buizza R (2002) Neural network load forecasting with weather ensemble predictions. IEEE Trans Power Syst 17:626–632
32. Wang NY, Chen SM (2009) Temperature prediction and TAIFEX forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series. Expert Syst Appl 36:2143–2154
33. Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q, Motoda H, McLachlan G, Ng A, Liu B, Yu P, Zhou ZH, Steinbach M, Hand D, Steinberg D (2008) Top 10 algorithms in data mining. Knowl Inf Syst 14:1–37
34. Xiong Y, Yeung DY (2002) Mixtures of ARMA models for model-based time series clustering. In: IEEE international conference on data mining. Los Alamitos, USA, pp 717–720
35. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353
36. Zadeh LA (1971) Similarity relations and fuzzy orderings. Inf Sci 3:177–200
37. Zadeh LA (1973) Outline of a new approach to the analysis of complex system and decision process. IEEE Tran Syst Man Cybern 3:28–44
38. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning. Inform Sci 8:199–249

Author Biographies

Pritpal Singh received a his Master's degree in Computer Applications from Dibrugarh University, Assam (India), in 2008. He is currently a Ph.D. student at the Department of Computer Science and Engineering, Tezpur University, Tezpur (India). His research interests include soft computing, data mining and time series analysis.

Bhogeswar Borah received his MS degree from BITS, Pilani, and Ph.D. degree from Tezpur University (India). He is currently an Associate Professor at the Department of Computer Science and Engineering, Tezpur University, Tezpur (India). His research interests include soft computing and data mining.