

Forecasting Using Point-valued Time Series and Fuzzy-valued Time Series Models

Rosnalini Mansor^{1*}, Bahtiar Jamili Zaini²

^{1,2}*School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia.*

E-mail: rosnalini@uum.edu.my

Abstracts: The point-valued time series (PTS) is simply about one value in each time or period of the data, but when the data have two values at each time, the suitable time series is called the interval-valued time series (ITS). An example of ITS is the daily close and open stock prices. If the time series data are of typed linguistic data, for example, "low increase", "medium increase" and "high increase", the time series data are called a fuzzy-valued time series or being well-known as fuzzy time series (FTS). The aim of this study is to compare the PTS and FTS models for forecasting stock prices in the stock market. The stock market is one of the essential investments in the economy. The movement of stock price can be leading or declining which respectively expands or contracts a country's economy. The movement may also represent the scenario or event happening in a company. Therefore, forecasting stock price movement is highly important since it will assist investors and sellers to make planning in their investment's decision. Therefore, the objectives of this study are to identify the best forecasting models in point-valued time series (PTS) and fuzzy time series (FTS) based on forecasting measurement error. Eight stock price forecasting models consisting of four models from PTS and four models from FTS. Meanwhile, four forecasting measurement errors were discussed in the analysis as a criterion for choosing the best forecasting model. A set of daily historical data from the Bursa Malaysia website was used as a basis for analysis. The finding shows the simple exponential smoothing model is the best PTS model. In the meantime, the Cheng model based on Sturges' rule is the best FTS model. However, among these two types of models, the Cheng model is found to be the best model with forecasting measurement errors of 0.0001 (MSFE), 0.0108 (RMSFE) and 1.1918 (MAPFE). The results reveal that besides the PTS model, the FTS model is an alternative model to forecast stock market movement. Moreover, these FTS and PTS models can also be applied to solving other forecasting problems.

Keywords: Forecasting, Fuzzy Time Series, Point-Valued Time Series, Stock Price.

1. INTRODUCTION

Time series forecasting is a process of predicting future values or situations based on historical data. Two types of time series forecasting can be generalized; univariate and multivariate time forecasting techniques. The univariate time series forecasting technique is the most widely used for short-term forecast horizons. It is due to some of the decision-makers wanting to obtain the results in time shortly. However, very limited previous studies explain the type of time series data values. It is important to understand the type of time series data value since different forecasting method are for different time series data value. A stock market is a type of exchange where a seller offers a share of the stock to the buyer representing the investment in the partial ownership of the business. Trading activities in stock market public companies is an important part of the economy. Many advantages that stock trading has in a certain business. For example, trading could encourage and allow businesses to increase their capital and solve some liabilities. From the trading outcome, a company has an opportunity to expand its business activities such as launching new products and expanding its operations. The above-mentioned is a benefit that a stock market exchange could bring to a company or a seller. Meanwhile, investors, have a chance to get profit from gains in stock value. Some companies also give dividends or bonuses to their investors. However, stock price movements affect investors and business confidence. This will affect the economy inclusively in the company and the country. The relationship also works in the other direction. Where the economic circumstances often impact stock markets. Hence, these factors and implications of the current environment affect the decision that has significant risk and uncertainty, and often decision-makers do not obtain enough information for that purpose [1]. Therefore, stock price forecasting is important and relevant in economic and financial planning. Over the years, stock price forecasting attracts many investors and researchers. The intention is mostly to develop and test the stock price model [2]. However, analysing the stock market movements is highly challenging since the market is influenced by many factors besides the behaviour of the stock market data itself. For instance, the markets are dynamic [3] and the historical data are large of uncertainty [4] such as unusual past shocks [5].

The stock market time series data can be categorised into point-valued time series (PTS), interval-valued time

series (ITS) [6-7] and fuzzy time series [8,9]. The main difference between these three methods is as follows. The PTS is simply about one value in the sequence time or period of the data, but when the data have two values at each time in sequence, the suitable time series is called the ITS. An example of ITS is the daily close and open stock prices. If the time series data are of typed linguistic data in sequence time, for example, “low increase”, “medium increase” and “high increase”, the time series data are called FTS. A number of FTS applications in forecasting such as in electricity load demand [8, 10-14], energy-water efficiency [15,16], the stock price [9] and students enrolment [8]. However, the scope of this study is illustrated how PTS and FTS models only were used to forecast stock price movement.

2. TIME SERIES DATA

In General, The Time Series Data is the Data Value in Sequence Time. Following Are Some Definitions of PTS, Its and FTS Data.

2.1. Interval-valued time series (ITS)

ITS is a chronological sequence of interval-valued variables; the numerical value of the variable in each instant of time t , ($t = 1 \dots n$) is expressed as a two-dimensional vector $[X_t^L, X_t^U]$ with the elements in n representing the lower bound X_t^L and upper bound X_t^U , with $X_t^L \leq X_t^U$. Therefore, an ITS is $[X_t] = [X_t^L, X_t^U]$ for $t = 1, \dots, n$ where n denotes the number of intervals of the time series (sample size) [17,18]. Since X^{ITS} is the variable in a time series, the ITS of X^{ITS} can be represented in Equation (1):

$$X^{ITS} = [X_1^L, X_1^U], [X_2^L, X_2^U], \dots, [X_n^L, X_n^U] \quad (1)$$

The ITS can correspond to the variability such as minimum and maximum temperature in a certain area from a certain time series period, and open and close price of the stock market in a certain company based on a certain period of trading time series period [19]. Figure 1 display the example graph of ITS of historical stock prices from Maia et. al (2008) study [19]. The graph consists of open and close price over the trading days.

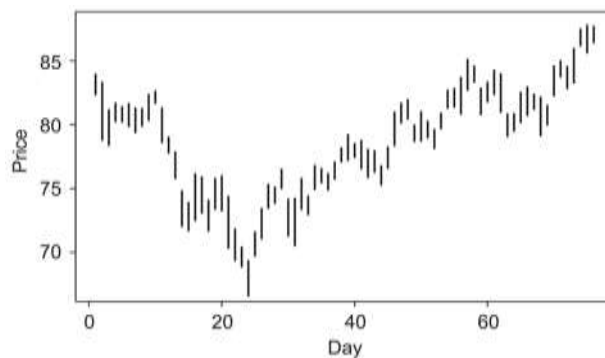


Figure 1. Example of ITS stock price graph.

Source [19].

2.2. Point-valued time series (PTS)

PTS is a sequence of single numerical data points occurring in successive order over time. Let X^{PTS} be the variable in a time series, for $t = 1, \dots, n$ where n is the total number of data points. Hence, PTS of X^{PTS} is represented in Equation (2):

$$X^{PTS} = X_1, X_2, X_3, \dots, X_n \quad (2)$$

In this paper, we named the numerical data value in PTS as a crisp value. The example of PTS are electricity load demand [20] and stock price [21]. Figure 2 display the example graph of PTS of historical stock prices from Zaini et. al (2020) study [21]. The graph consists monthly stock price.

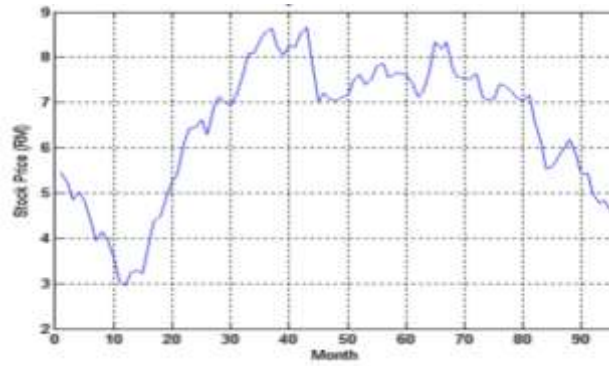


Figure 2. Example of PTS stock price graph.

Source [21].

2.3. Fuzzy-valued time series (FTS)

Fuzzy-valued time series or being well-known as Fuzzy Time Series (FTS) models have become increasingly popular in recent years because of their ability to deal with time series data without the need for validating any theoretical assumptions. In general, the FTS is for linguistic variable. However, the PTS can convert to FTS by fuzzify each PTS crisp value by replacing each value with the equivalent linguistic term to obtain the FTS [22]. Fuzzification is the process of converting crisp values into fuzzy values by identifying possible uncertainties or variations in the crisp values. This conversion is represented by the membership functions of fuzzy sets. Therefore, FTS is a time series of linguistic term in a sequence of time.

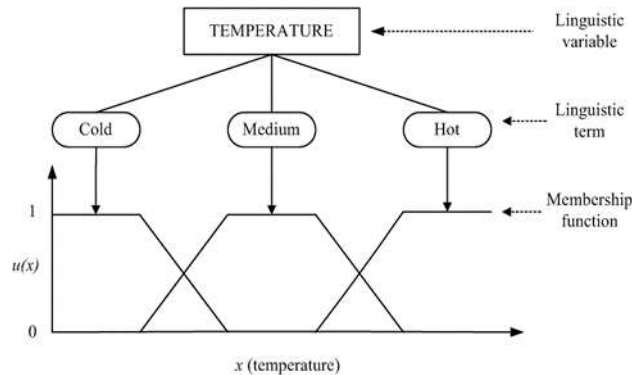


Figure 3. Linguistic variable in FTS

Figure 3 explains the simple relationship between linguistic variables, linguistic terms, and membership function in FTS. *Temperature* is the linguistic variable with three linguistic terms; ‘*Cold*’, ‘*Medium*’, and ‘*Hot*’. Let x be the crisp temperature value, then fuzzify x using a certain membership function to form the linguistic term. The membership function $\mu(x)$ represents the degree of membership corresponding to the linguistic terms accordingly.

Let Y_t^{FTS} , where $t = 1, 2, 3, \dots$, a subset of real numbers, be the universe of discourse by which fuzzy sets $f_i(t), (i = 1, 2, 3, \dots)$. If X_t^{FTS} is a collection of $f_1(t), f_2(t), \dots$ then X_t^{FTS} is called a FTS, defined by universal discourse $Y_t^{FTS}, t = 1, 2, 3, \dots$.

3. METHODOLOGY

This study applies the PTS and FTS forecasting model into Bursa Malaysia data. This section describes the data used in this study, the procedure and the forecasting model. Three phases in the forecasting procedure were performed; data pre-processing, models development and models evaluation.

3.1. Phase 1: Data Pre-processing

Bursa Malaysia is an exchange holding company established in 1973 and listed in 2005. Bursa Malaysia is one of the largest bourses in ASEAN. It is hosting more than 900 companies across 60 economic activities (<http://www.bursamalaysia.com/corporate/about-us>). This study adopts a set of historical data from Bursa Malaysia website. The daily data of two month-period data were collected from 6 August 2018 to 25 September 2018. From the 33 data points, 22 first data points were used in modelling part and the rest of the data were for evaluation part.

3.2. Phase 2: Model development

Eight models were developed based on modelling part of the historical data. Table 1 shows the models used in this study. Model A, B, C and D are PTS models while D, E, F and G are FTS models. There are many types of PTS and FTS forecasting models. However, this study chooses Naïve, Moving Average and Simple Exponential Smoothing since the data shows stationary pattern. Meanwhile, Chen and Cheng forecasting models is selected since these FTS forecasting model are well-known and widely used models in FTS research area.

Table 1. List of forecasting models in this study.

| Model | Model | Forecasting model |
|-------|------------------------------|--|
| A | Simple Naïve | $y_{t+1} = y_t$ |
| B | Moving Average (3) | $y_{t+1} = \frac{\sum(\text{most recent 3 data values})}{3}$ |
| C | Moving Average (5) | $y_{t+1} = \frac{\sum(\text{most recent 5 data values})}{5}$ |
| D | Simple Exponential Smoothing | $y_{t+1} = \alpha y_t + (1 - \alpha)y_{t-1}, 0 \leq \alpha \leq 1$ |
| E | Chen model (5 intervals) | Chen model based on Stuges' rule |
| F | Cheng model (5 intervals) | Cheng model based on Stuges' rule |
| G | Chen model (6 intervals) | Chen model based on 6 equal interval lengths |
| H | Cheng model (6 intervals) | Cheng model based on 6 equal interval lengths |

The general steps in Chen model are [23-24]:

Step 1: Define fuzzy sets based on the universe of discourse and fuzzify the historical data.

Step 2: Define the universe of discourse and intervals for rules abstraction.

Step 3: Fuzzify observed rules.

Step 4: Establish fuzzy logical relationships and group them based on the current states of the data of the fuzzy logical relationships.

Step 5: Forecast

Step 6: Defuzzify

The general steps in Cheng model are [24]:

Step 1: Define fuzzy sets based on the universe of discourse and fuzzify the historical data.

Step 2: Define the universe of discourse and intervals for rules abstraction.

Step 3: Fuzzify observed rules.

Step 4: Establish fuzzy logical relationships and group them based on the current states of the data of the fuzzy logical relationships.

Step 5: Assign weight

Step 6: Forecast

Step 7: Defuzzify

From the above-mentioned model steps, we can see the main different between Chen and Cheng models is that, Cheng model included weights in their model. The Stuges' rule in Model E and F is the method to identify the number of intervals in FTS model. Stuges' rule is a familiar rule in constructing histogram. One of the steps in Stuges' rule is to identify the number of classes in histogram chart. Therefore, the number of intervals in this study is presented in Equation (3).

$$k = 1 + 3.3\log(n) \quad (3)$$

where, k is the number of intervals, and n is the number of modelling data points.

3.3. Phase 3: Model evaluation

This study applied four well known measurements forecasting error criteria for best model selection [25]. The measurements of forecasting error are absolute percentage (APFE), mean absolute percentage (MAPFE), mean squared (MSFE) and root mean squared (RMSFE) of forecasting error. The respective formula of the forecasting error are shown in Equation (4) – (7).

$$MSFE = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n} \quad (4)$$

$$RMSFE = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}} \quad (5)$$

$$MAPFE = \frac{\sum_{t=1}^n \left| \frac{(y_t - \hat{y}_t)}{y_t} \right| \times 100}{n} \quad (6)$$

$$APFE = \left| \frac{(y_t - \hat{y}_t)}{y_t} \right| \times 100 \quad (7)$$

where y_t is the close price actual value

\hat{y}_t is the forecasted close price value

n is the number of evaluation data points

4. RESULTS AND DISCUSSION

This section focuses on the results from model evaluation part only. Three measurement forecast error were presented in Table 2. Among PTS models, model D was the best model with 0.0002 (MSFE), 0.0129 (RMSFE) and 1.3206 (MAPFE). Meanwhile, model F was the best FTS model with 0.0001 (MSFE), 0.0108 (RMSFE) and 1.1918 (MAPFE). However, by comparing all the models in this study, model F is the best model with the lowest error values. Table 3 shows the comparison between the actual and forecasted value of stock price in evaluation part of the data for model F, where the range of absolute percent error is from 0% to 3.26%. Figure 2 illustrates the graphs of actual, forecast by Chen Model, and forecast by Cheng Model.

Table 2. Measurement forecasting error in evaluation part of data.

| Model | Model's name | MSFE | RMSFE | MAPFE |
|-------|------------------------------|---------------|---------------|---------------|
| A | Simple Naïve | 0.0002 | 0.0141 | 1.6199 |
| B | Moving Average (3) | 0.0002 | 0.0141 | 1.5027 |
| C | Moving Average (5) | 0.0002 | 0.0147 | 1.5933 |
| D | Simple Exponential Smoothing | 0.0002 | 0.0129 | 1.3206 |
| E | Chen model (5 intervals) | 0.0003 | 0.0178 | 2.1556 |
| F | Cheng model (5 intervals) | 0.0001 | 0.0108 | 1.1918 |
| G | Chen model (6 intervals) | 0.0004 | 0.0188 | 2.1233 |
| H | Cheng model (6 intervals) | 0.0003 | 0.0178 | 2.0240 |

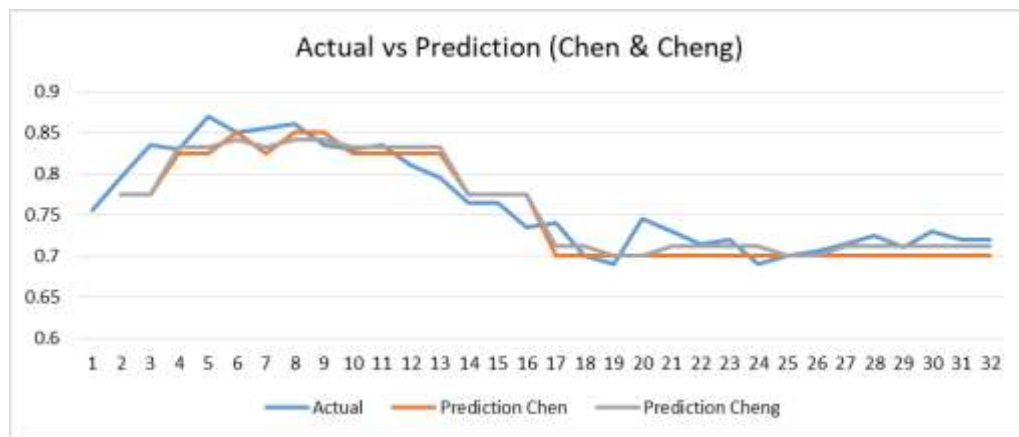


Figure 2. Comparison of actual and Sturges' Rule Chen and Cheng models.

Table 3. Evaluation part of measurement error based on Model F.

| Date | time (t) | Close price | | Error | APFE |
|-------------|----------|------------------|------------------------------|---------|--------|
| | | Actual value (y) | Forecast value (\hat{y}) | | |
| 7-Sep-2018 | 23 | 0.7200 | 0.7125 | 0.0075 | 1.0417 |
| 12-Sep-2018 | 24 | 0.6900 | 0.7125 | -0.0225 | 3.2609 |
| 13-Sep-2018 | 25 | 0.700 | 0.7000 | 0.0000 | 0.0000 |
| 14-Sep-2018 | 26 | 0.7050 | 0.7000 | 0.0050 | 0.7092 |
| 18-Sep-2018 | 27 | 0.7150 | 0.7125 | 0.0025 | 0.3497 |
| 19-Sep-2018 | 28 | 0.7250 | 0.7125 | 0.0125 | 1.7241 |
| 20-Sep-2018 | 29 | 0.7100 | 0.7125 | -0.0025 | 0.3521 |
| 21-Sep-2018 | 30 | 0.7300 | 0.7125 | 0.0175 | 2.3973 |
| 24-Sep-2018 | 31 | 0.7200 | 0.7125 | 0.0075 | 1.0417 |
| 25-Sep-2018 | 32 | 0.7200 | 0.7125 | 0.0075 | 1.0417 |

5. CONCLUSION

There is no one forecasting model is the best for all pattern of the time series data. This study discusses three type of time series data application in forecasting area of study; interval-valued time series (ITS), point-valued time series (PTS) and fuzzy time series (FTS). However, the aim of this study is exploring the performance of forecasting error using two different nature of time series; PTS and FTS only. Three phases in the procedure were discussed as a framework of stock market forecasting methodology. The forecasting methodology shows that it is an acceptable procedure to compare different type of time series data (PTS and FTS) with the objective to find out which time series models would give the accurate stock price forecast value.

However, each model has its advantages and disadvantages. Both PTS and FTS models in this study are simple to use, but FTS model is easier especially in interpreting the results as this model uses natural human language by describing something in linguistic value. Moreover, FTS model also has the ability to forecast PTS data by converting the PTS data into FTS data through the fuzzification step. After that, the fuzzy data could be converted to crisp data using through the defuzzification step. For future study, we suggest to implement FTS model using ITS data in solving a suitable forecasting problem.

6. ACKNOWLEDGEMENT

This research was supported by the Ministry of Higher Education (MoHE) of Malaysia through Fundamental Research Grant Scheme (FRGS/1/2020/STG06/UUM/03/1). The authors also want to thank Research and

Innovation Management Centre (RIMC), Universiti Utara Malaysia (UUM) for the assistances (S/O Code: 14878).

REFERENCES

- [1] Zaini, B.J, Sarkawi, M.N, Mansor, R., Bohari, A.M & Mustafa, M.A (2022). Pembuatan keputusan dan konteks syarikat industry kecil dan sejahtera in Kelestarian Perniagaan dan Cabaran Abad ke21 (Cetakan Pertama), Kedah. UUM Press, 194-201.
- [2] Shah, D., Isah, H., & Zulkernine, F. (2019). Stock market analysis: A review and taxonomy of prediction techniques. *International Journal of Financial Studies*, 7(2), 26.
- [3] Selvamuthu, D., Kumar, V., & Mishra, A. (2019). Indian stock market prediction using artificial neural networks on tick data. *Financial Innovation*, 5(1), 1-12.
- [4] Ntemi, M., & Kotropoulos, C. (2020). A dynamic dyadic particle filter for price prediction. *Signal Processing*, 167, 107334.
- [5] Bui, T. (2020). Stock holding decisions of foreign investors in emerging stock markets: A case study in Vietnam. *Management Science Letters*, 10(3), 625-630.
- [6] Maciel, L. (2019). Financial interval time series modelling and forecasting using threshold autoregressive models. *International Journal of Business Innovation and Research*, 19(3), 285-303.
- [7] Sun, Y., Han, A., Hong, Y., & Wang, S. (2018). Threshold autoregressive models for interval-valued time series data. *Journal of Econometrics*, 206(2), 414-446.
- [8] Mansor, R., Othman, M., & Kasim, M. M. (2017). Fuzzy reasoning based time series forecasting using weighted subsethood model. *Advanced Science Letters*, 23(9), 9094-9097.
- [9] Mansor, R., Zaini, B. J., & Yusof, N. (2019, August). Prediction stock price movement using subsethood and weighted subsethood fuzzy time series models. In *AIP Conference Proceedings* (Vol. 2138, No. 1, p. 050018). AIP Publishing LLC.
- [10] Mansor, R., Mat Kasim, M., & Othman, M. (2016). A weighted subsethood mamdani fuzzy rules based system rule extraction (MFRBS-WSBA) for forecasting electricity load demand—a framework. *Journal of Telecommunication, Electronic and Computer Engineering*, 8(8), 97-102.
- [11] Mansor, R., Kasim, M. M., & Othman, M. (2016, August). The framework of weighted subset-hood Mamdani fuzzy rule based system rule extraction (MFRBS-WSBA) for forecasting electricity load demand. In *AIP Conference Proceedings* (Vol. 1761, No. 1, p. 020061). AIP Publishing LLC.
- [12] Mansor, R., Kasim, M. M., & Othman, M. (2020). Weighted subsethood segmented fuzzy time series for moving holiday electricity load demand forecasting. *Journal of Advanced Research in Dynamical and Control Systems*, 12(Specia), 686-700.
- [13] Mansor, R., Zaini, B. J., Othman, M., & Kasim, M. M. (2018, June). The effect of interval length in weighted subsethood fuzzy time series. In *AIP Conference Proceedings* (Vol. 1974, No. 1, p. 040017). AIP Publishing LLC.
- [14] Mansor, R., Zulkifli, M., Yusof, M. M., Ismail, M. I., Ismail, S., & Yin, Y. C. (2014, December). Performance of fuzzy approach in Malaysia short-term electricity load forecasting. In *AIP Conference Proceedings* (Vol. 1635, No. 1, pp. 817-824). American Institute of Physics.
- [15] Othman, S., Mansor, R., & Ahmad, F. (2022). Weighted Subsethood Fuzzy Time Series towards Energy-Water Efficiency for Water Treatment Plant.
- [16] Othman, S., & Mansor, R. (2022, August). A comparison of chen and cheng fuzzy time series methods in energy-water efficiency forecasting at water treatment plant. In *AIP Conference Proceedings* (Vol. 2472, No. 1, p. 050021). AIP Publishing LLC.
- [17] Zhang, D., Li, Q., Muger, A. W., & Ling, L. (2020). A hybrid model considering cointegration for interval-valued pork price forecasting in China. *Journal of Forecasting*, 39(8), 1324-1341.
- [18] Xiong, T., Bao, Y., & Hu, Z. (2014). Interval forecasting of electricity demand: A novel bivariate EMD-based support vector regression modeling framework. *International Journal of Electrical Power & Energy Systems*, 63, 353-362.
- [19] Maia, A. L. S., de Carvalho, F. D. A., & Ludermir, T. B. (2008). Forecasting models for interval-valued time series. *Neurocomputing*, 71(16-18), 3344-3352.
- [20] Ismail, S., Zulkifli, M., Mansor, R., Mat Yusof, M. and Ismail, M. I. (2016) The role of exploratory data analysis (EDA) in electricity forecasting. *Pertanika Journal of Social Science and Humanities*, 93-100.
- [21] Zaini, B. J., Mansor, R., Yusof, Z. M., Gabda, D., & Seng, W. K. (2020). Comparison of double exponential smoothing for Holt's method and artificial neural network in forecasting the Malaysian banking stock markets. *J. Appl. Stat.*, 48, 2231-2238.
- [22] Selim, K. S., & Elanany, G. A. (2013). A new method for short multivariate fuzzy time series based on genetic algorithm and fuzzy clustering. *Advances in Fuzzy Systems*, 2013, 10-10.
- [23] Chen, S. M. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy sets and systems*, 81(3), 311-319.
- [24] Cheng, C. H., Chen, T. L., Teoh, H. J., & Chiang, C. H. (2008). Fuzzy time-series based on adaptive expectation model for TAIEX forecasting. *Expert systems with applications*, 34(2), 1126-1132.
- [25] Lazim, M. A. (2014) *Introductory business forecasting- a practical approach* (3rd Edition). Kuala Lumpur: UiTM Press.

DOI: <https://doi.org/10.15379/ijmst.v10i2.1168>

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.