# Fuzzy Time Series Forecasting: Chen, Markov Chain and Cheng Models[1]

**Dr. Mona Mahmoud Samy Abo El Nasr**

Lecturer in Faculty of Commerce
Mansoura University

monamahmoud332015@gmail.com
m_eldeep@mans.edu.eg

## ABSTRACT

*This paper studies and reviews several procedures for Fuzzy Time Series analysis. Even though forecasting methods have advanced applications in the last few decades, Fuzzy Time Series are common and have a lot of interest because they do not require any statistical assumptions on time series data. Previous research has employed Fuzzy Time Series models to forecast enrollment statistics, stock prices, exchange rates, etc. The major goal of This work is a comparative study of some different methods of forecasting the Fuzzy Time Series among which are the Markov Chain, Chen, and Cheng for Ghabbour Autocars data. Seven statistical criteria have been used for investigating the accuracy of the models. All the calculations were performed using the R software system using the AnalyzeTS R package. The Markov-chain fuzzy time series model showed the highest performance (in all metrics); for instance, in RMSE, MAPE, and U-statistics are 0.013, 0.116, and 1.05 respectively.*

**Keywords:** Fuzzy Time Series, Markov Chain methods, Chen methods, Cheng methods, MSE, RMSE and R software.

[1] Received in 3/1 /2023, accepted in 18/2/2023.

# 1. INTRODUCTION

Forecasting is one of the most popular research topics, especially time series. Because time Series analysis uses statistical techniques to model and explain a time-dependent series of data points. In other meaning the process of making predictions about the future based on data from the past (Zulfikar et al., 2018; & Afnisah & Marpaung, 2020). Forecasting is an important and indispensable tool nowadays as it is used in many fields including business, industry, politics, economics, environmental sciences, medicine and social work. There are several popular conventional time series models, for example, autoregressive moving average (ARMA) and Box-Jenkins. These conventional statistical time series models are unable to anticipate data with linguistic values, but they can forecast problems resulting from new patterns. Also these models do not try to simulate nonlinear dynamics either (Kafi et al., 2018).

As a result, several unconventional models were developed, one of which is the fuzzy time series proposed by Song and Chissom (1994) (Chen and Hsu, 2004). An advantage of fuzzy time series is that it relies on simple and uncomplicated calculations compared to genetic algorithms, including Neural network models, and also can more efficiently use historical data since they can associate trend or cyclic components in fuzzy logical relationships. However, the neural network models are better at handling non-linear problems, but they are less effective due to lengthy training periods and their predicted values are less precise as a result of their inability to handle non-stationary data (Li and Cheng 2007; Vamitha et al., 2012). Several methods of fuzzy time series among them, the Chen method, the Cheng method, the Markov chain method, and others, each of these methods has a different methodology for forecasting (Rachim et al., 2020). The rest of this paper is outlined as follows: Section II presents the types of fuzzy time series methods. A real data set is analyzed in Section III and concluding remarks are included in Section IV.

## 2. FUZZY TIME SERIES

Fuzzy time series is a notion that can be employed to forecast issues or situations where historical data are produced in linguistic values. This also indicates that the prior data in fuzzy time series is linguistic data, while the real data as a result is made up of real numbers (Afnisah and Marpaung, 2020).

### 2.1 Fuzzy time series Chen

Steps in Fuzzy time series Chen strategy forecast, there are expressed in terms (Hariyanto and Udjiani, 2021):

**Step.1:** Define the historical data's minimal and maximum data ($Dmin$ and $Dmax$), as well as the universal discourse $\boldsymbol{U}$. The commonplace discourse is characterised by equation (1) as following:

$$\boldsymbol{U} = (Dmin - D1, Dmax + D2) \qquad (1)$$

where $D1$ and $D2$ are the constant defined by researcher.

**Step.2:** Creating the pause divide the discourse universe into many equal-length segments. To determine how many intervals are required can be using equation (2):

$$m = (Dmax + D2 - Dmin - D1) / i \qquad (2)$$

**Step.3:** Fuzzification of historical data and the definition of fuzzy sets.

**Step.4:** Establishing Fuzzy Logic Relations and Fuzzy Logic Relations Group. Step.5: When fuzzy logical relationships are gained from state $A2$, any state $A\,j$, where $j = 1, 2, \dots, n$, undergoes a change because $A2 \rightarrow A3, A2 \rightarrow A2, \dots, A2 \rightarrow A1$; as a result, the fuzzy logical relationships are immediately gathered into a fuzzy logical relationships group. Step.6: Defuzzyfication, As an illustration, For example, $F(t) = A1, A2, A3, \dots \dots \dots . An$ , the equation (3) to determine the final forecast value as follows:

$$\hat{Y}(t) = \frac{\sum_{i=1}^{n} \mathrm{mi}}{k} \qquad (3)$$

where $\hat{Y}$ is defuzzification and $\boldsymbol{mi}$ is the median of $Ai$.

## 2. 2 Fuzzy Time Series Markov Chain

Fuzzy Time Series and Markov Chains are combined. Using the transition probability matrix that is applied to the Markov Chain, this is done to acquire the highest likelihood. Fuzzy Time Series Markov Chain, provides good accuracy. The steps for fuzzy time series Markov Chain's prediction are as described in the following Tsaur et al., (2005), Tsaur (2012) and Hariyanto and Udjiani (2021): **Step.1:** Identify the discourse universe as per step 1's recommendation. **Step.2:** The discourse universe U is split up into many periods of the same duration as those described in Step. 2. **Step. 3:** Indicate how many fuzzy intervals there are. **Step.4:** The Fuzzy Ai set defines the Fuzzy set in the discourse universe U and declares the linguistic fuzzy set by a factor of one the share price. **Step.5:** data from the past is fuzzy. When a time series data is part of the U i interval, the data is fuzzified into the $Ai$. **Step.6:** Determine the fuzzy logic's relationship based on previous data and Identify groups from fuzzy logic relations and convert them to groups of fuzzy logic relations. **Step.7**: Using FLRG, it is possible to calculate a probability from one state moving on to the next for time series data. Therefore, Markov probability transition matrix with a dimension of n was employed to calculate the predicting value. We shall reap FLRG if state Ai changes to state $Aj$ and passes state $Ak, i\ , j\ =\ 1\ , 2\ , \dots \dots \dots ., n$. The following is the transition probability formula (Jasim et al., 2012; and Ramadani and Deviant, 2020):

$$Pi\,j \quad = \frac{Mij}{Mi} \quad , i\ , j \ = \ 1, 2\ , \dots \dots \dots . . , n. \qquad (4)$$

Where $Pij$ probability of a one-step transition from state $Ai$ to state $Aj$, $Mij$ number of one-step transitions from state $Ai$ to state $Aj$, and $Mi$ is equal to the amount of data in the $Ai$. **Step. 8:** Determine the forecasting value, $f\ (t)$, in accordance with the regulations Ramadani and Deviant, 2020:

If $Y(t-1) = \ Aj$ and there is a one-to-one fuzzy logical relationship, then $Aj \ \rightarrow Ak$ , the predicted future value is:

$$f(t) \ = \ m_k \qquad (5)$$

If the fuzzy logical relationship group $A_i$ is one to many (assuming instance, $Aj \ \rightarrow \ A1, A2, \dots \dots \dots \dots . . An, j = 1, 2, \dots \dots \dots \dots, n,$) while $Y(t -$

1) at the time of $(t-1)$ wrapped within side the Aj then predicted future value of F(t) is

$$F(t) = m1Pj1 + m2\,Pj2 +, \dots, +m_{j-1}\,P\,j(j-1) + Y(t-1)P\,jj + m_{j+1}\,P\,j(j+1) +, \dots, +mn\,\mathrm{P}_n \quad (6)$$

where: $mi$: the middle of the interval $U_i$, $Pjn$: possibility that state j will lead to state $n$. **Step.9:** Modify the predicting value's trend in accordance with the regulations.

If Ai converses with $Aj$ and makes a rising transition, where $Y\,(t-1) = Ai, Y\,(t) = A_j$, and $(Ai \rightarrow Aj, i > j)$, then the adjusted trend value will be:

$$\mathrm{Dt1} = \frac{l}{2} \quad (7)$$

If $Ai$ converses with $Aj$ and transitions to a decreasing state, where $Y\,(t-1) = A_i, Y(t) = Aj, (Ai \rightarrow Aj, i > j)$, the modified trend value will be as follows

$$\mathrm{Dt1} = \frac{-l}{2} \quad (8)$$

The modified trend value will be as follows if $Y(t-1) = Ai$ and at time $t, Y(t) = Ai + s, (1 \leq s \leq n - i)$ there is a jump forward transition, then

$$\mathrm{Dt2} = \frac{l}{2}\,s \quad (9)$$

The modified trend value will be: If $Y(t-1) = Ai$ and performs a jump backward transition at time t where $Y(t) = Ai - v, (1 \leq v \leq i)$, then

$$\mathrm{Dt2} = \frac{-l}{2}\,(\mathrm{v}) \quad (10)$$

**Step.10:** Determine the revised forecasting values [Song and Chissom, 1994]:

$$\acute{F}\,(t) = F\,(t) \pm Dt1 \pm Dt2 \quad (11)$$

## 2. 3 Fuzzy Time Series Cheng

The following list outlines the stages of forecasting using Cheng's Fuzzy Time Series model procedures (Ismail and Efendi 2011; and Afnisah and Marpaung; 2020): **Step.1:** Define the discourse universe and then divide it into many equal-sized intervals. The interval can be divided into smaller intervals by dividing 2 if the quantity of data in an interval exceeds the average value of the amount of data in each interval. **Step.2:** Create a Fuzzy set and Fuzzy the collected series data. **Step.3:** Create fuzzy logic relationships using past information. Two sets of Fuzzy patterns $Ai\ (t-1)\ , Aj(t)$ could be described as FLR $Ai\ \rightarrow\ Aj$ in data that has been fuzzified. **Step.4:** Identification of all relationships in Fuzzy logic relations. **Step.5:** In the set of Fuzzy logic relations, determine weights. **Step.6:** After that, add the weight to the standardized weighted matrix $Wn\ (t)$, which has the following equation.

$$Wn\ (t) = [\ w_1, \ldots \ldots ., w_k] = [\frac{\mathrm{w}\ 1}{\sum_{h=1}^{L} \mathrm{w}\ h}, \frac{\mathrm{w}\ 2}{\sum_{h=1}^{L} \mathrm{w}\ h}, \ldots, \frac{\mathrm{w}\ k}{\sum_{h=1}^{L} \mathrm{w}\ h}] \quad (12)$$

**Step.7:** Determining forecast outcomes. The weighted matrix W(t) which was standardized $Wn$ (t), is multiplied by the fuzzified matrix, where $L\ df\ = [m1\ , m2\ , \ldots \ldots \ldots \ldots \ldots\ , mk\ ]$, where $m_k$is the median value of each interval, to obtain the forecast value. The forecast is calculated as follows:

$$Ft = \ L\ df\ (t-1)\ \ Wn\ (t-1) \quad (13)$$

**Step.8:** Use the formula to adapt forecasting by implementing flexible forecasting: Adjustable forecasting

$$Y\ (t)\ =\ Y\ (t-1)\ +\ h\ * [Ft - Y\ (t-1)] \quad (14)$$

## 3. APPLICATION: GHABBOUR AUTO CARS DATA

Several methods were applied to the time series of the Ghabbour Auto cars data from the 1st of January 2021 to the 28th of August 2021 to forecast the Auto Ghabbour cars and all the calculations were performed using the R software system using the AnalyzeTS R package. The data used in this study available at https://www.mubasher.info/countries/eg/stock-prices. Fig.1 represents the method proposed time series plots the Ghabbour Auto cars values. Analysis of the model in order to choose the optimum model with the lowest error, it is crucial to evaluate predicting performance. To assess the validity of a model, various statistical tests can be used. The three statistical standards utilized in this study to verify the suggested model's predicting accuracy are mean error (ME), mean absolute error (MAE), Mean Percentage Error (MPE), mean square error (MSE), root of mean square error (RMSE), Mean Absolute Percentage Error (MAPE) and Thiel's U-statistics, which are given in Equations. (15–17):

$$\text{ME} = \frac{\sum_{i=1}^{N} E_i}{N}, \text{MAE} = \frac{\sum_{i=1}^{N}(|E_i|)}{N}, \text{MPE} = \frac{\sum_{i=1}^{N}(\frac{E_i}{Y_i}\times 100)}{N}; \quad (15)$$

$$\text{MSE} = \frac{\sum_{i=1}^{N}(E_i * E_i)}{N}, RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - F_i)^2}{N}}; \quad (16)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Y_i - F_i}{Y_i}\right| \times 100, \text{Thiel'sU} = \frac{\sqrt{\sum_{i=1}^{N}(Y_i - F_i)^2}}{\sqrt{\sum_{i=1}^{N} {Y_i}^2} + \sqrt{\sum_{i=1}^{N} {F_i}^2}}. \quad (17)$$

where the $Y_i$is 'observation series'. The $F_i$ is 'Forecasting series'. The $E_i$ is 'residual series'. The $N$ is size of sample.

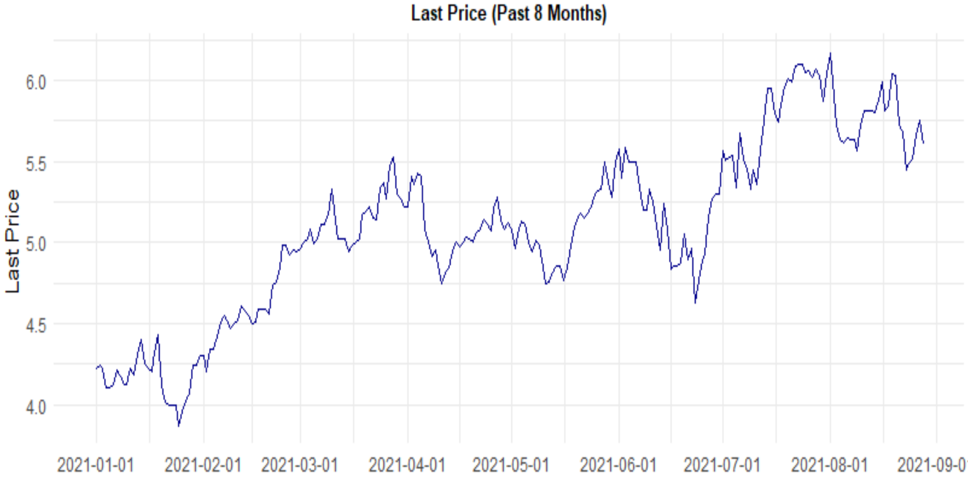


**Fig. 1: Time series plots the Ghabbour Auto cars.**

The step-by-step procedure proposed by Chen is listed as: a) Partition the universe of discourse into equally lengthy intervals. b) Define fuzzy sets on the universe of discourse. c) Fuzzify historical data. e) Identify fuzzy relationships (FLR's).

**Table 1: Summarize of Ghabbour Auto cars.**

| N Observation | Date | Last Price | Fuzzy set relationships |
|---|---|---|---|
| **1** | 1-1-2021 | 4.22 | A2- |
| **2** | 2-1-2021 | 4.24 | A2-->A2 |
| **3** | 3-1-2021 | 4.22 | A2-->A2 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| **238** | 26-8-2021 | 5.67 | A5-->A5 |
| **239** | 27-8-2021 | 5.75 | A5-->A5 |
| **240** | 28-8-2021 | 5.61 | A5-->A5 |
| **D.min** | | 3.87 | |
| **D.max** | | 6.17 | |
| $\boldsymbol{D_1}$ | | 0.37 | |
| $\boldsymbol{D_2}$ | | 0.33 | |

f)Establish fuzzy relationship groups (FLRG's). g) Defuzzify the forecasted output. Transform the Ghabbour Auto cars data into fuzzy numbers and determine the fuzzy logic relationships (FLRs) as can be observed in Table 2, which reveals the alterations of the observed Ghabbour Auto cars to be the linguistic values. Define the universe of discourse $U$ from Ghabbour Auto cars data. Since $U = [D.min - \boldsymbol{D_1}, D.max + \boldsymbol{D_2}]$, then, $U = [3.87 - 0.37, 6.17 + 0.33]$, thus, $U = [3.5, 6.5]$. Particularly, the U has been partitioned into 6 intervals with equal lengths as follows table 2.

**Table 2: Information about fuzzy sets.**

| N of Set | Set | Dow. | up | Mid-Point | Number |
|---|---|---|---|---|---|
| 1 | $u_1$ | 3.5 | 4 | 3.75 | 5 |
| 2 | $u_2$ | 4 | 4.5 | 4.25 | 34 |
| 3 | $u_3$ | 4.5 | 5 | 4.75 | 62 |
| 4 | $u_4$ | 5 | 5.5 | 5.25 | 84 |
| 5 | $u_5$ | 5.5 | 6 | 5.75 | 42 |
| 6 | $u_6$ | 6 | 6.5 | 6.25 | 13 |

Table 3 demonstrates that the Markov chain model outperforms the models that were examined, delivering superior results with smaller error values across all metrics than the error values of the competing models.

**Table 3: Statistical criteria of different models and Forecasting.**

| Model | ME | MAE | MPE | MSE | RMSE | MAPE | U-statistics |
|---|---|---|---|---|---|---|---|
| Chen | -0.017 | 0.156 | -0.43 | 3.079 | 0.035 | 0.188 | 1.721 |
| Cheng | -0.035 | 0.15 | -0.809 | 2.97 | 0.034 | 0.183 | 1.677 |
| Markov-chain | -0.008 | 0.089 | -0.247 | 1.756 | 0.013 | 0.116 | 1.057 |

**Forecasted Ghabbour Auto cars by different models**

| N Observation | Real data | Chen | Cheng | Markov-chain |
|---|---|---|---|---|
| 1 | 4.22 | - | - | - |
| 2 | 4.24 | 4.25 | 4.26 | 4.236856 |
| 3 | 4.22 | 4.25 | 4.26 | 4.255142 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 238 | 5.67 | 5.75 | 5.717 | 5.542 |
| 239 | 5.75 | 5.75 | 5.717 | 5.656 |
| 240 | 5.61 | 5.75 | 5.717 | 5.717 |

According to Table 3, the Markov-chain model yields the most accurate results and meets all statistical minimum standards. The RMSE, MAPE, and U statistics are, respectively, 0.013, 0.116, and 1.05. They show that the Markov chain model is efficient and allowed to make predictions with a respectable degree of accuracy.

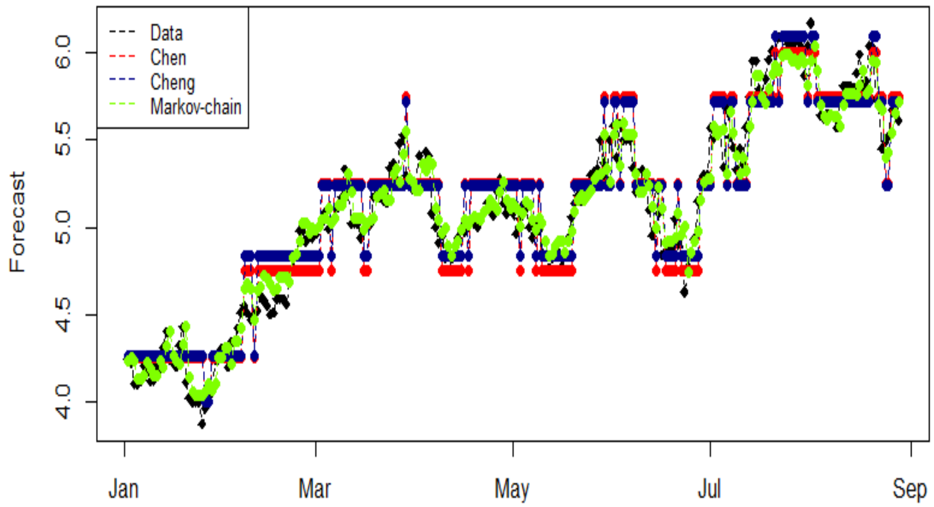


**Fig. 2. Actual Ghabbour Auto Data VS. Forecasted Ghabbour Auto data**

## 4.CONCLUSION

The Fuzzy Time Series Markov-chain model, as illustrated in Table 3, offers the most better measurement and is completely compatible over all statistical models. The respective RMSE, MAPE, and U-statistics are 0.013, 0.116, and 1.05. They suggest that the Markov-chain model is sufficient and capable of offering a respectable level of prediction accuracy. As shown in Fig. 2, the results are further supported by the fact that the original Ghabbour Auto Data and the predicted values based on the Fuzzy Time series Markov-chain model are quite better compared to each other. Figure 2 shows the forecasting patterns using the previously discussed techniques. the analysis of various evaluations in comparison to each other and a graphical display of the anticipated values produced by various methods. The Markov-chain fuzzy time series model is clearly better to other fuzzy time series models.

## REFERENCES

Afnisah, N., & Marpaung, F. (2020). A Comparison of the Fuzzy Time Series Methods of Chen, Cheng and Markov Chain in Predicting Rainfall in Medan. *In Journal of Physics: Conference Series*, 1462, (1), 012044. IOP Publishing.

Chen, S. M., & Hsu, C. C. (2004). A new method to forecast enrollments using fuzzy time series. *International Journal of Applied Science and Engineering*, 2(3), 234-244.

Hariyanto, S., & Udjiani, T. (2021). Fuzzy time series Markov Chain and Fuzzy time series Chen & Hsu for forecasting. *In Journal of Physics: Conference Series*, 1943(1), 012128. IOP Publishing.

Ismail, Z., & Efendi, R. (2011). Enrollment forecasting based on modified weight fuzzy time series. *Journal of Artificial Intelligence,* 4(1), 110-118.

Jasim, H. T., Jasim Salim, A. G., & Ibraheem, K. I. (2012). A novel algorithm to forecast enrollment based on fuzzy time series. Applications and Applied Mathematics: *An International Journal* (AAM), 7(1), 25.

Kafi, R. A., Safitri, Y. R., Widyaningsih, Y., & Handari, B. D. (2019). Comparison of weighted Markov chain and fuzzy time series Markov chain in forecasting stock closing price of company X. In AIP Conference Proceedings, 2168(1), 020033. AIP Publishing LLC.

Li, S. T., & Cheng, Y. C. (2007). Deterministic fuzzy time series model for forecasting enrollments. Computers & Mathematics with Applications, 53(12), 1904-1920.

Rachim, F., Tarno, T., & Sugito, S. (2020). Perbandingan Fuzzy Time Series dengan Metode Chen dan Metode SR Singh (Studi Kasus: Nilai Impor di Jawa Tengah Periode Januari 2014–Desember 2019). *Jurnal Gaussian*, 9(3), 306-315.

Ramadani, K. and Deviant, D. (2020). The Forecasting Model of Bitcoin price with Fuzzy Time Series Markov Chain and Chen logical Method, AIP Conference proceedings 2296.

Song, Q., and Chissom, B. S. (1994). Forecasting enrollments with fuzzy time series—Part II. Fuzzy sets and systems, 62(1), 1-8.

Tsaur, R. C. (2012). A fuzzy time Series-Markov chain model with an application to forecast the exchange rate between the Taiwan and US dollar. *International journal of innovative computing*, information and control, 8(7), 4931-4942.

Tsaur, R. C., Yang, J. C. O., & Wang, H. F. (2005). Fuzzy relation analysis in fuzzy time series model. Computers & Mathematics with Applications, 49(4), 539-548.

Vamitha, V., Jeyanthi, M., Rajaram, S., and Revathi, T. (2012). Temperature prediction using fuzzy time series and multivariate Markov chain. *International Journal of Fuzzy Mathematics and Systems*, 2(3), 217-230.

Zulfikar, R., Mayvita, A. P., and Purboyo, P. (2018). Econometric: Modification Technique of Fuzzy Time Series First Order and Time-Invariant Chen and Hsu to Increase the Forecasting Accuracy of Value Stock Index in Indonesia.

# التنبؤ بالسلاسل الزمنية الضبابية باستخدام نموذج Chen ونموذج سلاسل ماركوف ونموذج Cheng

د. مني محمود سامي أبو النصر

## ملخص البحث باللغة العربية

تدرس هذه الورقة وتستعرض عدة إجراءات لتحليل السلسله الزمنية الضبابية. على الرغم من أن طرق التنبؤ لها تطبيقات متقدمة في العقود القليلة الماضية ، إلا أن السلاسل الزمنية الضبابية شائعة ولديها الكثير من الاهتمام لأنها لا تتطلب أي افتراضات إحصائية بشأن بيانات السلاسل الزمنية. استخدمت الأبحاث السابقة نماذج السلاسل الزمنيه الضبابيه للتنبؤ بإحصائيات التسجيل وأسعار الأسهم وأسعار الصرف وما إلى ذلك. الهدف الرئيسي من هذا العمل هو دراسة مقارنة لبعض الطرق المختلفة للتنبؤ بالسلسلة الزمنية الضبابية ومن بينها Markov Chain و Chen و Cheng لبيانات شركة سيارات غبور و تم استخدام سبعة معايير إحصائية للتحقق من دقة النماذج. تم إجراء جميع الحسابات باستخدام نظام برنامج R باستخدام حزمة ( AnalyzeTS R) وقد أظهر نموذج السلسلة الزمنية الضبابية لسلسلة ماركوف أعلى أداء (في جميع المقاييس) ؛ على سبيل المثال، في إحصائيات RMSE و MAPE و U هي 0.013 و0.116 و1.05 على التوالي.

**الكلمات الدالة**: السلاسل الزمنيه الضبابيه، سلاسل ماركوف، متوسط مربعات الخطأ ولغة R.

### Suggested Citation according to APA Style

Abo El Nasr, M.M. (2022). Fuzzy Time Series Forecasting: Chen, Markov Chain and Cheng Models, Journal *of Alexandria University for Administrative Sciences,* Faculty of Commerce, Alexandria University 60(2),33-45.