Modeling and Forecasting of Time Series Data Set using Fuzzy Set

A thesis submitted in partial fulfillment of the requirements for award of the degree of

Master of Technology

in

Computer Science and Applications

Submitted By

Alok Tripathi

Roll No. 601403004

Under the Supervision of:

Dr. H.S. Pannu

Lecturer, CSED

Computer Science and Engineering Department

Thapar University

Patiala-147004

JUNE 2016

CERTIFICATE

I hereby certify that the work which is being presented in the thesis entitled, **"Modeling and Forecasting of Time Series Data Set using Fuzzy Set "**, in partial fulfillment of the requirements for the award of degree of Master of Technology (M.Tech) in *Computer Science And Applications* (CSA) submitted in Computer Science and Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of *Dr. H.S. Pannu* and refers other researcher's work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

Alok Tripathi

Alok Tripathi

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

Dr. H.S. Pannu
Lecturer
CSED
Thapar University

Countersigned by

Dr. Maninder Singh
Head
Computer Science and Engineering Department
Thapar University Patiala

Dr. S. S. Bhatia
Dean (Academic Affairs)
Thapar University
Patiala

i

ACKNOWLEDGEMENT

The successful completion of any task would be incomplete without acknowledging the people who made it possible and whose constant guidance and encouragement secured the success. First of all, I would like to express my gratitude to **Dr. Pritpal Singh** and **Dr. H.S. Pannu,** CSED, Thapar University for giving me an opportunity to work under their guidance and for all their support through the course of this research. This thesis work was enabled and sustained by their vision and ideas. I have been amazingly fortunate to have advisors who gave me the freedom to explore on my own and at the same time the guidance to recover when my steps faltered. Their patience and support helped me overcome many crisis situations and finish this dissertation. They have been a source of inspiration for me.

I would like to thank **Dr. S.S. Bhatia**, Professor and Dean of Academic Affairs, Thapar University, Patiala, for giving provisions of the entire required infrastructure such as computer labs, library facilities, immensely useful for learners to equip themselves with the latest in the field.

I am grateful to **Dr. Maninder Singh**, Head of Department, and **Dr. Sanmeet Kaur,** P.G. Coordinator, M.Tech CSA, entire faculty and staff of Computer Science and Engineering Department and all my friends who devoted their valuable time and helped me in all possible ways towards successful completion of this work. I thank all those who have contributed directly or indirectly to this research work.

I would also like to thank my parents and family for their years of unyielding love and encouragement. They have always wanted the best for me and I admire their determination and sacrifice. Last but not least, I would like to thank almighty for everything.

Date: 15/06/2016

Thapar University, Patiala

Alok Tripathi

(Alok Tripathi)

ii

ABSTRACT

Forecasting future outcomes has always been a subject of keen interest in real world. In the present era various government and private sectors uses time series data which requires forecasting the future outcome within time using past and present information. To do so there is a need of efficient forecasting model which can predict future outcome with high accuracy. Conventional time series model failed to provide accurate forecasting result if time series data consist of linguistic variables like high, low, and medium, etc., because of vagueness. The concept of soft computing was introduced to deals with imprecise and vague data. A new time series model was introduced to overcome the problem of conventional time series model using an important element of soft computing, i.e., *fuzzy set*.

In this thesis, we have introduced a new time series model. The concept of high-order fuzzy time series is used to develop forecasting model. Historical data set is partitioned in to effective interval lengths using an existing Re-Partitioning Discretization (RPD) approach and a novel defuzzification technique based on *weighted-mean* defuzzification approach is introduced to obtain forecasting outcome. Historical temperature data set of Taipei, student enrollment data set of Alabama University and stock price of TAIFEX (Taiwan Future Exchange) is used to evaluate the forecasting accuracy of proposed model.

Statistical parameters like mean, Standard Deviation, Root Mean Square Error (RMSE) and Average Forecasting Error Rate (AFER) are used to calculate the accuracy of proposed model. RMSE and AFER obtained using proposed model over different data set are compared with other existing models. This comparison shows that proposed model outperform over existing models. Forecasting accuracy of proposed model over historical temperature data set of Taipei is better than other existing model, i.e., model based upon Artificial Neural Network and Multi-variant Markov Chain Model. Also forecasting accuracy of proposed model over student enrollment data set of Alabama University is better than Song and Chissom's 4^{th} order model. Forecasting accuracy of proposed model over stock price of TAIFEX is as good as Chen's method.

Keywords: *Fuzzy set, Fuzzy time series, Interval, Defuzzification, Forecasting, Root Mean Square Error (RMSE) Average Forecasting Error Rate (AFER).*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

In this thesis, we have introduced a time series model using an important constituent of soft computing, i.e., fuzzy set. The concept of soft computing was introduced by Professor L.A Zadeh [1] in year 1981.The conventional computing technique also known as hard computing requires well stated precise problem statements. Unlike hard computing soft computing deals with imprecise data, partial truth and approximate reasoning. Hard computing deals with bi-valued logic whereas soft computing deals with multi-valued logic. Unlike hard computing which allows sequential computing soft computing allows parallel computing. Hard computing requires a pre-defined program for modeling whereas soft computing can evolve its own program. Soft computing aimed to develop model to solve real life problem which consist of fuzzy information and which were unable to solve using conventional mathematical model. It also aims to exploit imprecise, partial belief and approximate reasoning so that real life problems can be model easily. Human thinking involves lot of fuzzy knowledge e.g. if Ram say to Shyam that I will visit your home around 09:30 a.m. tomorrow. Word “around” in this statement is fuzzy because to define around there is no single quantitative value. Around 09:30 a.m. may be 09:00 a.m. or 10:00 a.m. and so on. Similarly, if someone says that Ratul is an intelligent student. Here word “intelligent” is fuzzy as a person can be intelligent up to a particular extent. Soft computing can be defined as combination of evolutionary computing e.g. Genetic Algorithm (GA), neural network and fuzzy logic. Evolutionary computing in soft computing used for optimization purpose, neural network used for learning and adapting, and to represent knowledge if-then else rule of fuzzy set is used. Since past two decades various researchers proposed many time series model using the elements of soft computing, i.e., artificial neural network, fuzzy logic, evolutionary algorithm and so on.

The very first time series model using the concept of fuzzy set was introduced by Song and Chissom [2, 3] in year 1993. Conventional time series model cannot be applicable if time series data consist of linguistic values like very low, low, medium, high, very high and so on because of vagueness. In this article we have proposed a time series model using the concept of fuzzy set. In mathematics a set is defined as collection of similar elements. The idea of fuzzy set [1] was given by Professor L.A Zadeh in year 1965. The idea of fuzzy set was based upon the philosophy of partial belief which allows existence of set of elements with varying degree of membership. The concept of membership function was introduced to distinguish between crisp set and fuzzy set. In classical set (crisp set) an element either belongs to the set or not. There is no concept of partial existence of an element in classical set. There are so many sets exist in real life which allows partial existence of elements e.g. set of intelligent students, set of young students, set of basketball players having medium height and so on. In these examples the words "intelligent", "young", "medium" is fuzzy. Membership function maps every elements of fuzzy set to continuous interval [0, 1]. This mapped value is known as degree of membership of element which tells about up to what extent an element lies in the set. In classical set degree of membership value for each element that belongs to set is 1 whereas 0 for the elements not belonging to set. Thus we can say that fuzzy set is super set of classical set. A big drawback of classical set is that it cannot deal with linguistic variables because of imprecision e.g. set of tall men; in this word "tall" is imprecise. Linguistic variables are those variables whose values are in natural language form. For example, if "very low", "low", "medium", "high", "very high" are the values of speed then speed is linguistic variable. Fuzzy set allows varying degree of membership so one can say a person is tall up to some degree of grades.

1.1 Crisp Set

Crisp set also known as classical set is defined as collection of elements having similar property. Let $S = \{e_1, e_2, e_3, \ldots\ldots e_n\}$is a crisp set. In classical set an element either belongs to set or not. Classification of each individual element e_i ($\forall$ i=1 to n) of crisp set is decided by characteristic function f defined as:

$$f: S \rightarrow \{0,1\}$$

$$f(e_i) = \begin{cases} 1, ei \in S \\ 0, ei \notin S \end{cases} \quad (1)$$

$$\forall \ i=1 \text{ to } n.$$

For example S= {set of prime numbers less than 10}

$$S=\{2,3,5,7\}$$

$$f(2)=1 \text{ as } 2 \in S$$

$$f(1)=0 \text{ as } 1 \notin S, \text{ etc. .}$$

Boundary of crisp set is precise (Figure 1.2) means no concept of partial existence of an elements in classical set. But in real life there are so many examples which allow partial belief; to deals with such kind of problems the idea of fuzzy set was introduced.

Example: Consider the speed of 5 motor cars C1 C2,C3,C4 and C5 in kilo meter per hour are 40,45,50,55,60 respectively. Let speed of motor car having 50 kilo meter per hour is considered as high speed. If we talk about set of high speed then according to classical set there is no difference between the speed 45 kilo meter per hour and 55 kilo meter per hour because characteristic function of classical set map value 0 to both speed. Means according to classical set there is no difference between speed 45 kilo meter per hour and 55 kilo meter per hour as both should not be present in the set. But this is not applicable in real life because 55 kilo meter is greater than 50 kilo meter per hour.

1.2 Fuzzy Set

Fuzzy set is a super set of crisp set as crisp set mapped each element of set to extreme point of closed interval [0,1] only where as membership function of fuzzy set maps each elements of fuzzy set to interval [0,1] (see Figure 1.1). In case of crisp set there are only two values (bi-valued) for linguistic variable speed which are (slow=0) and (fast=1) whereas in case of fuzzy set linguistic variable speed can have more than two values (multi-valued) which are “slowest”, “slow”, “fast” and “fastest” (see Figure 1.1).

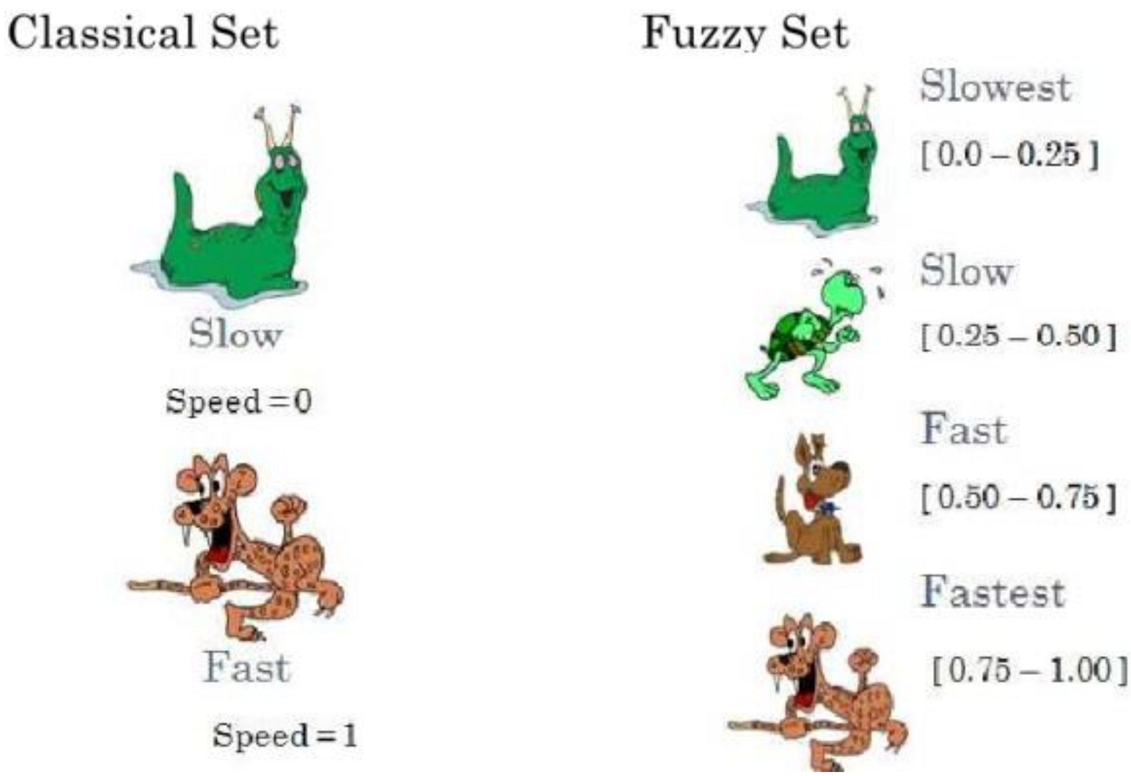

Figure 1.1 Crisp Set vs. Fuzzy Set

Crisp set is characterized by crisp boundary (outline) whereas fuzzy set by fuzzy or imprecise boundary (outline) as shown in Figure 1.2. In case of crisp set element ‘y’ lies completely inside crisp set ‘A’ whereas element ‘x’ lies completely outside crisp set ‘A’, i.e., no concept of partial existence in case of crisp set. But fuzzy set allow partial existence of elements (see Figure 1.2). Element ‘t’ partially exists in fuzzy set ‘B’ .

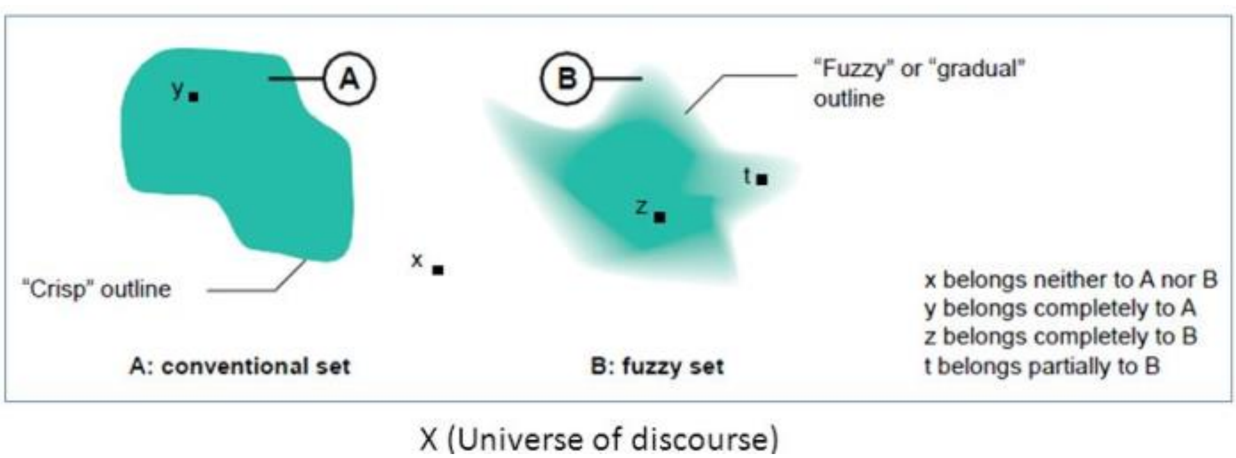

Figure 1.2 Crisp and Fuzzy Boundaries

In general a fuzzy set U can be defined as set of ordered pairs

$$U=\{(u_1,\mu_u(u_1)),(u_2,\mu_u(u_2)),(u_3,\mu_u(u_3)),\ldots\ldots\ldots.,(u_n, \mu_u(u_n))\}. \quad (2)$$

Here element u_i is the member of the fuzzy set U, $\mu_u(u_i)$ is the membership values of element u_i ($\forall i=1$ to n) and μ_u is the membership function of U which maps each elements of set U to closed interval [0,1] and defined as:

$$\mu_u:U\rightarrow[0,1] \quad (3)$$

1.3 Operations on Fuzzy Set

Consider two fuzzy sets $\underline{P}$ and $\underline{Q}$ over the universe of discourse U (here underlined alphabet is used to distinguish fuzzy set with crisp set). Let u be the member element of U, the fuzzy set operations complement, union and intersection for fuzzy sets $\underline{P}$ and $\underline{Q}$ over U is defined as:

Union:

$$\mu_{(A\cup B)}(u)=\max[\mu_A(u),\mu_B(u)] \quad (4)$$

Intersection:

$$\mu_{(A\cap B)}(u)=\min[\mu_A(u),\mu_B(u)] \quad (5)$$

Complement:

$$\mu_A^c(u)=1-\mu_A(u) \quad (6)$$

Where μ_A, μ_B, $\mu_{(A\cup B)}$ and μ_A^c are membership function of set A, B, $(A\cup B)$,$(A\cap B)$ and A^c respectively and defined as follows:

$$\mu_A: A\rightarrow[0,1]$$

$$\mu_{(A\cup B)}: (A\cup B)\rightarrow[0,1]$$

$$\mu_{(A\cap B)}: (A\cap B)\rightarrow[0,1]$$

$$\mu_A^c: \quad A\rightarrow(1- \mu_A)$$

Venn diagram of AUB, A∩B and A^c are represented in Figure 1.3, 1.4, 1.5 respectively. Consider two fuzzy sets $\underline{P}$ and $\underline{Q}$ defined over universe of discourse U={2,3,4,5} using notation (A) as follows:

$\underline{P}$ = {(2,1),(3,0.5),(4,0.3),(5,0.2)} and $\underline{Q}$ = {(2,0.5),(3,0.7),(4,0.2),(5,0.4)}

Fuzzy set operations Complement, Union, and Intersection over above two fuzzy set is discussed in the following sub-sections:

1.3.1 Union

$\underline{A} \cup \underline{B} = \{(u, \max[\mu_A(u), \mu_B(u)])\}$; $\forall$ u ∈ U

={(2,1),(3,0.7),(4,0.3),(5,0.4)}

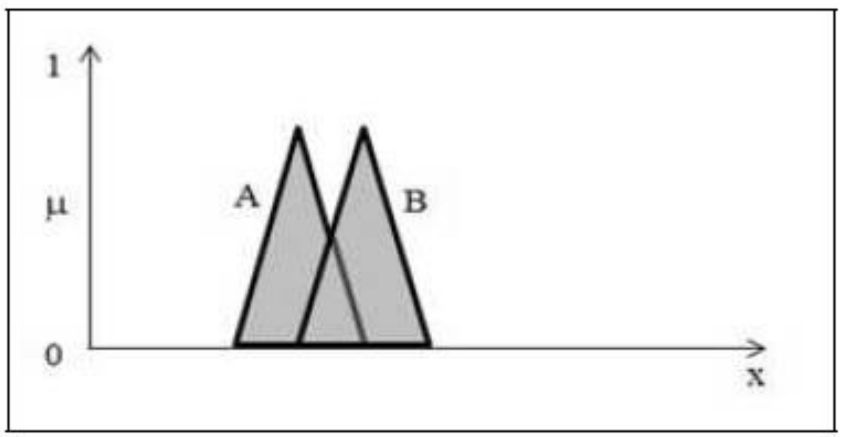

Figure 1.3 Fuzzy Union Operations

1.3.2 Intersection

$\underline{A} \cup \underline{B} = \{(u, \min[\mu_A(u), \mu_B(u)])\}$; $\forall$ u ∈ U

={(2,0.5),(3,0.5),(4,0.2),(5,0.2)}

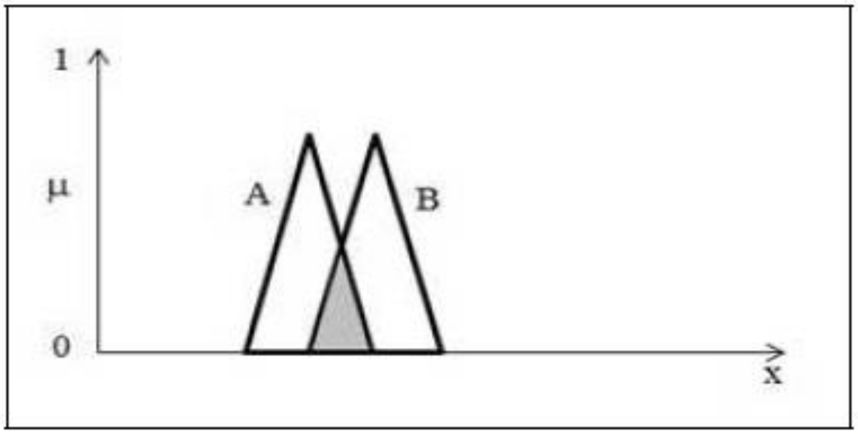

Figure 1.4 Fuzzy Intersection Operations

1.3.3 Complement

$A^c = \{(u, 1-\mu_A(u))\}\ ;\ \forall\ u \in U$

$= \{(2,0),(3,0.5),(4,0.7),(5,0.8)$

$B^c = \{(u, 1-\mu_B(u))\}\ ;\ \forall\ u \in U$

$= \{(2,0.5),(3,0.3),(4,0.8),(5,0.6)$

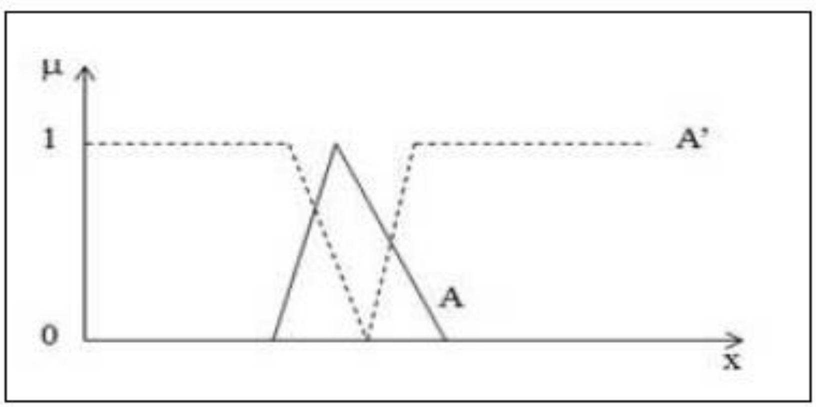

Figure 1.5 Fuzzy Complement Operations

1.4 Fuzzification

Fuzzification is defined as the process of converting crisp value into fuzzy value (linguistic value) with help of membership function as shown in Figure 1.6.

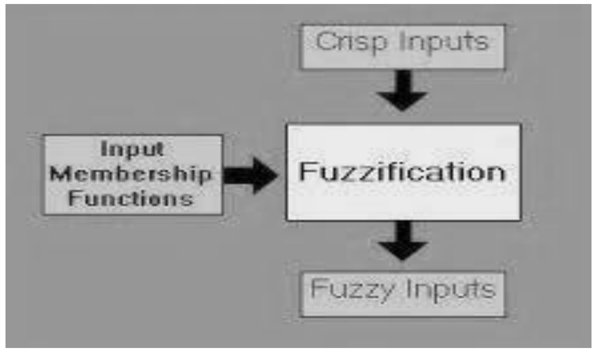

Figure 1.6 Process Outline of Fuzzification

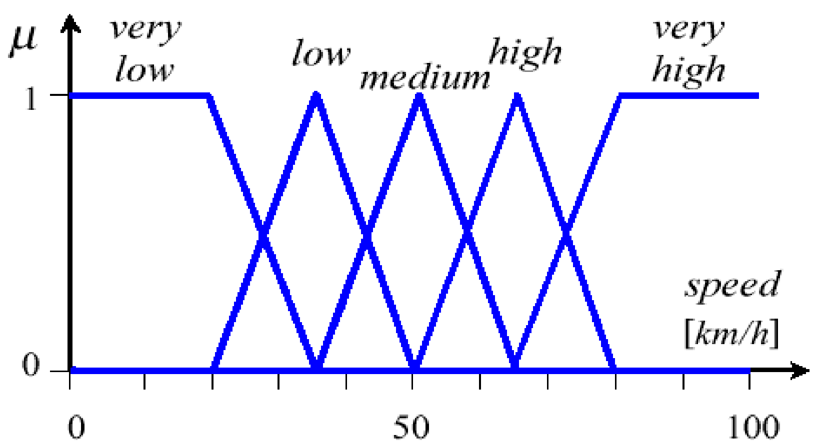

Figure 1.7 Fuzzification Example

For example consider the speed a bike taken from closed interval [0,100] in kilo meter per hour (km/h). This crisp data set of speed is converted into fuzzy (linguistic) values like very low, low, medium, high and very high as shown in Figure 1.7.

Various methods are available to assign values to linguistic variables which are given below:

i. Intuition method
ii. Inference
iii. Rank ordering
iv. Neural networks
v. Genetic Algorithms (GA)
vi. Inductive method

vii. Reasoning approach

We have presented a fuzzification technique in this thesis which is based on fuzzy inference (if-then-else rule) and discussed further.

1.5 Defuzzification

Defuzzification is defined as process of converting fuzzy value in to crisp value. There are various techniques of defuzzification which are given below:

i. Centroid method
ii. Max-membership principle
iii. Weighted-mean method
iv. Mean-max membership principle
v. Center of sum method
vi. Center of largest area
vii. First of maxima

In this thesis we have presented a new defuzzification rule for forecasting which is based on weighted-mean method. Working principle and example of weighted-mean method are shown in Figure 1.8 and 1.9 respectively.

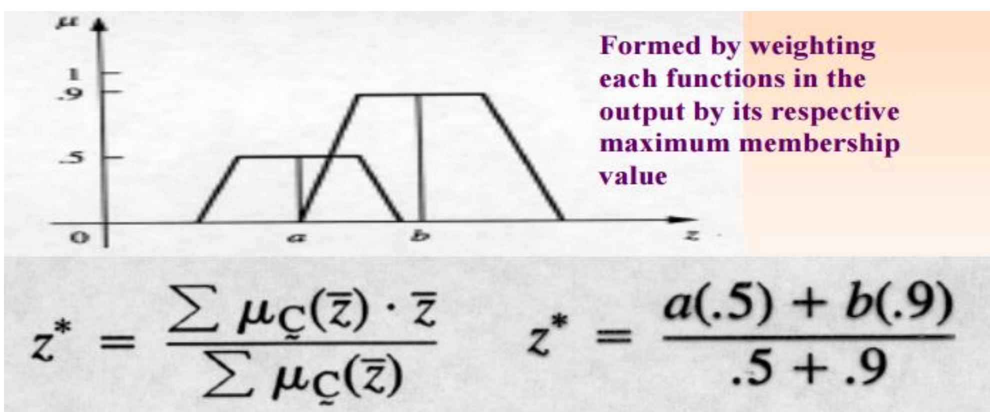

Figure 1.8 Working Principle of Weighted-Mean Defuzzification Technique

Consider the example presented in Figure 1.9. In this the maximum degree of membership for fuzzified value “Slow” and “Fast” occurs at speed 25 mph and 75 mph respectively. The defuzzified value for speed is 20% slow and 70% fast is calculated using weighted-mean defuzzification technique.

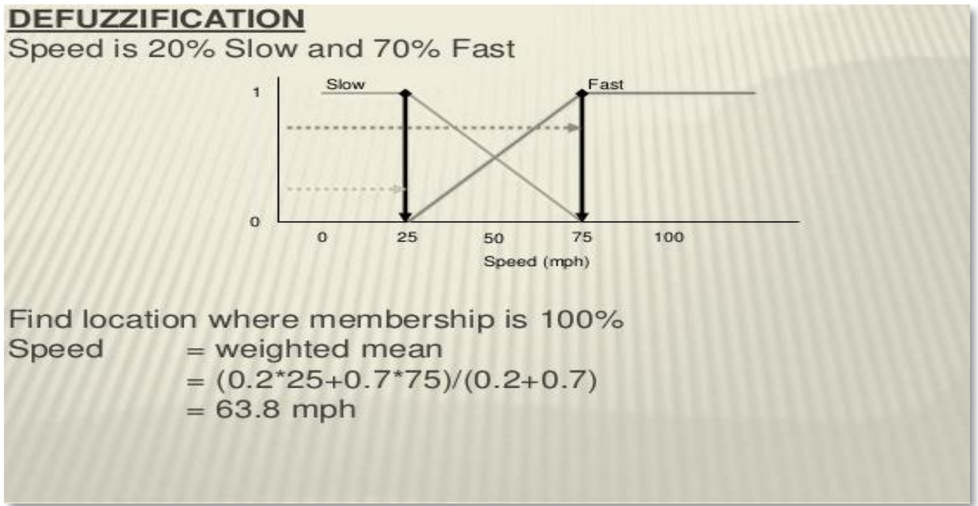

Figure 1.9 Weighted-Mean Defuzzification Technique Example

CHAPTER 2
LITERATURE SURVEY

This Chapter is further divided into two sub-chapters. In sub-chapter one work done by various researchers in field of fuzzy time series model is discussed. In second sub-chapter the role of effective length of intervals and defuzzification technique in time series forecasting is discussed.

2.1 Related Work

After the introduction of fuzzy time series model in year 1993 by Song and Chissom [2,3] many researchers introduced various time series model based on the concept of fuzzy set. Song and Chissom introduced a fuzzy time series model using the max-min composition operation and validated proposed model on historical Alabama University student enrollment data set. Kuo *et al.* [4] proposed a new hybrid fuzzy time series model using the concept of voting scheme combined with Particle Swarm Optimization (PSO) technique and to evaluate the performance of this model the historical student enrollment of Alabama University is used. Chen *et al.* [5] proposed a new time series model using the concept of high-order fuzzy logical relationship combined with Genetic Algorithm for student enrollment data set of Alabama University.

Chen and Hsu [6] proposed a first-order time variant forecasting model using max-min composition technique for efficient forecasting of Alabama university student enrollment data set. Li and Cheng [7] introduced a new deterministic fuzzy time series model using concept of longest subsequence to forecast student enrollment of Alabama University. Singh [8] introduced a new fuzzy time series model using the concept of difference between parameter (like year, Day and so on) for enrollment forecasting and this model is also validating on crop (wheat) production. Tsaur *et al.* [9] introduced a new fuzzy time series prediction model for Alabama University student enrollment data set using the concept of entropy combined with time-invariant relation matrix. Saxena and Easo [10] introduced a new fuzzy time series model for advance prediction of student enrollment of Alabama University using percentage change in universe of discourse combined with mean based partitioning and centroid method of defuzzification. Singh [11] proposed a new computational algorithm for fuzzy time series model combined with concept of high-order fuzzy logic and validated this model on Alabama university student enrollment data set. Huang *et al.* [12] introduced a hybrid time series model based on fuzzy concept combined with PSO technique for advance prediction of enrollment. Student enrollment of Alabama is used to validate proposed model. Liu *et al.* [13] proposed a new time series model based on the concept of fuzzy logic combined with the concept of approximate reasoning and rough set. For performance evaluation historical student enrollment of Alabama University is used.

Gangwar and Kumar [14] presented a new time series predicting model based upon high-order fuzzy logical relationship combined with a new partitioning algorithm to partition universe of discourse. Performance of model evaluated by applying proposed model on student enrollment of Alabama University and market price of State Bank of India (SBI). Egrioglu *et al.* [15] proposed a new hybrid fuzzy time series model combined with Artificial Neural Network (ANN) and c-means clustering. The student enrollment of Alabama University is predicted using this model and better result found. Liu [16] introduced a new time series model based on fuzzy set concept using trapezoidal fuzzy number to overcome the problem of time series model based on single- valued fuzzy number. Singh [17] proposed a new time series model based on high-order fuzzy set combined with a linear order computational method for student enrollment data set of Alabama University. This model was also validated on historical data set of crop (Lahi) production.

Cheng *et al.* [18] introduced a new time series model based on concept of fuzzy set combining with rough-set and probabilistic approach. Historical time series exchange price of TAIFEX is taken to validate the accuracy of model. Zarandi *et al.* [19] introduced a time series model for temperature data set of Taipei using fuzzy logical relationship combined with double interval division and centroid defuzzification technique. Forecasting accuracy of this model is also validated on historical data set TAIFEX (1996). Wang and Chen [20] proposed a time series model for advance prediction of temperature using automatic clustering technique combined with two-factor and high-order fuzzy time series. Cross validation of forecasting accuracy of the model was performed using TAIFEX data set. In year 2008 Li *et al.* [21] introduced a new deterministic time series model for temperature prediction using fuzzy concept combined with Fuzzy C-Mean (FCM) clustering. Performance evaluation of this model validated using historical temperature data set of Taipei. Singh and Borah [22] introduced a new hybrid two-factor (viz temperature and cloud density) time series model based on fuzzy logic concept combining with ANN approach. Historical temperature data set of Taipei along with cloud density of Taipei is taken to validate the accuracy of presented model. Vamitha *et al.* [23] presented a time series predicting methodology for advance prediction of temperature based on fuzzy concept combined with multi-variant Markov chain model and performance of model is evaluated using historical temperature data set of Taipei. Chen and Hwang [24] proposed two different algorithms based on 2-factor time-variant time series model using fuzzy concept for temperature forecasting. Historical temperature data set of Taipei is used to evaluate the performance of this model.

K Huarng [25] proposed a hybrid time series predicting model combining heuristic knowledge. This work was the extension of Chen's model of fuzzy time series using arithmetic computations. Performance analysis of proposed model is evaluated by applying model on historical data set of TAIFEX. Lee *et al.* [26] introduced a new hybrid forecasting model for advance prediction of temperature and future exchange of TAIFEX using GA combined with Simulated Annealing (SA). Kuo *et al.* [27] proposed a new hybrid time series model to forecast future exchange of TAIFEX using concept of fuzzy logic combined with PSO technique. Singh and Borah [28] introduced a new hybrid time series model based on the type-2 fuzzy set concept combining PSO technique. To evaluate the forecasting accuracy of proposed model daily stock index price of State SBI is used. Daily stock index price of Google is also taken for cross checking of forecasting accuracy of proposed model. P Singh [29] proposed a new hybrid time series model based on the concept of high-order fuzzy logical relationship concept combined with fuzzy-neuro-entropy integration approach. Daily stock index price of State Bank of India (SBI) is used to check the forecasting accuracy of proposed model.

Singh and Borah [30] developed a hybrid time series model based on fuzzy set concept combining feed-forward neural network to forecast Indian Summer Mansoon Rainfall (ISMR). Egrioglu *et al.* [31] presented a new hybrid time series model to forecast seasonal time series data based on SARIMA (Seasonal Autoregressive Integrated Moving Average) model combined with the concept of high-order bi-variant fuzzy time series. Yu and Huarng [32] presented a time series forecasting model to forecast future exchange of TAIFEX using the concept of type-2 fuzzy set. The forecasting accuracy of this model outperform over type-1 fuzzy time series model. Chen *et al.* [49] introduced a new hybrid forecasting model for stock index data set based on high-order fuzzy time series concept combined with multi-period adapting model.

2.2 Importance of Effective Interval Length and Defuzzification

There are two key parameters upon which the accuracy of fuzzy time series model much depends. These are effective interval length used to divide universe of discourse in to sub-intervals and defuzzification technique used to convert forecasted fuzzy value in to crisp value. In year 2000 K. Huarng [25] proved that effective interval length improve forecasting accuracy but how interval length affect the accuracy of time series model was not answered. The Huarng's study of length of intervals was extension of Chen's work that was extension of fuzzy time series model of Song and Chissom. Huarng took 6,10,12,14-time series values defined over some period of times to study the length of intervals. He considered the range of universe of discourse S as S=[3,13]. He first selected interval-length 5 for which time series data set is divided in to two intervals, i.e., I_1=[3,8] and I_2 =[8,13]. The root mean square error according to Chen's model is 3.16. Now if length of interval is 2 then time series data set is divided in to five intervals, i.e., I_1=[3,5] , I_2=[5,7], I_3=[7,9], I_4=[9,11] and I_5=[11,13]. The root mean square error for these intervals using Chen's model was found 2.12. Therefore, using these results of Chen's model Huarng proposed their work on effective length of intervals.

The study of Huarng upon length of intervals was divided in to two parts, i.e., distribution-based length and average-based length. Historical student enrollment data set of Alabama University and future exchange data set of TAIFEX are used by Huarng to validate proposed work. After this work various researchers introduced various techniques for creating effective length of intervals and defuzzification technique to compute forecasting values.

Singh and Borah [34] presented a new data Discretization technique to partition universe of discourse in to effective length of intervals. A new defuzzification technique is used as forecasting rule by assigning the weights to fuzzy logical relationships. Historical student enrollment of Alabama University is used to check the forecasting accuracy of model. Egrioglu *et al.* [35] introduced a MATLAB function based on parabolic interpolation and golden-section search to create effective length of intervals and a new defuzzification technique based on centroid method to convert fuzzy forecasting value in to crisp value. Leung *et al.* [36] uses Ant Colony Optimization (ACO) for tuning original intervals to create effective length of intervals. For defuzzification an auto-regression high-order fuzzy time series model was introduced. Singh and Borah [37] introduced an ANN based algorithm to generate effective length of intervals for two-factor time series data set. For defuzzification a high-order fuzzy time series forecasting rule was used.

Avazbeigi *et al.* [38] introduced a new time series model to forecast auto-industry production based upon the concept of fuzzy set. Tabu search optimization technique is used to find effective length of intervals in the proposed model. A new forecasting rule based on the concept of N-factor high-order fuzzy logical relationship is introduced for defuzzification. Bahrepour *et al.* [39] introduced an adaptive order selection technique to partition the historical data set in to effective interval length. This automatic order selection technique partitions the historical data set in to unequal length of intervals using the strategy of fast self-organizing. The concept of high-order fuzzy logical relationship is used for defuzzification. Historical FOREX market exchange price was taken to validate the forecasting accuracy of proposed model. Yu and Huarng [40] presented a new ratio-based method to partition universe of discourse in to sub-intervals and found that intervals generated using this technique is better than other existing techniques. To defuzzify the fuzzy forecasted value a forecasting rule based on fuzzy logical relationship is used. Historical future price exchange of TAIFEX is used to validate the performance evaluation of presented model.

Yolcu *et al.* [41] introduced a new technique to generate effective length of intervals using single-variable constrained optimization approach. Well known centroid method of defuzzification is used to obtain forecasting result. Alabama University student enrollment data set is used for forecasting and forecasting accuracy is found better. Davari *et al.* [42] presented a new approach to generate effective length of intervals by modifying PSO technique. PSO is used to adjust the length of intervals. Li and Chen [43] presented a recursive natural partitioning technique to partition the universe of discourse in to effective interval length and a heuristic rule is used as defuzzification process to obtain forecasting value. Chen and Kao [44] introduced a new hybrid fuzzy time series model combining Support Vector Machine (SVM) with PSO. PSO is used to generate optimal length of intervals. SVM is used to classify the data. Future price exchange of TAIFEX is forecasted using proposed model and forecasting accuracy of model found outperform over other existing forecasting model. Singh and Borah [45] presented a new Re-partitioning Discretization (RPD) approach to partition universe of discourse in to optimal length of intervals. The concept of ANN is used as defuzzification technique to obtain forecasted results. Historical temperature data set of Taipei is used for the validation of forecasting accuracy of proposed model.

CHAPTER 3
PROBLEM STATEMENT

3.1 Introduction

Time series forecasting is defined as predicting future outcome in advance based on past and present information which are available. A time series data is defined as set of data points defined over a specific period of time. Time series data consist of two fields one is time and another is value corresponding to time e.g. Temperature data set of 30 days of a particular month, Student enrollment of a University in a particular year, stock price of stock exchange, rain fall in centimeter over a specific time period. Many areas need a forecasting tool to help people for making decisions in advance for better future plan such as stock markets, student enrollment of a university, temperature forecasting, predicting air pollution, forecasting rainfall and so on. Therefore developing an efficient time series model is a challenging task now a day. Biggest problem arise in conventional time series forecasting is that it cannot provide accurate result if time series data set consist of linguistic variable. To overcome this problem the concept of fuzzy time series was introduced in year 1993 by Song and Chissom [2, 3]. The performance accuracy of fuzzy time series model much depends on two key topics which are effective interval length of universe of discourse and defuzzification technique used to obtain forecasting values.

3.2 Thesis Objective

The main objectives of this thesis are to develop an efficient time series model which can forecast future outcome accurately. We have also provided brief introduction of works done by various researchers in the area of time series forecasting based on fuzzy set along with the importance of effective interval length and defuzzification technique in forecasting accuracy. We have selected the concept of fuzzy set to develop forecasting model to overcome the problem of conventional time series models. In this thesis we have presented a new model of forecasting based on high-order fuzzy time series concept. To partition the universe of discourse a RPD approach is used and to find final forecasting outcome a new forecasting rule is introduced based on weighted-mean defuzzification technique. The proposed model is validated using some historical data sets and the forecasting result was compared with some existing model and found better.

CHAPTER 4
PROPOSED WORK

4.1 Preliminaries

Some standard definitions which are used in this thesis are explained in the following sub-sections:

4.1.1 Fuzzy Set

Let X is universe of discourse. A real function μ_Y: X→[0,1] is known as membership function of Y and defines the fuzzy set Y of X. If X is finite and discrete then:

$$Y= \{\mu_Y(x_1)/x_1 + \mu_Y(x_2)/x_2 + \mu_Y(x_3)/x_3 + \ldots\ldots\ldots\}=\sum_i \mu_Y(x_i)/x_i,\ x_i \in X \quad (7)$$

Where "+" indicates fuzzy addition (union) operation and "/" used for separator and not for algebraic division.

If X is infinite and continuous then:

$$Y=\{\int \mu_Y(x_i)/x_i\},x_i \in X \quad (8)$$

Where symbol "∫" indicates fuzzy union and "/" used for separator.

4.1.2 Fuzzy Time Series

Let R(t) (t=0,1,2,…..), is a subset of real numbers and is the universe of discourse on which fuzzy set $y_i(t)$(i=0,1,2,…..)is defined.

If Y(t) =$y_1(t) \cup y_2(t) \cup y_3(t) \cup$………. , then Y(t) is called as fuzzy time series defined over R(t)(t=0,1,2,…).(9)

4.1.3 Fuzzy Time-variant and Time-invariant Series

Suppose Y(t) is a fuzzy time series and let S(t,t-1) is a first order fuzzy logical relationship model of Y(t) which represent relationship between time series value y(t) and y(t-1) (t=0,1,2,….) . If S(t,t-1)=S(t-1,t-2) and Y(t) has finite elements only, then Y(t) is said to be time-invariant fuzzy time series otherwise it is said to be time-variant fuzzy time series.

4.1.4 Fuzzy Logical Relationship (FLR)

Let $Y(t-1)=L_i$ and $Y(t)=L_k$, a fuzzy logical relationship between Y(t) and Y(t-1) is denoted as

$$L_i \rightarrow L_k \quad (10)$$

Here L_i is antecedent of fuzzy logical relationship and L_k the consequent.

4.1.5 High-order Fuzzy Time Series

Assume Y(t) is a fuzzy time series. If Y(t) is depend on multiple fuzzy time series i.e., Y(t-1), Y(t-2), Y(t-3)…… Y(t-n), then fuzzy logical relationship is defined as:

$$Y(t-n),\ldots\ldots,Y(t-3),Y(t-2),Y(t-1) \rightarrow Y(t) \quad (11)$$

and it is called as high-order fuzzy time series.

4.1.6 Fuzzy Logical Relationship Group (FLRG)

Chen [49] suggested that fuzzy logical relationship having same premises can be combined together. Consider the following fuzzy logical relationships:

$$L_i \rightarrow L_{k1}$$

$$L_i \rightarrow L_{k2} \ldots\ldots\ldots\ldots$$

Then these can be grouped together as following:

$$L_i \rightarrow L_{k1},L_{k2},\ldots\ldots.. \quad (12)$$

4.2 Overview of Data Sets

In this sub-chapter the historical time series data sets on which proposed model is validated are discussed. In Table 4.1 historical daily average temperature data set of Taipei from June 1996 to September 1996 [45] is presented.

Table 4.1 Historical Daily Average Temperature Data Set of Taipei

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

Historical student enrollment data set of Alabama University from year 1971 to 1992 [2, 3] is presented in Table 4.2.

Table 4.2 Historical Student Enrollment Data Set of Alabama University

Year	*Enrollments*	*Year*	*Enrollments*
1971	13055	1982	15433
1972	13563	1983	15497
1973	13867	1984	15145

1974	14696	1985	15163
1975	15460	1986	15984
1976	15311	1987	16859
1977	15603	1988	18150
1978	15861	1989	18970
1979	16807	1990	19328
1980	16919	1991	19337
1981	16388	1992	18876

In Table 4.3 stock price of TAIFEX of Year 1998 [25] is presented.

Table 4.3 Historical TAIFEX Data Set of Year 1998

Date	*Stock Price*	*Date*	*Stock Price*
8/3/1998	7552	9/2/1998	6430
8/4/1998	7560	9/3/1998	6200
8/5/1998	7487	9/4/1998	6403.2
8/6/1998	7462	9/5/1998	6697.5
8/7/1998	7515	9/7/1998	6722.3
8/10/1998	7365	9/8/1998	6859.4
8/11/1998	7360	9/9/1998	6769.6
8/12/1998	7330	9/10/1998	6709.8
8/13/1998	7291	9/11/1998	6726.5
8/14/1998	7320	9/14/1998	6774.6
8/15/1998	7300	9/15/1998	6762
8/17/1998	7219	9/16/1998	6952.8
8/18/1998	7220	9/17/1998	6906

8/19/1998	7285	9/18/1998	6842
8/20/1998	7274	9/19/1998	7039
8/21/1998	7225	9/21/1998	6861
8/24/1998	6955	9/22/1998	6926
8/25/1998	6949	9/23/1998	6852
8/26/1998	6790	9/24/1998	6890
8/27/1998	6835	9/25/1998	6871
8/28/1998	6695	9/28/1998	6840
8/29/1998	6728	9/29/1998	6806
8/31/1998	6566	9/30/1998	6787
9/1/1998	6409		

4.3 Proposed Algorithms

4.3.1 Partition Algorithm

The proposed partitioning algorithm based on existing RPD [45] approach is explained as follows:

i. Let $U=[Min_u, Max_u]$ is the universe of discourse taken into consideration.

Consider the temperature data set of June month as universe of discourse shown in Table 1.

ii. Identify the maximum and minimum value of universe of discourse as Max_u, Min_u respectively.

$$Min_u=26.1, Max_u=30.9.$$

iii. Identify the mid-value (U_{mid}) of universe of discourse U as:

$$U_{mid}=(Min_u+Max_u)/2 \tag{13}$$

$= (26.1+30.9)/2$

$= 28.5.$

iv. Divide universe of discourse into two sub-sets as:

$X= \{x|\ x\leq U_{mid}\},$

$X=\{26.1,27.1,27.4,27.5,27.6,27.7,27.8,27.8,28.4,28.5\}$

and $Y= \{x|\ x> U_{mid}\}$

$Y=\{28.7,28.7,28.8,28.8,29,29,29,29.3,29.4,29.4,29.5,29.5,29.7,30,30.2,30.2,30.3,30.5,30.8,30.9\}$

v. Define sub-boundaries for X and Y as:

$U_X = [Min_x, Max_x]$

$U_Y = [Min_y, Max_y]$

Where,

Min_x, Max_x, are the minimum and maximum values of the sub-set X respectively and Min_y and Max_y are the minimum and maximum values of the sub-set Y respectively.

$Min_x=26.1$, $Max_x=28.5$;

$Min_y=28.7$, $Max_y=30.9$.

Therefore,

$U_X= [26.1, 28.5]$

$U_Y= [28.7, 30.9]$

vi. Define the decision parameters for X and Y as:

$$DP_X=(Max_x-Min_x)/N_X, \quad (14)$$

$= (28.5-26.1)/10$

$$= 0.24$$

$$DP_Y=(Max_y-Min_y)/N_Y, \quad (15)$$

$$= (30.9-28.7)/20$$

$$=0.11$$

Here DP_X and DP_Y are the decision parameters for X and Y respectively and N_X and N_Y denotes the number of elements of X and Y respectively.

vii. Sub-division of the sub-boundaries U_X and U_Y into sub-intervals

$$P_k= [L(k),U(k)],\ k=1, 2 ,3 ,\ldots\ldots\ldots.;\ U(k) \leq Max_x ,P_k \in U_X,$$

Where,

$$\left.\begin{aligned} L(k) &= Min_x+(k-1)*DP_X, \\ U(k) &= Min_x+k*DP_X. \end{aligned}\right\} \quad (16)$$

$$Q_k= [M(k),V(k)],\ k=1, 2 ,3 ,\ldots\ldots\ldots.;\ V(k) \leq Max_y ,Q_k \in U_Y$$

Where

$$\left.\begin{aligned} M(k) &= Min_y+(k-1)*DP_Y , \\ V(k) &= Min_y+k*DP_Y. \end{aligned}\right\} \quad (17)$$

Find out intervals corresponding to each element and assign each element to their respective intervals.

- Intervals with corresponding elements, rank and center-point for June month of historical temperature data set using proposed partitioning and rank finding algorithm is presented in Table 4.4.

- Similarly we have identified intervals with corresponding elements, rank and center-point for July, August, September months of historical temperature data sets and are presented in Table 4.10, 4.11, 4.12 of Appendix 4.1 respectively.

Table 4.4 Intervals with Corresponding Elements, Rank and Center-point for June Month of Temperature Data Set

Interval	**Corresponding elements**	**Center-point**	**Rank**
U_X Interval			
[26.1,26.3]	<26.1>	26.22	1
[27.06,27.3]	<27.1>	27.18	1
[27.3,27.54]	<27.4,27.5>	27.42	2
[27.54,27.78]	<27.6,27.7>	27.66	2
[27.78,28.02]	<27.8,27.8>	27.9	2
[28.26,28.5]	<28.4,28.5>	28.38	2
U_Y Interval			
[28.7,28.81]	<28.7,28.7,28.8,28.8>	28.76	4
[28.92,29.03]	<29,29,29>	28.98	3
[29.25,29.36]	<29.3>	29.31	1
[29.36,29.47]	<29.4,29.4>	29.42	2
[29.47,29.58]	<29.5,29.5>	29.53	2
[29.69,29.8]	<29.7>	29.75	1
[29.91,30.02]	<30>	29.97	1
[30.13,30.24]	<30.2,30.2>	30.19	2
[30.24,30.35]	<30.3>	30.3	1
[30.46,30.57]	<30.5>	30.52	1
[30.79,30.9]	<30.8,30.9>	30.85	2

- Intervals with corresponding elements, rank and center-point for historical student enrollment data set of Alabama University is presented in Table 4.5.

Table 4.5 Intervals with Corresponding Elements, Rank and Center-point of Student Enrollment Data Set

Interval	**Corresponding elements**	**Center-point**	**Rank**
U_X Interval			
[13055,13293]	<13055>	13174	1
[13531,13769]	<13563>	13650	1
[13769,14007]	<13867>	13888	1
[14483,14721]	<14696>	14602	1
[14959,15197]	<15145,15163>	15078	2
[15197,15435]	<15311,15433>	15316	2
[15435,15673]	<15460,15497,15603>	15554	3
[15673,15911]	<15861>	15792	1
[15911,16149]	<15984>	16030	1
U_Y Interval			
[16387,16625]	<16388>	16506	1
[16807,17123]	<16807,16859,16919>	16965	3
[18071,18387]	<18150>	18229	1
[18703,19019]	<18876,18970>	18861	2
[19019,19335]	<19328>	19177	1
[19335,19337]	<19337>	19336	1

- Intervals with corresponding elements, rank and center-point for TAIFEX data set is presented in Table 4.6.

Table 4.6 Intervals with Corresponding Elements, Rank and Center-point of TAIFEX Data Set

Interval	**Corresponding elements**	**center-point**	**Rank**
U_X Interval			
[6200,6227.95]	<6200>	6213.98	1
[6395.65,6423.6]	<6403.2,6409>	6409.63	2
[6423.6,6451.55]	<6430>	6437.58	1
[6557.5,6585]	<6566>	6571.25	1
[6667.5,6695]	<6695>	6681.25	1
[6695,6722.5]	<6697.5,6709.8,6722.3>	6708.75	3
[6722.5,6750]	<6726.5,6728>	6736.25	2
[6750,6777.5]	<6762,6769.6,6774.6>	6763.75	3
[6777.5,6805]	<6787,6790>	6791.25	2
[6805,6832.5]	<6806>	6818.75	1
[6832.5,6860]	<6835,6840,6842,6852,6859.4>	6846.25	5
[6860,6871]	<6861,6871>	6865.5	2
U_Y Interval			
[6890,6919.13]	<6890,6906>	6904.57	2
[6919.13,6948.26]	<6926>	6933.7	1
[6948.26,6977.39]	<6949,6952.8,6955>	6962.83	3
[7035.65,7064.78]	<7039>	7050.22	1
[7210.43,7239.56]	<7219,7220,7225>	7225	3
[7268.69,7297.82]	<7274,7285,7291>	7283.26	3

[7297.82,7326.95]	<7300,7320>	7312.39	2
[7326.95,7356.08]	<7330>	7341.52	1
[7356.08,7385.21]	<7360,7365>	7370.65	2
[7443.47,7472.6]	<7462>	7458.04	1
[7472.6,7501.73]	<7487>	7487.17	1
[7501.73,7530.86]	<7515>	7516.3	1
[7530.86,7560]	<7552,7560>	7545.43	2

4.3.2 Forecasting Rule

The proposed forecasting rule is explained as follows:

a. Partition the historical data set into varying length of intervals. Discard those intervals which do not contain any element from universe of discourse. Let historical data set is divided into 'n' valid varying length of intervals, i.e., i_1, i_2, i_3, …… ,i_n.

b. Assign linguistic term to each valid interval. For 'n' varying length of intervals, i.e., i_1, i_2, i_3, …… ,i_n define 'n' linguistic term I_1, I_2,…….. , I_n ; that can be denoted by fuzzy set as follows:

$$\left.\begin{aligned}
I_1&=1/i_1+0.5/i_2+0/i_3+0/i_4+0/i_5+\ldots\ldots\ldots+0/i_{(n-2)}+0/i_{(n-1)}+0/i_n,\\
I_2&=0.5/i_1+1/i_2+0.5/i_3+0/i_4+0/i_5+\ldots\ldots\ldots+0/i_{(n-2)}+0/i_{(n-1)}+0/i_n,\\
I_3&=0/i_1+0.5/i_2+1/i_3+0.5/i_4+0/i_5+\ldots\ldots\ldots+0/i_{(n-2)}+0/i_{(n-1)}+0/i_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
I_n&=1/i_1+0.5/i_2+0/i_3+0/i_4+0/i_5+\ldots\ldots\ldots+0/i_{(n-2)}+0.5/i_{(n-1)}+1/i_n.
\end{aligned}\right\} \quad (18)$$

Get the degree of membership for every day's temperature value corresponding to every I_i. Where highest degree of membership for fuzzy set I_i occurs at interval i_i and $1 \le i \le n$ (see equation 18).

- Apply Fuzzification on time series data using fuzzy inference method. If a particular day's temperature value assigned to interval i_i, then the corresponding fuzzified temperature value is I_i.
- Determine the high-order fuzzy logical relationships among fuzzified daily temperature value using equation 11.
- Apply defuzzification on fuzzified temperature data set. We have introduced a new high-order fuzzy time series model based on high-order fuzzy logical relationship combined with weighted-mean defuzzification technique. Therefore to discuss the defuzzification method, we have considered n^{th} order fuzzy logical relationships, here n≥5. Following steps are involved in defuzzification method:
    a. Get the n^{th} order fuzzy logical relationship corresponding to forecasting day Y(t) as follows:

$$I_{tn}, I_{t(n-1)}, \ldots\ldots\ldots\ldots, I_{t2}, I_{t1} \rightarrow ?(t); \quad (19)$$

Here t denotes a forecasting day; 'n' represent the order of fuzzy logical relationships (n≥5) and I_{tn}, $I_{t(n-1)}$,…………., I_{t1} are the fuzzified values for days Y(t-1), Y(t-2)…….,Y(t-n) respectively.

   b. Obtain the intervals corresponding to highest membership values of the fuzzy sets I_{tn}, $I_{t(n-1)}$,……..,I_{t1} and assume these intervals are i_n, $i_{(n-1)}$,…….,i_1 respectively.

   c. Obtain the center points of each interval. Let m_1, m_2,……,m_n are the center points of intervals i_1, i_2,…… i_n, respectively.

   d. Obtain the rank of each interval which represents the number of elements present in the corresponding interval. Let r_1, r_2,………,r_n are the rank of intervals i_1, i_2,…… i_n respectively.

   e. Use following forecasting formula to obtain the desired output '?'

$$F(t) = \frac{[(m_1 * r_1) + (m_2 * r_2) + \ldots\ldots + (m_n * r_n)]}{(r_1 + r_2 + \ldots\ldots\ldots + r_n)} = \frac{\sum m_i * r_i}{\sum r_i} \quad (20)$$

Where F(t) is predicted temperature value of day '?(t)'.

4.3.3 Algorithm for Finding Rank of Interval

Start

a. Let U[m] is the array of historical data set where 'm' is number of elements available.
b. Suppose historical data set U[m] is divided into 'n' effective interval length as: $I_i=[L_{bi},U_{bi}]$; (∀ i=0 to n-1) , here L_{bi} and U_{bi} are the lower and upper bound of interval I_i respectively.
c. Let r[n] is the array consisting of rank of interval I_i; ∀ i=0to (n-1)
d. for i=0 to (n-1)

```
    for k=0 to (m-1)
        if (Lbi<U[k]<=Ubi)
            r[i]++;
        else
            printf(" element not present in this interval");
```

End

4.4 Implementation

- RPD and algorithm to find rank are implemented using Java Language
- Formula of Forecasting rule is implemented on Excel sheet.
- Empirical analysis of proposed model using statistical parameters like mean, standard deviation AFER and RMSE is implemented using Excel sheet.

4.5 Architecture of Proposed Model

Architecture of proposed model is shown below in Figure 4.1.

The basic steps which are involved in the proposed model are given below:

- Take one-factor time series data set as Input.
- Partition the input time series data set into effective interval length using partitioning algorithm of proposed model.
- Define linguistic terms corresponding to each interval and find out center-point of each interval.

- Find out rank of each interval using proposed rank finding algorithm.
- Fuzzify time series data set, i.e., convert crisp value in to fuzzy value using equation 18 of proposed forecasting rule.
- Establish high-order Fuzzy Logical Relationship (FLR) among fuzzified time series data set using equation 11.
- Defuzzify the fuzzified time series data set using equation 20 of proposed forecasting rule to find forecasting outcome.

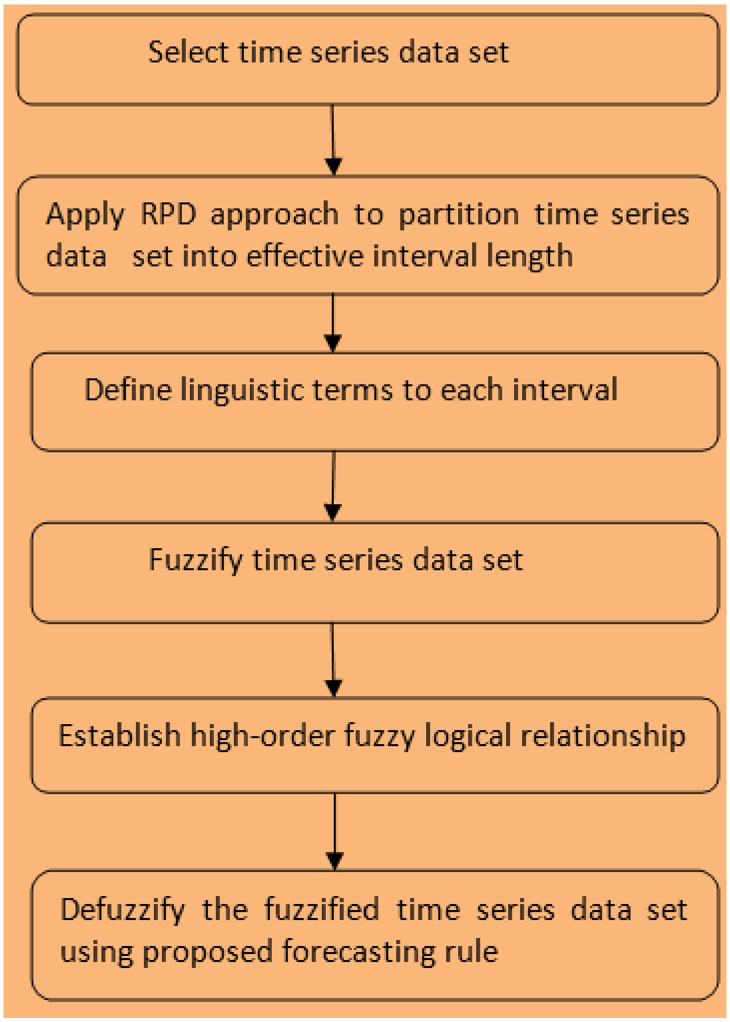

Figure 4.1 Architecture of Proposed Model

- 5^{th} order Fuzzy Logical Relationship (FLR) for June month of historical temperature data set of Taipei using equation 19 of proposed forecasting rule is presented in Table 4.7.

- Similarly we have obtained 5th order Fuzzy Logical Relationship (FLR) for July, August, September months of temperature data set of Taipei and are presented in Table 4.13,4.14,4.15 of Appendix 4.2 respectively.

Table 4.7 5th Order FLR for June Month of Temperature Data Set

$I_1,I_4,I_8,I_{16},I_{13}$→?(6)	$I_3,I_{11},I_7,I_8,I_{15}$→?(19)
$I_4,I_8,I_{16},I_{13},I_{11}$→?(7)	$I_{11},I_7,I_8,I_{15},I_{14}$→?(20)
$I_8,I_{16},I_{13},I_{11},I_{12}$→?(8)	$I_7,I_8,I_{15},I_{14},I_{17}$→?(21)
$I_{16},I_{13},I_{11},I_{12},I_{10}$→?(9)	$I_8,I_{15},I_{14},I_{17},I_{17}$→?(22)
$I_{13},I_{11},I_{12},I_{10},I_7$→?(10)	$I_{15},I_{14},I_{17},I_{17},I_7$→?(23)
$I_{11},I_{12},I_{10},I_7,I_{10}$→?(11)	$I_{14},I_{17},I_{17},I_7,I_5$→?(24)
$I_{12},I_{10},I_7,I_{10},I_9$→?(12)	$I_{17},I_{17},I_7,I_5,I_3$→?(25)
$I_{10},I_7,I_{10},I_9,I_6$→?(13)	I_{17},I_7,I_5,I_3,I_4→?(26)
I_7,I_{10},I_9,I_6,I_7→?(14)	I_7,I_5,I_3,I_4,I_2→?(27)
I_{10},I_9,I_6,I_7,I_3→?(15)	I_5,I_3,I_4,I_2,I_6→?(28)
I_9,I_6,I_7,I_3,I_{11}→?(16)	I_3,I_4,I_2,I_6,I_5→?(29)
I_6,I_7,I_3,I_{11},I_7→?(17)	I_4,I_2,I_6,I_5,I_8→?(30)
I_7,I_3,I_{11},I_7,I_8→?(18)	

- 5th order Fuzzy Logical Relationship (FLR) of historical student enrollment of Alabama University using equation 19 of proposed forecasting rule is presented in Table 4.8.

Table 4.8 5th Order FLR of Student Enrollment Data Set of Alabama University

I_1,I_2,I_3,I_4,I_7→?(1976)	$I_{11},I_{10},I_6,I_7,I_5$→?(1985)
I_2,I_3,I_4,I_7,I_6→?(1977)	I_{10},I_6,I_7,I_5,I_5→?(1986)
I_3,I_4,I_7,I_6,I_7→?(1978)	I_6,I_7,I_5,I_5,I_9→?(1987)
I_4,I_7,I_6,I_7,I_8→?(1979)	I_7,I_5,I_5,I_9,I_{11}→?(1988)

I_7,I_6,I_7,I_8,I_{11}→?(1980)	$I_5,I_5,I_9,I_{11},I_{12}$→?(1989)
$I_6,I_7,I_8,I_{11},I_{11}$→?(1981)	$I_5,I_9,I_{11},I_{12},I_{13}$→?(1990)
$I_7,I_8,I_{11},I_{11},I_{10}$→?(1982)	$I_9,I_{11},I_{12},I_{13},I_{14}$→?(1991)
$I_8,I_{11},I_{11},I_{10},I_6$→?(1983)	$I_{11},I_{12},I_{13},I_{14},I_{15}$→?(1992)
$I_{11},I_{11},I_{10},I_6,I_7$→?(1984)	

- 5th order Fuzzy Logical Relationship (FLR) of historical TAIFEX data set using equation 19 of proposed forecasting rule is presented in Table 4.9.

Table 4.9 5th Order FLR of TAIFEX Data Set

$I_{25},I_{25},I_{23},I_{22},I_{24}$→?(8/10/1998)	I_7,I_4,I_2,I_3,I_1→?(9/4/1998)
$I_{25},I_{23},I_{22},I_{24},I_{21}$→?(8/11/1998)	I_4,I_2,I_3,I_1,I_2→?(9/5/1998)
$I_{23},I_{22},I_{24},I_{21},I_{21}$→?(8/12/1998)	I_2,I_3,I_1,I_2,I_6→?(9/7/1998)
$I_{22},I_{24},I_{21},I_{21},I_{20}$→?(8/13/1998)	I_3,I_1,I_2,I_6,I_7→?(9/8/1998)
$I_{24},I_{21},I_{21},I_{20},I_{18}$→?(8/14/1998)	I_1,I_2,I_6,I_7,I_{11}→?(9/9/1998)
$I_{21},I_{21},I_{20},I_{18},I_{19}$→?(8/15/1998)	I_2,I_6,I_7,I_{11},I_8→?(9/10/1998)
$I_{21},I_{20},I_{18},I_{19},I_{19}$→?(8/17/1998)	I_6,I_7,I_{11},I_8,I_6→?(9/11/1998)
$I_{20},I_{18},I_{19},I_{19},I_{17}$→?(8/18/1998)	I_7,I_{11},I_8,I_6,I_7→?(9/14/1998)
$I_{18},I_{19},I_{19},I_{17},I_{17}$→?(8/19/1998)	I_{11},I_8,I_6,I_7,I_8→?(9/15/1998)
$I_{19},I_{19},I_{17},I_{17},I_{18}$→?(8/20/1998)	I_8,I_6,I_7,I_8,I_6→?(9/16/1998)
$I_{19},I_{17},I_{17},I_{18},I_{18}$→?(8/21/1998)	I_6,I_7,I_8,I_6,I_{15}→?(9/17/1998)
$I_{17},I_{17},I_{18},I_{18},I_{17}$→?(8/24/1998)	$I_7,I_8,I_6,I_{15},I_{13}$→?(9/18/1998)
$I_{17},I_{18},I_{18},I_{17},I_{15}$→?(8/25/1998)	$I_8,I_6,I_{15},I_{13},I_{11}$→?(9/19/1998)
$I_{18},I_{18},I_{17},I_{15},I_{15}$→?(8/26/1998)	$I_6,I_{15},I_{13},I_{11},I_{16}$→?(9/21/1998)
$I_{18},I_{17},I_{15},I_{15},I_9$→?(8/27/1998)	$I_{15},I_{13},I_{11},I_{16},I_{12}$→?(9/22/1998)
$I_{17},I_{15},I_{15},I_9,I_{11}$→?(8/28/1998)	$I_{13},I_{11},I_{16},I_{12},I_{13}$→?(9/23/1998)
$I_{15},I_{15},I_9,I_{11},I_5$→?(8/29/1998)	$I_{11},I_{16},I_{12},I_{13},I_{11}$→?(9/24/1998)

$I_{15},I_9,I_{11},I_5,I_7 \rightarrow ?$(8/31/1998)	$I_{16},I_{12},I_{13},I_{11},I_{13} \rightarrow ?$(9/25/1998)
$I_9,I_{11},I_5,I_7,I_4 \rightarrow ?$(9/1/1998)	$I_{12},I_{13},I_{11},I_{13},I_{12} \rightarrow ?$(9/28/1998)
$I_{11},I_5,I_7,I_4,I_2 \rightarrow ?$(9/2/1998)	$I_{13},I_{11},I_{13},I_{12},I_{11} \rightarrow ?$(9/29/1998)
$I_5,I_7,I_4,I_2,I_3 \rightarrow ?$(9/3/1998)	$I_{11},I_{13},I_{12},I_{11},I_{10} \rightarrow ?$(9/30/1998)

4.6 Statistical Parameters Used

Statistical parameters mean, standard deviation, Average Forecasting Error Rate (AFER) and Root Mean square Error (RMSE) are used to validate the performance accuracy of proposed model and are discussed in the following subsections:

4.6.1 Mean

It is an average and computed as sum of total observation divide by total number of observations. Formula for mean (μ) is given as:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} AV_i \tag{21}$$

4.6.2 Standard Deviation

Standard Deviation (SD) of a group of data tells about up to which degree data differs with the mean value of the group. For good forecasting actual mean and standard deviation of group of data should be close to forecasted mean and standard deviation. Formula to calculate SD is given below:

$$SD=\sqrt{\frac{\sum_{i=1}^{n}(AV_i-\mu)}{n}} \tag{22}$$

4.6.3 Root Mean Square Error (RMSE)

RMSE is used to measure up to which degree forecasted values are differ with actual values. For good forecasting RMSE value should be small. Formula to calculate RMSE is given below:

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}(AV_i-FV_i)^2}{n}} \tag{23}$$

4.6.4 Average Forecasting Error Rate (AFER)

In statistics forecasting error rate is used to find out difference between actual outcome and predicted outcome of time series data at any point of time. AFER is defined as average of such differences over a specific period of time and is defined as:

$$\text{AFER (in \%)} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{FV_i - AV_i}{AV_i}\right| * 100 \qquad (24)$$

Note: In above equations (21-24), AV_i and FV_i are actual values and forecasted values of observation i respectively (∀ i=1 to n) where 'n' is total number of observations taken in to consideration.

CHAPTER 5
EXPERIMENTAL RESULTS

In this chapter the experimental results using proposed forecasting model is presented. This chapter is further divided into two sub-chapters: in the former one we have presented forecasting values of time series data sets taken into consideration whereas in later one we have presented empirical analysis of model using RMSE.

5.1 Forecasting Results

In this sub-chapter forecasting results of time series data sets are presented along with actual one.

- The forecasting temperature along with actual temperature for temperature data set of Taipei using 5th order of proposed model and equation 20 of proposed forecasting rule is presented in Table 5.1.

- Similarly we have obtained forecasting temperature along with actual temperature for temperature data set using 6th, 7th, and 8th order and are presented in Table 5.8, 5.9, 5.10 of Appendix 5.1 respectively.

Table 5.1 Forecasted Temperature of Temperature Data Set Using 5th Order of Proposed Model

Day	June Actual	June Predicted	July Actual	July Predicted	August Actual	August Predicted	September Actual	September Predicted
1	26.1	NA	29.9	NA	27.1	NA	27.5	NA
2	27.6	NA	28.4	NA	28.9	NA	26.8	NA
3	29	NA	29.2	NA	28.9	NA	26.4	NA
4	30.5	NA	29.4	NA	29.3	NA	27.5	NA
5	30	NA	29.9	NA	28.8	NA	26.6	NA
6	29.5	28.62	29.6	29.39	28.7	28.6	28.2	26.93
7	29.7	29.09	30.1	29.3	29	28.9	29.2	27.22

8	29.4	29.53	29.3	29.55	28.2	28.95	29	27.46
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
14	27.5	28.85	29.1	28.63	27.6	28.37	30.3	30.13
15	29.5	28.62	27.7	28.85	26.8	28.34	29.5	30.13
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
30	30.2	28.2	27.9	28.54	26	28.17	23.5	25.25
31	-		26.9	28.36	27.7	27.94	-	

- Forecasting student enrollment along with actual student enrollment of student enrollment data set of Alabama University using 2ndand 5thorder of proposed model and equation 20 of forecasting rule is presented in Table 5.2.
- Forecasting stock price along with actual stock price of TAIFEX data set using 2ndand 5thorder of proposed model and equation 20 of forecasting rule is presented in Table 5.3.

Table 5.2 Forecasted Enrollments of Student Enrollment Using Various Order of Proposed Model

Year	Actual Enrollments	Predicted Enrollments	
		2nd Order	5th Order
1971	13055	NA	NA
1972	13563	NA	NA
1973	13867	13412	NA
1974	14696	13769	NA
1975	15460	14245	NA

+1976	15311	15316	14568
1977	15603	15458	14929
1978	15861	15458	15244
1979	16807	15613	15435
1980	16919	16671	15886
1981	16388	16965	16239
1982	15433	16850	16431
1983	15497	15712	16472
1984	15145	15458	16299
1985	15163	15363	15895
1986	15984	15078	15411
1987	16859	15395	15363
1988	18150	16731	15809
1989	18970	17281	16162
1990	19328	18650	17003
1991	19337	18966	17756
1992	18876	19335	18189

Table 5.3 Forecasted Stock Price of TAIFEX Data Set Using Various Order of Proposed Model

Date	Actual Price	Predicted Price	
		2nd Order	5th Order
8/3/1998	7552	NA	NA
8/4/1998	7560	NA	NA
8/5/1998	7487	7545.43	NA
8/6/1998	7462	7526.01	NA
8/7/1998	7515	7472.61	NA

-	-	-	
-	-	-	
8/24/1998	6955	7254.13	18120.76
8/25/1998	6949	7093.92	17989.68
-	-	-	
-	-	-	
9/1/1998	6409	6681.25	10648.39
9/2/1998	6430	6463.5	12295.92
9/3/1998	6200	6418.95	15327.28
-	-	-	
-	-	-	
9/29/1998	6806	6851.75	15687.4
9/30/1998	6787	6841.67	14688.77

5.2 Empirical Analysis

In this sub-chapter the empirical analysis of forecasting result using proposed model is presented.

- Root Mean Square Error (RMSE) for forecasting result of temperature data set of Taipei using various order of proposed model and equation 23 is presented in Table 5.4.

Table 5.4 RMSE Using Various Order of Proposed Model for Temperature Data Set

Order	Months			
	June	July	August	September
5^{th}	1.17	1.38	1.1	1.38
6^{th}	1.22	1.47	1.11	1.49
7^{th}	1.24	1.56	1.11	1.58
8^{th}	1.24	1.64	1.13	1.71

- Average Forecasting Error Rate (AFER) for forecasting result of temperature data set of Taipei using various orders of proposed model and equation 24 is presented in Table 5.5.

Table 5.5 AFER (in %) Using Various Order of Proposed Model for Temperature Data Set

Order	Months			
	June	July	August	September
5^{th}	0.08	0.88	0.88	1.48
6^{th}	0.2	1.12	0.88	1.7
7^{th}	0.26	1.45	1	2.4
8^{th}	0.54	1.47	0.91	2.7

- AFER (in %) of June, July, August, September months using 5^{th} order proposed model and equation 24 are 0.08, 0.88, 0.88, 1.48 respectively which are better than Lee et al. model [46] .
- Average RMSE of proposed model is 1.21 which is better than model [45] with average RMSE 1.24 and [23] with average RMSE 1.91.
- RMSE and AFER for forecasting result of student enrollment of Alabama University using proposed model and equation 23 & 24 are presented in Table 5.6.

Table 5.6 RMSE and AFER Using Various Order of Proposed Model for Student Enrollment Data Set

Order	RMSE	AFER
2^{nd}	876	2 %
5^{th}	1382	4 %

- RMSE for student enrollment using 2^{nd} order of proposed model and equation 23 is 876 which is better than Song and Chissom's method [2] with RMSE 881.

- AFER (in %) for student enrollment using 2^{nd} order of proposed model and equation 24 is 2 which is better than Cheng et al. model[47] with AFER(in %) 2.08 .
- RMSE and AFER for TAIFEX data set using various order of proposed model and equation 23& 24 is presented in Table 5.7.

Table 5.7 RMSE and AFER Using Various Order of Proposed Model for TAIFEX Data Set

Order	RMSE	AFER
2^{nd}	120	0.45 %
5^{th}	12521	17 %

- RMSE for TAIFEX data set using 2^{nd} order of proposed model is 120 which is as good as Chen's model [48].

CHAPTER 6
CONCLUSION AND FUTURE WORK

This chapter is further divided into two sub-chapters. In the former we have discussed about our contribution in the research area of time series model based on the concept of high-order fuzzy time series whereas the future scope of proposed model is discussed into later one.

6.1 Research Contribution

In this thesis we have presented a new time series model based on the concept of high-order fuzzy time series. Historical data sets are partitioned in to effective interval lengths using existing RPD approach. A novel defuzzification technique is introduced based on the concept of "weighted-average" defuzzification approach. A new mathematical formula is introduced in this thesis as forecasting rule. Concept of ranking of intervals is introduced in the forecasting rule which tells about number of elements present in a particular interval. Forecasting accuracy of proposed model is evaluated by applying model on some historical time series data sets which are: Temperature data set of Taipei, Student enrollment data set of Alabama University and stock price of TAIFEX. The relative comparison of forecasting result of proposed model with other existing techniques shows that proposed model outperform over existing models.

6.2 Future Work

In this thesis we have introduced an efficient time series model for one-factor time series data set. The proposed forecasting model can be used to forecast any one-factor time series data like temperature data set, student enrollment data set, stock price of stock market, etc., As we have discussed in *Literature Review* chapter that the forecasting accuracy of forecasting model based on fuzzy concept much depends on effective interval length and defuzzification technique. Therefore forecasting accuracy of proposed model can be improve using meta-heuristic optimization techniques like PSO, ACO, SA, Tabu Search and so on. By using these optimization techniques the length of intervals can be tuned to generate new effective interval length.

Appendix 4.1

In this section intervals with corresponding element, rank of intervals, and center-point of intervals are presented as discussed in Chapter 4.

- Intervals with corresponding elements, rank and center-point for July month of historical temperature data set using proposed partitioning and rank finding algorithm is presented in Table 4.10.

Table 4.10 Intervals with Corresponding Elements, Rank and Center-point for July Month of Temperature Data Set

Interval	**Corresponding elements**	**center-point**	**Rank**
U_X Interval			
[26.9,27.04]	<26.9>	26.97	1
[27.6,27.74]	<27.7>	27.67	1
[27.74,27.88]	<27.8>	27.81	1
[27.88,28.02]	<27.9,28,28>	27.95	3
[28.02,28.16]	<28.1,28.1>	28.09	2
[28.3,28.44]	<28.4,28.4>	28.37	2
[28.58,28.72]	<28.6,28.7>	28.65	2
[28.86,29]	<28.9,28.9>	28.93	2
[29,29.2]	<29.1,29.2>	29.1	2
U_Y Interval			
[29.3,29.45]	<29.3,29.3,29.4>	29.38	3
[29.45,29.6]	<29.6,29.6>	29.53	2
[29.75,29.9]	<29.9,29.9,29.9>	29.83	3
[30.05,30.2]	<30.1>	30.13	1
[30.65,30.8]	<30.8>	30.73	1

[31.25,31.4]	<31.3,31.3,31.3,31.4>	31.33	4
[31.55,31.6]	<31.6>	31.58	1

- Intervals with corresponding elements, rank and center-point for August month of historical temperature data set using proposed partitioning and rank finding algorithm is presented in Table 4.11.

Table 4.11 Intervals with Corresponding Elements, Rank of intervals and Center-points of Intervals for August Month of Temperature Data Set

Interval	**Corresponding elements**	**center-point**	**Rank**
U_X Interval			
[26,26.19]	<26>	26.1	1
[26.19,26.38]	<26.2>	26.29	1
[26.76,26.95]	<26.8>	26.86	1
[26.95,27.14]	<27,27.1>	27.05	2
[27.52,27.71]	<27.6,27.6,27.7,27.7>	27.62	4
[27.71,27.9]	<27.9>	27.81	1
U_Y Interval			
[28,28.09]	<28>	28.05	1
[28.09,28.18]	<28.1>	28.14	1
[28.18,28.27]	<28.2>	28.23	1
[28.27,28.36]	<28.3,28.3>	28.32	2
[28.54,28.63]	<28.6>	28.59	1
[28.63,28.72]	<28.7,28.7>	28.68	2
[28.72,28.81]	<28.8>	28.77	1
[38.81,28.9]	<28.9,28.9,28.9>	33.86	3
[28.99,29.08]	<29,29,29>	29.04	3
[29.17,29.26]	<29.2>	29.22	1

[29.26,29.35]	<29.3,29.3>	29.31	2
[29.53,29.62]	<29.6>	29.58	1
[29.71,29.8]	<29.8>	29.76	1
[29.89,29.9]	<29.9>	29.89	1

- Intervals with corresponding elements, rank and center-point for September month of historical temperature data set using proposed partitioning and rank finding algorithm is presented in Table 4.12.

Table 4.12 Intervals with Corresponding Elements, Rank of Intervals and Center-points of Intervals for September month of Temperature data set

Interval	**Corresponding elements**	**center-point**	**Rank**
U_X Interval		23.45	
[23.3,23.59]	<23.3,23.5>	24.32	2
[24.17,24.46]	<24.2>	24.9	1
[24.75,25.04]	<25>	25.48	1
[25.33,25.62]	<25.6>	25.77	1
[25.62,25.91]	<25.7,25.8>	26.35	2
[26.2,26.49]	<26.4,26.4,26.4>	26.64	3
[26.49,26.78]	<26.6>	26.79	1
[26.78,26.8]	<26.8>	23.45	1
U_Y Interval			
[27,27.19]	<27>	27.1	1
[27.38,27.57]	<27.5,27.5>	27.48	2
[27.95,28.14]	<28.1>	28.05	1
[28.14,28.33]	<28.2,28.3,28.3>	28.24	3
[28.33,28.52]	<28.4>	28.43	1
[28.52,28.71]	<28.6>	28.62	1

[28.9,29.09]	<29>	29	1
[29.09,29.28]	<29.2>	29.19	1
[29.47,29.66]	<29.5>	29.57	1
[29.85,30.04]	<29.9,29.9>	29.95	2
[30.04,30.23]	<30.2>	30.14	1
[30.23,30.42]	<30.3,30.3>	30.33	2
[30.42,30.5]	<30.1>	30.46	1

Appendix 4.2

In this section Fuzzy Logical Relationships (FLR) for Temperature data set using various Order of proposed model as discussed in Chapter 4 is presented.

- 5th order Fuzzy Logical Relationship (FLR) for July month of historical temperature data set of Taipei using equation 19 of proposed forecasting rule is presented in Table 4.13.

Table 4.13 5th Order FLR for July Month of Temperature Data Set

$I_{12},I_6,I_9,I_{10},I_{12} \rightarrow ?(6)$	$I_9,I_2,I_5,I_7,I_{12} \rightarrow ?(19)$
$I_6,I_9,I_{10},I_{12},I_{11} \rightarrow ?(7)$	$I_2,I_5,I_7,I_{12},I_{14} \rightarrow ?(20)$
$I_9,I_{10},I_{12},I_{11},I_{13} \rightarrow ?(8)$	$I_5,I_7,I_{12},I_{14},I_{16} \rightarrow ?(21)$
$I_{10},I_{12},I_{11},I_{13},I_{10} \rightarrow ?(9)$	$I_7,I_{12},I_{14},I_{16},I_{15} \rightarrow ?(22)$
$I_{12},I_{11},I_{13},I_{10},I_5 \rightarrow ?(10)$	$I_{12},I_{14},I_{16},I_{15},I_{15} \rightarrow ?(23)$
$I_{11},I_{13},I_{10},I_5,I_8 \rightarrow ?(11)$	$I_{14},I_{16},I_{15},I_{15},I_{15} \rightarrow ?(24)$
$I_{13},I_{10},I_5,I_8,I_6 \rightarrow ?(12)$	$I_{16},I_{15},I_{15},I_{15},I_{15} \rightarrow ?(25)$
$I_{10},I_5,I_8,I_6,I_{11} \rightarrow ?(13)$	$I_{15},I_{15},I_{15},I_{15},I_8 \rightarrow ?(26)$
$I_5,I_8,I_6,I_{11},I_3 \rightarrow ?(14)$	$I_{15},I_{15},I_{15},I_8,I_4 \rightarrow ?(27)$
$I_8,I_6,I_{11},I_3,I_9 \rightarrow ?(15)$	$I_{15},I_{15},I_8,I_4,I_7 \rightarrow ?(28)$
$I_6,I_{11},I_3,I_9,I_2 \rightarrow ?(16)$	$I_{15},I_8,I_4,I_7,I_4 \rightarrow ?(29)$
$I_{11},I_3,I_9,I_2,I_5 \rightarrow ?(17)$	$I_8,I_4,I_7,I_4,I_{10} \rightarrow ?(30)$
$I_3,I_9,I_2,I_5,I_7 \rightarrow ?(18)$	$I_4,I_7,I_4,I_{10},I_4 \rightarrow ?(31)$

- 5th order Fuzzy Logical Relationship (FLR) for August month of historical temperature data set of Taipei using equation 19 of proposed forecasting rule is presented in Table 4.14.

Table 4.14 5th Order FLR for August Month of Temperature Data Set

$I_4,I_{14},I_{14},I_{17},I_{13} \rightarrow ?(6)$	$I_3,I_5,I_6,I_{15},I_{16} \rightarrow ?(20)$
$I_{14},I_{14},I_{17},I_{13},I_{12} \rightarrow ?(7)$	$I_5,I_6,I_{15},I_{16},I_{19} \rightarrow ?(21)$
$I_{14},I_{17},I_{13},I_{12},I_{15} \rightarrow ?(8)$	$I_6,I_{15},I_{16},I_{19},I_{18} \rightarrow ?(22)$
$I_{17},I_{13},I_{12},I_{15},I_9 \rightarrow ?(9)$	$I_{15},I_{16},I_{19},I_{18},I_{17} \rightarrow ?(23)$
$I_{13},I_{12},I_{15},I_9,I_4 \rightarrow ?(10)$	$I_{16},I_{19},I_{18},I_{17},I_7 \rightarrow ?(24)$
$I_{12},I_{15},I_9,I_4,I_{10} \rightarrow ?(11)$	$I_{19},I_{18},I_{17},I_7,I_{10} \rightarrow ?(25)$
$I_{15},I_9,I_4,I_{10},I_{14} \rightarrow ?(12)$	$I_{18},I_{17},I_7,I_{10},I_{11} \rightarrow ?(26)$
$I_9,I_4,I_{10},I_{14},I_8 \rightarrow ?(13)$	$I_{17},I_7,I_{10},I_{11},I_{12} \rightarrow ?(27)$
$I_4,I_{10},I_{14},I_8,I_{20} \rightarrow ?(14)$	$I_7,I_{10},I_{11},I_{12},I_{15} \rightarrow ?(28)$
$I_{10},I_{14},I_8,I_{20},I_5 \rightarrow ?(15)$	$I_{10},I_{11},I_{12},I_{15},I_5 \rightarrow ?(29)$
$I_{14},I_8,I_{20},I_5,I_3 \rightarrow ?(16)$	$I_{11},I_{12},I_{15},I_5,I_2 \rightarrow ?(30)$
$I_8,I_{20},I_5,I_3,I_5 \rightarrow ?(17)$	$I_{12},I_{15},I_5,I_2,I_1 \rightarrow ?(31)$
$I_{20},I_5,I_3,I_5,I_6 \rightarrow ?(18)$	
$I_5,I_3,I_5,I_6,I_{15} \rightarrow ?(19)$	

- 5th order Fuzzy Logical Relationship (FLR) for September month of historical temperature data set of Taipei using equation 19 of proposed forecasting rule is presented in Table 4.15.

Table 4.15 5th Order FLR for September Month of Temperature Data set

$I_{10},I_8,I_6,I_{10},I_7 \rightarrow ?(6)$	$I_{20},I_{17},I_{12},I_{14},I_{11} \rightarrow ?(19)$
$I_8,I_6,I_{10},I_7,I_{12} \rightarrow ?(7)$	$I_{17},I_{12},I_{14},I_{11},I_{13} \rightarrow ?(20)$
$I_6,I_{10},I_7,I_{12},I_{16} \rightarrow ?(8)$	$I_{12},I_{14},I_{11},I_{13},I_{12} \rightarrow ?(21)$
$I_{10},I_7,I_{12},I_{16},I_{15} \rightarrow ?(9)$	$I_{14},I_{11},I_{13},I_{12},I_6 \rightarrow ?(22)$
$I_7,I_{12},I_{16},I_{15},I_{20} \rightarrow ?(10)$	$I_{11},I_{13},I_{12},I_6,I_5 \rightarrow ?(23)$
$I_{12},I_{16},I_{15},I_{20},I_{18} \rightarrow ?(11)$	$I_{13},I_{12},I_6,I_5,I_3 \rightarrow ?(24)$

$I_{16}, I_{15}, I_{20}, I_{18}, I_{18} \rightarrow ?(12)$	$I_{12}, I_6, I_5, I_3, I_9 \rightarrow ?(25)$
$I_{15}, I_{20}, I_{18}, I_{18}, I_{21} \rightarrow ?(13)$	$I_6, I_5, I_3, I_9, I_5 \rightarrow ?(26)$
$I_{20}, I_{18}, I_{18}, I_{21}, I_{19} \rightarrow ?(14)$	$I_5, I_3, I_9, I_5, I_6 \rightarrow ?(27)$
$I_{18}, I_{18}, I_{21}, I_{19}, I_{20} \rightarrow ?(15)$	$I_3, I_9, I_5, I_6, I_4 \rightarrow ?(28)$
$I_{18}, I_{21}, I_{19}, I_{20}, I_{17} \rightarrow ?(16)$	$I_9, I_5, I_6, I_4, I_2 \rightarrow ?(29)$
$I_{21}, I_{19}, I_{20}, I_{17}, I_{12} \rightarrow ?(17)$	$I_5, I_6, I_4, I_2, I_1 \rightarrow ?(30)$
$I_{19}, I_{20}, I_{17}, I_{12}, I_{14} \rightarrow ?(18)$	

Appendix 5.1

In this section the forecasting outcomes along with actual values using various Order of proposed model and equation 20 of forecasting rule as discussed in Chapter 5 are presented.

- The forecasting temperature along with actual temperature for temperature data set of Taipei using 6th order of proposed model and equation 20 of proposed forecasting rule is presented in Table 5.8.

Table 5.8 Forecasted Temperature of Temperature Data Set Using 6th Order of Proposed Model

Day	June Actual	June Predicted	July Actual	July Predicted	August Actual	August Predicted	Sep. Actual	Sep. Predicted
1	26.1	NA	29.9	NA	27.1	NA	27.5	NA
2	27.6	NA	28.4	NA	28.9	NA	26.8	NA
3	29	NA	29.2	NA	28.9	NA	26.4	NA
4	30.5	NA	29.4	NA	29.3	NA	27.5	NA
5	30	NA	29.9	NA	28.8	NA	26.6	NA
6	29.5	NA	29.6	NA	28.7	NA	28.2	NA
7	29.7	28.8	30.1	29.41	29	28.62	29.2	27.26
8	29.4	29.16	29.3	29.37	28.2	28.93	29	27.4
9	28.8	29.51	28.1	29.51	27	28.89	30.3	27.6
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
14	27.5	28.92	29.1	28.82	27.6	28.35	30.3	30.01
15	29.5	28.66	27.7	28.71	26.8	28.14	29.5	30.17
-	-	-	-	-	-	-	-	-

-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
29	29	27.79	29.3	29.64	26.2	28.33	23.3	25.82
30	30.2	28.07	27.9	29.2	26	28.19	23.5	25.44
31	NA	28.8	26.9	28.43	27.7	28	NA	27.26

- The forecasting temperature along with actual temperature for temperature data set of Taipei using 7th order of proposed model and equation 20 of proposed forecasting rule is presented in Table 5.9.

Table 5.9 Forecasted Temperature of Temperature Data Set Using 7th Order of Proposed Model

Day	June Actual	June Predicted	July Actual	July Predicted	August Actual	August Predicted	Sep. Actual	Sep. Predicted
1	26.1	NA	29.9	NA	27.1	NA	27.5	NA
2	27.6	NA	28.4	NA	28.9	NA	26.8	NA
3	29	NA	29.2	NA	28.9	NA	26.4	NA
4	30.5	NA	29.4	NA	29.3	NA	27.5	NA
5	30	NA	29.9	NA	28.8	NA	26.6	NA
6	29.5	NA	29.6	NA	28.7	NA	28.2	NA
7	29.7	NA	30.1	NA	29	NA	29.2	NA
8	29.4	28.89	29.3	29.45	28.2	28.7	29	27.41
9	28.8	29.2	28.1	29.37	27	28.88	30.3	27.53
10	29.4	29.29	28.9	29.33	28.3	28.63	29.9	28.02
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
14	27.5	28.97	29.1	28.92	27.6	28.51	30.3	29.93

15	29.5	28.75	27.7	28.86	26.8	28.14	29.5	30.07
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
29	29	28.05	29.3	29.95	26.2	28.46	23.3	25.81
30	30.2	28.05	27.9	29.6	26	28.18	23.5	25.39
31	NA	28.89	26.9	29.01	27.7	28.04	NA	27.41

- The forecasting temperature along with actual temperature for temperature data set of Taipei using 8th order of proposed model and equation 20 of proposed forecasting rule is presented in Table 5.10.

Table 5.10 Forecasted Temperature of Temperature Data Set Using 8th Order of Proposed Model

Day	June Actual	June Predicted	July Actual	July Predicted	August Actual	August Predicted	Sep. Actual	Sep. Predicted
1	26.1	NA	29.9	NA	27.1	NA	27.5	NA
2	27.6	NA	28.4	NA	28.9	NA	26.8	NA
3	29	NA	29.2	NA	28.9	NA	26.4	NA
4	30.5	NA	29.4	NA	29.3	NA	27.5	NA
5	30	NA	29.9	NA	28.8	NA	26.6	NA
6	29.5	NA	29.6	NA	28.7	NA	28.2	NA
7	29.7	NA	30.1	NA	29	NA	29.2	NA
8	29.4	NA	29.3	NA	28.2	NA	29	NA
9	28.8	28.97	28.1	29.44	27	28.67	30.3	27.52
10	29.4	29.09	28.9	29.23	28.3	28.67	29.9	27.93
11	29.3	29.31	28.4	29.29	28.9	28.59	29.9	28.27
-	-	-	-	-	-	-	-	-

-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
15	29.5	28.8	27.7	28.94	26.8	28.3	29.5	29.99
16	28.8	28.83	28.1	28.78	27.6	28.06	28.3	30.02
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-
29	29	28.38	29.3	30.16	26.2	28.53	23.3	25.93
30	30.2	28.2	27.9	29.88	26	28.32	23.5	25.45
31	NA	28.97	26.9	29.4	27.7	28.04	NA	27.52

BIBLIOGRAPHY

[1] L.A. Zadeh, "Fuzzy sets Information and Control" 8(1965) 338-353

[2] Q, Song, B.S. Chissom,"Forecasting enrollments with fuzzy time series-Part I", Fuzzy sets and systems 54(1993)1-9

[3] Q, Song, B.S. Chissom,"Forecasting enrollments with fuzzy time series-Part II", Fuzzy sets and systems 54(1993)1-9

[4] Kuo, I-Hong, Shi-Jinn Horng, Tzong-Wann Kao, Tsung-Lieh Lin, Cheng-Ling Lee, and Yi Pan. "An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization." *Expert Systems with Applications* 36, no. 3 (2009): 6108-6117.

[5] Chen, Shyi-Ming, and Nien-Yi Chung. "Forecasting enrollments using high-order fuzzy time series and genetic algorithms." International Journal of Intelligent Systems 21, no. 5 (2006): 485-501.

[6] Chen, Shyi-Ming, and Chia-Ching Hsu. "A new method to forecast enrollments using fuzzy time series." International Journal of Applied Science and Engineering 2, no. 3 (2004): 234-244. Harvard

[7] Li, Sheng-Tun, and Yi-Chung Cheng. "Deterministic fuzzy time series model for forecasting enrollments." Computers & Mathematics with Applications 53, no. 12 (2007): 1904-1920.

[8] Singh, S. R. "A simple method of forecasting based on fuzzy time series." Applied mathematics and computation 186, no. 1 (2007): 330-339.

[9] Tsaur, Ruey-Chyn, Jia-Chi O. Yang, and Hsiao-Fan Wang. "Fuzzy relation analysis in fuzzy time series model." Computers & Mathematics with Applications 49, no. 4 (2005): 539-548.

[10] Saxena, Preetika, and Santhosh Easo. "A New Method for Forecasting Enrollments based on Fuzzy Time Series with Higher Forecast Accuracy Rate."

[11] Singh, S. R. "A simple method of forecasting based on fuzzy time series." Applied mathematics and computation 186, no. 1 (2007): 330-339.

[12] Huang, Yao-Lin, Shi-Jinn Horng, Mingxing He, Pingzhi Fan, Tzong-Wann Kao, Muhammad Khurram Khan, Jui-Lin Lai, and I-Hong Kuo. "A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimization." Expert Systems with Applications 38, no. 7 (2011): 8014-8023.

[13] Liu, Tung-Kuan, Yeh-Peng Chen, and Jyh-Horng Chou. "Extracting fuzzy relations in fuzzy time series model based on approximation concepts." Expert Systems with Applications 38, no. 9 (2011): 11624-11629.

[14] Gangwar, Sukhdev Singh, and Sanjay Kumar. "Partitions based computational method for high-order fuzzy time series forecasting." Expert Systems with Applications 39, no. 15 (2012): 12158-12164.

[15] Egrioglu, Erol, Cagdas Hakan Aladag, and Ufuk Yolcu. "Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks." Expert Systems with Applications 40, no. 3 (2013): 854-857.

[16] Liu, Hao-Tien. "An improved fuzzy time series forecasting method using trapezoidal fuzzy numbers." Fuzzy Optimization and Decision Making 6, no. 1 (2007): 63-80.

[17] Singh, Shiva Raj. "A computational method of forecasting based on high-order fuzzy time series." Expert Systems with Applications 36, no. 7 (2009): 10551-10559.

[18] Teoh, Hia Jong, Ching-Hsue Cheng, Hsing-Hui Chu, and Jr-Shian Chen. "Fuzzy time series model based on probabilistic approach and rough set rule induction for empirical research in stock markets." *Data & Knowledge Engineering* 67, no. 1 (2008): 103-117.

[19] Zarandi, M. H. F., A. Molladavoudi, and MH Ali Beigi. "A new method for temperature prediction and the TAIFEX Forecasting based on fuzzy logical

relationship and double Interval division." In Industrial Engineering and Engineering Management, 2009. IEEM 2009. IEEE International Conference on, pp. 1543-1547. IEEE, 2009. Harvard

[20] Wang, Nai-Yi, and Shyi-Ming Chen. "Temperature prediction and TAIFEX forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series." Expert Systems with Applications 36, no. 2 (2009): 2143-2154.

[21] Li, Sheng-Tun, Yi-Chung Cheng, and Su-Yu Lin. "A FCM-based deterministic forecasting model for fuzzy time series." Computers & Mathematics with Applications 56, no. 12 (2008): 3052-3063.

[22] Singh, Pritpal, and Bhogeswar Borah. "An effective neural network and fuzzy time series-based hybridized model to handle forecasting problems of two factors." Knowledge and information systems 38, no. 3 (2014): 669-690.

[23] Vamitha, V., M. Jeyanthi, S. Rajaram, and T. Revathi. "Temperature Prediction Using Fuzzy Time Series and Multivariate Markov Chain." International Journal of Fuzzy Mathematics and Systems 2, no. 3 (2012): 217-230.

[24] Chen, Shyi-Ming, and Jeng-Ren Hwang. "Temperature prediction using fuzzy time series." Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 30, no. 2 (2000): 263-275.

[25] Huarng, Kunhuang. "Heuristic models of fuzzy time series for forecasting." Fuzzy sets and systems 123, no. 3 (2001): 369-386.

[26] Lee, Li-Wei, Li-Hui Wang, and Shyi-Ming Chen. "Temperature prediction and TAIFEX forecasting based on high-order fuzzy logical relationships and genetic simulated annealing techniques." Expert Systems with Applications 34, no. 1 (2008): 328-336.

[27] Kuo, I-Hong, Shi-Jinn Horng, Yuan-Hsin Chen, Ray-Shine Run, Tzong-Wann Kao, Rong-Jian Chen, Jui-Lin Lai, and Tsung-Lieh Lin. "Forecasting TAIFEX based on

fuzzy time series and particle swarm optimization." Expert Systems with Applications 37, no. 2 (2010): 1494-1502.

[28] Singh, Pritpal, and Bhogeswar Borah. "Forecasting stock index price based on M-factors fuzzy time series and particle swarm optimization." International Journal of Approximate Reasoning 55, no. 3 (2014): 812-833.

[29] Singh, Pritpal. "High-order fuzzy-neuro-entropy integration-based expert system for time series forecasting." Neural Computing and Applications (2016): 1-18.

[30] Singh, Pritpal, and Bhogeswar Borah. "Indian summer monsoon rainfall prediction using artificial neural network." Stochastic environmental research and risk assessment 27, no. 7 (2013): 1585-1599.

[31] Egrioglu, Erol, Cagdas Hakan Aladag, Ufuk Yolcu, Murat A. Basaran, and Vedide R. Uslu. "A new hybrid approach based on SARIMA and partial high order bivariate fuzzy time series forecasting model." Expert Systems with Applications 36, no. 4 (2009): 7424-7434.

[32] Huarng, Kunhuang, and Hui-Kuang Yu. "A type 2 fuzzy time series model for stock index forecasting." Physica A: Statistical Mechanics and its Applications 353 (2005): 445-462.

[33] Huarng, Kunhuang, and Hui-Kuang Yu. "A type 2 fuzzy time series model for stock index forecasting." Physica A: Statistical Mechanics and its Applications 353 (2005): 445-462.

[34] Singh, Pritpal, and Bhogeswar Borah. "An efficient time series forecasting model based on fuzzy time series." Engineering Applications of Artificial Intelligence 26, no. 10 (2013): 2443-2457.

[35] Egrioglu, Erol, Cagdas Hakan Aladag, Ufuk Yolcu, Vedide R. Uslu, and Murat A. Basaran. "Finding an optimal interval length in high order fuzzy time series." Expert Systems with Applications 37, no. 7 (2010): 5052-5055.

[36] Cai, Qisen, Defu Zhang, Wei Zheng, and Stephen CH Leung. "A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression." Knowledge-Based Systems 74 (2015): 61-68.

[37] Singh, Pritpal, and Bhogeswar Borah. "An effective neural network and fuzzy time series-based hybridized model to handle forecasting problems of two factors." Knowledge and information systems 38, no. 3 (2014): 669-690.

[38] Avazbeigi, Milad, Seyed Hossein Hashemi Doulabi, and Behrooz Karimi. "Choosing the appropriate order in fuzzy time series: A new N-factor fuzzy time series for prediction of the auto industry production." Expert Systems with Applications 37, no. 8 (2010): 5630-5639.

[39] Bahrepour, Majid, Mohammad-R. Akbarzadeh-T, Mahdi Yaghoobi, and Mohammad-B. Naghibi-S. "An adaptive ordered fuzzy time series with application to FOREX." Expert Systems with Applications 38, no. 1 (2011): 475-485.

[40] Huarng, Kunhuang, and Tiffany Hui-Kuang Yu. "Ratio-based lengths of intervals to improve fuzzy time series forecasting." Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 36, no. 2 (2006): 328-340.

[41] Yolcu, Ufuk, Erol Egrioglu, Vedide R. Uslu, Murat A. Basaran, and Cagdas H. Aladag. "A new approach for determining the length of intervals for fuzzy time series." Applied Soft Computing 9, no. 2 (2009): 647-651.

[42] Davari, Soheil, Mohammad Hossein Fazel Zarandi, and I. Burhan Turksen. "An improved fuzzy time series forecasting model based on particle swarm intervalization." In Fuzzy Information Processing Society, 2009. NAFIPS 2009. Annual Meeting of the North American, pp. 1-5. IEEE, 2009.

[43] Li, Sheng-Tun, and Yeh-Peng Chen. "Natural partitioning-based forecasting model for fuzzy time-series." In Fuzzy Systems, 2004. Proceedings. 2004 IEEE International Conference on, vol. 3, pp. 1355-1359. IEEE, 2004.

[44] Chen, Shyi-Ming, and Pei-Yuan Kao. "TAIEX forecasting based on fuzzy time series, particle swarm optimization techniques and support vector machines." information Sciences 247 (2013): 62-71. Harvard

[45] Singh, Pritpal, and Bhogeswar Borah. "High-order fuzzy-neuro expert system for time series forecasting." Knowledge-Based Systems 46 (2013): 12-21.

[46] Lee, Li-Wei, Li-Hui Wang, Shyi-Ming Chen, and Yung-Ho Leu. "Handling forecasting problems based on two-factors high-order fuzzy time series."*IEEE Transactions on Fuzzy Systems* 14, no. 3 (2006): 468-477.

[47] Cheng, Ching-Hsue, Tai-Liang Chen, and Chen-Han Chiang. "Trend-weighted fuzzy time-series model for TAIEX forecasting." In *International Conference on Neural Information Processing*, pp. 469-477. Springer Berlin Heidelberg, 2006.

[48] Chen, Shyi-Ming. "Forecasting enrollments based on fuzzy time series."*Fuzzy sets and systems* 81, no. 3 (1996): 311-319.

[49] Chen, Tai-Liang, Ching-Hsue Cheng, and Hia-Jong Teoh. "High-order fuzzy time-series based on multi-period adaptation model for forecasting stock markets." *Physica A: Statistical Mechanics and its Applications* 387, no. 4 (2008): 876-888.

List of Publications

Details of list of publications are given below:

1. A. Tripathi, H.S. Pannu, "A Review: Fuzzy Time Series Forecasting Model", in *IEEE International Conference on Innovations in Information, Embedded and Communication System (ICIIECS'16)*, vol. 4, pp. 2039-2041, March 2016.

   (Published in Conference Proceedings)

2. A. Tripathi, H.S. Pannu, " High-order Fuzzy Time Series Forecasting Model for Advance Prediction of Temperature", in *IEEE International Conference on Inventive Computation Technologies (ICICT) 2016*, Coimbatore, Tamilnadu, India.

   (Accepted)

3. A. Tripathi, H.S. Pannu, "Fuzzy Time Series Forecasting Model for One-factor Time Series Data Set", in *International Conference on Recent Trend in Computer Science and Information Technology (RTCSIT) 2016, Amritsar, Punjab, India.*

   (Communicated)

Video Presentation Link

Below is the link of my video presentation:

https://www.youtube.com/channel/UCgXrHIulJmE2gSbfM6ad0Mw