Swarm and Evolutionary Computation 15 (2014) 19–26

Contents lists available at ScienceDirect

Swarm and Evolutionary Computation

journal homepage: www.elsevier.com/locate/swevo

Regular Paper

A fuzzy time series approach based on weights determined by the number of recurrences of fuzzy relations

Vedide Rezan Uslu [a], Eren Bas [b], Ufuk Yolcu [b], Erol Egrioglu [a,*]

[a] Department of Statistics, Ondokuz Mayis University, Samsun 55139, Turkey
[b] Department of Statistics, Giresun University, Giresun 28000, Turkey

ARTICLE INFO

Article history:
Received 8 February 2013
Received in revised form
13 October 2013
Accepted 24 October 2013
Available online 31 October 2013

Keywords:
Fuzzy time series
Recurrence
Weight scheme
Forecasting
Defuzzification

ABSTRACT

Fuzzy time series approaches, which do not require the strict assumptions of traditional time series approaches, generally consist of three stages. These are called as the fuzzification of crisp time series observations, the identification of fuzzy relationships and the defuzzification. All of these stages play a very important role on the forecasting performance of the model. Although there are many studies contributing to the stages of fuzzification and determining fuzzy relationships, the number of the studies about the defuzzification stage, which is very important at least as much as the others, is limited. None of them considered the number of recurrence of the fuzzy relationships in the stage of defuzzification. However it is very reasonable to take into account since fuzzy relations and their recurrence number are reflected the nature of the time series. Then the information obtained from the fuzzy relationships can be used in the defuzzification stage. In this study, we take into account the recurrence number of the fuzzy relations in the stage of defuzzification. Then this new approach has been applied to the real data sets which are often used in other studies in literature. The results are compared to the ones obtained from other techniques. Thus it is concluded that the results present superior forecasts performance.

1. Introduction

Fuzzy time series procedures do not require the assumptions such as the large sample and that the model is true. Recent studies are about fuzzy time series procedures since they do not require the strict assumptions and generally provide remarkable forecasting performances.

The fuzzy set was first introduced by Ref. [56] and this concept has found many application areas since then. Fuzzy time series was first introduced in the literature [47–49]. These proposed fuzzy time series techniques in literature generally consist of the three stages; these are fuzzification, determining of fuzzy relations and defuzzification.

In the literature, the decomposition of universe of discourse was mostly used in the fuzzification stage and intervals of it was determined arbitrarily in the studies of Refs. [47–49,6,7]. In addition, Ref. [29] has put forward the importance of the interval length on the forecasting performance and proposed two new techniques based on the mean and the distribution in order to find intervals. Refs. [18,19] have suggested forming the problem of finding intervals as an optimization problem. And also, Refs. [30] used heuristic algorithms for forecasting of time series. Refs. [8,39] used the different interval lengths, instead of the fixed interval length, found by using the genetic algorithm and also in the literature [36,37,13,42,27] used the particle swarm optimization. And also, Refs. [23,28] used particle swarm optimization technique to determine the dynamic length of interval and also the other studies in this stage were proposed by Refs. [9,41,53,19]. Besides these studies, Refs. [10,41] used fuzzy c-means clustering method in their studies and also Ref. [20] used Gustafson–Kessel fuzzy clustering method in this stage.

The contribution of some studies in literature is to the stage of determining of fuzzy relations. In this stage, while the matrix operations were used in Refs. [47–49,6] and some others were used the fuzzy logic relations group table and also the artificial neural networks for determining the fuzzy relations were used in Refs. [32,1,15–17,31,29]. The other studies in this stage were proposed in Refs. [15,16,50,1,2,54].

In this study, we want to point out the importance of the stage of the defuzzification on the model performance. There are quite few contributions on this issue in literature. Studies in the literature mostly used the centroid method. This method was used in the studies of Refs. [6,29,32]. Refs. [11,2] preferred to use adaptive expectation method in the defuzzification process and also while [49] used artificial neural networks. Different techniques in this stage were used in Refs. [55,33,34]. Ref. [55] proposed a new technique based on chronologically-determined weights. Ref. [55] emphasized the chronologic order of the fuzzy relations.

* Corresponding author. Tel.: +90 312 1919.
E-mail address: erole@omu.edu.tr (E. Egrioglu).

2210-6502/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.swevo.2013.10.004

Ref. [55] puts forward that the last occurred fuzzy relation should have more impact on the defuzzyfied forecast but we believe that the general trend or behavior of the fuzzy time series should be specified by the recurrence number of all fuzzy relations in whole series. Therefore we studied on this subject and we finally proposed a modified method for getting defuzzyfied forecasts. In this study we used the genetic algorithm in the fuzzification stage and fuzzy logic relations (*FLR*) and fuzzy logic group relations (*FLGR*) tables in the stage of the determination of fuzzy relations.

The most basic element in the structure of the model is fuzzy relations. In this study, the interval lengths have been determined by avoiding subjective decisions because of using genetic algorithm in fuzzification stage and also it is aimed to obtain more realistic forecasts by using fuzzy relations recurrence numbers. Because using recurrence numbers of fuzzy relations is important as well as fuzzy relations occur or not.

The rest part of the paper can be outlined as below: The second section of the article is about a revision of genetic algorithm. The fundamental definitions of fuzzy time series are given in Section 3. In Section 4 the proposed method is introduced. In Section 5 presents the results from the application of the proposed method to three real life data sets and finally Section 6 presents conclusions and discussions.

2. Genetic algorithms (GA)

The GA was first proposed by Ref. [26]. GA has been successfully solving many complex optimization problems. It does not require the specific mathematical analysis of optimization problems. The first question to ask when starting a problem in genetic algorithm is encoding. There are many types of encoding in genetic algorithm such as binary encoding, value encoding and permutation encoding. And also, another encoding type is real-coded encoding, so we used it in our proposed method. A genetic algorithm structure consists of chromosomes and the correct coding of these chromosomes is very important for genetic algorithm process. Genetic algorithms imitate the evolutionary process for solving the problems. An initial population is chosen randomly in the beginning and the fitness of initial population members (chromosomes) is evaluated. Then, the main genetic algorithm parameters are applied to produce new generation.

These parameters are; population size, crossover rate, mutation rate and also if necessary the other genetic operators can be used such as repairing operator. The loop of genetic algorithm is repeated until the best solution is found. The best advantage of genetic algorithm is that it can be applied to discrete, integer and continuous variables. There are many studies in the literature about the parameters of genetic algorithms. Refs. [11,24,41] studied on these parameters and suggested some parameter values about these parameters. Ref. [22] also studied on the parameters of the genetic algorithm for optimizing these parameters. Ref. [21] studied on the population size to find optimal population size and also genetic algorithms give successful results in the areas of artificial intelligence applications, control systems, robotics, image and voice recognition, engineering design, planning. Ref. [21] proposed an approach based on genetic algorithm for forecasting of fuzzy time series and found the best genetic algorithm parameters for his study. And also, Ref. [14] proposed that the population size should be taken between 50 and 100 and the mutation rate should be taken 0.001. Ref. [44] proposed that the population size should be taken between 20 and 30 and the mutation rate should be taken between 0.005 and 0.01. Ref. [25] proposed that the population size should be taken 30 and the mutation rate should be taken 0.01. Besides these studies, there are some studies in the literature about optimization of these parameters [4,12,52]. The GA search optimal solution with many chromosomes. In one chromosome, there are many gens. The gens are decision variables when GA is used for optimization. GA is generally starting from a random population. The population size can be determined by researchers due to the characteristic of the problem. After first population is randomly generated, various techniques can be utilized in the production of the next generation. Some of these techniques are crossover, mutation and natural selection which can be summarized as follows:

Crossover: The system randomly selects two chromosomes from a population and randomly picks a crossover point from the two selected chromosomes in order to swap genes after this crossover point. This operation is called as the crossover operation. And, the crossover operation is depending on the crossover rate. A random number is generated from the uniform distribution. Then, the crossover operation is performed if random number is bigger than the crossover rate. At the end of the crossover operation, two different chromosomes are obtained and therefore, it is provided diversity for the population. Besides these properties, there are some crossover methods using in the literature. Some of them are; single-point crossover, multi-point crossover, and uniform crossover. But the most popular method is single-point crossover method among these methods.

Mutation: First of all, the researcher has to determine the mutation rate. Then, one chromosome is randomly selected. If a real value generated from the interval (0,1) is smaller than or equal to the mutation rate, the mutation operation will be performed with a randomly selected gene from the chromosomes. There are various mutation operations based on the characteristic of the problem. But the most popular mutation method is used in binary encoding system. The chromosomes get values 0 or 1 in this system. When applying mutation operation in this system, 0 is changed as 1 and 1 is changed as 0 in the chromosome. But using mutation operation differs according to the types of problem.

Natural Selection: Each chromosome of any generation is evaluated by using an evaluation function. All chromosomes are ordered according to their corresponding evaluation function values. The best chromosomes are transferred the next generation. Some chromosomes among the worst ones are discarded from generations. Then, the new chromosomes instead of discarded chromosomes are placed to the new generation. Fig. 1 presents the population structure in a genetic algorithm.

3. Fuzzy time series

The definition of fuzzy time series was first introduced by Refs. [47,48] and also there were some studies in the literature that used fuzzy time series and genetic algorithm together. For example, Ref. [43] suggested a method to optimize fuzzy time series using genetic algorithms. Ref. [32] suggested an efficient nonlinear time series prediction systems using GA and fuzzy time series. Ref. [8] proposed a forecasting model by using genetic algorithm with fuzzy time series and also, Ref. [40] used genetic algorithm in fuzzification stage which is the first stage of fuzzy

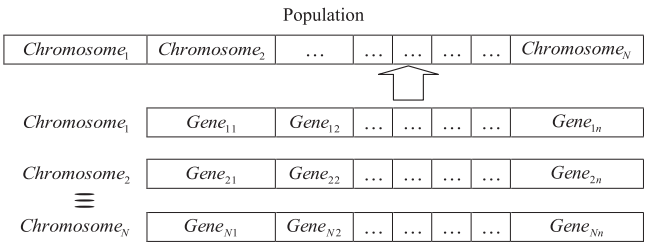

Fig. 1. The population structure in a genetic algorithm.