

Use of fuzzy time series to predict the numbers of students enrolled in the Private University of Ebla (case study at the Faculty of Engineering in Aleppo)

¹Alaa Alden Al Mohamed,²Ali Jiblak

^{1,2} Department of business administration, Aleppo University, Aleppo, Syria.

Article Information

Received : July 18 2021
Revised : July 30 2021
Accepted : Aug 03 2021
Published : Aug 20 2021

Corresponding Author:

Alaa Alden Al Mohamed Email:
aldeenalaa53@gmail.com

Abstract— the fuzzy time series model was used to eliminate the ambiguity and volatility of the number of students enrolled in the Faculty of Engineering. At The Private University of Ebla from 2009 to 2020 in order to predict the number of students expected to be enrolled from 2021 to 2032, the program has been used (QM for windows, MAT LAB, Microsoft excel for conducting the necessary mathematical and statistical analyses, and the study concluded that the results of the prediction obtained after performing the fuzzy treatment are closer to the actual numbers recorded than the results of the prediction resulting without performing the fuzzy treatment process from 2016 to 2020, which showed the importance of applying the theory of misty groups and their efficiency in reducing the effects of environmental fluctuations faced by the university studied by controlling the level of the recorded and this prompted the researcher to use the entire period of 2009-2020) in order to predict a 12-year period because of its role in the quality of decisions related to human and financial resources, supporting the process of minimizing the university's costs and providing proposed solutions for the decision maker.

Keywords: *fuzzy time series, fuzzy inference, theories of fuzzy sets.*

Copyright © 2021: Alaa Alden Al Mohamed and Ali Jiblak. This is an open access distribution, and reproduction in any medium, provided Access article distributed under the Creative Commons Attribution License the original work is properly cited License, which permits unrestricted use.

Citation: Alaa Alden Al Mohamed and Ali Jiblak. "Use of fuzzy time series to predict the numbers of students enrolled in the Private University of Ebla (case study at the Faculty of Engineering in Aleppo)", Journal of Science, Computing and Engineering Research, 2(4), 184-189, 2021.

I. INTRODUCTION

Forecasting is the cornerstone that most companies rely on in the strategic planning process and the preparation of the necessary material and financial resources, investment, and others to be able to meet the expected demand [1]. The time series is considered as a sequence of data points, measured typically at a successive time spaced at uniform time intervals [2], in recent researches the use of non-stochastic models has become widespread, especially fuzzy time series which is distinguished from other traditional methods and that because of many reasons, first of all, it does not require assumptions that stochastic models have [3], secondly it provides semantic meaning for uncertain and fluctuating data [4], most of the time series encountered in real life should be considered as fuzzy time series due to the uncertainty that they contain and they should be analyzed with models appropriate to fuzzy set theory[5].

Fuzzy sets were proposed by Zadeh in 1965[6]. In 1993, Song and Chissom introduced a new prediction model based on the concept of fuzzy time series based on University of Alabama historical records. Song and Chissom introduced a time-constant fuzzy time-series model and a time-varying fuzzy time-series model based on fuzzy set theory to predict enrollment at the University of Alabama [4] Chen proposed a new prediction model based on highly ordered fuzzy time

series, and used it to predict enrollment at the University of Alabama. since then, several other studies have appeared [7,8,9,10].

The rest of this paper is ordered as follows: Section 2 A brief description of the problem and objectives of the research paper, Section 3 An overview of the time series and its model. Section 4 Practical study by applying to Ebla Private University. Finally, the conclusion is given in Section 5.

II. DESCRIPTION OF RESEARCH PROBLEM AND ITS OBJECTIVES

The problem of the research is that there is a discrepancy between the number of students accepted and the students planned to be admitted to the Faculty of Engineering at Ebla Private University from year to year, but often the number of students accepted is much less than the planned admission, which indicates that there is no clear scientific mechanism to predict student numbers They are expected to be accepted in each academic year, which in turn leads to a lack of knowledge of the material and financial resources that faculty and faculty members need to meet the demand for appropriate academic ability and determine them in proportion to the change in the number of students admitted each year.

This paper aims to predict the number of students who will enroll in the Faculty of Engineering in its two branches (Architecture - Informatics) at Ebla Private University in

Aleppo, so that it is more realistic and reduces the discrepancy between the number of students. Students and their planned admission for subsequent academic years were made using the Chen's Work method in a Fuzzy time series and applied to research data using the two programs (MATLAB and Microsoft Excel).

The importance of this paper lies in the following points:

- Addressing the ambiguity related to the data of the numbers of students enrolled in the College of Engineering using the fuzzy time series method (FTS, Chen's work).
- Attempting to contribute to raising the level of efficiency and accuracy in estimating the number of students expected to be registered in the coming years, which leads to better management of the college, which is reflected in the costs associated with the size of the numbers expected to be registered

III. PROPOSED FUZZY TIME SERIES METHOD

It is a set of values observed for a particular phenomenon during a specific period and tracking a specific pattern and the most common pattern is the increasing, decreasing, periodic, seasonality, and irregular fluctuations.

In many cases, especially in economics, the length of time series specified for analysis is short and limited to a hundred observations. For such time series, fuzzy models have been developed and successfully applied. According to the model proposed in [6], we will use a fuzzy logical relationship as a model and assume that the observations at time t are the cumulative results of the observation(s) in previous times.

Determine the number of time intervals by using the method of progress proposed by Huarng [6]

- Determine the number of equal-length intervals using advance method which is proposed by Huarng [6].
- Use Trapezoidal fuzzy numbers to define the fuzzy sets in fuzzy time series.
- Establish the Second Order Fuzzy logical relationship.

After mentioning the modifications that the researchers added, they summarized the steps of this algorithm into nine steps instead of six:

Step 1: Collect historical data for the time series.

Step 2: Set the minimum and upper limit of the historical data series for the request through the following two equations:

$$U = D_{min} - D_1, D_{max} + D_2$$

Where:

D_{min} = The lowest value in the data.

D_{max} = The highest value of the data.

D_1 And D_2 = Proper values to make universe of discourse.

Step 3: Determine the calculated interval length using the "average-based length" and can be determined as follows: The average difference between the first two periods' data is taken in historical time series data.

- Half of the average difference between the first two periods' data is taken in historical time series data.
- Select the length specified in table 1 defined by the Huarng researcher.
- Approximate the length of the interval according to the base value in the third step.

Table I: Average existing length method

Range	Base
0.1-1.0	0.1
1.1-10	10
11-100	100
101-1000	1000
1001-10000	10000

Step 4: Select the number of "m" commas using the appropriate length of the comma using the following formula:

$$m = (D_{max} + D_2 - D_{min} - D_1) / i$$

This step is intended to divide the historical data of the demand into equal-length intervals (u_1, u_2, \dots, u_m) we assume that the periods " m " $u_1 = [d_1, d_2]$, $u_2 = [d_2, d_3]$... $u_m - 2 = [d_{m-2}, d_{m-1}]$ are $\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_m$ therefore fuzzy sets can be determined as follows:

$$A_1 = \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \dots + \frac{0}{u_{m-1}} + \frac{0}{u_m}$$

$$A_2 = \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \dots + \frac{0}{u_{m-1}} + \frac{0}{u_m}$$

.

.

$$A_m = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \dots + \frac{0.5}{u_{m-1}} + \frac{0}{u_m}$$

Where:

U_i = Midpoints of intervals.

Step 5: In this step we replace the discrete fuzzy sets by fuzzy numbers with Trapezoidal. $\tilde{A} = a, b, c, d$

$$\mu A(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & b \leq x \leq c \\ \frac{c-x}{c-b}, & c \leq x \leq d \\ 0, & x > d \end{cases}$$

The Trapezoidal membership function consists of parameters include "a", "b", "c" and "d". Fuzzy numbers can be defined as follows:

$$A_1 = (d_0, d_1, d_2, d_3)$$

$$A_2 = (d_1, d_2, d_3, d_4)$$

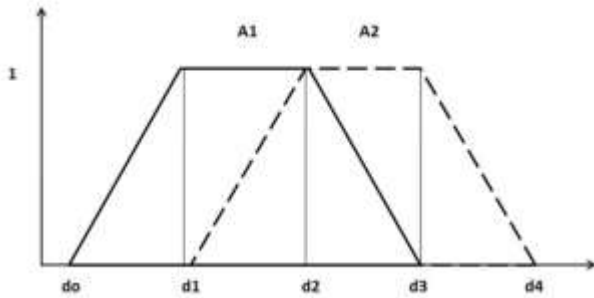
.

.

$$A_m = (d_{m-2}, d_{m-1}, d_m, d_{m+1})$$

The fuzzy numbers can be illustrated A_1 using A_2 the Trapezoidal membership functions of belonging in the following form:

Figure 1: Fuzzy numbers A_1 and A_2 using semi-deviant functions



Step 6: fuzzify time series data:

Fuzziness within this field means the process of determining the common links between the historical values of the demand data set and the fuzzy sets defined in the previous step, as each historical value of the request is fuzzy according to the fuzzy sets to which it belongs, if $f(t-1)$ is a specific historical time variable belonging to fuzzy number \tilde{A}_j , the $F_{A_k}(t-1)$ variable is quantified as \tilde{A}_j .

Step 7: Define fuzzy logical relationships:

First order fuzzy relationships can be generated using the following relationship:

$$\tilde{A}_j \rightarrow \tilde{A}_k$$

\tilde{A}_j Represents the left side and (\tilde{A}_k represents the right side of the fuzzy relationship, if the value of the time period belongs to the current period if it \tilde{A}_k represents the symbol), but if the value of the time period \tilde{A}_k belongs to the previous period it represents the symbol).

Step 8: Establish fuzzy relationship groups:

If a specific fuzzy group is associated with more than one set, the right side of it is merged to form so-called fuzzy relationship groups.

Step 9: Defuzzify the forecasted output:

In this step, two steps are included:

- Calculate the midpoints for length intervals.
- After determining the midpoints of the length intervals, we treat the fuzzy using the Median method by applying its own equation:

$$x_{med} = \frac{m_1 + m_2 + \dots + m_n}{n}$$

IV. PRACTICAL STUDY

After the researcher was briefed on several studies, including the study [1] and [11] Show that Chen model is the best method of a range of methods in this section the researcher will apply the proposed model to the students enrolled in the Faculty of Engineering (Architecture -Informatics) at the Private University of Ebla in the time series from 2009 to 2019.

Step 1: Collect historical data for students enrolled in the Faculty of Engineering at The Private University of Ebla in a time series from 2009 to 2019 as shown in the following table:

Table II: Actual student registrations in a time series at the Faculty of Engineering

Registered student numbers	The government's support
104	2009-2008
295	2010-2009
397	2011-2010
445	2012-2011
350	2013-2012
448	2014-2013
502	2015-2014
506	2016-2015
601	2017-2016
568	2018-2017
577	2019-2018
672	2020-2019

Step 2: Set the minimum and upper limit of the historical data series for the request through the following two equations:

$$U = [D_{min} - D_1, D_{max} + D_2]$$

Where:

$$D_{min} = 104$$

$$D_{max} = 672$$

$$D_1 = 4, D_2 = 28$$

$$U = [100, 700]$$

Step 3: Determine the calculated interval length using the "average-based length":

- The average difference between the first two periods data is taken in historical time series data (i.e. between 104-295) and the value equal to 191
- Half of the average difference between the first two periods data is taken in historical time series data and therefore the value is equal to 95.5
- Select the length specified in table 1 defined by huarng [6] and refer to the value of half the average difference between the data of the first two periods of 95.5 and therefore is located in the domain between 11 and 100 so the base value is 10.
- By reference to the value of half the average difference and taking into account the previous area, the value can be rounded to 100.

Step 4: Determine the number of periods "m" using the appropriate length of the period:

$$m = (D_{max} + D_2 - D_{min} - D_1) / i$$

$$= (700 - 100) / 100 = 6$$

Note that periods "m = 6" represent periods and u_1, u_2, \dots, u_6 knowledge as follows:

$$u_1 = [100,200]$$

$$u_2 = [200,300]$$

$$u_6 = [600,700]$$

This number of periods represents fuzzy sets from A_1, A_2 and A_6 using the following equations:

$$A_1 = \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \dots + \frac{0}{u_{6-1}} + \frac{0}{u_6}$$

$$A_2 = \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \dots + \frac{0}{u_{6-1}} + \frac{0}{u_6}$$

$$A_6 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \dots + \frac{0.5}{u_{6-1}} + \frac{0}{u_6}$$

Step 5: the fuzzy numbers use the Trapezoidal membership functions and can be represented as follows:

$$\tilde{A}_1 = (0,100,200,300)$$

$$\tilde{A}_2 = (100,200,300,400)$$

$$\tilde{A}_3 = (200,300,400,500)$$

$$\tilde{A}_4 = (300,400,500,600)$$

$$\tilde{A}_5 = (400,500,600,700)$$

$$\tilde{A}_6 = (500,600,700,800)$$

Step 6: In this step, we fuzzify time series data by setting data values corresponding to fuzzy numbers. For example, when we consider the number of student registrations in 2009 at 104, which falls within the field, the corresponding number of $u_1 = [100,200]$ 2009 is A_1 similarly the same as we have to determine the scope of each data value and the vague numbers of the data listed in the following table:

Table III: Fuzzy numbers

Fuzzy numbers	The government's support	Registered student numbers	The government's support
\tilde{A}_1	(100,200)	104	2009-2008
\tilde{A}_2	(200,300)	295	2010-2009
\tilde{A}_3	(300,400)	397	2011-2010
\tilde{A}_4	(400,500)	445	2012-2011
\tilde{A}_3	(300,400)	350	2013-2012
\tilde{A}_4	(400,500)	448	2014-2013
\tilde{A}_5	(500,600)	502	2015-2014
\tilde{A}_5	(500,600)	506	2016-2015
\tilde{A}_6	(600,700)	601	2017-2016
\tilde{A}_5	(500,600)	568	2018-2017
\tilde{A}_5	(500,600)	577	2019-2018
\tilde{A}_6	(600,700)	672	2020-2019

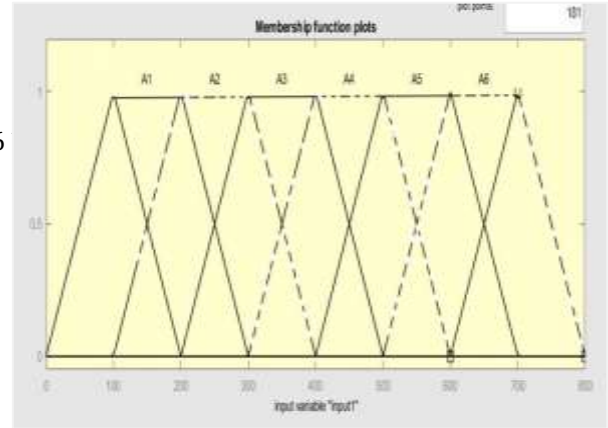


Figure 2: Fuzzy numbers using semi-perverted functions of belonging

Step 7: In this step we identify fuzzy logical relationships using Table 3.

Table IV: First-class fuzzy relations

$\tilde{A}_5 \rightarrow \tilde{A}_5$	$\tilde{A}_1 \rightarrow \tilde{A}_2$
$\tilde{A}_5 \rightarrow \tilde{A}_6$	$\tilde{A}_2 \rightarrow \tilde{A}_3$
$\tilde{A}_6 \rightarrow \tilde{A}_5$	$\tilde{A}_3 \rightarrow \tilde{A}_4$
$\tilde{A}_5 \rightarrow \tilde{A}_5$	$\tilde{A}_4 \rightarrow \tilde{A}_3$
$\tilde{A}_5 \rightarrow \tilde{A}_6$	$\tilde{A}_3 \rightarrow \tilde{A}_4$
.....	$\tilde{A}_4 \rightarrow \tilde{A}_5$

Step 8: Establish fuzzy relationship groups:

If a specific fuzzy group is associated with more than one group, the right side of it is merged to form the so-called fuzzy relationship groups "FIRG" and shows in table 4 that it is A_4 associated with each of the (A_5, A_3) and thus they are merged into a single fuzzy relationship group (G4), as in Table 5, which shows the sum of fuzzy relationships as follows:

Table V: 2nd Degree Hazy Logical Relationships Totals

Second-rate hazy relationships	Groups
$\tilde{A}_1 \rightarrow \tilde{A}_2$	1G
$\tilde{A}_2 \rightarrow \tilde{A}_3$	2G
$\tilde{A}_3 \rightarrow \tilde{A}_4$	3G
$\tilde{A}_4 \rightarrow \tilde{A}_5, \tilde{A}_3$	4G
$\tilde{A}_5 \rightarrow \tilde{A}_6, \tilde{A}_5$	5G
$\tilde{A}_6 \rightarrow \tilde{A}_5$	6G

Step 9: Defuzzify the forecasted output:

In this step, two steps are included:

- Calculate the midpoints for length intervals: We note from table 3 that the fuzzy set for 2009 is represented and we note from table (5) that the degree of affiliation of the misty group occurs in the period and that the midpoint of the period is equal, i.e. the midpoint of 2009 is equal to (250) as shown in table 6 $A_1 A_2 A_2 u_2 u_2 = [200,300] u_2 200 + 300/2 = 250$.

ii) After determining the midpoints for time periods, we treat the fuzzy using the Median method by applying its own equation mentioned above, for example, the fuzzy treatment for 2012:

$$X_{med} = \frac{350 + 550}{2} = 450$$

Table 6 shows the results of the midpoint calculation and the results of the fuzz treatment process.

Table VI: Remove fuzzy and set midpoints

years	Registered student numbers	Fuzzy relationship groups	Fuzzy relationship groups	Fuzzy Processing
2009	104	$\tilde{A}_1 \rightarrow \tilde{A}_2$	250
2010	295	$\tilde{A}_2 \rightarrow \tilde{A}_3$	350	250
2011	397	$\tilde{A}_3 \rightarrow \tilde{A}_4$	450	350
2012	445	$\tilde{A}_4 \rightarrow \tilde{A}_5, \tilde{A}_3$	350	550
2013	350	$\tilde{A}_3 \rightarrow \tilde{A}_4$	450	450
2014	448	$\tilde{A}_4 \rightarrow \tilde{A}_5, \tilde{A}_3$	350	550
2015	502	$\tilde{A}_5 \rightarrow \tilde{A}_6, \tilde{A}_5$	550	650
2016	506	$\tilde{A}_5 \rightarrow \tilde{A}_6, \tilde{A}_5$	550	650
2017	601	$\tilde{A}_6 \rightarrow \tilde{A}_5$	550	600
2018	568	$\tilde{A}_5 \rightarrow \tilde{A}_6, \tilde{A}_5$	550	650
2019	577	$\tilde{A}_5 \rightarrow \tilde{A}_6, \tilde{A}_5$	550	650
2020	672		550	600

To ensure that Post-Fuzzy Processing data shows results closer to real data than those extracted in the absence of Fuzzy Processing, the period between 2009-2014 (6 years) was taken to predict the other half of the time period between 2015-2020 as described in the following table:

Table VII: Comparison of post-processing and other fuzzy prediction results

year	The numbers are already registered	Post-Fuzzy Processing result	Results without the fuzzy Processing
2015	502	505.6	573.6
2016	506	538.9	553.0
2017	601	618.0	647.9
2018	568	605.6	752.3
2019	577	638.9	709.1
2020	672	726.0	815.0
	MSE	1591	13567
	RMSE	39.9	116.4
	MAPE	6%	18%

From the previous form, we note that post-fuzzy results are closer to the actually recorded numbers of results without fuzzy Processing, and it was confirmed by the results of the prediction accuracy measures, which are the MSE measure (the standard for measuring mean square error), and the RMSE

scale, which represents the standard for measuring root mean square error, MAPE, which represents the standard for measuring the absolute percentage, based on the following two equations [12]:

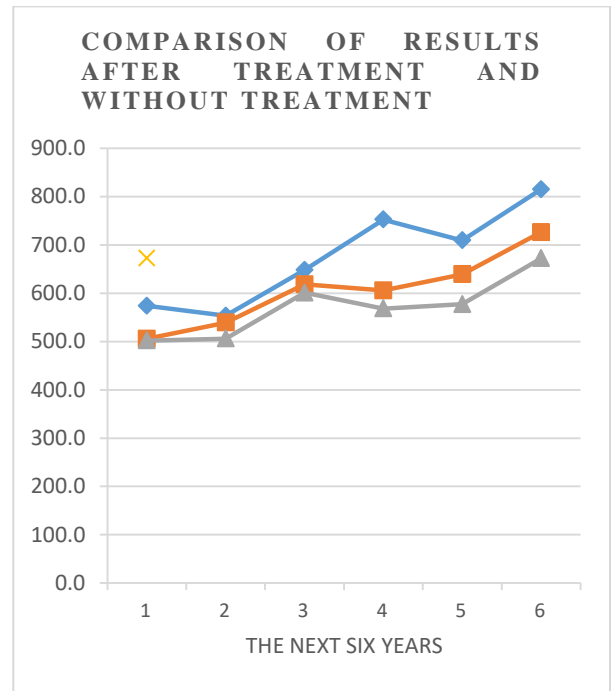


Figure 3: Comparison of the results of post-processing and other fuzzy predictions

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|forecast_t - actual_t|}{actual_t} * 100$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (forecast_t - actual_t)^2$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (forecast_t - actual_t)^2}$$

This confirms that the effect of the fuzzy Processing is very necessary for the prediction processes, especially if it is in a fuzzy environment dominated by uncertainty. After ensuring the effectiveness of fuzzy Processing, seasonal adjusted trend line method was used to predict the number of students enrolled for 12 years divided into three classes due to the fact that the application or enrollment in college is done quarterly during the year, using the Program (QM for Windows) Through which the best value of a and b was determined after the program was implemented for many tests, and by applying the equations mentioned in the ninth step of the seasonally adjusted trend line method, the results of the prediction process by determining the numbers expected to be recorded over the next 12 years were shown in the following table:

Table VIII: Numbers expected to be recorded in 12 years.

year	Expected recordings	year	Expected recordings
2020-2021	691.571	2027-2028	868.391

Use of fuzzy time series to predict the numbers of students enrolled in the Private University of Ebla (case study at the Faculty of Engineering in Aleppo)

<https://www.jscer.org>

2021-2022	702.271	2028-2029	860.739
2022-2023	701.17	2029-2030	947.026
2023-2024	776.723	2030-2031	951.451
2024-2025	785.331	2031-2032	940.524
2025-2026	780.955		
2026-2027	861.874	Total	9868.026

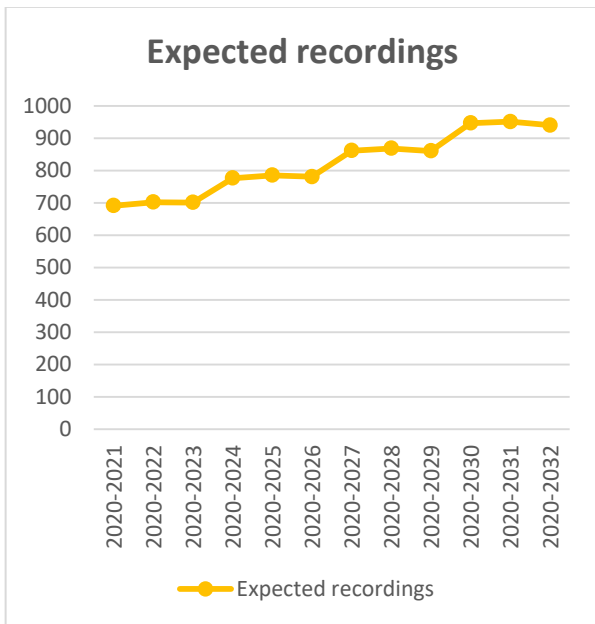


Figure 4: Chart the time series of numbers expected to be recorded within 12 years

We note from the previous figure that there is a stable increase in the number of students expected to be registered during the next time series, and this is corresponding to the increase in human resources and financial costs and profits in addition to the need to secure what is needed to accommodate the expected numbers to be recorded to a minimum.

V. CONCLUSION AND FUTURE WORK

It is clear from the previously inferred values of the numbers that can be enrolled in the Faculty of Engineering after fuzzy Processing that they are much closer to the values inferred without processing, and this demonstrates the importance of the method used in the forecasting process, namely, the hazy logic and the choice of Chen's work model as the best model among other models (Song Chissom, Hwan Chen, Chen, Burney) based on a study [1] stands out from other traditional forecasts and the need to apply it to all forecasts in other companies for their role in supporting decision-making and decision-making.

The data inferred from the numbers that can be recorded at the Faculty of Engineering in the next 12 years shows that the numbers are gradually increasing slightly, which means that

there is a slight increase in all administrative matters of human resources (staff and teachers) and the capacity of teaching facilities and other things, this In the light of the current variables and the absence of any new variable that can appear, such as the closure of engineering faculties in non-universities, advertising condensation or the financial capacity of the student and others that can increase or decrease the number of students, in this case, can be used sensitivity analysis after the fuzzy process to address such variables.

The need to expand the use of the logic of misty groups in various research, especially in terms of forecasting processes, both production and storage costs, or in determining the quantities of production expected to be produced in future periods, especially seasonal in production. Relying on Chen's work test, which is the best method in terms of accuracy standard based on previous studies in fuzzy treatment to remove uncertainty in subsequent studies. Developing the capacity of the Faculty of Engineering in proportion to the number of students planned to be admitted to the college for the coming years. Conducting a corresponding survey between private universities distributed to all provinces as they are subject to the same circumstances and comparing them and studying the factors that affect more closely the numbers of universities than other universities.

REFERENCES

- [1] Burney, S. A., Ali, S. M., & Khan, M. S. (2018). A novel high order Fuzzy Time Series forecasting method with higher accuracy rate. *IJCSNS*, 18(5), 9, pp.13-20.
- [2] Raza, S. Akhter, and S. M. Burney. *Time Series Analysis of High Speed Wireless Networks: Forecasting network load using time series models*. LAP Lambert Academic Publishing
- [3] Yolcu, O. C., Yolcu, U., Egrioglu, E., & Aladag, C. H. (2016). High order fuzzy time series forecasting method based on an intersection operation. *Applied Mathematical Modelling*, 40(19-20), pp.8750-8765.
- [4] Song, Q., & Chissom, B. S. (1993). Fuzzy time series and its models. *Fuzzy sets and systems*, 54(3), pp. 269-277.
- [5] L. A. Zadeh, *Fuzzy Sets*, Inform and Control, vol.8, pp.338-353, 1965
- [6] Huang, K. (2001). Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy sets and systems*, 123(3), pp. 387-394.
- [7] Chen, S. M. (2002). Forecasting enrollments based on high-order fuzzy time series. *Cybernetics and Systems*, 33(1), pp. 1-16.
- [8] Yu, H. K. (2005). A refined fuzzy time-series model for forecasting. *Physica A: Statistical Mechanics and its Applications*, 346(3-4), pp. 657-681.
- [9] Chen, S., & Chung, N. (2006). Forecasting enrollments of students by using fuzzy time series and genetic algorithms. *International journal of information and management sciences*, 17(3), pp.1-17.
- [10] Liu, H. T. (2007). An improved fuzzy time series forecasting method using trapezoidal fuzzy numbers. *Fuzzy Optimization and Decision Making*, 6(1), pp. 63-80.
- [11] Poulsen, J. R. (2009). *Fuzzy time series Forecasting*. Aalborg University Esbjerg.pp.
- [12] Carvalho J.G., Costa J.C.T. (2017) "Identification method for fuzzy forecasting models of time series." *Applied Soft Computing* 50, pp.166-182.