Weighted Subsethood Fuzzy Time Series towards Energy-Water Efficiency for Water Treatment Plant

Shahidah Othman^{1,*}, Rosnalini Mansor¹, Fakhrurazi Ahmad²

¹School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia

²Syarikat Air Darul Aman Sdn. Bhd., No. 892, Jalan Sultan Badlishah Bandar Alor Setar, 05000 Alor Setar, Kedah, Malaysia

Received August 27, 2021; Revised December 12, 2021; Accepted December 22, 2021

Cite This Paper in the following Citation Styles

(a): [1] Shahidah Othman, Rosnalini Mansor, Fakhrurazi Ahmad, "Weighted Subsethood Fuzzy Time Series towards Energy-Water Efficiency for Water Treatment Plant," *Environment and Ecology Research*, Vol. 10, No. 2, pp. 182 - 192, 2022. DOI: 10.13189/eer.2022.100207.

(b): Shahidah Othman, Rosnalini Mansor, Fakhrurazi Ahmad (2022). *Weighted Subsethood Fuzzy Time Series towards Energy-Water Efficiency for Water Treatment Plant*. *Environment and Ecology Research*, 10(2), 182 - 192. DOI: 10.13189/eer.2022.100207.

Copyright©2022 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract Forecasting is the process that uses statistical tools to make predictions based on past and present data. In the water treatment plant systems, forecasting is useful to increase the efficiency of energy-water units by forecasting the amount of energy used in the current and future time. Fuzzy time-series (FTS) employs an algorithm or statistical model to predict the future condition. By analysing the patterns of data, fuzzy time-series forecasting can forecast the future condition of a system. The FTS approaches have been increasingly popular in recent years, as there are so many methods available for a researcher to apply to forecast the future based on the needs of the user. This paper proposed the modification of the classical fuzzy time series by applying weighted subsethood Fuzzy Time Series (WeSuFTS) in the water management industry. 2 sets of monthly data, which are water production and energy consumption data from January 2017 to December 2020, are used in this study. To measure the performance of the proposed model, the forecast results of the proposed model were compared with the forecast results of Chen and Cheng models. This study compares 3 different models using Absolute Percentage Error (APE) and Mean Absolute Percentage Error (MAPE). The WeSuFTS model delivers a good performance in the evaluation part with MAPE 0.9252% and APE value between 0.02226% and 1.7877%.

Keywords Energy-Water Efficiency, Forecasting, Fuzzy Time Series, Water Treatment Plant, Weighted Subsethood

1. Introduction

Water treatment is one of the most important topics in the field of sustainable development. The largest water scarcity usually happens in developing countries and becomes a threat to countries with lower economic status. Water treatment is the process of enhancing the quality of water to make it suitable for a given end-use application. Each process in water treatment consumes high energy and is high cost [1]. Water treatment is a complex and high-energy-consuming process. This is an important part of the trade-off between capital and energy. Therefore, the water treatment process and energy usage are top priorities for sustainable development. The cost and energy consumption of water treatment are the main challenges for water quality and the environment. Hence, various methods are applied and studied to overcome the impact of water treatment plants on the environment [1]. Moreover, water treatment is a vital process that can be challenging for communities and countries with limited resources.

In Malaysia, water is a state matter, and the supply of clean water is limited due to insufficient water treatment plants. The conservation and preservation of water have been prioritised as critical components for the country's development. According to Saimy & Yusof [2], there are numerous options for achieving the goal of strengthening the federal government's engagement in state water governance and making the country's water sector more efficient. Malaysia has rapidly expanded its electricity supply and efficiency, but has lagged in developing a water

supply and treatment infrastructure. The cost of water treatment is also high, and there is a limited capacity for treatment centres. The water treatment system in Malaysia is not sufficient to meet the needs of a growing population. Consequently, the water quality of public water sources is not assured. In less developed countries, because of the lack of clean water, water treatment infrastructure, and energy, the use of water and energy are leads to high fees for the user, making water and energy provision expensive. The ongoing repair, upgrade, and expansion of existing water supply systems to increase treatment efficiency are the primary drivers of the Malaysian water and wastewater equipment market's growth [3].

In this study, the term energy-water efficiency refers to the overall energy efficiency in water treatment plant. Improving energy-water efficiency can increase the availability of clean water and reduce the need for energy usage. In Malaysia, water treatment plants need to be upgraded to meet the demand for clean water. The recycling of water in Malaysia is not sufficient. Achieving energy-water efficiency is challenging and requires a coordinated effort across sectors. Achieving energy-water efficiency is complicated, and many factors can affect the energy-water efficiency of a system. Water is a sensitive resource; any changes in the system can affect the water quality and the ability to fulfill the demand. Therefore, researchers should explore and investigate issues in water plants such as forecasting activity to assist decision-makers to minimise energy usage and maximise water production.

The researcher will explore the impact of energy on water treatment plants, and the objective should be to develop forecasting activity to minimise energy usage and maximise water production. Prediction of energy-water efficiency is essential to evaluate and forecast the activity of water treatment plants. Forecasting is critical for future planning in many technological fields since it is a tool that can help decision-makers make better future judgments [4]. The prediction of water treatment plant can also be achieved by using forecasting models and systems that are performed by the Fuzzy Time Series (FTS) forecasting.

The FTS is a novel approach that was born out of the need to solve linguistic time series data issues. This approach combines linguistic variables with the application of fuzzy logic to time series analysis in order data uncertainty or fuzziness [5]. Many forecasting techniques use fuzzy sets in their algorithms. Fuzzy logic was first used and introduced by Zadeh in 1965, and it has evolved rapidly over the past decade. Fuzzy logic is a mathematical language to express something as it has grammar, syntax, and semantic like a language for communication. In 1993, Song and Chissom first introduced the notion of fuzzy time series and suggested a fuzzy time series forecasting method [6-8]. Because of its better performance, the FTS model has attracted much attention from the research community. The researcher began to simulate the model in a real-life application using

historical data and realised that the model is able to work in many new interesting sectors such as in forecasting university enrolments [7-9, 16,19], stock index [4, 18], electricity [5, 17], rice production [11], crude palm oil price [14], temple visitor [15], the number of passengers [13] and many more.

2. Fuzzy Time Series

The creation of a fuzzy system from time series data is known as fuzzy time series (FTS). The concept of FTS is introduced by Song and Chissom [6-8] and since then many authors adopted it with some modifications to make the prediction more accurate. The FTS approach becomes popular because it is able to deal with small data, which makes the process of calculation straight forward and easy to apply in many problems [10-16]. This study proposed the Weighted Fuzzy Time Series (WeSuFTS) for energy-water efficiency. The WeSuSFTS procedure is as follows:

Step 1: Define the universe of discourse U .

In this step, the minimum and maximum values of data are defined and the universe of discourse, $U = [min-d_1, max+d_2]$ where d_1 and d_2 2 proper positive numbers. The U will be divided into subgroup or interval.

Step 2: Define fuzzy sets on the U .

In this step, each interval will present the linguistic variable for each data as defined on the universal set U .

Step 3: Fuzzifying

The fuzzification process is the process of changing numerical values into fuzzy variables. The corresponding linguistic variable needs to be assigned in order to fuzzify the data. The easier way is to assign the value with the highest membership degree for each interval [11].

Step 4: Determine the fuzzy logical relationship (FLR) and fuzzy logical relationship group (FLRG)

Determining the FLR is usually based on previous or Left-hand side (LHS) and current time or Right-hand side (RHS) in time series [13]. The FLRG is determined by grouping the outcome of RHS from LHS [14].

Step 5: Calculate the fuzzy subsethood value.

The Fuzzy subsethood is defined as [17-20]:

Let A and B be 2 fuzzy sets of the universe of discourse U with membership functions μ_A and μ_B , respectively. The fuzzy subsethood value $S(A, B)$ measures the degree in which A is a subset of B , which can be expressed as follows:

$$S(A, B) = \frac{\sum_{u \in U} \min\{\mu_A(u), \mu_B(u)\}}{\sum_{u \in U} \mu_A(u)} \quad (1)$$

where $S(A, B) \in [0,1]$.

Step 6: Calculate the weighted fuzzy subsethood value

The weighted fuzzy subsethood value is the relative weight created based on fuzzy subsethood value to provide multiplication factors for each input linguistic variable's value [17-20]. The weighted fuzzy subsethood value is defined as follows. For classification problem with K classes and M input linguistic variable's values (or attributes), the weight $w(A_k, B_{ij}) \in [0,1]$ between the j -th input linguistic value B_{ij} of linguistic variable $B_i, i = 1, \dots, M$ and class $A_k, k = 1, \dots, K$ can be estimated by

$$w(A_k, B_{ij}) = \frac{S(A_k, B_{ij})}{\max_{t=1, \dots, l} S(A_k, B_{it})} \quad (2)$$

given that the input linguistic variable B_i has l possible input linguistic value of $B_{i1}, \dots, B_{il}$. This value represents the degree of "importance" of input linguistic value B_{ij} to the corresponding A_k .

Step 7: Defuzzification

Defuzzification is the process of changing back the fuzzy variable into numerical values or real number.

The proposed model in this study can consider the extended Chen and Cheng models. The difference between Chen and Cheng models is the weight added in Cheng's model. Meanwhile, the proposed model, WeSuFTS, adds weighted subsethood value to the fuzzy reasoning model.

3. Methodology

In general, there are 3 phases in this study: data gathering, WeSuFTS modelling, and model evaluation. Although the main element of WeSuFTS modelling is in the second phase, the forecasting process will be incomplete without the first and third stages in this methodology.

Figure 1. The detailed steps of WeSuFTS

Phase 1: Data collection

The water treatment plant (WTP) in Jenun Baru, Kedah is selected in this research. According to report by The International Bank for Reconstruction and Development, the overall energy efficiency of the water sector can be measured by the amount of electricity used per unit of water delivered to the end-user (kWh/m³) [22]. These 2 portions can be formulated as energy-water efficiency (EE) as shown in Equation (3), where EC is energy consumption (kWh) and WP is water production (m³). These 2 data sets are collected from Syarikat Air Darul Aman (SADA) Sdn. Bhd, Kedah. The data be divided into 2 parts: modelling part (year 2017 – 2019) and evaluation part (year 2020). The modelling part of the data is used for Weighted Subsethood Fuzzy Time Series (WeSuFTS) model development in Phase 2. Meanwhile, the implementation of the model and evaluation of the model performance will proceed using the evaluation part of the data.

$$EE = \frac{EC}{WP} \quad (3)$$

Figure 2. Plots of Energy-Water Efficiency 2017-2020

Figure 3. Plots of Energy-Water Efficiency 2017-2020.

The line graph is displayed to detect the pattern as the first step of univariate forecasting processes according to [23], which is to plot data and identify the existence of time

series components based on data patterns. Based on Figure 2, since the WTP in Jenun Baru is upgraded, there is a rapidly increase of energy-water efficiency value in the

early year of 2018. 34 data sets of energy-water efficiency after the system upgrading from March 2018 to December 2020 are considered in this research to ensure a more accurate picture. Therefore, the line graph of energy-water efficiency after upgrading the WTP is shown in Figure 3.

According to Azizah [12], stationary data is a necessary assumption in time series analysis in order to acquire the correct model. As a result, it is necessary to examine the

stationary data.

Based on Figure 4, it demonstrates that data do not have any upward or downward trend or seasonal effects. If the data do not indicate the present of trend components, the data is considered as a stationary pattern. However, for more confirmation, the monthly data had also been tested by autocorrelation function and partial autocorrelation functions using SPSS.

Figure 4. Line graph of time against Energy-Water Efficiency

Figure 5. Autocorrelation function and partial autocorrelation function

Figure 5 (a) shows the stationary time series notice how the autocorrelations decline to non-significant levels quickly when the autocorrelation function declines to near zero rapidly for a stationary time series. At the same time, Figure 5 (b) shows a moderately large positive spike followed by correlations that bounce between positive and negative at lag 1, 2 and 6, all of which are not statistically significant. According to its unsystematic nature, it can be assumed that this time series is stationary.

Phase 2: Model development

Phase 2 of WeSuFTS energy efficiency in water treatment plant has been obtained from the results by following these steps:

Step 1: Defining the universal of discourse, U and partition U into some length interval.

The lowest data point in the stationary data set is 0.3719, while the greatest data point is 0.38253. $U = [0.37, 0.383]$ is the universal discourse for this data. The universal of discourse U was divided into 5 portions, each with a lengthy interval (0.0026) in this study.

Step 2: Defining fuzzy sets on the U .

Fuzzy sets are defined based on the universal set U in this step. The fuzzy values are presented using the triangle membership function in this study. In this study, energy-water efficiency at current month (C_t) has five fuzzy linguistic terms: $A_1 = \text{very little}$, $A_2 = \text{a little}$, $A_3 = \text{moderate}$, $A_4 = \text{much}$, $A_5 = \text{very much}$. Next month energy-water efficiency (C_{t+1}) also has five fuzzy linguistic terms: $B_1 = \text{very little}$, $B_2 = \text{a little}$, $B_3 = \text{moderate}$, $B_4 = \text{much}$, $B_5 = \text{very much}$. Then fuzzy sets A_1, A_2, A_3, A_4, A_5 and B_1, B_2, B_3, B_4, B_5 , on the universal of discourse U :

$$A_1 \text{ and } B_1 = \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \frac{0}{u_5}$$

$$A_2 \text{ and } B_2 = \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \frac{0}{u_4} + \frac{0}{u_5}$$

$$A_3 \text{ and } B_3 = \frac{0}{u_1} + \frac{0.5}{u_2} + \frac{1}{u_3} + \frac{0.5}{u_4} + \frac{0}{u_5}$$

$$A_4 \text{ and } B_4 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0.5}{u_3} + \frac{1}{u_4} + \frac{0.5}{u_5}$$

$$A_5 \text{ and } B_5 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \frac{0.5}{u_4} + \frac{1}{u_5}$$

Step 3: Fuzzifying

This study used a triangle membership function to fuzzify crisp data into fuzzy linguistic terms, according to the fuzzy sets. The following step makes a fuzzy set out of the intervals.

Step 4: Determining fuzzy logical relationship (FLR) and fuzzy logical relationship group (FLRG)

The fuzzy logical relationships (FLR) are produced by splitting the universal discourse and defining fuzzy sets, which are then used to forecast. The observations with the same linguistic value for the dependent variable are grouped. The fixed number of subgroups is based on the fixed number of fuzzy linguistic terms dependent on linguistic variable. This is done by grouping observations with the same dependent fuzzy linguistic terms created by FLR and FLRG.

Step 5: Calculating fuzzy subsethood value

In the FLRG, this phase involves determining the fuzzy subsethood value for each subgroup. The fuzzy subsethood value can be determined using Equation (1). The relationship between the dependent variable linguistic term and each independent linguistic term is described by these values.

Step 6: Calculating weighted fuzzy subsethood value

The weighted fuzzy subsethood value can be determined using Equation (2) from the value in Step 5. The information from weighted subsethood is used to generate the WeSuFTS model in the form of fuzzy reasoning WeSuFTS rules. In other words, the WeSuFTS model is composed of a set of fuzzy rules in the following form:

IF (condition) **THEN** (conclusion)

with operator **OR** using max operations and **AND** using min operations. The rule set is simplified as any linguistic word values with a weight of 0 are automatically eliminated from the model.

Step 7: Defuzzifying

This study uses the midpoint of dependent fuzzy set based on the highest calculation value from the WeSuFTS model. If more than one rule has the same maximum value, the average midpoint of corresponding dependent fuzzy set may have a maximum value to be taken as value of defuzzification [24-25]. For example, let calculation value from condition part of WeSuFTS model as follows: Rule 1 = 3.5, Rule 2 = 1.2, Rule 3 = 5.4, Rule 4 = 0, Rule 5 = 0, and the mid-point of dependent fuzzy set as $m_{A1} = a$, $m_{A2} = b$, $m_{A3} = c$, $m_{A4} = d$, and $m_{A5} = e$. Hence, the defuzzification value is c due to Rule 3 having the maximum value. For example, let Rule 1 = 0, Rule 2 = 0.13, Rule 3 = 3.7, Rule 4 = 4.5, Rule 5 = 4.5. Hence, the defuzzification value is the average of m_{A4} and m_{A5} equal to $(d + e)/2$ due to Rules 4 and 5 having the same highest value.

Phase 3: Model evaluation

Forecasting error calculation is a way to determine the accuracy of the models that have been obtained to measure the model's performance. To assess the forecasting

performance and accuracy of the models, they are measured using the Absolute Percentage Error (APE) and Mean Absolute Percentage Error (MAPE) as shown in Equations (4) and (5).

$$APE = \frac{y_t - \hat{y}_t}{y_t} \times 100 \quad (4)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{(y_t - \hat{y}_t)}{y_t} \times 100 \quad (5)$$

4. Results and Discussion

This chapter shows the results of the proposed WeSuFSTS models for energy-water efficiency. Figure 6 reveals that the WeSuFSTS model consists of 5 fuzzy reasoning rules related to five fuzzy sets in *next month energy-water efficiency* (C_{t+1}). The weighted subsethood value and linguistic terms of independent and dependent linguistic variables are included in each rule. For the independent linguistic term, the weighted subsethood value acts as a multiplication factor. The degree of

'importance' for the independent linguistic terms towards the dependent linguistic terms is represented by the multiplicative factor value.

For example, **Rule 1** consists of linguistic term B_1 , B_2 , B_3 , and B_4 that contribute to A_1 . It is shown that only 4 out of 5 independent linguistic terms only are important to A_1 . The interpretation is that as if the *current month energy-water efficiency* is *very little*, or *a little*, or *moderate*, or *much*, hence the *next month energy-water efficiency* is *very little*.

Therefore, by referring to the WeSuFSTS model in Figure 6, we can interpret the results as

- B_1 , B_2 , B_3 , and B_4 are important to A_1 .
- B_1 , B_2 , B_3 , B_4 , and B_5 are important to A_2 .
- B_1 , B_2 , and B_3 are important to A_3 .
- B_5 are important to A_4 .
- B_2 , and B_3 are important to A_5 .
- B_2 is the most important to A_1 , A_2 and A_3 .
- B_3 is the most important to A_5 .
- B_5 is the most important to A_4 .

Rule 1: IF C_t is $0.715366B_1$ OR B_2 OR $0.3351B_3$ OR $0.012938B_4$ THEN C_{t+1} is A_1

Rule 2: IF C_t is $0.353473B_1$ OR B_2 or $0.630691B_3$ OR $0.098081B_4$ OR $0.251158B_5$ then C_{t+1} is A_2

Rule 3: IF C_t is $0.701259B_1$ OR B_2 OR $0.546818B_3$ THEN C_{t+1} is A_3

Rule 4: IF C_t is B_5 THEN C_{t+1} is A_4

Rule 5: IF C_t is $0.959741B_2$ OR B_3 THEN C_{t+1} is A_5

Figure 6. WeSuFSTS model

Table 1. Comparison of the Defuzzification and APE for WeSuFSTS, Chen and Cheng models

t	Actual value, C_t	Defuzzification			APE		
		WeSuFSTS	Chen	Cheng	WeSuFSTS	Chen	Cheng
1	0.37516	0.3739			0.3022		
2	0.37778	0.3817	0.3735	0.37382	0.2478	1.1279	1.04883
3	0.37618	0.3817	0.3779	0.37615	0.4289	0.4596	0.00647
4	0.37523	0.3817	0.3779	0.37615	0.3020	0.7138	0.24663
5	0.38253	0.3791	0.3779	0.37615	0.0946	1.2082	1.66653
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
18	0.37314	0.3739	0.3735	0.37382	0.4154	0.1012	0.18129
19	0.37229	0.3713	0.3735	0.37382	0.2312	0.3318	0.41208
20	0.37663	0.3817	0.3735	0.37382	0.4655	0.8246	0.74527
21	0.37294	0.3817	0.3779	0.37615	0.3687	1.3320	0.86197
22	0.3733	0.3817	0.3735	0.37382	0.4505	0.0594	0.13943

Table 1 shows the results of defuzzification and forecasting error analysis for 3 models which are the proposed WeSuFTS model, Chen's model and Cheng's model for the modelling part of data set. Then, the performance of the 3 different models was measured by using the APE, and MAPE. Table 2 shows the forecasting error analysis for modelling part. From Table 2, the best performance of FTS model is Cheng's model, with the lowest MAPE value, which is 0.4351% compared to the WeSuFTS model with a MAPE value of 0.732% and Chen's model with a MAPE value of 0.75047%. MAPE error is good for measuring the result that is obtained because the result is considered better if the MAPE value is

closer to zero. Cheng's model also has the lowest APE value which ranges around 0.0065% to 1.6666% error compared to other models. Figure 7 shows the line graph for comparison of actual data and other models. The graph line of Cheng's model in line chart is the nearest to the actual value compared to other models.

Table 2. Model performance analysis in modelling part

Model	WeSuFTS	Chen	Cheng
APE (%)	0.0565-2.3489	0.0104-1.332	0.0065-1.6666
MAPE (%)	0.732	0.75047	0.4351

Figure 7. The line chart for modelling part

Figure 8. The line chart for evaluation part

Table 3. Model performance in evaluation part

Model	WeSuFSTS	Chen	Cheng
APE (%)	0.2226-1.7877	0.251-3.2323	0.3313-3.1549
MAPE (%)	0.9252	1.3765	1.3579

Table 3 and Figure 8 show the forecasting error analysis and line chart for the evaluation part. Based on the error analysis in Table 3, it shows that the proposed model, which is WeSuFSTS, is the best model compared to Chen's and Cheng's models. The MAPE value for the WeSuFSTS is the lowest, which is 0.9252% followed by the Cheng's model 1.3579% and Chen's model at 1.3765%. The prediction of WeSuFSTS is better when it has the lowest MAPE value because it considers results of WeSuFSTS are more acceptable compared to others. The APE value of the WeSuFSTS model is also the lowest in the evaluation part, which is only 0.02226% to 1.7877% compared to Cheng's model 0.251% to 3.2323% and Chen's model 0.3313% to 3.1549% range of error. The line chart in Figure 8 also denotes that the line of the WeSuFSTS model is the nearest to the line chart for actual value of energy-water efficiency compared to others models. According to Mierswa [26], the model performance result in the evaluation part is a far better predictor of how well the model will perform in the future for new and unknown scenarios. Therefore, the results in the evaluation part are more accurate as a decision of the best model selection to forecast future EE.

5. Conclusions

Fuzzy Time Series method has grown in popularity because this method is considered as an effective tool for forecasting time series data with uncertainty. Since its creation, several researchers have modified the FTS approach to increasing forecasting accuracy and obtaining beneficial results in both theory and application. This study proposed the Weighted Subsethood Fuzzy Time Series (WeSuFSTS) model for energy-water efficiency in Water Treatment Plant in Jenun Baru. The proposed model has more meaningful information because WeSuFSTS can identify which factors are the most important to the results based on the rule set obtained. Moreover, the simplification steps in WeSuFSTS model construction can be effective technique to reduce the number of rules in fuzzy reasoning application. Hence, the WeSuFSTS method is able to reduce the time and complexity of fuzzy reasoning application.

The investigation on the results is shown based on the forecasting errors, which are APE and MAPE, where Cheng's model is better compared to the WeSuFSTS and Chen's model in the modelling part of data set, but in the evaluation part, the WeSuFSTS model performed very well. The results obtained in this paper are also consistent with the result from Mansor et al., [21] in which the WeSuFSTS model is better compared to Chen's and Cheng's models in forecasting application problem. This supports the fact that

WeSuFSTS is indeed a better model for achieving high forecast accuracy. Therefore, the application of the WeSuFSTS model is suggested because the result is more accurate to evaluate as forecasting performance and to decide the best model to predict future energy-water efficiency.

However, this study considers energy consumption in water production only as a factor in the general EE calculation, as it does not discuss embedded energy in the water production system such as energy for making the chemical used for water treatment. Therefore, for future works, energy usage at each water treatment process is suggested to be taken into account to describe the specific EE in each water treatment process at water treatment plant. Then, the recommendation and planning for management will be carried out based on the specific EE results. Then, recommendations and plans for management will be carried out based on the results.

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 and Syarikat Air Darul Aman (SADA) Sdn. Bhd, Kedah for providing all the data needed in this study. (S/O Code: 14878).

REFERENCES

- [1] Copeland, C., & Carter, N. T. Energy-water Nexus: The Water Sector's Energy Use. Congressional Research Service. 2017
- [2] Saimy, I. S., & Yusof, N. A. M. The Need for Better Water Policy and Governance in Malaysia. *Procedia-Social and Behavioural Sciences*, 81, 371-375. 2013.
- [3] Liew, R. Malaysia: Water and Wastewater Treatment. *Asia-Pacific Economy Cooperation*. 2008.
- [4] Bas, E., Yolcu, U., & Egrioglu, E. Intuitionistic Fuzzy Time Series Functions Approach for Time Series Forecasting. *Granular Computing*, Volume 6, Issue 3, PP:619-629. 2021.
- [5] Ismail, Z., Efendi, R., & Deris, M. M. Application of Fuzzy Time Series Approach in Electric Load Forecasting. *New Mathematics and Natural Computation*, Volume 11, Issue 3, PP: 229-248. 2015.
- [6] Song, Q., & Chissom, B. S. Fuzzy Time Series and its Models. *Fuzzy sets and systems*, Volume 54, Issue 3, PP: 269-277. 1993.
- [7] Song, Q., & Chissom, B. S. Forecasting Enrollments with Fuzzy Time Series—Part I. *Fuzzy sets and systems*,

Volume 54, Issue 1, PP: 1-9. 1993.

- [8] Song, Q., & Chissom, B. S. Forecasting Enrollments with Fuzzy Time Series—Part II. Fuzzy sets and systems, Volume 62, Issue 1, PP: 1-8. 1994.
- [9] Tsaur, R. C., Yang, J. C. O., & Wang, H. F. Fuzzy Relation Analysis in Fuzzy Time Series Model. Computers & Mathematics with Applications, Volume 49, Issue 4, PP: 539-548. 2005.
- [10] Shafii, N. H., Ramli, N. E. D. M., Alias, R., & Fauzi, N. F. Fuzzy Time Series and Geometric Brownian Motion in Forecasting Stock Prices in Bursa Malaysia. Jurnal Intelek, Volume 14, Issue 2, PP: 240-250. 2019.
- [11] Van Tinh, N., & Dieu, N. C. Handling Forecasting Problems Based on Combining High-Order Time-Variant Fuzzy Relationship Groups and Particle Swarm Optimization Technique. International Journal of Computational Intelligence and Applications, Volume 17, Issue 2, PP: 1850009. 2018.
- [12] Azizah, N., A'yun, K., Septiarini, T. W., Wutsqa, D. U., & Abadi, A. M. Optimization of Fuzzy Inference System by using Table Look-up Method to Predict White Sugar Price in the International Market. Journal of Physics: Conference Series, Volume 1097, Issue 1, PP: 012074. 2018.
- [13] A'yun, K., Abadi, A. M., & Saptaningtyas, F. Y. Application of Weighted Fuzzy Time Series Model to Forecast Trans Jogja's Passengers. International Journal of Applied Physics and Mathematics, Volume 5, Issue 2, PP: 76. 2015.
- [14] Rahim, N. F., Othman, M., Sokkalingam, R., & Abdul Kadir, E. Type 2 Fuzzy Inference-based Time Series Model. Symmetry, Volume 11, Issue 11, PP: 1340. 2019.
- [15] Artika, R., Setiawan, A., Setyowisnu, G. E., & Uminasiah, S. Analysis Prediction of Prambanan Temple Visitors with Fuzzy Time Series Chen Model and Seasonal Auto Regressive Integrated Moving Average (SARIMA) Model. In International Conference of Science and Engineering, Volume 3, PP: 359-363. 2020.
- [16] Alam, N. M. F. H. N., Ramli, N., & Mohd, A. H. Intuitionistic fuzzy set-based time series forecasting model via delegeration of hesitancy degree to the major grade de-i-fuzzification and arithmetic rules based on centroid defuzzification. Journal of Physics: Conference Series, Volume 1988, PP: 46-53. 2021.
- [17] Mansor, R., Kasim, M. M., & Othman, M. Weighted Subsethood Segmented Fuzzy Time Series for Moving Holiday Electricity Load Demand Forecasting. Journal of Advanced Research in Dynamical and Control Systems, Volume 12, Issue 2S, PP: 686–700. 2020.
- [18] Mansor, R., Zaini, B. J., & Yusof, N. Prediction stock price movement using subsethood and weighted subsethood fuzzy time series models. In AIP Conference Proceedings Volume 2138, Number 1, PP: 050018. 2019.
- [19] Mansor, R., Zaini, B. J., Othman, M., & Kasim, M. M. The effect of interval length in weighted subsethood fuzzy time series. In AIP Conference Proceedings. Volume 1974, Number 1, PP: 040017. 2018.
- [20] Mansor, R., Kasim, M. M., Othman, M., Zaini, B. J., & Yusof, N. M-factors Fuzzy Time Series for Forecasting Stock Price Movement. Test Engineering & Management, Volume 81, PP: 832–838. 2019.
- [21] Mansor, R., Othman, M., & Kasim, M. M. Fuzzy Reasoning Based Time Series Forecasting using Weighted Subsethood Model. Advanced Science Letters, Volume 23, Issue 9, PP: 9094-9097. 2017.
- [22] Liu, Feng, Alain Ouedraogo, Seema Manghee, and Alexander Danilenko. A Primer on Energy Efficiency for Municipal Water and Wastewater Utilities. The International Bank for Reconstruction and Development. PP: 1-72. 2012.
- [23] Ismail, S., Zulkifli, M., Mansor, R., Mat Yusof, M. and Ismail, M. I. The Role of Exploratory Data Analysis (EDA) in Electricity Forecasting. Pertanika Journal Social Science & Humanities, 24, 111-118. 2016.
- [24] Chen, S. M. Forecasting Enrollments Based on Fuzzy Time Series. Fuzzy sets and systems, Volume 81, Issue 3, PP: 311-319. 1996.
- [25] Rasmani, K. A., A Data-driven Fuzzy Rule-based Approach for Student Academic Performance Evaluation. Centre for Intelligent Systems and Their Applications. 2002.
- [26] Mierswa, I., Why You Should Ignore the Training Error. Available at: <https://rapidminer.com/blog/validate-models-cross-validation/>. 2017.