

Chapter 2

Fuzzy Time Series Modeling Approaches: A Review

Although this may seem a paradox, all exact science is dominated by the concept of approximation.

By Bertrand Shaw (1872–1970)

Abstract Recently, there seems to be increased interest in time series forecasting using soft computing (SC) techniques, such as fuzzy sets, artificial neural networks (ANNs), rough set (RS) and evolutionary computing (EC). Among them, fuzzy set is widely used technique in this domain, which is referred to as “Fuzzy Time Series (FTS)”. In this chapter, extensive information and knowledge are provided for the FTS concepts and their applications in time series forecasting. This chapter reviews and summarizes previous research works in the FTS modeling approach from the period 1993–2013 (June). Here, we also provide a brief introduction to SC techniques, because in many cases problems can be solved most effectively by integrating these techniques into different phases of the FTS modeling approach. Hence, several techniques that are hybridized with the FTS modeling approach are discussed briefly. We also identified various domains specific problems and research trends, and try to categorize them. The chapter ends with the implication for future works. This review may serve as a stepping stone for the amateurs and advanced researchers in this domain.

Keywords Fuzzy time series (FTS) • Artificial neural networks (ANNs) • Rough set (RS) • Evolutionary computing (EC)

2.1 Soft Computing: An Introduction

The term “soft computing” is a multidisciplinary field which pervades from a mathematical science to computer science, information technology, engineering applications, etc. The conventional computing or hard-computing generally deals with precision, certainty and rigor (Zadeh 1994). However, the main desiderata of SC is

to tolerate with imprecision, uncertainty, partial truth, and approximation (Yardimci 2009). SC is influenced by many researchers. Among them, Zadeh's contribution is invaluable. Zadeh published his most influential work in SC in 1965 (Zadeh 1965). Later, he contributed in this area by publishing numerous research articles on the analysis of complex systems and decision processes (Zadeh 1973), approximate reasoning (Zadeh 1975a,b), knowledge representation (Zadeh 1989), design and deployment of intelligent systems (Zadeh 1997), etc. According to Jang et al. (1997), "SC is not a single methodology. Rather, it is a partnership in which each of the partners contributes a distinct methodology for addressing problems in its domain." Therefore, SC has been evolving as an amalgamated field of different methodologies such as *fuzzy sets*, *ANN*, *EC* and *probabilistic computing* (Dote and Ovaska 2001; Herrera-Viedma et al. 2014; Kacprzyk 2010; Szmidt et al. 2014). Later, *RS*, *chaos computing* and *immune network theory* have been included into SC (Castro and Timmis 2003; Mitra et al. 2002). The main objective of hybridizing these methodologies is to design an intelligent machine and find solution to nonlinear problems which can not be modeled mathematically (Zadeh 2002).

2.1.1 Time Series Events and Uncertainty

A time series represents a collection of values of certain events or tasks which are obtained with respect to time. Advance prediction of some significant time series events such as temperature, rainfall, stock price, population growth, economic growth, etc., are major scientific issues in the domain of forecasting. Imprecise knowledge or information cannot be overlooked in this domain. Because of the nature of the time series data, which is highly non-stationary and uncertain, the decision-making process becomes very tedious. For example, sudden rise and fall of daily temperature, sudden increase and decrease of daily stock index price, sudden increase and decrease of rainfall amount indicate that these events are very uncertain. The characteristics of all these events cannot be described accurately; therefore, it is referred to as "imprecise knowledge" or "incomplete knowledge". Due to these problems, mathematical or statistical models can not deal with this imprecise knowledge, thereby diluting the accuracy very significantly.

Future prediction of time series events has attracted people from the beginning of times. However, forecasting these events with 100 % accuracy may not be possible, their forecasting accuracy and the speed of forecasting process can be improved. To resolve this problem, Song and Chissom (1993a) developed a model in 1993 based on uncertainty and imprecise knowledge contained in time series data. They initially used the fuzzy sets concept to represent or manage all these uncertainties, and referred this concept as "Fuzzy Time Series (FTS)".

Forecasting the short term time series events are frequently attempted by the researchers, and its accuracy is better than long term predictions. From 1994 onwards, researchers have developed numerous models based on the FTS concept to deal with the forecasting problems of short term as well as long term events. This study focuses

on the application and use of fuzzy sets concept in forecasting of such events. The basic knowledge of ANNs, RS and EC are provided complimentary with the sound background of fuzzy sets, because in many cases a problem can be solved most effectively by hybridizing these techniques together rather independently. Hence, one of the objectives of this chapter is also to introduce the SC methodologies (such as ANNs, RS and EC) that are employed by the FTS modeling approach to represent and manage the imprecise knowledge in time series forecasting.

2.2 Definitions

In this section, we provide various definitions for the terminologies used throughout this book.

Definition 2.2.1 (*Time Series*) (Brockwell and Davis 2008). A time series is a set of observations x_t , each one being recorded at specific time t . When observations are made at fixed time intervals, then it is called a “discrete-time series”. If observations are recorded continuously over some time interval, then is called a “continuous-time series”.

The main objective of time series forecasting is reckoning the future values of the series. In literature, several time series forecasting models are available (Chatfield 2000). Forecasting model finds optimal forecasts based on the type of data and condition of the model. Suppose we have an observed time series x_1, x_2, \dots, x_N and wish to forecast future values such as x_{N+h} . Forecasting model can make the lead time forecasts (denoted as h), or make forecast h steps ahead of time N .

This survey is concerned with the study of SC techniques and its application in FTS modeling approach. Therefore, in the next, we will discuss initially the fuzzy sets concept and its application in time series forecasting.

In 1965, Zadeh (1965) introduced fuzzy sets theory involving continuous set membership for processing data in presence of uncertainty. He also presented fuzzy arithmetic theory and its application in these articles.¹

Definition 2.2.2 (*Universe of discourse*) (Song and Chissom 1993a). Let L_{bd} and U_{bd} be the lower bound and upper bound of the time series data, respectively. Based on L_{bd} and U_{bd} , we can define the universe of discourse U as:

$$U = [L_{bd}, U_{bd}] \quad (2.2.1)$$

Definition 2.2.3 (*Fuzzy Set*) (Zadeh 1965). 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:

¹References are: (Zadeh 1971, 1973, 1975a).

$$A = \{\mu_{A(x_1)}/x_1 + \mu_{A(x_2)}/x_2 + \dots\} = \sum_i \mu_A(x_i)/x_i \quad (2.2.2)$$

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:

$$A = \left\{ \int \mu_{A(x_i)}/x_i \right\}, \forall x_i \in U \quad (2.2.3)$$

where the integral sign stands for the union of the fuzzy singletons, $\mu_{A(x_i)}/x_i$.

Definition 2.2.4 (FTS).² Let $Y(t)$ ($t = 0, 1, 2, \dots$) be a subset of Z and the universe of discourse on which fuzzy sets $\mu_i(t)$ ($i = 1, 2, \dots$) are defined and let $F(t)$ be a collection of $\mu_i(t)$ ($i = 1, 2, \dots$). Then, $F(t)$ is called a FTS on $Y(t)$ ($t = 0, 1, 2, \dots$).

With the help of the following two examples, the notions of FTS can be explained:

[Example 1] The common observations of daily weather condition for certain region can be described using the daily common words “hot”, “very hot”, “cold”, “very cold”, “good”, “very good”, etc. All these words can be represented by fuzzy sets.

[Example 2] The common observations of the performance of a student during the final year of degree examination can be represented using the fuzzy sets “good”, “very good”, “poor”, “bad”, “very bad”, etc.

These two examples represent the processes, and conventional time series models are not applicable to describe these processes (Song and Chissom 1993b). Therefore, Song and Chissom (Song and Chissom 1993b) first time used the fuzzy sets notion in time series forecasting. Later, their proposed method has gained in popularity in scientific community as a “FTS forecasting model”.

Definition 2.2.5 (Fuzzification) (Zadeh 1975a). The operation of *fuzzification* transforms a nonfuzzy set (crisp set) into a fuzzy set or increasing the fuzziness of a fuzzy set. Thus, a *fuzzifier* R is applied to a fuzzy subset i of the universe of discourse U yields a fuzzy subset $R(i; T)$, which can be expressed as:

$$R(i; T) = \int_U \mu_i(u)T(u), \quad (2.2.4)$$

where the fuzzy set $T(u)$ is the kernel of R , i.e., the result of applying R to a singleton $1/u$:

$$T(u) = R(1/u; T) \quad (2.2.5)$$

where $\mu_i(u)T(u)$ represents the product of a scalar $\mu_i(u)$ and the fuzzy set $T(u)$; and \int is the union of the family of fuzzy sets $\mu_i(u)T(u)$, $u \in U$.

²References are: (Song and Chissom 1993a,b, 1994).

Definition 2.2.6 (*Defuzzification*) (Ross 2007). Defuzzification of a fuzzy set is the process of “rounding it off” from its location in the unit hypercube to the nearest vertex, i.e., it is the process of converting a fuzzy set into a crisp set.

Definition 2.2.7 (*FLR*).³ Assume that $F(t-1) = A_i$ and $F(t) = A_j$. The relationship between $F(t)$ and $F(t-1)$ is referred to as a FLR, which can be represented as:

$$A_i \rightarrow A_j, \quad (2.2.6)$$

where A_i and A_j refer to the left-hand side and right-hand side of the FLR, respectively.

Definition 2.2.8 (*FLRG*).⁴ Assume the following FLRs as follows:

$$\begin{aligned} A_i &\rightarrow A_{k1}, \\ A_i &\rightarrow A_{k2}, \\ A_i &\rightarrow A_{k3}, \\ A_i &\rightarrow A_{k4}, \\ &\dots \\ A_i &\rightarrow A_{km} \end{aligned} \quad (2.2.7)$$

Chen (1996) suggested that FLRs having the same fuzzy sets on the left-hand side can be grouped into a FLRG. So, based on Chen’s model (Chen 1996), these FLRs can be grouped into the FLRG as:

$$A_i \rightarrow A_{k1}, A_{k2}, A_{k3}, A_{k4} \dots, A_{km}. \quad (2.2.8)$$

Definition 2.2.9 (*High-order FLR*) (Chen and Chen 2011a). Assume that $F(t)$ is caused by $F(t-1)$, $F(t-2)$, \dots , and $F(t-n)$ ($n > 0$), then high-order FLR can be expressed as:

$$F(t-n), \dots, F(t-2), F(t-1) \rightarrow F(t) \quad (2.2.9)$$

Definition 2.2.10 (*M-factors FTS*). Let FTS $A(t)$, $B(t)$, $C(t)$, \dots , $M(t)$ be the factors/observations of the forecasting problems. If we only use $A(t)$ to solve the forecasting problems, then it is called a one-factor FTS. If we use remaining secondary-factors/secondary-observations $B(t)$, $C(t)$, \dots , $M(t)$ with $A(t)$ to solve the forecasting problems, then it is called M-factors FTS.

³References are: (Chen 1996; Song and Chissom 1993a, 1994).

⁴References are: (Chen 1996; Song and Chissom 1993a, 1994).

One-factor FTS models (referred to as Type-1 FTS models) employ only one variable for forecasting (Hsu et al. 2003; Huarng 2001). For example, researchers in these articles (Chen et al 2008; Hsu et al. 2003) consider only *closing* price in the forecasting of the stock index. However, the stock index price consists of many different observations, such as *opening*, *high*, *low*, etc. If these additional observations are used with one-factor variable, then it is referred to as M-factors FTS model. The model proposed by Huarng and Yu (Huarng and Yu 2005) is based on M-factors, because they use *high* and *low* as the secondary-observations to forecast the *closing* price of TAIEX.

Definition 2.2.11 (*Type-2 fuzzy set*) (Greenfield and Chiclana 2013). Let $A(U)$ be the set of fuzzy sets in U . A Type-2 fuzzy set A in X is fuzzy set whose membership grades are themselves fuzzy. This implies that $\mu_A(x)$ is a fuzzy set in U for all x , i.e., $\mu_A : X \rightarrow A(U)$ and

$$A = \{(x, \mu_A(x)) | \mu_A(x) \in A(U) \forall x \in X\} \quad (2.2.10)$$

The concept of Type-2 fuzzy set is explained with an example as follows:

[Explanation] When we cannot distinguish the degree of membership of an element in a set as 0 or 1, we use Type-1 fuzzy sets. Similarly, when the nature of an event is so fuzzy so that determination of degree of membership as a crisp number in the range $[0, 1]$ is so difficult, then we use Type-2 fuzzy sets (Mencattini et al. 2005). This Type-2 fuzzy sets concept was first introduced by Zadeh (1975a) in 1975. In Type-1 fuzzy set, the degree of membership is characterized by a crisp value; whereas in Type-2 fuzzy set, the degree of membership is regarded as a fuzzy set (Chen 2012). Thus, if there are more uncertainty in the event, and we have difficulty in determining its exact value, then we simply use Type-1 fuzzy sets, rather than crisp sets. But, ideally we have to use some finite-type sets, just like Type-2 fuzzy sets (Mencattini et al. 2005). Based on this explanation, we present an example which is based on article Huarng and Yu (2005) as follows:

Let us consider a fuzzy set for “Closing Price” of stock index, as shown in Fig. 2.1 (left). Here, we have a crisp degree of membership values 1.0 and 0.5 for the “Closing Price = 1000” and “Closing Price = 500”, respectively. Based on the above explanation, the “Closing Price = 1000” can have more than one degree of memberships. For example, in Fig. 2.1 (right), there are three degrees of memberships (0.4, 0.5 and 0.6) for the “Closing Price = 1000”. In other words, there can be multiple degrees of membership for the same “Closing Price = 1000”, as shown in Fig. 2.1 (right). In Fig. 2.1 (right), the highest degree of membership (0.6) indicates the positive view about the occurrence of event, whereas the lowest degree of membership (0.4) indicates the negative view about the occurrence of event. We can use these positive

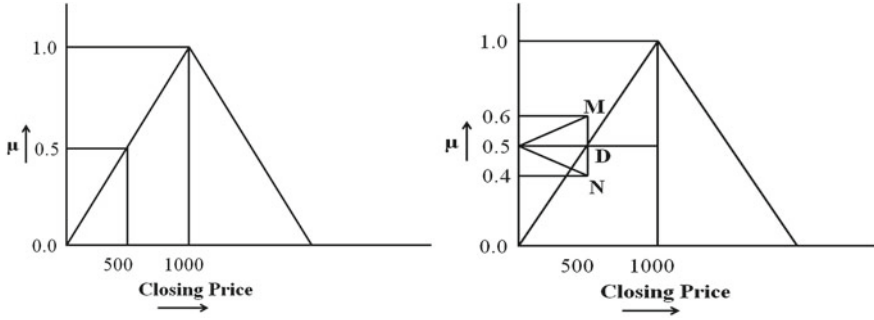


Fig. 2.1 A Type-1 (*left*) and Type-2 (*right*) fuzzy sets

and negative views together in FTS modeling approach. In summary, we can use more observations/information from the positive and negative views for forecasting in each time period.

Definition 2.2.12 (*Type-2 FTS model*) (Huarng and Yu 2005). A Type-2 FTS model can be defined as an extension of a Type-1 FTS model. The Type-2 FTS model employs the FLRs established by a Type-1 model based on Type-1 observations. Fuzzy operators such as union and intersection are used to establish the new FLRs obtained from Type-1 and Type-2 observations. Then, Type-2 forecasts are obtained from these FLRs.

2.3 FTS Modeling Approach

Chen (1996) proposed a simple calculation method to get a higher forecasting accuracy in FTS model. Still this model is used as the basis of FTS modeling. The basic architecture of this model is depicted in Fig. 2.2. This model employs the following five common steps to deal with the forecasting problems of time series, which are explained below. Contributions of various research articles in different phases of this model are also categorized in this section.

- step 1. *Partition the universe of discourse into intervals.* The universe of discourse can be defined based on Eq. 2.2.1. After determination of length of intervals, U can be partitioned into several equal lengths of intervals. For determining the universe of discourse and to partition them into effective lengths of intervals, many researchers provide various solutions in these articles (Chen and Tanuwijaya 2011; Cheng et al. 2008a; Lee and Chou 2004; Li and Cheng 2007; Liu and Wei 2010; Yu 2005a). Some recent advancement in this step can be found in these articles (Bang and Lee 2011; Chen and Wang 2010; Cheng et al. 2011; Egrioglu et al. 2010, 2011a,b; Gangwar and Kumar

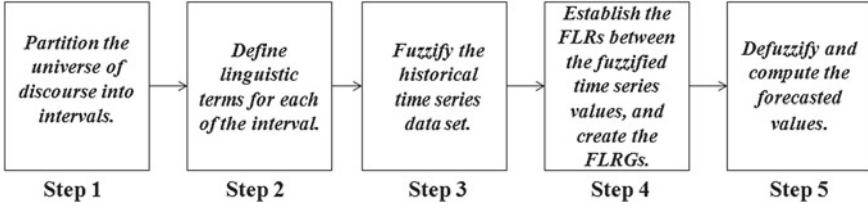


Fig. 2.2 Architecture of Chen's Model

2012; Huarng and Yu 2006b; Li et al. 2010, 2011; Liu et al. 2009; Wong et al. 2010).

step 2. *Define linguistic terms for each of the interval.* After generating the intervals, linguistic terms are defined for each of the interval. In this step, we assume that the historical time series data set is distributed among n intervals (i.e., a_1, a_2, \dots , and a_n). Then, define n linguistic variables A_1, A_2, \dots, A_n , which can be represented by fuzzy sets, as shown below:

$$\begin{aligned}
 A_1 &= 1/a_1 + 0.5/a_2 + 0/a_3 + \dots + 0/a_{n-2} + 0/a_{n-1} + 0/a_n, \\
 A_2 &= 0.5/a_1 + 1/a_2 + 0.5/a_3 + \dots + 0/a_{n-2} + 0/a_{n-1} + 0/a_n, \\
 A_3 &= 0/a_1 + 0.5/a_2 + 1/a_3 + \dots + 0/a_{n-2} + 0/a_{n-1} + 0/a_n, \\
 &\vdots \\
 A_n &= 0/a_n + 0/a_2 + 0/a_3 + \dots + 0/a_{n-2} + 0.5/a_{n-1} + 1/a_n.
 \end{aligned} \tag{2.3.1}$$

Then, we obtain the degree of membership of each time series value belonging to each A_i . Here, maximum degree of membership of fuzzy set A_i occurs at interval a_i , and $1 \leq i \leq n$. Then, each historical time series value is fuzzified. For example, if any time series value belongs to the interval a_i , then it is fuzzified into A_i , where $1 \leq i \leq n$.

For ease of computation, the degree of membership values of fuzzy set A_j ($j = 1, 2, \dots, n$) are considered as either 0, 0.5 or 1, and $1 \leq i \leq n$. In Eq. 2.3.1, for example, A_1 represents a linguistic value, which denotes a fuzzy set $= \{a_1, a_2, \dots, a_n\}$. This fuzzy set consists of n members with different degree of membership values $= \{1, 0.5, 0, \dots, 0\}$. Similarly, the linguistic value A_2 denotes the fuzzy set $= \{a_1, a_2, \dots, a_n\}$, which also consists of n members with different degree of membership values $= \{0.5, 1, 0.5, \dots, 0\}$. The descriptions of remaining linguistic variables, viz., A_3, A_4, \dots, A_n , can be provided in a similar manner.

Since each fuzzy set contains n intervals, and each interval corresponds to all fuzzy sets with different degree of membership values. For example, interval a_1 corresponds to linguistic variables A_1 and A_2 with degree of membership values 1 and 0.5, respectively, and remaining fuzzy sets with

degree of membership value 0. Similarly, interval a_2 corresponds to linguistic variables A_1 , A_2 and A_3 with degree of membership values 0.5, 1, and 0.5, respectively, and remaining fuzzy sets with degree of membership value 0. The descriptions of remaining intervals, viz., a_3, a_4, \dots, a_n , can be provided in a similar manner.

Liu (2007) introduced an improved FTS forecasting method in which the forecasted value is regarded as a trapezoidal fuzzy number instead of a single-point value. They replace the above discrete fuzzy sets (as discussed in Eq. 2.3.1) with trapezoidal fuzzy numbers. The main advantage of the proposed method is that the decision analyst can accumulate information about the possible forecasted ranges under different degrees of confidence.

- step 3. *Fuzzify the historical time series data set.* In order to fuzzify the historical time series data, it is essential to obtain the degree of membership value of each observation belonging to each A_j ($j = 1, 2, \dots, n$) for each day/year. If the maximum membership value of one day's/year's observation occurs at interval a_i and $1 \leq i \leq n$, then the fuzzified value for that particular day/year is considered as A_i .

In FTS model, each fuzzy set carries the information of occurrence of the historic event in the past. So, if these fuzzy sets would not be handled efficiently, then important information may be lost. Therefore, for fuzzification purpose, many researchers provided different techniques in these articles (Cheng et al. 2006; Hwang et al. 1998; Sah and Degtiarev 2005).

- step 4. *Establish the FLRs between the fuzzified time series values, and create the FLRGs.* After time series data is completely fuzzified, then FLRs have been established based on Definition 2.2.7. The first-order FLR is established based on two consecutive linguistic values. For example, if the fuzzified values of time $t - 1$ and t are A_i and A_j , respectively, then establish the first-order FLR as " $A_i \rightarrow A_j$ ", where " A_i " and " A_j " are called the previous state and current state of the FLR, respectively. Similarly, the n th-order FLR is established based on $n + 1$ consecutive linguistic values. For example, if the fuzzified values of time $t - 4, t - 3, t - 2, t - 1$ and t are $A_{ai}, A_{bi}, A_{ci}, A_{di}$ and A_{ej} , respectively, then the fourth-order FLR can be established as " $A_{ai}, A_{bi}, A_{ci}, A_{di} \rightarrow A_{ej}$ ", where " $A_{ai}, A_{bi}, A_{ci}, A_{di}$ " and " A_{ej} " are called the previous state and current state of the FLR, respectively.

Most of the existing FTS models⁵ use the first-order FLRs to get the forecasting results. In these articles,⁶ researchers show that the high-order FLRs (see Definition 2.2.9) can improve the forecasting accuracy. The main reason of obtaining high accuracy from these high-order FTS models is that it can consider more linguistic values that represent the high uncertainty

⁵References are: (Chang et al. 2007; Chen 1996; Cheng et al. 2006; Huarng 2001; Hwang et al. 1998; Song and Chissom 1993a, b, 1994).

⁶References are: (Aladag et al. 2009, 2010; Avazbeigi et al. 2010; Bahrepour et al. 2011; Chen 2002; Chen and Chen 2011a, b; Chen and Chung 2006b; Chen et al 2008; Gangwar and Kumar 2012; Jilani and Burney 2008; Own and Yu 2005; Singh 2007a, c, 2008, 2009; Tsai and Wu 2000).

involved in various dynamic processes. On the other hand, to extract rule from the fuzzified time series data set, Qiu et al. (2012) utilized C-fuzzy decision trees (Pedrycz and Sosnowski 2005) in FTS model. They introduced two major improvements in C-fuzzy decision trees, *viz.*, first a new stop condition is introduced to reduce the computational cost, and second weighted C-fuzzy decision tree (WCDDT) is introduced where weight distance is computed with information gain. In this approach, the forecast rule are expressed as “if input value is . . . then it can be labeled as . . .”.

Based on the same previous state of the FLRs, the FLRs can be grouped into a FLRG (see Definition 2.2.8). For example, the FLRG “ $A_i \rightarrow A_m, A_n$ ” indicates that there are following FLRs:

$$\begin{aligned} A_i &\rightarrow A_m, \\ A_i &\rightarrow A_n. \end{aligned}$$

Step 5. *Defuzzify and compute the forecasted values.* In these articles (Song and Chissom 1993a; Tsaur et al. 2005), researchers adopted the following method to forecast enrollments of the University of Alabama:

$$Y(t) = Y(t - 1) \circ R, \quad (2.3.2)$$

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 much time to compute the union of fuzzy relations R , especially when the number of fuzzy relations is more in Eq. 2.3.2 (Chen and Hwang 2000; Huarng et al 2007). Therefore, some researchers in these articles⁷ introduced various solutions for the defuzzification operation. One of the solution introduced by Chen (Chen 1996) is presented below.

This includes the following two principles, *viz.*, **Principle 1** and **Principle 2**. The procedure for **Principle 1** is given as follows:

- **Principle 1:** For forecasting $F(t)$, the fuzzified value for $F(t - 1)$ is required, where “ t ” is the current time which we want to forecast. The **Principle 1** is applicable only if there are more than one fuzzified values available in the current state. The steps under **Principle 1** are explained next.
- Step 1. Obtain the fuzzified value for $F(t - 1)$ as A_i ($i = 1, 2, 3 \dots, n$).
 - Step 2. Obtain the FLR whose previous state is A_i and the current state is $A_{j1}, A_{j2}, \dots, A_{jp}$, i.e., the FLR is in the form of “ $A_i \rightarrow A_{j1}, A_{j2}, \dots, A_{jp}$ ”.

⁷References are: (Chen 1996, 2002; Cheng et al. 2008b; Huarng 2001; Huarng et al 2007; Hwang et al. 1998; Jilani and Burney 2008; Kuo et al. 2009; Lee et al. 2006; Li et al. 2008; Qiu et al 2011; Singh and Borah 2012, 2013b; Singh 2007a, b, 2009; Yu 2005b).

- Step 3. Find the interval where the maximum membership value of the fuzzy sets $A_{j1}, A_{j2}, \dots, A_{jp}$ (current state) occur, and let these intervals be $a_{j1}, a_{j2}, \dots, a_{jp}$. All these intervals have the corresponding mid-values $C_{j1}, C_{j2}, \dots, C_{jp}$.
- Step 3. Compute the forecasted value as:

$$\text{Forecasted}_{value} = \left[\frac{C_{j1} + C_{j2} + \dots + C_{jp}}{p} \right] \quad (2.3.3)$$

Here, p represents the total number of fuzzy sets associated with the current state of the FLR.

- **Principle 2:** This principle is applicable only if there is only one fuzzified value in the current state. The steps under **Principle 2** are given as follows:

- Step 1. Obtain the fuzzified value for $F(t-1)$ as A_i ($i = 1, 2, \dots, n$).
- Step 2. Find the FLR whose previous state is A_i and the current state is A_j , i.e., the FLR is in the form of " $A_i \rightarrow A_j$ ".
- Step 3. Find the interval where the maximum membership value of the fuzzy set A_j occurs. Let these interval be a_j ($j = 1, 2, 3, \dots, n$). This interval a_j has the corresponding mid-value C_j . This C_j is the forecasted value for $F(t)$.

2.4 Hybridize Modeling Approach for FTS

Recently, several SC techniques have been employed to deal with the different challenges imposed by the FTS modeling approach. The main SC techniques for this purpose include ANN, RS, and EC. Each of them provides significant solution for addressing domain specific problems. The combination of these techniques leads to the development of new architecture, which is more advantageous and the expert, providing robust, cost effective and approximate solution, in comparison to conventional techniques. However, this hybridization should be carried out in a reasonable, rather than an expensive or a complicated, manner.

In the following, we describe the basics of individual SC techniques and their hybridization techniques, along with the several hybridized models developed for handling forecasting problems of the FTS modeling approach. It should be noted that still there is no any universally recognized method to select particular SC technique(s), which is suitable for resolving the problems. The selection of technique(s) is completely dependent on the problem and its application, and requires human interpretation for determining the suitability of a particular technique.

2.4.1 ANN: An Introduction

ANNs are massively parallel adaptive networks of simple nonlinear computing elements called *neurons* which are intended to abstract and model some of the functionality of the human nervous system in an attempt to partially capture some of its computational strengths (Kumar 2004). The neurons in an ANN are organized into different layers. Inputs to the network are entered in the input layer; whereas outputs are produced as signals in the output layer. These signals may pass through one or more intermediate or *hidden* layers which transform the signals depending upon the neuron signal functions.

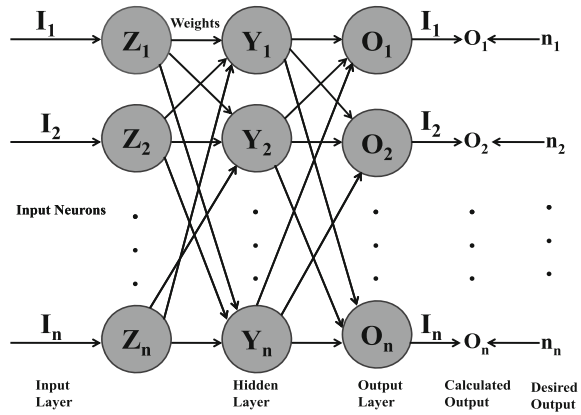
The neural networks are classified into either single-layer or multi-layer. In multi-layer networks hidden layers exist between input layer and output layer. A single-layer feed-forward (SLFF) neural network is formed when the nodes of input layer are connected with output nodes with various weights. A multi-layer feed-forward (MLFF) neural network architecture can be developed by increasing the number of layers in SLFF neural network. Feed-forward ANNs allow signals to travel from input to output. There is no feed-back loop. Feed-back networks can have signals travelling in both directions by introducing loops in the network. Feed-back networks are also referred to as interactive or recurrent networks.

Usually FFNN are used in time series forecasting. Recurrent networks are also used in some cases. Researchers employ ANN in various forecasting problems such as electric load forecasting (Taylor and Buizza 2002), short-term precipitation forecasting (Kuligowski and Barros 1998), credit ratings forecasting (Kumar and Bhat-tacharya 2006), tourism demand forecasting (Law 2000) etc., due to its capability to discover complex nonlinear relationships (Czibula et al. 2013; Donaldson and Kamstra 1996; Indro et al. 1999) in the observations. More detailed description on applications of ANN (especially BPNN) can be found in article written by Wilson et al. (2002).

Multi-layer FFNN uses back-propagation learning algorithm, therefore such networks are also known as back-propagation networks (BPNN). The main objective of using BPNN is to minimize the output error obtained from the difference between the calculated output (o_1, o_2, \dots, o_n) and target output (n_1, n_2, \dots, n_n) of the neural network by adjusting the weights (see Fig. 2.3). So in BPNN, each information is sent back again in the reverse direction until the output error is very small or zero. BPNN is trained under the process of three phases: (a) feed-forward of the input training pattern, (b) the calculation and back-propagation of the associated error, and (c) the adjustment of the weights.

Due to large number of additional parameters (e.g., initial weight, learning rate, momentum, epoch, activation function, etc.), an ANN model has great capability to learn by making proper adjustment of these parameters, in order to produce the desired output. During the training process, this output may fit the data very well, but it may produce poor results during the testing process. This implies that the neural network may not generalize well. This might be caused due to *overfitting* or *overtraining* of data (Weigend 1994), which can be controlled by monitoring the error

Fig. 2.3 A BPNN architecture with one hidden layer



during training process and terminate the process when the error reaches a minimum threshold with respect to the testing set (Sivanandam and Deepa 2007; Venkatesan et al. 1997). Another way to make the neural network generalize enough so that it performs well for training and testing data is to make small changes in the number of layers and neurons in the input space, without changing the output components. However, choosing the best neural network architecture is a heuristic approach. One solution to this problem is to keep the architecture of neural network relatively simple and small (Beale et al. 2010), because complex architectures are much more prone to overfitting (Gaume and Gosset 2003; Haykin 1999; Liu et al. 2008; Piotrowski and Napiorkowski 2012). Therefore, Hornik et al. (1989) suggested to design the neural network architecture with optimum number of neurons in the single hidden layer. Tan et al. (2009) also stated that one of the best ways to do this is to construct a fully connected neural network with a sufficiently large number of neurons in the hidden layer, and then iterate the architecture-building process with a smaller number of neurons.

Hybridization of ANN with FTS is a significant development in the domain of forecasting. It is an ensemble of the merits of ANN and FTS, by substituting the demerits of one technique by the merits of another technique. This includes various advantages of ANN, such as parallel processing, handling of large data set, fast learning capability, etc. Handling of imprecise/ uncertain and linguistic variables are done through the utilization of fuzzy sets. Besides these advantages, the FTS-ANN hybridization help in designing complex decision-making systems.

ANN can be used in different steps of FTS modeling approach. These steps are discussed in Sect. 2.3. In Fig. 2.4, three different hybridized architectures are presented, where applications of ANN are demonstrated in different steps of FTS modeling approach. In the first architecture, ANN is responsible for determination of FLRs (top); in the second architecture, ANN is responsible for partitioning the universe of discourse (middle); and in the third architecture, ANN is responsible

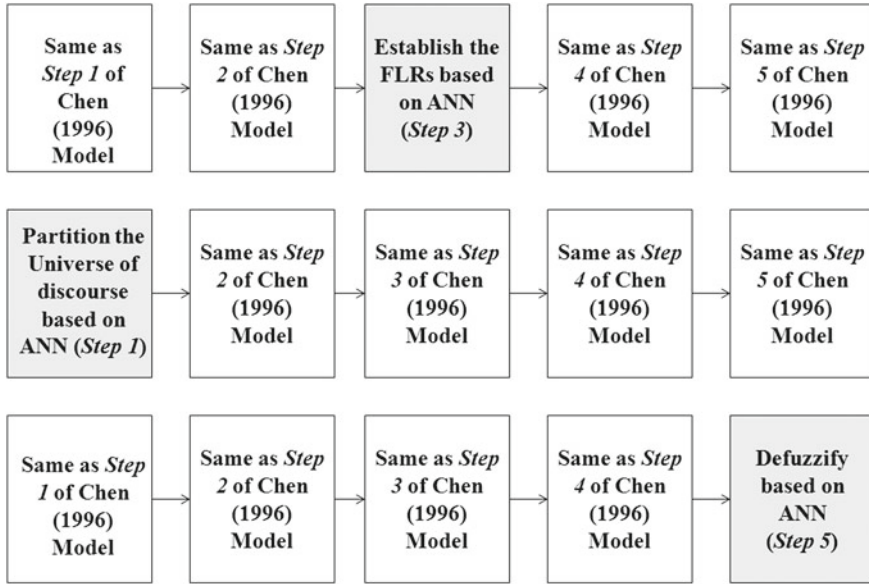


Fig. 2.4 Block diagrams of FTS-ANN hybridized models

for defuzzification operation (bottom). The roles of ANN in these architectures are explained below.

- (a) *For defining FLRs*: In this case, primary inputs for connection-oriented neural network are fuzzified time series values. The neural network is trained in terms of the number of input nodes, hidden nodes and desired outputs. One or more hidden layers are employed to automatically generate the FLRs, which may later be clustered into similar FLRGs.

In the articles (Aladag et al. 2009, 2010), researchers employ FFNN to define high-order FLRs in FTS model. Both these models are applied in forecasting the enrollments of the University of Alabama. Similar to these two approaches, many researchers (Egrioglu et al. 2009; Huarng and Yu 2006a, 2012; Yolcu et al 2011; Yu and Huarng 2010) use the ANN in FTS model to capture the FLRs for improving the forecasted accuracy.

For defining high-order FLRs, a neural network architecture for the n th-order FLRs is shown in Fig. 2.5. Here, each input node take the previous days $F(t - n), \dots, F(t - 2), F(t - 1)$ fuzzified time series values, e.g., A_l, \dots, A_m, A_n respectively to predict current day $F(t)$ fuzzified time series value, e.g., A_j . Here, each “ t ” represents the day for corresponding fuzzified time series values. Based on the input and output fuzzified values, the n th-order FLRs are established as: $A_l, \dots, A_m, A_n \rightarrow A_j$. During simulation, the indices of previous state fuzzy sets (e.g., l, \dots, m, n) are used as inputs, whereas index of current state fuzzy set (e.g., j) is used as target output.

- (b) *For partitioning the Universe of discourse*: Data clustering is a popular approach for automatically finding classes, concepts, or groups of patterns (Gondek and Hofmann 2007). Time series data are pervasive across all human endeavors, and their clustering is one of the most fundamental applications of data mining (Keogh and Lin 2005). In literature, many data clustering algorithms (Estivill-Castro 2002; Ordonez 2003; Wu et al. 2008) have been proposed, but their applications are limited to the extraction of patterns that represent points in multidimensional spaces of fixed dimensionality (Xiong and Yeung 2002). In these articles (Bahrepour et al. 2011; Singh and Borah 2012), researchers employ SOFM clustering algorithm for determining the intervals of the historical time series data sets by clustering them into different groups. This algorithm is developed by Kohonen (Kohonen 1990), 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 (Liao 2005). The SOFM converts the patterns of arbitrary dimensionality into response of one-dimensional or two-dimensional arrays of neurons, i.e., it converts a wide pattern space into a feature space. The neural network performing such a mapping is called feature map (Sivanandam and Deepa 2007).
- (c) *For defuzzification operation*: (Singh and Borah (2013b) develop an ANN based architecture and hybridize this architecture with FTS model to defuzzify the fuzzified time series values. The neural network architecture as shown in Fig. 2.5 can be employed for this purpose. In this case, the arrangement of nodes in input layer can be done in the following sequence:

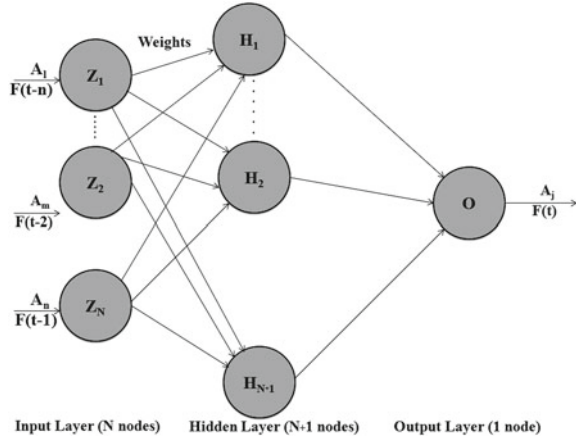
$$F(t-n), \dots, F(t-2), F(t-1) \rightarrow F(t) \quad (2.4.1)$$

Here, each input node take the previous days $(t-n), \dots, (t-2), (t-1)$ fuzzified time series values (e.g., A_l, \dots, A_m, A_n) to predict one day (t) advance time series value " A_j ". In Eq. 2.4.1, each " t " represent the day for considered fuzzified time series values.

2.4.2 RS: An Introduction

RS is a new mathematical tool proposed by Pawlak (Pawlak 1982). The RS concept (Cheng et al. 2010) is based on the assumption that with every associated object of the universe of discourse, some information objects characterized by the same information are indiscernible in the view of the available information about them. Any set of all indiscernible objects is called an elementary set and forms a basic granule of knowledge about the universe. Any union of elementary sets is referred to as a precise set; otherwise the set is rough. A fundamental advantage of RS theory is the ability to handle a category that cannot be sharply defined from a given knowledge base (Pattaraintakorn and Cercone 2008). Therefore, the RS theory is used

Fig. 2.5 ANN architecture for the n th-order FLRS



in attribute selection, rule discovery and various knowledge discovery applications as data mining, machine learning and medical diagnoses (Chen and Cheng 2013).

To understand the RS theory in-depth, we need to review some of the basic definitions as follows (Pawlak 1991):

U is a finite set of objects, i.e., $U = \{x_1, x_2, x_3, \dots, x_n\}$. Here, each $x_1, x_2, x_3, \dots, x_n$ represents the object.

Definition 2.4.1 (*Equivalence relation*). Let R be an equivalence relation over U , then the family of all equivalence classes of R is represented by U/R .

Definition 2.4.2 (*Lower approximation and upper approximation*). X is a subset of U , R is an equivalence relation, the lower approximation of X (i.e., $\underline{R}(X)$) and the upper approximation of X (i.e., $\overline{R}(X)$) is defined as follows:

$$\underline{R}(X) = \cup\{x \in U \mid [x]_R \subseteq X\} \quad (2.4.2)$$

$$\overline{R}(X) = \cup\{x \in U \mid [x]_R \cap X \neq \emptyset\} \quad (2.4.3)$$

The lower approximation comprises of all objects that completely belong to the set, and the upper approximation comprises all objects that possibly belong to the set.

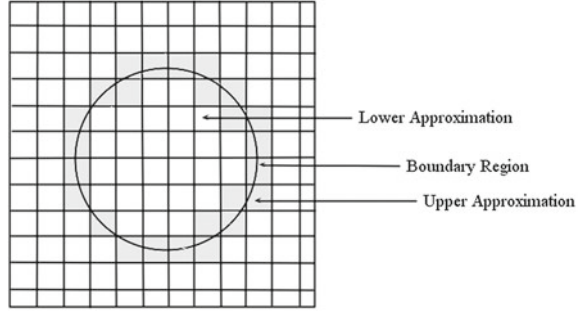
Definition 2.4.3 (*Boundary region*). The set of all objects which can be decisively classified neither as members of X nor as members of non- X with respect to R is called the boundary region of a set X with respect to R , and denoted by RS_B .

$$RS_B = \overline{R}(X) - \underline{R}(X) \quad (2.4.4)$$

Based on the notions shown in Fig. 2.6, we can formulate the definitions of crisp set and RS as follows:

Definition 2.4.4 (*Crisp set*). A set X is called crisp (exact) with respect to R if and only if the boundary region of X is empty.

Fig. 2.6 Basic notations of the rough set



Definition 2.4.5 (*RS*). A set X is called rough (inexact) with respect to R if and only if the boundary region of X is nonempty.

The role of RS in FTS modeling approach is discussed below.

- *For rule induction:* In FTS model, each fuzzy set carries the information of occurrence of the historic event in the past. So, if these fuzzy sets would not be handled efficiently, then important information may be lost. Therefore, after generating the intervals, the historical time series data set is fuzzified, and can be used to prepare an information table. To mine reasonable rules from the information table, the RS based rule induction technique can be used, because the RS (Pawlak 1982) acts as a powerful tool for analyzing data and information tables. Teoh et al. 2008, 2009 employ this concept in FTS modeling approach to generate rules from the FLRs. The rules produced by RS rule induction method are in the form of “if-then” by combining a condition value (A_i) with several decision values (A_j, A_k, \dots, A_n). For example, these decision values can be represented with “Then” as follows:

$$\text{If (condition} = A_i) \text{ Then (decision} = A_j, A_k, \dots, A_n) \quad (2.4.5)$$

2.4.3 EC: An Introduction

EC is a collection of problem solving techniques that includes paradigms such as Evolutionary Strategies, Evolutionary Programs and GAs (Bonissone 1997). GA concept was first proposed by (Holland (1975)). All GAs contain three basic operators: reproduction, crossover, and mutation, where all three are analogous to their namesakes in genetics (Ross 2007). In GAs, a population consists of chromosomes and a chromosome consists of genes, where the number of chromosomes in a population is called the population size (Lee et al. 2007). In the following, we briefly review the basic concept of GA (Gen and Cheng 1997; Goldberg 1989; Sivanandam and Deepa 2007).

- Step 1. *Create a random initial state.* An initial population is created from a random selection of solutions (chromosomes).
- Step 2. *Evaluate fitness.* A value for fitness is assigned to each solution depending on how close it actually is to solving the problem.
- Step 3. *Reproduce.* Those chromosomes with a higher fitness value are more likely to reproduce offspring.
- Step 4. *Next generation.* If the new generation contains a solution that produces an output that is close enough or equal to the desired answer then the problem has been solved. Otherwise, iterate the whole process with the new generation.

PSO is a new algorithm of EC, which is applied to solve the bilevel programming problem (Wan et al. 2013). To deal with complicated optimization problem, recently many researchers hybridize this optimization technique with FTS modeling approach. In the following, we briefly review the basic concept of the PSO (Jiang et al. 2013; Lee 2006; Montalvo et al. 2008).

The PSO algorithm was first introduced by Eberhart and Kennedy 1995. It is a population-based evolutionary computation technique, which is inspired by the social behavior of animals such as bird flocking, fish schooling, and swarming theory (Eberhart and Shi 2001; Lin et al. 2010a, b). The PSO can be employed to solve many of the same kinds of problems as genetic algorithms (Kennedy and Eberhart 1995). The PSO algorithm is applied to a set of particles, where each particle has been assigned a randomized velocity. Each particle is then allowed to move towards the problem space. At each movement, each particle keeps track of its own best solution (fitness) and the best solution of its neighboring particles. The value of that fitness is called “*pbest*”. Then each particle is attracted towards finding of the global best value by keeping track of overall best value of each particle, and its location (Trelea 2003). The particle which obtained the global fitness value is called “*gbest*”.

At each step of optimization, velocity of each particle is dynamically adjusted according to its own experience and its neighboring particles, which is represented by the following equations:

$$Vel_{id,t} = \alpha \times Vel_{id,t} + M_1 \times R_{and} \times (PB_{id} - CP_{id,t}) + M_2 \times R_{and} \times (PG_{best} - CP_{id,t}) \quad (2.4.6)$$

The position of a new particle can be determined by the following equation:

$$CP_{id,t} = CP_{id,t} + Vel_{id,t} \quad (2.4.7)$$

where i represents the i th particle and d represents the dimension of the problem space. In Eq. 2.4.6, α represents the inertia weight factor; $CP_{id,t}$ represents the current position of the particle i in iteration t ; PB_{id} denotes the previous best position of the particle i that experiences the best fitness value so far (*pbest*); PG_{best} represents the global best fitness value (*gbest*) among all the particles; R_{and} gives the random value in the range of $[0, 1]$; M_1 and M_2 represent the self-confidence coefficient and the

Algorithm 1 Standard PSO Algorithm

-
- Step 1: Initialize all particles with random positions and velocities in the d -dimensional problem space.
- Step 2: Evaluate the optimization fitness function of all particles.
- Step 3: For each particle, compare its current fitness value with its $pbest$. If current value is better than $pbest$, then update $pbest$ value with the current value.
- Step 4: For each particle, compare its fitness value with its overall previous best. If the current fitness value is better than $gbest$, then update $gbest$ value with the current best particle.
- Step 5: For each particle, change the movement (velocity) and location (position) according to Eqs. 2.4.6 and 2.4.7.
- Step 6: Repeat Step 2, until stopping criterion is met, usually a sufficiently $gbest$ value is obtained.
-

social coefficient, respectively; and $Vel_{id,t}$ represents the velocity of the particle i in iteration t . Here, $Vel_{id,t}$ is limited to the range $[-Vel_{max}, Vel_{max}]$, where Vel_{max} is a constant and defined by users. The steps for the standard PSO are presented in Algorithm 1.

The role of EC in FTS modeling approach is categorized below based on different functions.

- (a) *For determination of optimal interval lengths using GA*: GA used in FTS modeling approach to arrive optimal interval lengths using certain genetic operators. In this case, some chromosomes are defined as the initial population based on the number of intervals, where each chromosome consists of genes. Initially each chromosome is randomly generated by the system. Then, the system randomly selects chromosomes and genes from the population to perform the crossover and mutation operations, respectively. The whole process is repeated until optimal interval lengths are achieved. The achievement of optimality can be measured with the performance measure parameters (refer to Sect. 2.6), such as AFER, MSE, etc. Based on this concept, researchers in these articles (Chen and Chung 2006a, b) presented the methods for forecasting the enrollments by hybridizing GA technique with FTS modeling approach. However, the basic difference between the models presented in these articles (Chen and Chung 2006a, b) is that the first model (Chen and Chung 2006b) is based on high-order FLRs, whereas the second model (Chen and Chung 2006a) is based on first-order FLRs. Similar to above approach, Lee et al. (2007, 2008) presented new methods for temperature and the TAIEX forecasting based on two-factors high-orders FLRs.
- (b) *For finding best intervals using PSO*: Recently, many researchers⁸ show that appropriate selection of intervals also increases the forecasting accuracy of the model. Therefore, in order to get the optimal intervals, they used PSO algorithm in their proposed model.⁹ They signify that PSO algorithm is more efficient and powerful than GA as applied by the researcher (Chen and Chung 2006b) in selection of proper intervals.

⁸References are: (Huang et al. 2011a,b; Kuo et al. 2009, 2010).

⁹References are: (Huang et al. 2011a,b; Kuo et al. 2009, 2010).

Algorithm 2 Type-2 FTS Forecasting Model

- Step 1: Select Type-1 and Type-2 observations.
 Step 2: Determine the universe of discourse of time series data set and partition it into different/equal lengths of intervals.
 Step 3: Define linguistic terms for each of the interval.
 Step 4: Fuzzify the time series data set of Type-1 and Type-2 observations.
 Step 5: Establish the FLRs based on Definition 2.2.7.
 Step 6: Construct the FLRGs based on Definition 2.2.8.
 Step 7: Establish the relationships between FLRGs of both Type-1 and Type-2 observations, and map-out them to their corresponding day.
 Step 8: Apply fuzzy operators (such as union or intersection) on mapped-out FLRGs of Type-1 and Type-2 observations, and obtain the fuzzified forecasting data.
 Step 9: Defuzzify the forecasting data and compute the forecasted values.
-

- (c) *For determination of membership values using PSO*: The PSO technique is first time employed by the researcher Aladag et al. (2012) to obtain the optimal membership values of the fuzzy sets in the fuzzy relationship matrix “R” (refer to Eq. 2.3.2). In this approach, first FCM clustering algorithm is used for fuzzification phase of time series data set.

2.5 Financial Forecasting and Type-2 FTS Models

The application of FTS in financial forecasting has attracted many researchers’ attention in the recent years. Many any researchers focus on designing the models for the TAIEX¹⁰ and the TIFEX¹¹ forecasting. Their applications are limited to deal with either one-factor or two-factors time series data sets. However, forecasting accuracy of financial data set can be improved by including more observations (e.g., *close*, *high*, and *low*) in the models. In Type-2 FTS modeling approach, observation that is handled by Type-1 FTS model can be termed as “main-factor / Type-1 observation”, whereas observations that are handled by Type-2 FTS model can be termed as “secondary-factors / Type-2 observations”. Later, both these observations are combined together to take the final decision. But, due to involvement of Type-2 observations with Type-1 observation, massive FLRGs are generated in Type-2 model. For this reason, Type-2 FTS model suffers from the burden of extra computation. Therefore, most of the researchers still use to prefer Type-1 FTS modeling approach for forecasting. But, as far as accuracy of forecasting is concerned, Type-2 FTS models produce better result than Type-1 FTS models. Basic steps involve in Type-2 FTS modeling approach that can deal with multiple observations together are presented in Algorithm 2.

¹⁰References are: (Cheng et al. 2013; Huarng and Yu 2012; Wei et al 2011; Yu and Huarng 2010).

¹¹References are: (Aladag et al. 2012; Avazbeigi et al. 2010; Bai et al. 2011; Kuo et al. 2010).

Contributions of various researchers in Type-2 FTS models are presented below:

- *Huarng and Yu* (Huarng and Yu 2005) *model*: This model first time employs the Type-2 FTS concept in financial forecasting (TAIEX) by considering *close*, *high*, and *low* observations together. In this model, they suggested some improvement in Algorithm 2 as:
 - (i) Introduction of union (\vee) and intersection (\wedge) operators. This operators are applied in **Step 8** of Algorithm 2. Both these operators are used to include Type-1 and Type-2 observations., and (ii) For defuzzification operation, they employ **Principal 1** and **Principal 2** (as discussed in Sect. 2.3) in **Step 9** of Algorithm 2.
- *Bajestani and Zare* (Bajestani and Zare 2011) *model*: This model is the enhancement of the model proposed by Huarng and Yu (2005). In this model, researchers employ the four changes as:
 - (i) Using triangular fuzzy set with indeterminate legs and optimizing these triangular fuzzy sets. This improvement is applied in **Step 3** of Algorithm 2., (ii) Using indeterminate coefficient in calculating Type-2 forecasting. This improvement is applied in **Step 9** of Algorithm 2., (iii) Using center of gravity defuzzifier. This improvement is applied in **Step 9** of Algorithm 2., and (iv) Using 4-order Type-2 FTS. This improvement is applied in **Step 5** of Algorithm 2.
- *Lertworaprachaya et al.* (Lertworaprachaya et al. 2010) *model*: Based on these articles (Huarng and Yu 2005; Singh 2007b), a novel high-order Type-2 FTS model is proposed in this article (Lertworaprachaya et al. 2010). This model is divided into two parts: high-order Type-1 FTS forecasting and Type-2 FTS forecasting. The high-order Type-1 FTS model is employed to define the FLRs. This improvement is suggested in **Step 5** of Algorithm 2. The high-order FLRs can be defined based on Definition 2.2.9. Then the rules in the high-order Type-1 FTS is used in Type-2 FTS forecasting.
- *Singh and Borah* (Singh and Borah 2013a) *model*: This Type-2 FTS model can utilize multiple observations together in forecasting, which was the limitation of previous existing Type-2 FTS models. Detail discussion on this model is provided in Chap. 6.

2.6 Performance Measure Parameters

To assess the performance of the time series forecasting models (especially FTS models), researchers use numerous performance measure parameters, such as $AFER$, MSE , $RMSE$, \bar{A} , SD , U , TS , DA , δ_r , R , R^2 , PP , etc. All these parameters and their statistical significance are presented in Table 2.1. In this table, each F_i and A_i is the forecasted and actual value of day/year i , respectively, and N is the total number of days/years to be forecasted.

Table 2.1 Performance measure parameters and its statistical significance

Parameter	Significant
$A F E R = \frac{ F_i - A_i /A_i}{N} \times 100 \%$	Smaller value of $A F E R$ indicates good forecasting
$M S E = \frac{\sum_{i=1}^N (F_i - A_i)^2}{N}$	Smaller value of $M S E$ indicates good forecasting
$R M S E = \sqrt{\frac{\sum_{i=1}^N (F_i - A_i)^2}{N}}$	Smaller value of $R M S E$ indicates good forecasting
$\bar{A} = \frac{\sum_{i=1}^N A_i}{N}$	For a good forecasting, the observed mean should be close to the predicted mean
$S D = \sqrt{\frac{1}{N} \sum_{i=1}^N (A_i - \bar{A})^2}$	For a good forecasting, the observed $S D$ should be close to the predicted $S D$
$U = \frac{A}{B}$	<p>Here, $A = \sqrt{\sum_{i=1}^N (A_i - F_i)^2}$ and</p> <p>$B = \sqrt{\sum_{i=1}^N A_i^2} + \sqrt{\sum_{i=1}^N F_i^2}$</p> <p>The U is bounded between 0 and 1, with values closer to 0 indicating good forecasting accuracy</p>
$T S = \frac{R_{sfe}}{M_{ad}}$	<p>A $T S$ value between -4 and $+4$ indicates that the model is working correctly</p> <p>Here, $M_{ad} = \frac{\sum_{i=1}^N (F_i - A_i) }{N}$ and</p> <p>$R_{sfe} = \sum_{i=1}^N (F_i - A_i)$</p> <p>A $M_{ad} > 0$ indicates that forecasting model tends to under-forecast</p> <p>A $M_{ad} < 0$ indicates that forecasting model tends to over-forecast</p>
$D A = \frac{1}{N-1} \sum_{i=1}^{N-1} a_i$	<p>Here, $a_i = \begin{cases} 1, & (A_{i+1} - A_i)(F_{i+1} - A_i) > 0 \\ 0, & \text{Otherwise} \end{cases}$</p> <p>$D A$ value is measured in % and its value closer to 100 indicates good forecasting</p>
$\delta_r = \frac{ F_i - A_i }{S D}$	A value of δ_r less than 1 indicates good forecasting
$R = \frac{n \sum A_i F_i - (\sum A_i)(\sum F_i)}{\sqrt{n(\sum A_i^2) - (\sum A_i)^2} \sqrt{n(\sum F_i^2) - (\sum F_i)^2}}$	A value of R greater than equal to 0.8 is generally considered as strong
	The R^2 lies between $0 < R^2 < 1$, and indicates the strength of the linear association between A_i and F_i
$P P = 1 - (R M S E / S D)$	A $P P$ value greater than zero indicates good forecasting and vice-versa

2.7 Conclusion and Discussion

From 1994 onwards, numerous time series forecasting models have been proposed based on the FTS modeling approach.¹² Due to the uncertain nature of time series, scope of extensive applications in this domain raised simultaneously with the development of new algorithms and architectures. The FTS modeling approach is currently applied to a diverse range of fields from economy, population growth, weather forecasting, stock index price forecasting to pollution forecasting, etc. Various aspects of complexities arise in this research domain, if the number of factors in time series data sets is large. These complexities can be evolved in terms of (a) Determination of length of intervals, (b) Establishment of FLRs between different factors, and (c) Defuzzification of fuzzified time series values.

Present research in the FTS modeling approach mainly aims at designing algorithms for discretization of time series data set, rule generation from the fuzzified time series values, proposing techniques for defuzzification operation, and designing various hybridized based architectures for resolving complex decision making problems.

SC techniques comprise of ANN, RS, EC, and their hybridizations, have recently been employed to solve FTS modeling problems. They endeavor to provide us approximate results in a very cost effective manner, thereby reducing the time complexity. In this survey, a categorization has been presented based on utilization of different SC techniques with the FTS modeling approach along with basic architectures of different hybridized based FTS models.

Fuzzy sets are the oldest component of SC, which is known for representation of real time or uncertain events in a linguistic manner, and can take decisions very faster. ANNs are especially used in discovering the rules, and can establish a linear association between the inputs and outputs. RSs is mainly employed for extracting hidden patterns from the data in terms of rules. EC provides efficient search algorithms to select based intervals from the discretized time series data set, based on some evaluation criterion.

FTS-ANN hybridization exploits the features of both ANN and fuzzy sets in establishment of FLRs/linguistic rules, data discretization, and defuzzification of fuzzified time series data set. FTS-RS hybridization uses the features of both RS and fuzzy sets in discovering meaning full rules from the fuzzified time series data set, thereby employing these rules in defuzzification operation. FTS-EC hybridization utilizes the characteristics of both EC and fuzzy sets in the determination of optimal interval lengths of the discretized time series data set, which are further used to represent time series data set in terms of fuzzy sets/linguistic terms. From this survey, it is obvious that the research scope in FTS will be increased in the near future for its flexibility in representing real life problems in a very natural way. This study also describes elaborately different phases of the FTS modeling approach. Various research issues and challenges in the FTS modeling approach are presented in the subsequent section. All

¹²References are: (Egrioglu et al. 2010, 2011a; Sah and Degtiarev 2005; Wong et al. 2010; Yu 2005a).

these inclusions may help the researchers to identify: (a) What are the problems in the FTS modeling approach?, (b) How to resolve all these problems using heuristics approach?, and (c) How to employ different SC methodologies in the FTS modeling approach to improve its efficiency?

References

- Aladag CH, Basaran MA, Egrioglu E, Yolcu U, Uslu VR (2009) Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations. *Expert Syst Appl* 36(3):4228–4231
- Aladag CH, Yolcu U, Egrioglu E (2010) A high order fuzzy time series forecasting model based on adaptive expectation and artificial neural networks. *Math Comput Simul* 81(4):875–882
- Aladag CH, Yolcu U, Egrioglu E, Dalar AZ (2012) A new time invariant fuzzy time series forecasting method based on particle swarm optimization. *Appl Soft Comput* 12(10):3291–3299
- Avazbeigi M, Doulabi SHH, Karimi B (2010) Choosing the appropriate order in fuzzy time series: a new N-factor fuzzy time series for prediction of the auto industry production. *Expert Syst Appl* 37(8):5630–5639
- Bahrepour M, Akbarzadeh-T MR, Yaghoobi M, Naghibi-S MB (2011) An adaptive ordered fuzzy time series with application to FOREX. *Expert Syst Appl* 38(1):475–485
- Bai E, Wong WK, Chu WC, Xia M, Pan F (2011) A heuristic time-invariant model for fuzzy time series forecasting. *Expert Syst Appl* 38(3):2701–2707
- Bajestani NS, Zare A (2011) Forecasting TAIEX using improved type 2 fuzzy time series. *Expert Syst Appl* 38(5):5816–5821
- Bang YK, Lee CH (2011) Fuzzy time series prediction using hierarchical clustering algorithms. *Expert Syst Appl* 38(4):4312–4325
- Beale MH, Hagan MT, Demuth HB (2010) *Neural network Toolbox 7*. The MathWorks Inc, Natick, MA
- Bonissone PP (1997) Soft computing: the convergence of emerging reasoning technologies. *Soft Comput* 1:6–18
- Brockwell PJ, Davis RA (2008) *Introduction to time series and forecasting*, 2nd edn. Springer, New York
- Castro LN, Timmis JI (2003) Artificial immune systems as a novel soft computing paradigm. *Soft Comput* 7:526–544
- Chang JR, Lee YT, Liao SY, Cheng CH (2007) Cardinality-based fuzzy time series for forecasting enrollments. *New Trends Appl Artif Intell*, vol 4570. Springer, Berlin/Heidelberg, pp 735–744
- Chatfield C (2000) *Time-series forecasting*. Chapman and Hall, CRC Press, Boca Raton
- Chen SM (1996) Forecasting enrollments based on fuzzy time series. *Fuzzy Sets Syst* 81:311–319
- Chen SM (2002) Forecasting enrollments based on high-order fuzzy time series. *Cybern Syst: Int J* 33(1):1–16
- Chen SM, Chen CD (2011a) Handling forecasting problems based on high-order fuzzy logical relationships. *Expert Syst Appl* 38(4):3857–3864
- Chen SM, Chen CD (2011b) Handling forecasting problems based on high-order fuzzy logical relationships. *Expert Syst Appl* 38(4):3857–3864
- Chen SM, Chung NY (2006a) Forecasting enrollments of students by using fuzzy time series and genetic algorithms. *Int J Inf Manag Sci* 17(3):1–17
- Chen SM, Chung NY (2006b) Forecasting enrollments using high-order fuzzy time series and genetic algorithms. *Int J Intell Syst* 21(5):485–501
- Chen SM, Hwang JR (2000) Temperature prediction using fuzzy time series. *IEEE Trans Syst Man Cybern Part B: Cybern* 30:263–275
- Chen SM, Tanuwijaya K (2011) Multivariate fuzzy forecasting based on fuzzy time series and automatic clustering techniques. *Expert Syst Appl* 38(8):10,594–10,605

- Chen SM, Wang NY (2010) Fuzzy forecasting based on fuzzy-trend logical relationship groups. *IEEE Trans Syst Man Cybern Part B: Cybern* 40(5):1343–1358
- Chen TL, Cheng CH, Teoh HJ (2008) High-order fuzzy time-series based on multi-period adaptation model for forecasting stock markets. *Phys A: Stat Mech Appl* 387(4):876–888
- Chen TY (2012) A signed-distance-based approach to importance assessment and multi-criteria group decision analysis based on interval type-2 fuzzy set. *Knowl Inf Syst*
- Chen YS, Cheng CH (2013) Application of rough set classifiers for determining hemodialysis adequacy in ESRD patients. *Knowl Inf Syst* 34:453–482
- 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 Change* 73:524–542
- Cheng CH, Cheng GW, Wang JW (2008a) Multi-attribute fuzzy time series method based on fuzzy clustering. *Expert Syst Appl* 34:1235–1242
- Cheng CH, Wang JW, Li CH (2008b) Forecasting the number of outpatient visits using a new fuzzy time series based on weighted-transitional matrix. *Expert Syst Appl* 34(4):2568–2575
- Cheng CH, Chen TL, Wei LY (2010) A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting. *Inf Sci* 180(9):1610–1629
- Cheng CH, Huang SF, Teoh HJ (2011) Predicting daily ozone concentration maxima using fuzzy time series based on a two-stage linguistic partition method. *Comput Math Appl* 62(4):2016–2028
- Cheng CH, Wei LY, Liu JW, Chen TL (2013) OWA-based ANFIS model for TAIEX forecasting. *Econ Modell* 30:442–448
- Czibula G, Czibula IG, Găceanu RD (2013) Intelligent data structures selection using neural networks. *Knowl Inf Syst* 34:171–192
- Donaldson RG, Kamstra M (1996) Forecast combining with neural networks. *J Forecast* 15(1):49–61
- Dote Y, Ovaska SJ (2001) Industrial applications of soft computing: a review. *Proc IEEE* 89(9):1243–1265
- Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. In: *Proceedings of the sixth international symposium on micro machine and human science, Nagoya*, pp 39–43
- Eberhart R, Shi Y (2001) Particle swarm optimization: Developments, applications and resources. In: *Proceedings of the IEEE international conference on evolutionary computation, Brisbane, Australia*, pp 591–600
- Egrioglu E, Aladag CH, Yolcu U, Basaran MA, Uslu VR (2009) A new hybrid approach based on SARIMA and partial high order bivariate fuzzy time series forecasting model. *Expert Syst Appl* 36(4):7424–7434
- Egrioglu E, Aladag CH, Yolcu U, Uslu VR, Basaran MA (2010) Finding an optimal interval length in high order fuzzy time series. *Expert Syst Appl* 37(7):5052–5055
- Egrioglu E, Aladag CH, Basaran MA, Yolcu U, Uslu VR (2011a) A new approach based on the optimization of the length of intervals in fuzzy time series. *J Intell Fuzzy Syst* 22(1):15–19
- Egrioglu E, Aladag CH, Yolcu U, Uslu V, Erilli N (2011b) Fuzzy time series forecasting method based on Gustafson-Kessel fuzzy clustering. *Expert Syst Appl* 38(8):10355–10357
- Estivill-Castro V (2002) Why so many clustering algorithms: a position paper. *ACM SIGKDD Explor Newslett* 4(1):65–75
- Gangwar SS, Kumar S (2012) Partitions based computational method for high-order fuzzy time series forecasting. *Expert Systems with Applications* 39(15):12,158–12,164
- Gaume E, Gosset R (2003) Over-parameterisation, a major obstacle to the use of artificial neural networks in hydrology? *Hydrol Earth Syst Sci* 7(5):693–706
- Gen M, Cheng R (1997) *Genetic algorithms and engineering design*. Wiley, New York
- Goldberg DE (1989) *Genetic algorithm in search, optimization, and machine learning*. Addison-Wesley, Massachusetts
- Gondek D, Hofmann T (2007) Non-redundant data clustering. *Knowl Inf Syst* 12:1–24
- Greenfield S, Chiclana F (2013) Accuracy and complexity evaluation of defuzzification strategies for the discretised interval type-2 fuzzy set. *Int J Approximate Reasoning* 54(8):1013–1033

- Haykin S (1999) *Neural Networks, a comprehensive foundation*. Macmillan College Publishing Co., New York
- Herrera-Viedma E, Cabrerizo FJ, Kacprzyk J, Pedrycz W (2014) A review of soft consensus models in a fuzzy environment. *Information Fusion* 17:4–13
- Holland JH (1975) *Adaptation in natural and artificial systems*. MIT Press, Cambridge
- Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. *Neural Netw* 2(5):359–366
- Hsu YY, Tse SM, Wu B (2003) A new approach of bivariate fuzzy time series analysis to the forecasting of a stock index. *Int J Uncertain Fuzziness Knowl-Based Syst* 11(6):671–690
- Huang YL, Horng SJ, He M, Fan P, Kao TW, Khan MK, Lai JL, Kuo IH (2011a) A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimization. *Expert Syst Appl* 38(7):8014–8023
- Huang YL, Horng SJ, Kao TW, Run RS, Lai JL, Chen RJ, Kuo IH, Khan MK (2011b) An improved forecasting model based on the weighted fuzzy relationship matrix combined with a PSO adaptation for enrollments. *Int J Innov Comput Inf Control* 7(7A):4027–4046
- Huarng K (2001) Heuristic models of fuzzy time series for forecasting. *Fuzzy Sets Systems* 123:369–386
- Huarng K, Yu HK (2005) A Type 2 fuzzy time series model for stock index forecasting. *Phys A: Stat Mech Appl* 353:445–462
- Huarng K, Yu THK (2006a) The application of neural networks to forecast fuzzy time series. *Phys A: Stat Mech Appl* 363(2):481–491
- Huarng K, Yu THK (2006b) Ratio-based lengths of intervals to improve fuzzy time series forecasting. *IEEE Trans Syst Man Cybern Part B: Cybern* 36(2):328–340
- Huarng KH, Yu THK (2012) Modeling fuzzy time series with multiple observations. *Int J Innov Comput Inf Control* 8(10(B)):7415–7426
- Huarng KH, Yu THK, Hsu YW (2007) A multivariate heuristic model for fuzzy time-series forecasting. *IEEE Trans Syst Man Cybern Part B: Cybern* 37:836–846
- Hwang JR, Chen SM, Lee CH (1998) Handling forecasting problems using fuzzy time series. *Fuzzy Sets Syst* 100:217–228
- Indro DC et al (1999) Predicting mutual fund performance using artificial neural networks. *Omega* 27(3):373–380
- Jang JSR, Sun CT, Mizutani E (1997) *Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence*. Prentice-Hall, London
- Jiang Y, Li X, Huang C, Wu X (2013) Application of particle swarm optimization based on CHKS smoothing function for solving nonlinear bilevel programming problem. *Appl Math Comput* 219(9):4332–4339
- Jilani TA, Burney SMA (2008) A refined fuzzy time series model for stock market forecasting. *Phys A* 387(12):2857–2862
- Kacprzyk J (2010) *Advances in soft computing*, vol 7095. Springer, Heidelberg
- Kennedy J, Eberhart R (1995) Particle swarm optimization. In: *IEEE international conference on neural networks*, Perth, WA 4:1942–1948
- 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
- Kohonen T (1990) The self organizing maps. In: *Proceedings of the IEEE*, vol 78, pp 1464–1480
- Kuligowski RJ, Barros AP (1998) Experiments in short-term precipitation forecasting using artificial neural networks. *Mon Weather Rev* 126:470–482
- Kumar K, Bhattacharya S (2006) Artificial neural network vs. linear discriminant analysis in credit ratings forecast: A comparative study of prediction performances. *Review of Accounting and Finance* 5(3):216–227
- Kumar S (2004) *Neural networks: a classroom approach*. Tata McGraw-Hill Education Pvt. Ltd., New Delhi

- Kuo IH, Horng SJ, Kao TW, Lin TL, Lee CL, Pan Y (2009) An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. *Expert Syst Appl* 36(3, Part 2):6108–6117
- Kuo IH, Horng SJ, Chen YH, Run RS, Kao TW, Chen RJ, Lai JL, Lin TL (2010) Forecasting TAIEX based on fuzzy time series and particle swarm optimization. *Expert Syst Appl* 37(2):1494–1502
- Law R (2000) Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. *Tour Manag* 21(4):331–340
- Lee HS, Chou MT (2004) Fuzzy forecasting based on fuzzy time series. *Int J Comput Math* 81(7):781–789
- 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
- Lee LW, Wang LH, Chen SM (2007) Temperature prediction and TAIEX forecasting based on fuzzy logical relationships and genetic algorithms. *Expert Syst Appl* 33(3):539–550
- Lee LW, Wang LH, Chen SM (2008) Temperature prediction and TAIEX forecasting based on high-order fuzzy logical relationships and genetic simulated annealing techniques. *Expert Syst Appl* 34(1):328–336
- Lee ZY (2006) Method of bilaterally bounded to solution blasius equation using particle swarm optimization. *Appl Math Comput* 179(2):779–786
- Lertworapachaya Y, Yang Y, John R (2010) High-order Type-2 fuzzy time series. *International conference of soft computing and pattern recognition*. IEEE, Paris, pp 363–368
- Li ST, Cheng YC (2007) Deterministic fuzzy time series model for forecasting enrollments. *Comput Math Appl* 53(12):1904–1920
- Li ST, Cheng YC, Lin SY (2008) A FCM-based deterministic forecasting model for fuzzy time series. *Comput Math Appl* 56(12):3052–3063
- Li ST, Kuo SC, Cheng YC, Chen CC (2011) A vector forecasting model for fuzzy time series. *Appl Soft Comput* 11(3):3125–3134
- Li ST et al (2010) Deterministic vector long-term forecasting for fuzzy time series. *Fuzzy Sets Syst* 161(13):1852–1870
- Liao TW (2005) Clustering of time series data - a survey. *Pattern Recogn* 38(11):1857–1874
- Lin SY, Horng SJ, Kao TW, Huang DK, Fahn CS, Lai JL, Chen RJ, Kuo IH (2010a) An efficient bi-objective personnel assignment algorithm based on a hybrid particle swarm optimization model. *Expert Syst Appl* 37(12):7825–7830
- Lin TL, Horng SJ, Kao TW, Chen YH, Run RS, Chen RJ, Lai JL, Kuo IH (2010b) An efficient job-shop scheduling algorithm based on particle swarm optimization. *Expert Syst Appl* 37(3):2629–2636
- Liu HT (2007) An improved fuzzy time series forecasting method using trapezoidal fuzzy numbers. *Fuzzy Optim Decis Making* 6:63–80
- Liu HT, Wei ML (2010) An improved fuzzy forecasting method for seasonal time series. *Expert Syst Appl* 37(9):6310–6318
- Liu HT, Wei NC, Yang CG (2009) Improved time-variant fuzzy time series forecast. *Fuzzy Optim Decis Making* 8:45–65
- Liu Y, Starzyk JA, Zhu Z (2008) Optimized approximation algorithm in neural networks without overfitting. *IEEE Trans Neural Netw* 19(6):983–995
- Mencattini A, Salmeri M, Bertazzoni S, Lojacono R, Pasero E, Moniaci W (2005) Local meteorological forecasting by type-2 fuzzy systems time series prediction. *IEEE International conference on computational intelligence for measurement systems and applications*. Giardini Naxos, Italy, pp 20–22
- Mitra S, Pal SK, Mitra P (2002) Data mining in soft computing framework: a survey. *IEEE Trans Neural Netw* 13(1):3–14
- Montalvo I, Izquierdo J, Pérez R, Tung MM (2008) Particle swarm optimization applied to the design of water supply systems. *Comput Math Appl* 56(3):769–776

- Ordóñez C (2003) Clustering binary data streams with K-means. In: Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery. ACM Press, New York, USA, pp 12–19
- Own CM, Yu PT (2005) Forecasting fuzzy time series on a heuristic high-order model. *Cybern Syst: Int J* 36(7):705–717
- Pattaraintakorn P, Cercone N (2008) Integrating rough set theory and medical applications. *Appl Math Lett* 21(4):400–403
- Pawlak Z (1982) Rough sets. *Int J Comput Inf Sci* 11(5):341–356
- Pawlak Z (1991) Rough sets: theoretical aspects of reasoning about data. Kluwer Academic Publishers
- Pedrycz W, Sosnowski Z (2005) C-Fuzzy decision trees. *IEEE Trans Syst Man Cybern C Appl Rev* 35(4):498–511
- Piotrowski AP, Napiorkowski JJ (2012) A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. *J Hydrol*
- Qiu W, Liu X, Li H (2011) A generalized method for forecasting based on fuzzy time series. *Expert Syst Appl* 38(8):10446–10453
- Qiu W, Liu X, Wang L (2012) Forecasting shanghai composite index based on fuzzy time series and improved C-fuzzy decision trees. *Expert Syst Appl* 39(9):7680–7689
- Ross TJ (2007) Fuzzy Logic with engineering applications. John Wiley and Sons, Singapore
- Sah M, Degtiarev K (2005) Forecasting enrollment model based on first-order fuzzy time series. *Prec World Acad Sci Eng Technol* 1:132–135
- Singh P, Borah B (2012) An effective neural network and fuzzy time series-based hybridized model to handle forecasting problems of two factors. *Knowl Inf Syst* 38(3):669–690
- Singh P, Borah B (2013a) Forecasting stock index price based on M-factors fuzzy time series and particle swarm optimization. *International Journal of Approximate Reasoning*
- Singh P, Borah B (2013b) High-order fuzzy-neuro expert system for daily temperature forecasting. *Knowl-Based Syst* 46:12–21
- Singh SR (2007a) A robust method of forecasting based on fuzzy time series. *Appl Math Comput* 188(1):472–484
- Singh SR (2007b) A simple method of forecasting based on fuzzy time series. *Appl Math Comput* 186(1):330–339
- Singh SR (2007c) A simple time variant method for fuzzy time series forecasting. *Cybernetics and Systems: An International Journal* 38(3):305–321
- Singh SR (2008) A computational method of forecasting based on fuzzy time series. *Math Comput Simul* 79(3):539–554
- Singh SR (2009) A computational method of forecasting based on high-order fuzzy time series. *Expert Syst Appl* 36(7):10,551–10,559
- Sivanandam SN, Deepa SN (2007) Principles of soft computing. Wiley India (P) Ltd., New Delhi
- Song Q, Chissom BS (1993a) Forecasting enrollments with fuzzy time series - Part I. *Fuzzy Sets Syst* 54(1):1–9
- Song Q, Chissom BS (1993b) Fuzzy time series and its models. *Fuzzy Sets Syst* 54(1):1–9
- Song Q, Chissom BS (1994) Forecasting enrollments with fuzzy time series - Part II. *Fuzzy Sets Syst* 62(1):1–8
- Szmidt E, Kacprzyk J, Bujnowski P (2014) How to measure the amount of knowledge conveyed by atanassov's intuitionistic fuzzy sets. *Inf Sci* 257:276–285
- Tan PN, Steinbach M, Kumar V (2009) Introduction to data mining, 4th edn. Dorling Kindersley Publishing Inc, India
- Taylor JW, Buizza R (2002) Neural network load forecasting with weather ensemble predictions. *IEEE Trans Power Syst* 17:626–632
- Teoh HJ, Cheng CH, Chu HH, Chen JS (2008) Fuzzy time series model based on probabilistic approach and rough set rule induction for empirical research in stock markets. *Data Knowl Eng* 67(1):103–117

- Teoh HJ, Chen TL, Cheng CH, Chu HH (2009) A hybrid multi-order fuzzy time series for forecasting stock markets. *Expert Syst Appl* 36(4):7888–7897
- Trelea IC (2003) The particle swarm optimization algorithm: convergence analysis and parameter selection. *Inf Process Lett* 85(6):317–325
- Tsai CC, Wu SJ (2000) Forecasting enrolments with high-order fuzzy time series. 19th International conference of the North American. Fuzzy Information Processing Society, Atlanta, GA, pp 196–200
- Tsaur RC, Yang JCO, Wang HF (2005) Fuzzy relation analysis in fuzzy time series model. *Comput Math Appl* 49(4):539–548
- Venkatesan C, Raskar SD, Tambe SS, Kulkarni BD, Keshavamurthy RN (1997) Prediction of all India summer monsoon rainfall using error-back-propagation neural networks. *Meteorol Atmos Phys* 62:225–240
- Wan Z, Wang G, Sun B (2013) A hybrid intelligent algorithm by combining particle swarm optimization with chaos searching technique for solving nonlinear bilevel programming problems. *Swarm Evol Comput* 8:26–32
- Wei LY, Chen TL, Ho TH (2011) A hybrid model based on adaptive-network-based fuzzy inference system to forecast Taiwan stock market. *Expert Syst Appl* 38(11):13,625–13,631
- Weigend A (1994) An overfitting and the effective number of hidden units. In: Mozer MC, Smolensky P, Weigend AS (eds) *Proceedings of the 1993 Connectionist Models Summer School*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp 335–342
- Wilson ID, Paris SD, Ware JA, Jenkins DH (2002) Residential property price time series forecasting with neural networks. *Knowl-Based Syst* 15(5–6):335–341
- Wong WK, Bai E, Chu AWC (2010) Adaptive time-variant models for fuzzy-time-series forecasting. *IEEE Trans Syst Man Cybern Part B: Cybern* 40:453–482
- Wu X et al (2008) Top 10 algorithms in data mining. *Knowl Inf Syst* 14:1–37
- 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
- Yardimci A (2009) Soft computing in medicine. *Appl Soft Comput* 9(3):1029–1043
- Yolcu U et al (2011) Time-series forecasting with a novel fuzzy time-series approach: an example for Istanbul stock market. *Journal of Statistical Computation and Simulation*
- Yu HK (2005a) A refined fuzzy time-series model for forecasting. *Phys A* 346(3–4):657–681
- Yu HK (2005b) Weighted fuzzy time series models for TAIEX forecasting. *Phys A* 349(3–4):609–624
- Yu THK, Huarng KH (2010) A neural network-based fuzzy time series model to improve forecasting. *Expert Syst Appl* 37(4):3366–3372
- Zadeh LA (1965) Fuzzy sets. *Inf Control* 8(3):338–353
- Zadeh LA (1971) Similarity relations and fuzzy orderings. *Inf Sci* 3:177–200
- Zadeh LA (1973) Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans Syst Man Cybern SMC-3*:28–44
- Zadeh LA (1975a) The concept of a linguistic variable and its application to approximate reasoning - I. *Inf Sci* 8(3):199–249
- Zadeh LA (1975b) The concept of a linguistic variable and its application to approximate reasoning-III. *Inf Sci* 9(1):43–80
- Zadeh LA (1989) Knowledge representation in fuzzy logic. *IEEE Trans Knowl Data Eng* 1(1):89–100
- Zadeh LA (1994) Fuzzy logic, neural networks, and soft computing. *Commun ACM* 37(3):77–84
- Zadeh LA (1997) The roles of fuzzy logic and soft computing in the conception, design and deployment of intelligent systems. In: *Software agents and soft computing*, pp 183–190
- Zadeh LA (2002) From computing with numbers to computing with words - From manipulation of measurements to manipulation of perceptions. *Int J Appl Math Comput Sci* 12(3):307–324

<http://www.springer.com/978-3-319-26292-5>

Applications of Soft Computing in Time Series

Forecasting

Simulation and Modeling Techniques

Singh, P.

2016, XXI, 158 p. 24 illus., 14 illus. in color., Hardcover

ISBN: 978-3-319-26292-5