IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 21, 2021, accepted May 16, 2021, date of publication May 26, 2021, date of current version June 8, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3084048*

# A Novel Stochastic Fuzzy Time Series Forecasting Model Based on a New Partition Method

**YOUSIF ALYOUSIFI[1,4], MAHMOD OTHMAN[2], AND AKRAM A. ALMOHAMMEDI[3], (Member, IEEE)**

[1]Department of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
[2]Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS, Seri Iskandar 31750, Malaysia
[3]Automobile Transportation Department, South Ural State University, 454080 Chelyabinsk, Russia
[4]Department of Mathematics, Faculty of Applied Sciences, Thamar University, Dhamar, Yemen

Corresponding author: Yousif Alyousifi (yalyousifi@tu.edu.ye)

This work was supported by Yayasan Universiti Teknologi PETRONAS (YUTP) under Grant 015LC0-296.

**ABSTRACT** Fuzzy Time Series (FTS) models are commonly used in time series forecasting, where they do not require any statistical assumptions on time series data. FTS models can handle data sets with a small number of observations or with uncertainty. This is a general advantage of FTS as compared with other techniques. However, FTS models still have some criticisms, such as the optimal lengths of intervals and the proper weights, which always influence the model accuracy and still have been of many concerns in literature. The work in this paper proposes a novel FTS forecasting model based on a new tree partitioning method (TPM) and Markov chain (MC), called FTSMC-TPM, for determining the optimal partitions of intervals and the proper weights vectors respectively, and this will greatly improve the model accuracy. The efficiency of the FTSMC-TPM model is tested using two types of time series consisting of the air pollution index (API) data, which is collected from Kuala Lumpur, Malaysia and the benchmark data of the yearly enrollments for the University of Alabama. Three statistical criteria have been used for investigating the accuracy of the proposed model. The results indicate that the proposed model outperforms the existing classic and advanced time series models in terms of forecasting accuracy. In addition, the proposed model shows the ability to successfully deal with forecasting problems to obtain higher model accuracy, which is examined in comparison with the existing models to validate its superiority. Hence, this study demonstrates that the proposed model is more suitable for the accurate prediction of air pollution events as well as for forecasting any type of random time series.

**INDEX TERMS** Air pollution, Kuala Lumpur, fuzzy time series, fuzzy logical relation, Markov chain, tree partition method.

## I. INTRODUCTION

Forecasting is one of the most important topics of research that has attracted the concern of many researchers and scientists. Amongst various well-known and developed forecasting models, fuzzy time series models are successfully employed better than traditional forecasting methods. Fuzzy time series forecasting model is the application of linguistic mathematical reasoning to model and predicts the future from a time series of linguistic historical observations. In contrast to traditional time series forecasting, fuzzy time forecasting models consider uncertainties in observations over time, do not require restrictive assumptions and too much background knowledge of the observations [1]. Also, fuzzy time series forecasting methods are able to work with a very small set of observations [2].

Air pollution is a serious environmental problem that has drawn worldwide attention. It can be described as the presence of harmful substances in the air at higher concentrations than their normal ambient rates, which can cause damage or undesirable changes to human health and the environment. From these substances, particulate matter ($PM_{10}$), ozone ($O_3$), sulphur dioxide ($SO_2$), nitrogen dioxide ($NO_2$), and carbon monoxide (CO), which are the most materials that could impact negatively on human life, ecosystems, and other living creatures in the natural environment. Thus, monitoring and forecasting the parameters of air quality are necessary for assessing the air pollution level where the results found can be

The associate editor coordinating the review of this manuscript and approving it for publication was Nikhil Padhi.

used for managing the air quality. Accordingly, air pollution forecasting is one of the most powerful ways, providing information on air quality and developing early warning systems, which help in preventing negative effects on human health and improving life quality that produces from air pollution. In addition, it considers a fundamental tool for the anticipated implementation of strategies in public health.

Several statistical and soft computing methods are used for predicting the transition of air pollutants in the lower atmosphere. Up to now, statistical methods and soft computing techniques are the main categories that have been used and developed in order to forecast air pollution in order to improve the accuracy of air quality prediction. The major purpose of these techniques is to analyze the past behavior of air pollutant concentration to predict future values, assuming they will be similar in the future. For instance, regression analysis [3], autoregressive integrated moving average (ARIMA) [4], [7]–[11], artificial neural network (ANN) [5], [6], [15] and Markov chain model [13], [14] are applied for air pollution forecasting. However, these techniques have some drawbacks on generalization issues, satisfy the statistical assumptions, and treating with subjectivity and uncertainty data, that can extant in the issues of environmental, such as air pollution.

Accordingly, fuzzy time series (FTS) models are most appropriate for the data that consists of a small number of observations, incomplete, include uncertainty, and do not require statistical assumptions. Particularly, forecasting air pollution data that generally collects in the form of time series doesn't satisfy such assumptions. Thus, it is believed that the FTS model is suitable for modeling air pollution. However, very few studies have been done using FTS models as in [16]–[18] and others. For instance, the authors in [19] proposed an FTS model for predicting the daily O3 pollutant in Taiwan. A seasonal FTS model for forecasting air pollution in Ankara was applied in the study by [20]. In [21], the authors conducted an analysis in order to evaluate the forecasting performance of five models, which are ARIMA, ANN and three models of FTS for predicting air pollution in Malaysia. In [22], the authors proposed fuzzy time series based on the Markov transition probability matrix, which is used for obtaining the largest probability using a transition probability matrix to establish weights of fuzzy relationships between the observations of the stochastic pattern of time series. They also used a random length of interval for the universe of discourse, which leads to a negative effect by abnormal observations and outliers. In [16], the authors proposed a fuzzy time series model using clustering techniques, which is used for minimizing the negative effects of abnormal observations and outliers on the model accuracy. According to the literature, it is observed that the fuzzy time series have some issues in determining the appropriate length of intervals, and various methods have been presented to investigate the length of the interval and partition number of the universe of discourse in fuzzy time series to achieve high-level forecasting accuracy. Nevertheless, the issue of interval length is still a matter of concern and thus, influencing the forecasting accuracy. Thus, FTS models in air quality forecasting may present interesting issues where it impacts forecasting accuracy.

In order to increase accuracy and to decrease the level of complexity and the overload of calculation in forecasting, this study proposes a novel fuzzy time series Markov chain model based on a new Tree Partition Method (FTSMC-TPM). The FTSMC-TPM model improves the interval length along with the partition number of the universe of discourse, simplifies the partitioning of the universe of discourse and minimizes the undesirable influences of abnormal data points on the performance of forecasting. TPM is a novel linguistic partition method developed as a re-partitioning approach based on the average length of first sub-intervals (the average inter-quartile range). This method is simpler and more effective than clustering methods in determining the proper length of intervals in order to provide an optimal partition of the universe of discourse. In addition, in the forecasting method, the first-differencing data is also considered to obtain a better forecast. The proposed FTSMC-TPM model in this study is an extension and development of the existing fuzzy time series Markov chain model proposed by [22]. For validating and evaluating the performance of the proposed method, we demonstrate an application of the proposed forecasting model using two different types of real-world data set, including 1) the hourly air pollution index (API) data from Kuala Lumpur, Malaysia. 2) the benchmark data of enrollment at the University of Alabama. Finally, we perform a comparison between the FTSMC-TPM model and the existing fuzzy time series models. The main advantage of the proposed model is to decrease the level of complexity in the partition method and the overload of calculation in forecasting.

The rest of this paper is outlined as follows: Section II presents the preliminary of the study, which includes a brief introduction, definitions related to fuzzy time series and reviews of existing FTS models. Section III introduces the research framework and algorithm used for analysis. Section IV gives an implementation of the proposed model and validates the effectiveness of the proposed model with a real dataset, and by comparison with the prediction accuracy of some existing various models that have been developed. Section V offers the model evaluation. Finally, Section VI presents the conclusion of the study.

## II. PRELIMINARY

This section involves a brief introduction, main definitions of fuzzy time series and reviews its related literature

### A. FUZZY TIME SERIES

A fuzzy set [23] is a class of objects with a continuum of the grade of membership. Let $U$ be the universe of discourse with $U = \{u_1, u_2, \ldots, u_n\}$, where $u_i$ are possible linguistic values of $U$. Fuzzy time series is a combination of the concept of fuzzy sets with the time series model. The aim of the essential steps for designing fuzzy time series models is to i) determine the universe of discourse $U$, ii) split $U$ into a fixed number of intervals, and fuzzification, iii) determine fuzzy logic relation, iv) define forecasted values and defuzzification. Fuzzy time series definitions have been presented in the past works as follows.

*Definition 1:* [24] Let $X(t), (t = 1, 2, \ldots)$ a subset of real numbers, be the universe of discourse in which fuzzy sets $f_j(t)(j = 1, 2, \ldots)$ are defined. If $F(t)$ is a collection of $f_1(t), f_2(t), \ldots,$ then $F(t)$ is called a fuzzy time series of $X(t)$.

*Definition 2:* [24] A fuzzy set is a class of objects with a continuum of the grade of membership. Let $U$ be the Universe of discourse with $U = \{u_1, u_2, u_3, \ldots, u_n\}$, where $u_i$ are possible linguistic values of $U$, then a fuzzy set of linguistic variables $A_i$ of $U$ is defined by

$$A_i = \frac{f_{A_i}(u_1)}{u_1} + \frac{f_{A_i}(u_2)}{u_2} + \ldots + \frac{f_{A_i}(u_n)}{u_n} \tag{1}$$

where $f_{A_i}$ is the membership function of fuzzy set $A_i$; $f_{A_i}$: $U \rightarrow [0, 1].f_{A_i}(u_r) \in [0, 1]$ and $1 \leq r \leq n$. If $u_j$ is a member of $Ai$, then $f_{A_i}\left(u_j\right)$ is the degree of belonging of $u_j$ to $A_i$.

*Definition 3:* [20] Let us consider the fuzzy logical relationship (FLR) $R(t, t-1)$ such that $F(t) = F(t) = F(t-1) * R(t-1, t)$, where $*$ represents an operation, then $F(t)$ is said to be caused by $F(t)$. Then, the logical relationship between $F(t)$ and $F(t1)$ is denoted as

$$F(t-1) \rightarrow F(t) \tag{2}$$

*Definition 4:* [20] Let $F(t) = A_i$ and $F(t) = A_j$. The relationship between two consecutive observations , $F(t)$ and $F(t)$, referred to as an FLR, can be denoted by $A_i \rightarrow A_j$, where $A_i$ is the left-hand side of the FLR and $A_j$ is the right-hand side of the FLR.

### B. REVIEW AND FRAMEWORK OF THE FUZZY TIME SERIES

In 1993, the Fuzzy time series model was proposed by [24], [25] by replacing the values of time series with fuzzy sets, to establish fuzzy relationships among observations. As creating fuzzy relations need a complex matrix process, therefore, it takes a long time of computation to execute it. Song and Chissom's model was modified by [26] employing a simpler arithmetic process rather than a complex max-min process to simplify the use of the model.

Fuzzy time series models consist of three steps, and these are: i) split the universe of discourse $U$ into fixed intervals and fuzzification of classical time series, ii) generate fuzzy logical relations between fuzzy sets ($A_i$), and iii) defuzzification and forecasting. However, in the FTS models mentioned above, no information has been offered on how the length of intervals is set, so, it is determined arbitrarily. Therefore, several FTS models have been introduced to determine the length of the interval and enhance the forecasting performance of [26] model. For instance, two heuristic FTS models depend on the mean of differences and distribution of differences were presented by [27]. An approach that depends on ratios rather than equivalent lengths of intervals was proposed by [28]. The authors in [29], [30] employed genetic algorithms. The particle swarm optimization was applied by [12], [31]–[34]. An optimization technique with a single-variable restriction was employed by [35], [36]. In addition, [37] applied the sliding window technique with the type-2 fuzzy time-series model for determining the appropriate length of intervals. Also, an interval type-2 fuzzy logic system (IT2FLS) based on a fuzzy time series was proposed by [38]. A fuzzy time series model has been developed by combining the model with some techniques such as network traffic anomaly detection algorithm [39], complex network analysis [40] and C-Means clustering algorithm [41].

In most studies on fuzzy time series, the universe of discourse is split into equivalent-length intervals. However, when the distribution of the universe of discourse is not uniform, fuzzy time series models may not generate pretty forecasting outcomes. Thus, many studies used clustering techniques to set the distribution of universe discourse from itself and discovered a different fuzzy partition at a different interval length [16], [33], [35], [36], [42]–[44] in order to partition the universe of discourse. Nevertheless, the clustering techniques are difficult in determining an optimal partition number in many cases. According to the literature reviewed above, it is believed that determining an optimal and appropriate partition method for the universe of discourse presents interesting issues.

## III. METHOD

In this section, an algorithm of the proposed partition method called Tree Partition Method (TPM) and the proposed model building algorithm are presented in the next sub-section.

### A. ALGORITHM OF THE PROPOSED PARTITION METHOD

This sub-section involves an algorithm of the proposed partition method called the Tree Partition Method (TPM). This method can be applied for determining the optimal partition number for the partitioning of the universe of discourse that may improve the model accuracy. A pseudo-code of the algorithm is presented in Table1. The steps of the proposed partition method (TPM) algorithm are in detail as following:

1) Define the universe of discourse, $U$ for the historical data, based on the range of available historical time series data, by the formula $U = [D_{min} - D_1, D_{max} + D_2]$, where $D_{min}$ denotes the minimum value in the universe of discourse $U$, $D_{max}$ denotes the maximum value in the universe of discourse $U$, $D_1$ and $D_2$ represent positive values, which are used for extending the length on both sides of $U$ to preserve a variation space and to ensure that the future results fall in $U$.
2) Partition the universe $U$ into at most five equal-length intervals and not less than three, as it is the smallest number of partitions in many previous studies. In addition, usually, we consider applying the inter-quartile

range partition method to guarantee that each interval consists of a number of observations and its length is greater than zero especially for the dataset of a small number of observations (small sample size).

3) Calculate the average of the sub-intervals found in step 3.
4) Re-divide the interval in Step 2. If there are one or more of the sub-intervals that have a size greater than the average found in step 3, then the particular sub-interval with a size greater than the average should be further partitioned into half.
5) If there is any of the sub-interval found in step 4 still has a size greater than the first average value found in step 3, then the particular sub-interval with a size greater than the average should be further partitioned into half.
6) Repeat the operation in step 3 until all sub-intervals' size becomes less than the first average amount.
7) The latest partitions with sub-intervals size less than average length are considered in the calculation.

### B. PROPOSED MODEL BUILDING ALGORITHM

In this subsection, the simplified arithmetic operations introduced by [22] are used in the proposed model algorithm. The steps of the proposed model algorithm can be described in detail as follows:

*Step 1:* Define the universe of discourse, $U$ for the historical data.

*Step 2:* Partition $U$ for the historical data following the algorithm's steps of the proposed partition method (TPM) as presented in Table 1.

**TABLE 1. Pseudo-code of the proposed partition method (TPM.)**

```
Algorithm 1 A Pseudo-code of the Tree Partition Method (TPM)
Begin
    1. Read dataset.
    2. Define $U = [D_{min} - D_1, D_{max} + D_2]$.
    3. Partition $U$ into n intervals $u_i, i = 1, 2, \ldots, n$,
       s.t. $3 \leq n \leq 5$.
    4. For $i = 1\ to\ n$.
    5. Find the length of intervals $L_i$
    6. Find the average of sub-intervals' length, i.e.,
       $ALFP = \frac{sum(L_i)}{n} \quad \forall i, i = 1, 2, \ldots, n$.
    7. If $\exists\, i : L_i > ALFP$, then re-divide $L_i$ to half
       $\forall i\ (i = i + 1)$.
    8. For $i = 1\ to\ (n + p)$, ($p$ is the new sub-intervals).
    9. If $\exists\, i : L_i > ALFP$, go to step 7, else break the
       loop [i.e., $L_i < ALFP, \forall i$].
End
```

*Step 3:* Define the fuzzy sets $A_k$ on the universe of discourse $U$. Fuzzy sets $A_k$ are determined based on the interval $u_k$ that already have been formed using the grid method in the previous step with the function membership as follows.

$$A_k = \begin{cases} \dfrac{1}{u_1} + \dfrac{0.5}{u_2} & k = 1 \\ \dfrac{0.5}{u_1} + \dfrac{1}{u_2} + \dfrac{0.5}{u_3} & 2 \leq k \leq n-1 \\ \dfrac{0.5}{u_{n-1}} + \dfrac{1}{u_n} & k = n \end{cases} \tag{3}$$

*Step 4:* Fuzzify the observations into linguistic values based on the maximum membership value. Then, the observations are mapped into a fuzzy set.

*Step 5:* Construct the fuzzy logical relationships (FLRs) and establish fuzzy logical relation groups (FLRGs) to build frequencies (count) matrix of fuzzy relation between observations.

*Step 6:* Establish transitions count matrix and calculate the Markov transition probability matrix based on the frequencies of the fuzzy relationship groups found in Step 5. State transition probability $P_{ij}$, from state $A_i$ to state $A_j$ is the probability of observing $y_{t+1}$ given $y_t$, i.e., $P_{ij} = Pr\,(y_{t+1} = j | y_t = i)$, which can be calculated as follows

$$P_{ij} = \frac{N_{ij}}{Ni.}, \quad i, j = 1, 2, \ldots, n \tag{4}$$

where $N_{ij}$ and $N_i$. represent transition time in one step from state $A_i$ to state $A_j$ and the number of observations in the state $A_i$, respectively. Thus, Markov transition probabilities matrix $P$ can be given as follows

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{12} & p_{22} & & p_{2n} \\ \vdots & \ddots & \vdots & \\ p_{n1} & p_{n1} & \cdots & p_{nn} \end{bmatrix} \tag{5}$$

where $P_{ij} \geq 0$ and $\sum_{j=1}^{n} P_{ij} = 1$.

*Step 7:* Defuzzify the linguistic value and calculate forecasted values. More specifically, in the calculations of forecasts, there are two cases, which are one-to-one, and one-to-many are considered.

*Case 1:* In the case of the fuzzy logical relationship group of $A_i$ is one-to-one, in which there only one transition for $A_i$ (i.e., $A_i \to A_k$, with $P_{ik} = 1$ and $P_{ij} = 0, j \neq k$), then the forecasting of $F\,(t)$ is $c_k$, the center of $u_k, k = 1, 2, \ldots n$, which can be calculated according to (6) below

$$F\,(t+1) = c_k \quad P_{ik} = c_k \tag{6}$$

*Case 2:* In the case of the fuzzy logical relationship group of $Ai$ is one-to-many, in which there are more than one transitions for $A_i$ (i.e., $A_i \to A_1, A_2, \ldots, A_n, i = 1, 2, \ldots, n$). Thus, if the state is $A_i$ for the observation $Y\,(t)$ at time $t$, the forecast value $F\,(t+1)$ can be calculated by using (7) below

$$\begin{aligned} F\,(t+1) = {} & c_1 p_{i1} + c_1 p_{12} + \ldots + c_{i-1} p_{i(i-1)} \\ & + Y\,(t)\, p_{ii} + c_{i+1} p_{i(i+1)} \\ & + \ldots + c_n p_{in} \end{aligned} \tag{7}$$

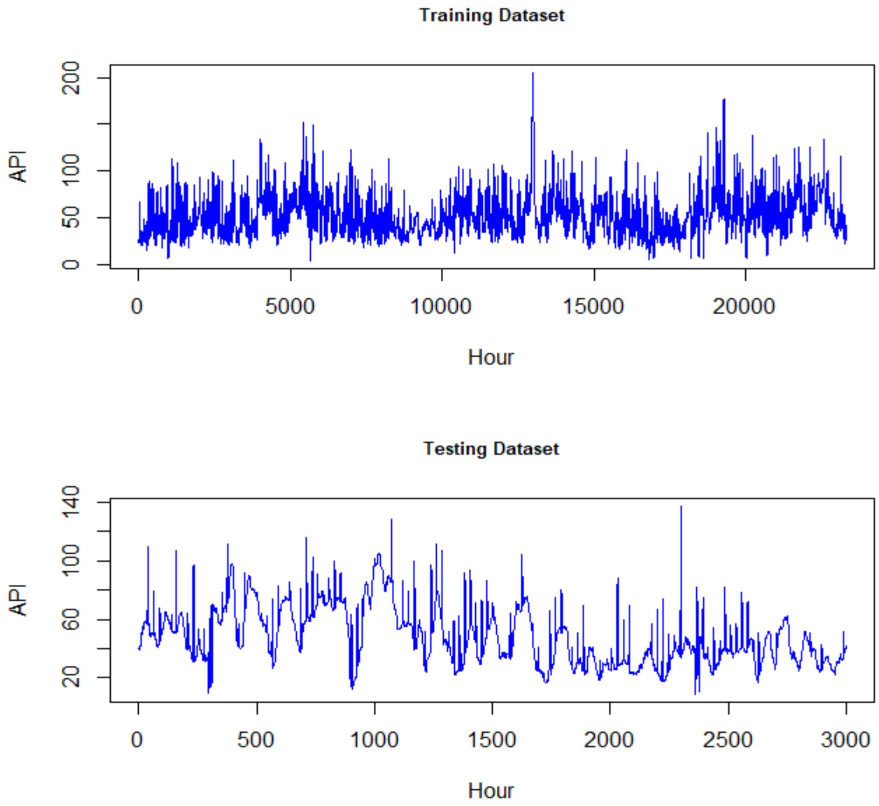


**FIGURE 1.** Time series plots the API values of training and testing datasets.

**TABLE 2.** Description of the procedures of the tree partition method (TPM) for training API dataset.

| The universe of discourse $U = [0, 250]$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| **1st Partition**<br>Using any common partition method based on the type of dataset (at least 3-5 intervals). | | **2ed Partition**<br>Via divide each interval into half | | **3rd Partition**<br>Via divide each interval into half | | **4th Partition**<br>Via divided each interval into half | |
| **Interval** | **length** | **Interval-R1** | **length** | **Interval-R2** | **length** | **Interval-R3** | **length** |
| [0, 50] | **14893** | [0, 25] | 1677 | [0, 25] | 1677 | [0, 25] | 1677 |
| [50,100] | **8059** | [25,50] | **13216** | [25, 37.5] | **6294** | [25, 31,25] | 2700 |
| [100,150] | 279 | [50,75] | **6907** | [37.5, 50] | **6922** | [31.27, 37.5] | 3594 |
| [150,200] | 70 | [75, 100] | 1152 | [50, 62.6] | **5072** | [37.5, 43.75] | 3540 |
| [200,250] | 3 | [100,150] | 279 | [62.5, 75] | 1835 | [43.75, 50] | 3382 |
| - | - | [150,200] | 70 | [75, 100] | 1152 | [50, 56.25] | 3200 |
| - | - | [200,250] | 3 | [100,150] | 279 | [56.25, 62.5] | 1872 |
| - | - | - | - | [150,200] | 70 | [62.5, 75] | 1835 |
| - | - | - | - | [200,250] | 3 | [75, 100] | 1152 |
| - | - | - | - | - | - | [100,150] | 279 |
| - | - | - | - | - | - | [150,200] | 70 |
| - | - | - | - | - | - | [200,250] | 3 |
| Repartition | > 4661 | Re-partition | > 4661 | Re-partition | > 4661 | Stop | < 4661 |
| * Average length of the first partition (ALFP) = $\frac{14893+8059+279+70+3}{5} \approx \mathbf{4661}$<br>Since all the sub-interval lengths are less than 4662, then stop repartition. Otherwise, continue repartition. | | | | | | | |

where $c_1, c_2, \ldots, c_n$ are the centers of $u_1, u_2, \ldots, u_n$ and $c_i$ replaced by $Y(t)$ to has more information from the state $A_i$ at time $t$.

*Step 8:* Adjust the forecasted values by considering the first difference of actual values $Y(t)$. Then it is necessary to adjust the trend of the pre-obtained forecasting value in order

**TABLE 3.** Sub-intervals after partition with corresponding fuzzy numbers.

| Number of intervals | Interval $u_i$ | Mid-point $c_i$ | Interval Code | Fuzzy Number |
|---|---|---|---|---|
| 1 | [0, 25] | 12.50 | $u_1$ | $A_1$ |
| 2 | [25, 31,25] | 28.125 | $u_2$ | $A_2$ |
| 3 | [31.27, 37.5] | 34.375 | $u_3$ | $A_3$ |
| 4 | [37.5, 43.75] | 40.625 | $u_4$ | $A_4$ |
| 5 | [43.75, 50] | 46.875 | $u_5$ | $A_5$ |
| 6 | [50, 56.25] | 53.125 | $u_6$ | $A_6$ |
| 7 | [56.25, 62.5] | 59.375 | $u_7$ | $A_7$ |
| 8 | [62.5, 75] | 68.750 | $u_8$ | $A_8$ |
| 9 | [75, 100] | 87.50 | $u_9$ | $A_9$ |
| 10 | [100,150] | 112.50 | $u_{10}$ | $A_{10}$ |
| 11 | [150,200] | 137.50 | $u_{10}$ | $A_{11}$ |
| 12 | [200,250] | 162.50 | $u_{12}$ | $A_{12}$ |

to reduce the estimated error. The adjusted forecasted values can be written by

$$\hat{F}\,(t+1) = F\,(t+1) + \mathrm{Diff}\,(Y\,(t)) \tag{8}$$

*Step 9:* Validate the performance of the proposed model.

### C. PERFORMANCE EVALUATION OF THE PROPOSED MODEL

The statistical criteria used to evaluate models are the Root Mean Squared Error (RMSE), Mean Average Percent Error (MAPE), and Thiels' U statistic, which are described in (9), (10) and (11) respectively, where $Y_i$ means the real data, $F_i$ the forecasted values and $N$ is the total number of the data set [17], [18].

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - F_i)^2}{N}} \tag{9}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Y_i - F_i}{Y_i}\right| \times 100 \tag{10}$$

$$Theil's\ U = \frac{\sqrt{\sum_{i=1}^{N}(Y_i - F_i)^2}}{\sqrt{\sum_{i=1}^{N} Y_i^2} + \sqrt{\sum_{i=1}^{N} F_i^2}} \tag{11}$$

## IV. RESULTS AND DISCUSSION

This section involves implementing the proposed model using two different types of datasets in order to provide results that show the superiority of the model used. The first dataset used is the air pollution index data collected from Kuala Lumpur, Malaysia for the period of three years 2012-2014 whose plots for training data and testing data are given in Fig.1. The second dataset is the benchmark data of enrollment of Alabama University. The results of implementing the model using the datasets mentioned are presented in the next sections.

**TABLE 4.** Linguistic time series values.

| No | Linguistic time series values *Ak* |
|---|---|
| 1 | $A_1 = \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \cdots + \frac{0}{u_{10}} + \frac{0}{u_{11}} + \frac{0}{u_{12}}$ |
| 2 | $A_2 = \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \frac{0}{u_4} + \cdots + \frac{0}{u_{10}} + \frac{0}{u_{11}} + \frac{0}{u_{12}}$ |
| 3 | $A_3 = \frac{0}{u_1} + \frac{0.5}{u_2} + \frac{1}{u_3} + \frac{0.5}{u_4} + \cdots + \frac{0}{u_{10}} + \frac{0}{u_{11}} + \frac{0}{u_{12}}$ |
| | .<br>.<br>. |
| | $A_{11} = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \cdots + \frac{0.5}{u_9} + \frac{1}{u_{10}} + \frac{0.5}{u_{11}} + \frac{0}{u_{12}}$ |
| 12 | $A_{12} = \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \frac{0}{u_4} + \cdots + \frac{0}{u_{10}} + \frac{0.5}{u_{11}} + \frac{1}{u_{12}}$ |

**TABLE 5.** API values are expressed as fuzzy numbers.

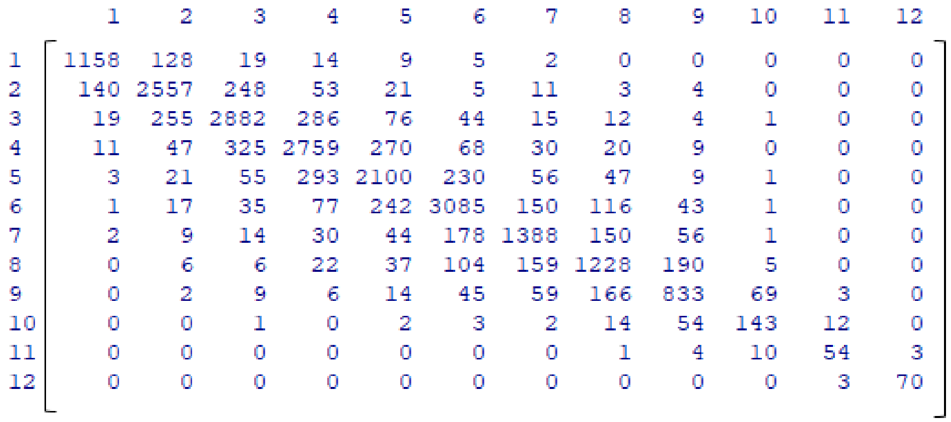
| Number of Hours | Date/Time | API | Fuzzy number | Fuzzy logic relationships |
|---|---|---|---|---|
| 1 | 2012/1/1 01:00 | 24 | $A_1$ | - |
| 2 | 2012/1/1 02:00 | 25 | $A_2$ | $A_1 \rightarrow A_2$ |
| 3 | 2012/1/1 03:00 | 26 | $A_2$ | $A_2 \rightarrow A_2$ |
| 4 | 2012/1/1 04:00 | 26 | $A_2$ | $A_2 \rightarrow A_2$ |
| 5 | 2012/1/1 04:00 | 26 | $A_2$ | $A_2 \rightarrow A_2$ |
| 6 | 2012/1/1 06:00 | 26 | $A_2$ | $A_2 \rightarrow A_2$ |
| 7 | 2012/1/1 07:00 | 25 | $A_2$ | $A_2 \rightarrow A_2$ |
| : | : | : | : | : |
| : | : | : | : | : |
| 23300 | 2014/8/31 18:00 | 40 | $A_4$ | $A_3 \rightarrow A_4$ |
| 23301 | 2014/8/31 19:00 | 41 | $A_4$ | $A_4 \rightarrow A_4$ |
| 23302 | 2014/8/31 20:00 | 41 | $A_4$ | $A_4 \rightarrow A_4$ |
| 23303 | 2014/8/31 22:00 | 42 | $A_4$ | $A_4 \rightarrow A_4$ |
| 23304 | 2014/8/31 23:00 | 40 | $A_4$ | $A_4 \rightarrow A_4$ |
| 23304 | 2014/8/31 24:00 | 40 | $A_4$ | $A_4 \rightarrow A_4$ |

### A. API FORECASTING

The air pollution index (API) data is categorized based on the highest index value of five main air pollutants namely $PM_{10}$, $CO_2$, $O_3$, $NO_2$ and $SO_2$ [45]. The API values are determined by the average indices for these five pollutant variables and then the maximum value among these five sub-indices is chosen as the API value [18], [19], [46].

**TABLE 6.** Fuzzy logical relationship groups for the tree partition method.

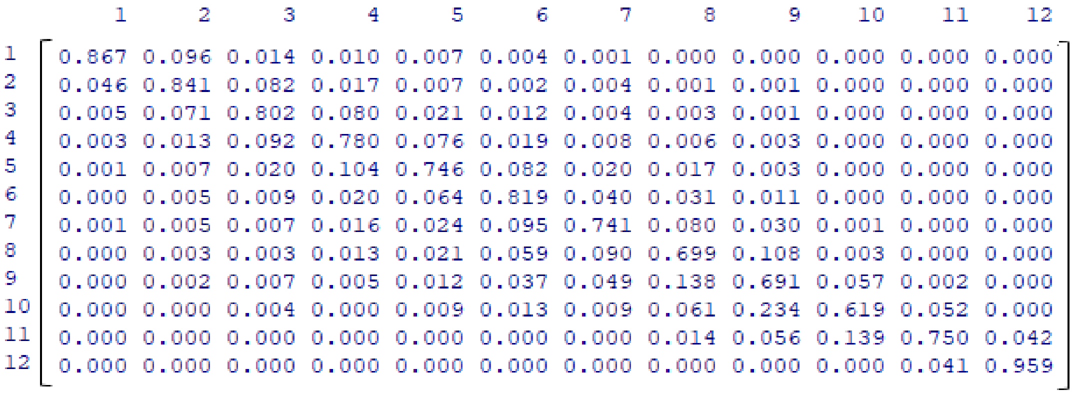
| Group | Fuzzy logical relationships (FLRG) |
|---|---|
| G1 | A1 → (1158)A1, (128)A2, (19)A3, (14)A4, (9)A5, (5)A6, (2)A7 |
| G2 | A2 → (140)A1, (2557)A2, (248)A3, (53)A4, (21)A5, (5)A6, (11)A7, (3)A8, (4)A9 |
| G3 | A3 → (19)A1, (255)A2, (2882)A3, (286)A4, (76)A5, (44)A6, (15)A7, (12)A8, (4)A9, (1)A10 |
| G4 | A4 → (11)A1, (47)A2, (325)A3, (2759)A4, (270)A5, (68)A6, (30)A7, (20)A8, (9)A9 |
| G5 | A5 → (3)A1, (21)A2, (55)A3, (293)A4, (2100)A5, (230)A6, (56)A7, (47)A8, (9)A9, (1)A10 |
| G6 | A6 → (1)A1, (17)A2, (35)A3, (77)A4, (242)A5, (3085)A6, (150)A7, (116)A8, (43)A9, (1)A10 |
| G7 | A7 → (2)A1, (9)A2, (14)A3, (30)A4, (44)A5, (178)A6, (1388)A7, (150)A8, (56)A9, (1)A10 |
| G8 | A8 → (6)A2, (6)A3, (22)A4, (37)A5, (104)A6, (159)A7, (1228)A8, (190)A9, (5)A10 |
| G9 | A9 → (2)A2, (9)A3,(6)A4, (37)A5, (104)A6, (159)A7, (1228)A8, (190) A9, (5)A10 |
| G10 | A10 →(2)A5, (3)A6, (2)A7, (14)A8, (54)A9, (143)A10, (12)A11 |
| G11 | A11 → (1)A8, (4)A9, (10)A10, (54)A11, (3)A12 |
| G12 | A12 → (3)A11,  (70) A12 |



```
        1     2     3     4     5     6     7     8     9    10    11    12

 1   1158   128    19    14     9     5     2     0     0     0     0     0
 2    140  2557   248    53    21     5    11     3     4     0     0     0
 3     19   255  2882   286    76    44    15    12     4     1     0     0
 4     11    47   325  2759   270    68    30    20     9     0     0     0
 5      3    21    55   293  2100   230    56    47     9     1     0     0
 6      1    17    35    77   242  3085   150   116    43     1     0     0
 7      2     9    14    30    44   178  1388   150    56     1     0     0
 8      0     6     6    22    37   104   159  1228   190     5     0     0
 9      0     2     9     6    14    45    59   166   833    69     3     0
10      0     0     1     0     2     3     2    14    54   143    12     0
11      0     0     0     0     0     0     0     1     4    10    54     3
12      0     0     0     0     0     0     0     0     0     0     3    70
```

**FIGURE 2.** Frequencies (count matrix) based on fuzzy logic relationship groups.



```
       1     2     3     4     5     6     7     8     9     10    11    12

 1  0.867 0.096 0.014 0.010 0.007 0.004 0.001 0.000 0.000 0.000 0.000 0.000
 2  0.046 0.841 0.082 0.017 0.007 0.002 0.004 0.001 0.001 0.000 0.000 0.000
 3  0.005 0.071 0.802 0.080 0.021 0.012 0.004 0.003 0.001 0.000 0.000 0.000
 4  0.003 0.013 0.092 0.780 0.076 0.019 0.008 0.006 0.003 0.000 0.000 0.000
 5  0.001 0.007 0.020 0.104 0.746 0.082 0.020 0.017 0.003 0.000 0.000 0.000
 6  0.000 0.005 0.009 0.020 0.064 0.819 0.040 0.031 0.011 0.000 0.000 0.000
 7  0.001 0.005 0.007 0.016 0.024 0.095 0.741 0.080 0.030 0.001 0.000 0.000
 8  0.000 0.003 0.003 0.013 0.021 0.059 0.090 0.699 0.108 0.003 0.000 0.000
 9  0.000 0.002 0.007 0.005 0.012 0.037 0.049 0.138 0.691 0.057 0.002 0.000
10  0.000 0.000 0.004 0.000 0.009 0.013 0.009 0.061 0.234 0.619 0.052 0.000
11  0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.014 0.056 0.139 0.750 0.042
12  0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.041 0.959
```

**FIGURE 3.** Markov transition probability matrix based on fuzzy logic relationship groups.

In this study, the hourly API values, which were gathered from an air monitoring station located in Kuala Lumpur, Malaysia are used in the analysis to validate the proposed model. The API dataset is divided into a training dataset which is from the 1st of January 2012 to 31st of August 2014 and the testing dataset which is from the 1st

**TABLE 7.** Fuzzy logical relationship groups for the TPM method.

| Hour | Actuals (API) | Forecasts | Differences | Adjusted Forecast |
|---|---|---|---|---|
| 2012/1/1 1:00 | 24 | - | - | - |
| 2012/1/1 2:00 | 25 | 25.03380 | 1 | 26.03380 |
| 2012/1/1 3:00 | 26 | 25.90812 | 1 | 26.90812 |
| 2012/1/1 4:00 | 26 | 26.74869 | 0 | 26.74869 |
| 2012/1/1 5:00 | 26 | 26.74869 | 0 | 26.74869 |
| 2012/1/1 6:00 | 26 | 26.74869 | 0 | 26.74869 |
| 2012/1/1 7:00 | 25 | 26.74869 | -1 | 25.74869 |
| : | : | : | : | : |
| 2014/12/31 19:00 | 40 | 39.68487 | 1 | 40.68487 |
| 2014/12/31 20:00 | 41 | 41.24407 | 1 | 42.24407 |
| 2014/12/31 21:00 | 41 | 41.24407 | 0 | 41.24407 |
| 2014/12/31 22:00 | 42 | 41.24407 | 1 | 42.24407 |
| 2014/12/31 23:00 | 40 | 42.02366 | -2 | 40.88306 |
| 2014/12/31 24:00 | 40 | 40.46447 | 0 | 40.46447 |

of September 2014 to 31st of December 2014. To verify the proposed method, the API dataset is considered to be used for evaluating the performance and compare with other methods.

The implementation of the proposed model and the detailed computation processes are demonstrated step by step as follow:

*Step 1:* Define the universe of discourse $U$ from data (API data).

Let $U = [D_{min}-D_1, D_{max}+D_2]$, then $U = [9-5, 205+5]$, hence $U = [4, 210]$.

*Step 2:* Partitioning the universe of discourse $U$ based on the algorithm of the proposed partition method (TPM), as shown in Table 1. TPM is applied for partitioning the universe of discourse $U$ of air pollution index data as shown in Table 2, which partition numerical data into equal and unequal intervals. For further explanation, in this study, we have the universe of discourse $U = [4, 210]$ for the historical data of API that gathered from Kuala Lumpur, Malaysia. The partitioning has been conducted as follows: First, partition $U$ into five equal intervals of linguistic values according to the

API classifications such as; $u_1 = [0, 50]$, $u_2 = [50,100]$, $u_3 = 100,150]$, $u_4 = [150,200]$, $u_5 = [200,250]$, which are indicating good, moderate, unhealthy, very unhealthy, and very unhealthy state, respectively, according to the classifications of API data.

Second, re-partition the intervals that its number of historical is larger than the average length, meaning that the original linguistic value should be farther partitioned in half.

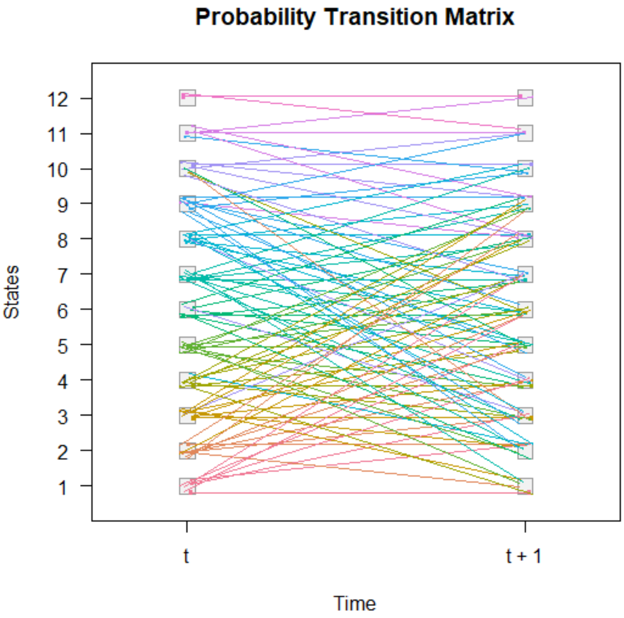


**FIGURE 4.** Visualization of the transition probability matrix.

**TABLE 8.** Comparison of statistical criteria of the proposed model with another benchmark FTS models using the training dataset.

| Model | Statistical Criterions | | | Rank |
|---|---|---|---|---|
| | MAPE | RMSE | Theil's U | |
| **Song & Chosom [25]** | 65.766 | 27.339 | 4.2364 | 10 |
| **Chen [26]** | 52.266 | 21.678 | 3.2302 | 8 |
| **Heuristic [52]** | 53.467 | 24.376 | 3.6322 | 9 |
| **Singh [45]** | 5.584 | 4.648 | 0.6293 | 4 |
| **AN. model [47]** | 27.018 | 10.999 | 1.6373 | 7 |
| **AM. model [47]** | 16.556 | 10.131 | 1.0934 | 6 |
| **Chen-Hsu [51]** | 14.131 | 7.306 | 1.0886 | 5 |
| **Efendi [49]** | 6.148 | 4.984 | 0.6789 | 3 |
| **Tsaur [22]** | 5.323 | 3.815 | 0.5194 | 2 |
| **Proposed Model** | **1.781** | **1.412** | **0.2754** | **1** |

As a result, we have found that the length of $u_1$ and $u_2$ are larger than the average length. Then, we should re-partition $u_1$ and $u_2$ in half. After that, we check again if there is any of the intervals has a length larger than the first average length. At that point, we repeat the re-partitioning until all the sub-intervals have a length less than the first average length. Once all lengths are less than the average length, then we stop and consider the final partition in the analysis as an optimal partition.

*Step 3:* Define fuzzy sets. Based on the intervals $u_k$, $(k = 1, 2, \ldots, n)$ found in the previous step, fuzzy sets $A_k$, $(k = 1, 2, \ldots, n)$ are determined with the function membership by

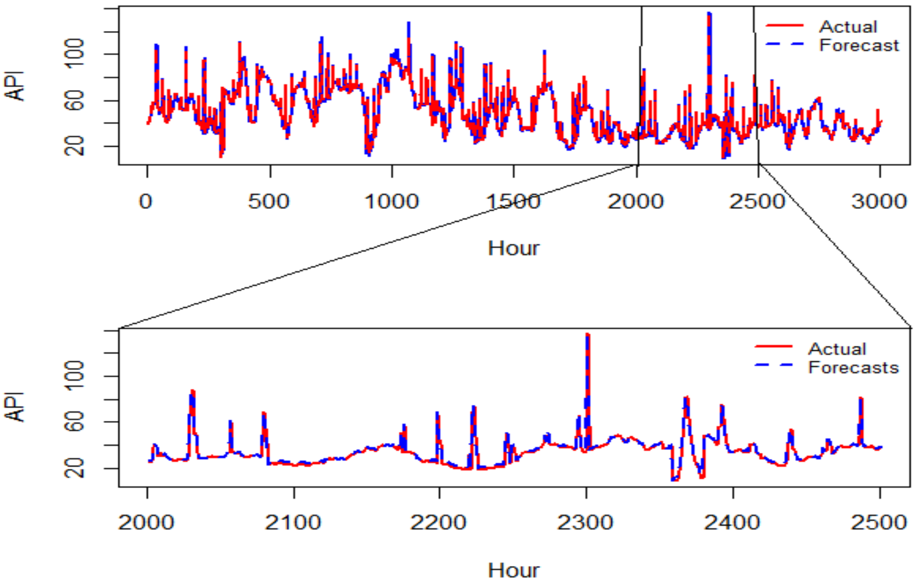


**FIGURE 5.** Comparison between the actual values and predicted values based on the proposed model.

**TABLE 9.** Comparison of statistical criteria of the proposed model with another benchmark FTS models using the testing dataset.

| Model | Statistical Criterions | | | Rank |
|---|---|---|---|---|
| | MAPE | RMSE | Theil's U | |
| Song & Chosom [25] | 67.951 | 27.952 | 4.6344 | 10 |
| Chen [26] | 56.622 | 24.439 | 3.7831 | 8 |
| Heuristic [52] | 54.529 | 24.903 | 3.8553 | 9 |
| Singh [50] | 5.7152 | 4.822 | 0.6945 | 3 |
| AN. model [47] | 27.992 | 11.406 | 1.6863 | 7 |
| AM. model [47] | 17.735 | 10.836 | 1.0934 | 6 |
| Chen-Hsu [51] | 14.849 | 8.006 | 1.1066 | 5 |
| Efendi [49] | 5.688 | 5.227 | 0.6903 | 2 |
| Tsaur [22] | 5.737 | 4.041 | 0.5422 | 4 |
| Proposed Model | **1.801** | **1.464** | **0.2573** | **1** |

using (3) as given in Table 3. Table 3 presents the fuzzy sets $A_k, (k = 1, 2, 3, \ldots, n)$. It could be seen that the greater the value of $k$ indicating that the fuzzy set of API values will move from the lowest to the highest fuzzy set of API values.

*Step 4:* Fuzzify the dataset into linguistic values. Transform the API data into the linguistic time series values and establish the fuzzy logic relationships (FLRs) as shown in Tables 3 and 5 respectively, which demonstrates the transformations of the actual API to be the linguistic time series values. The FLRs among these values are also established. Since $u_1$ has the maximum membership degree in fuzzy set $A_1$, observation 24 is mapped into a fuzzy set $A_1$. Similarly, the other values of the air pollution index are fuzzified. The actual observations and the corresponding fuzzified observations obtained from the fuzzification process have presented in Table 3. The established fuzzy logic relationships (FLRs) are shown in Table 5.

*Step 5:* Establish fuzzy logical relationships groups (FLRGs) and frequencies (count) matrix of fuzzy relation between observations. This step shows that the fuzzy logic relationship group (FLRGs) with the same left-hand side (LHSs) can be grouped into the FLRGs. The groups are given as in Table 6 presents twelve groups of the linguistic time series values, which have been found with various FLRs. Based on Table 6, the Markov transition frequency matrix or frequencies (count) matrix of fuzzy relation between observations is determined, which could be a matrix $\boldsymbol{N}_{12\times 12}$ as shown in Fig. 2.

*Step 6:* Calculate the Markov transition probability matrix $\boldsymbol{P}$ based on the matrix of frequencies from step 5 by using (4) as shown in Fig. 3. This figure shows that Markov transitions of the linguistic time series values that are used for establishing the Markov transition probability matrix $\boldsymbol{P}_{12\times 12}$ using (4), which can be used for calculating the forecasting values in the next step. Furthermore, the transition process diagram could be established for visualization of the transitions of the fuzzy sets based on Markov transition probability matrix $\boldsymbol{P}$ which is given in Fig. 4.

*Step 7:* Forecast values are calculated by using (6) or (7) based on Markov weights. For example, the forecast value for the hour (2012/1/1 2:00) is calculated by using (7) as follows:

$$\begin{aligned} F(t+1) &= c_1 p_{i1} + c_1 p_{12} + \ldots + c_{i-1} p_{i(i-1)} + Y(t) p_{ii} \\ &\quad + c_{i+1} p_{i(i+1)} + \ldots + c_n p_{in} \\ F(2) &= Y(t) p_{11} + c_2 p_{12} + c_3 p_{13} + c_4 p_{14} + c_5 p_{15} \\ &\quad + c_6 p_{16} + c_6 p_{16} = 25.033 \end{aligned}$$

*Step 8:* The forecasted values are adjusted by using (8). In the same way, calculate the forecast values based on

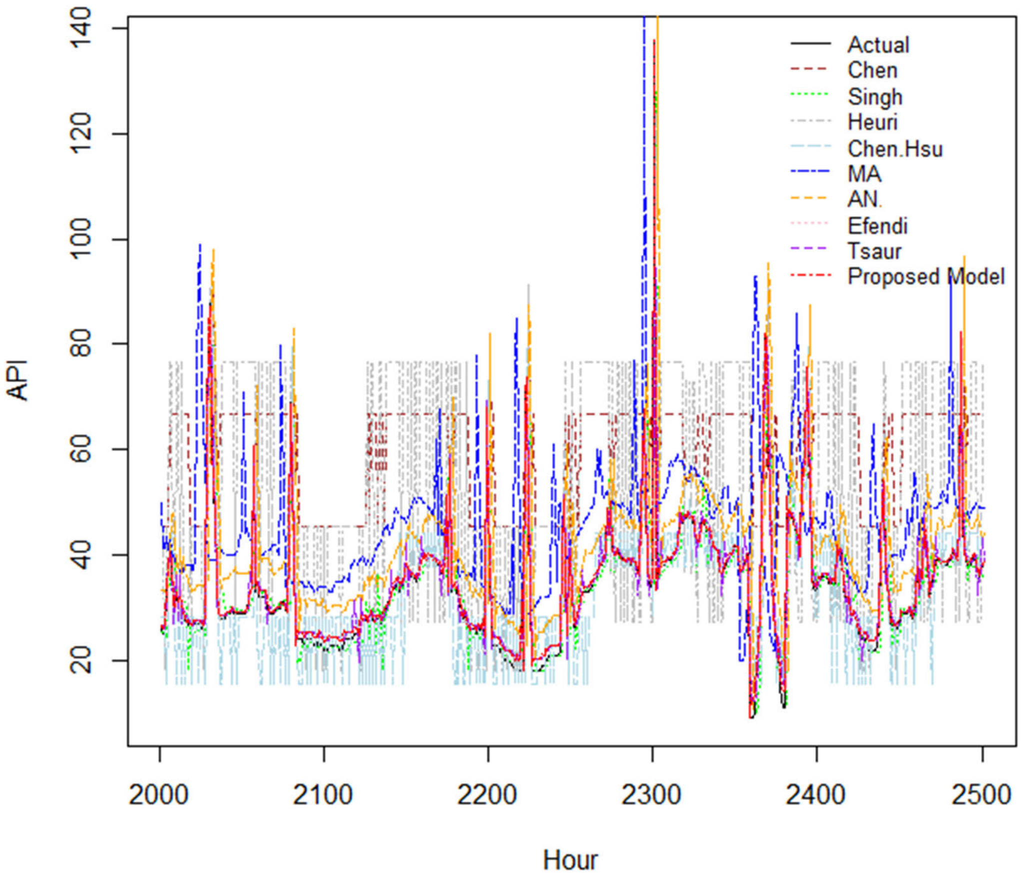


**FIGURE 6.** Comparison of the proposed model and some benchmark FTS models using the testing data.

**TABLE 10.** Comparison of the FTSMC model with a different number of intervals based on several partition methods.

| Partition Method | Number of Intervals | Training Dataset |  | Testing Dataset |  |
|---|---|---|---|---|---|
|  |  | MAPE | RMSE | MAPE | RMSE |
| **Based on API Classification [33]** | 5 | 4.1537 | 2.1855 | 3.2692 | 2.2490 |
| **Random (Tsaur's Model) [21]** | 7 | 2.5841 | 1.6488 | 2.8803 | 2.2213 |
| **Grid Partition [18]** | 18 | 2.0026 | 1.2384 | 2.441 | 1.3713 |
| **Random selecting** | 10 | 2.7762 | 1.7512 | 2.7507 | 2.2467 |
| **Based on Sterling's Formula [30]** | 15 | 2.4515 | 1.6488 | 2.6756 | 1.9991 |
| **Average Based Method [27]** | 18 | 2.7762 | 1.7512 | 2.7507 | 2.2467 |
| **Automatic Clustering [56]** | 11 | 2.5324 | 1.7094 | 2.5223 | 1.9923 |
| **K-Means Clustering [17]** | 14 | 2.2167 | 1.6021 | 2.1230 | 1.7822 |
| **Tree partition Method (TPM)** | **12** | **1.8112** | **1.4997** | **1.9034** | **1.5199** |

the obtained results of each partition method in order to fit the optimum partition method that provides the best results. For example, in step 7 we have found the forecast value is 56.66 which is move up then using (8) as follows:

$$\hat{F}\,(\mathrm{t}+1) = F\,(\mathrm{t}+1) - [Y\,(t) - Y\,(t-1)]$$
$$\hat{F}\,(3) = F\,(3) - [Y\,(t) - Y\,(t-1)]$$
$$\hat{F}\,(3) = 25.03380 - (25-24) = 26.03380$$

In general, the results of the forecasted values are adjusted and presented in Table 7.

*Step 9:* Performance evaluation of the proposed model. The evaluation of the proposed model is presented in subsection C using (9)–(11).

### B. MODEL EVALUATION

The performance of the proposed model is assessed in this subsection. In addition, a comparison of the proposed model

**TABLE 11.** Comparison of the proposed model and some existing models.

| Dataset | Model | Statistical Criteria | | |
|---|---|---|---|---|
| | | MAPE | RMSE | Theil's U |
| Training | ARIMA(2,1,3) [53] | 6.5284 | 6.2413 | 0.8271 |
| | ARIMA-GARCH [54] | 5.8582 | 5.1156 | 0.7184 |
| | ANN [21] | 5.0316 | 4.9735 | 0.6547 |
| | WANN [57] | - | 3.3474 | 0.5248 |
| | WARIMA [55] | 9.2750 | 6.1471 | 0.8032 |
| | **FTSMC-TPM** | **1.7814** | **1.4122** | **0.2571** |
| Testing | ARIMA(3,1,1) [53] | 6.843 | 6.5034 | 0.7843 |
| | ARIMA-GARCH [54] | 6.054 | 5.8183 | 0.7553 |
| | ANN [21] | 5.858 | 5.1156 | 0.7271 |
| | WANN [57] | - | 5.1904 | 0.7094 |
| | WARIMA [55] | 10.303 | 7.2691 | 0.8483 |
| | **FTSMC-TPM** | **1.8013** | **1.4641** | **0.2753** |

**TABLE 12.** Comparison between actual and predicted values in terms of the model accuracy.

| Values | Statistical Descriptive | | | | | | Accuracy | |
|---|---|---|---|---|---|---|---|---|
| | Min | Max | Mean | SD | kurtosis | skewness | $R^2$ | MAE |
| **Actual** | 9 | 137 | 48.44 | 19.316 | -0.0383 | 0.6644 | **0.9932** | **0.8682** |
| **Predicted** | 9.16 | 137.6 | 48.53 | 18.842 | -0.287 | 0.6029 | | |

and several conventional fuzzy time series models offered by [22], [25], [26], [47]–[51] is presented in order to further examine the performance of the proposed model.

It can be seen that the proposed forecasting model based on the testing dataset of the API has modeled the air pollution very well as shown in Fig. 5, where the actual values of API are very similar to the forecasted values. It also observed from Table 8 and Fig. 6 that the performance of the proposed model using the training dataset is very well, where the proposed model produces the smallest values of three statistical criteria, which are RMSE, MAPE and U statistic as compared to the other existing FTS models. In general, based on the results, the proposed method is considered adequate in determining the length of intervals of the universe of discourse, which produces better forecasting accuracy.

Besides, Table 9 reveals that the proposed model using the testing dataset has performed very well as compared to the existing fuzzy time series models, indicating that the proposed model outperforms the existing forecasting models. This implies that the proposed model is powerful for predicting air pollution occurrences.

To have a comparison of accuracy in forecasted values of the proposed model with another benchmark fuzzy time series models, the statistical criteria (MAPE, RMSE and Thali's U statistic) in the forecast have been computed, and the values of these statistical criteria of above models are placed in Tables 8 and 9.

**TABLE 13.** Paired t-test for actual and predicted values in terms of the level of significance.

| Model | P-value | Remark of Difference |
|---|---|---|
| **Actual vs Predicted** | 0.59821 | Not significant |

The comparative study of statistical criteria, as could be seen from Tables 8 and 9, exhibits that the forecasts of the proposed model are more accurate than other models. The trends in the forecast of the other models are being illustrated in Fig. 6. Moreover, a comparison of the proposed partition method (TPM) and other partition methods is given in Table 10. It is found from the results in Table 10 that the proposed partition method produces the smallest errors as compared to the other partition methods, indicating that the proposed partition method provides a very good portion of API data. This implies that the proposed partition method (TPM) is superior in partitioning the universe of discourse of any random data such as air pollution index (API).

Apart from that ARIMA [53] models with different lags are applied to the API data and the results are presented in Table 16 in Appendix. Likewise, Fig. 8 in the Appendix shows the ACF and partial ACF plots for the training dataset to demonstrate the autocorrelation of the data. It can be seen from Table 16 that the models ARIMA (2,1,3) and ARIMA (3,1,1) using the training dataset and testing

**TABLE 14. Description of the procedures of the tree partition method (TPM) for enrollment of the University of Alabama.**

| The universe of discourse is defined as $U$ =[13000, 20000] | | | | | |
|---|---|---|---|---|---|
| **1st Partition**<br>Using any common partition method based on the type of dataset (at least 3-5 intervals). | | **2ed Partition**<br>Via divide each interval to half | | **3rd Partition**<br>Via divide each interval to half | |
| ***A. Interval*** | **length** | ***B. Interval-R1*** | **length** | ***C. Interval-R2*** | **length** |
| [13000, 14750] | 4 | [0, 25] | 4 | [0, 25] | 4 |
| [14750, 16500] | **10** | [14750, 15625] | **7** | [14750, 15187.5] | 2 |
| | | | | [15187.5, 15625] | 5 |
| | | [15625, 16500] | 3 | [50, 62.6] | 3 |
| [16500, 18250] | 4 | [150,200] | 4 | [75, 100] | 4 |
| [18250, 20000] | 4 | [200,250] | 4 | [100,150] | 4 |
| - | - | - | - | - | - |
| Repartition | > 5.5 | Re-partition | > 5.5 | Re-partition | < 5.5 |
| * Average length of the first partition (ALFP) = $\frac{4+10+4+4}{4} \approx$ **5.5**<br>Since all the sub-interval lengths are less than 4662 then stop repartition. Otherwise, continue repartition. | | | | | |

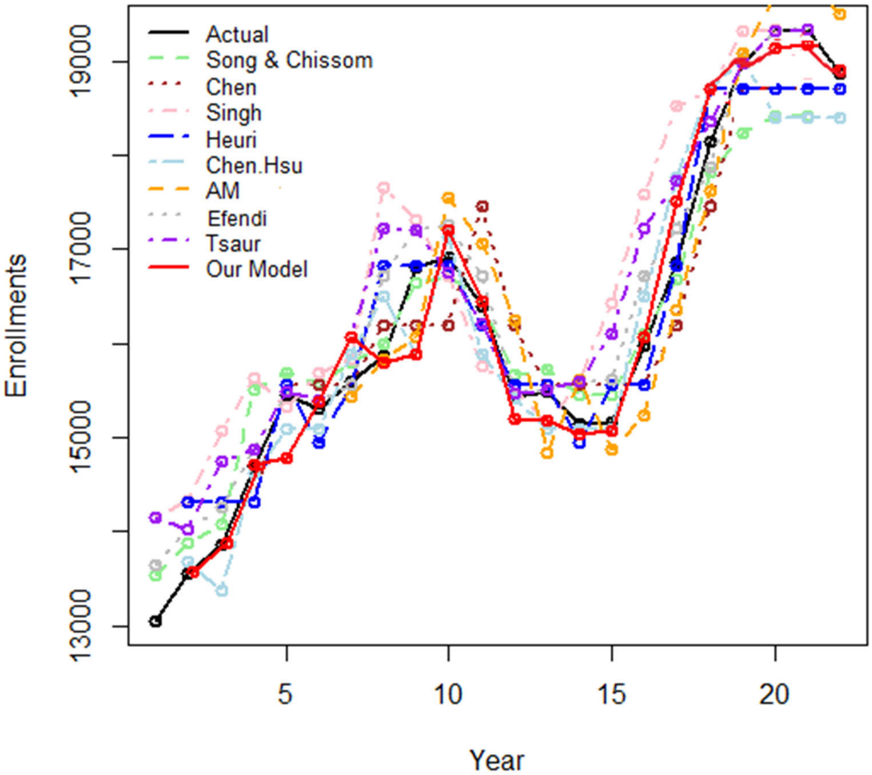


**FIGURE 7. Comparison of the proposed model and fuzzy time series models using enrollments data for the University of Alabama.**

dataset respectively are the best ARIMA models as compared to the other models. Accordingly, the proposed model is compared with the best models of ARIMA and some other models such as ARIMA-GARCH [54], the artificial neural

**TABLE 15.** Comparison of the proposed model and fuzzy time series models using enrollments data for the University of Alabama.

| Data | S&C [25] | Chen [26] | Singh [50] | Heuristic [52] | Chen.Hsu [51] | Efendi [49] | AM [47] | Tsaur [22] | FTSMC-TPM |
|---|---|---|---|---|---|---|---|---|---|
| **13055** | - | - | - | - | - | - | - | - | - |
| **13563** | 14150.8 | 14311.4 | - | 14311.4 | 13683.2 | 13642.81 | - | 13537 | 13563.7 |
| **13867** | 14327.8 | 14311.4 | - | 14311.4 | 13369.1 | 14023.81 | - | 13875 | 13866.7 |
| **14696** | 14078 | 14311.4 | 14709.3 | 14311.4 | 14625.5 | 14251.81 | - | 14578 | 14690.4 |
| **15460** | 15500 | 15567.8 | 14792.07 | 15567.8 | 15096.65 | 14873.56 | - | 16000 | 15637.56 |
| **15311** | 15691.2 | 15567.8 | 15383.15 | 14939.6 | 15096.65 | 15482.25 | - | 15691 | 15333.25 |
| **15603** | 15575.2 | 15567.8 | 16070.37 | 15567.8 | 15881.9 | 15392.85 | 15429.53 | 15575 | 15684.85 |
| **15861** | 15802.3 | 16196 | 15797.78 | 16824.2 | 16510.1 | 15568.05 | 15835.45 | 15802 | 15826.05 |
| **16807** | 16003 | 16196 | 15886.77 | 16824.2 | 15881.9 | 16718.75 | 16070.58 | 16503 | 16664.75 |
| **16919** | 16653.5 | 16196 | 17199.83 | 16824.2 | 17138.3 | 17200.38 | 17549.54 | 16653 | 17312.37 |
| **16388** | 16709.5 | 17452.4 | 16454.7 | 16196 | 15881.9 | 17256.38 | 17067.64 | 16709 | 16725.37 |
| **15433** | 16444 | 16196 | 15206.86 | 15567.8 | 15410.75 | 16718.75 | 16246.16 | 15944 | 15763.75 |
| **15497** | 15670.1 | 15567.8 | 15185.85 | 15567.8 | 15096.65 | 15466.05 | 14834.54 | 15670 | 15530.05 |
| **15145** | 15719.9 | 15567.8 | 15043.97 | 14939.6 | 15096.65 | 15504.45 | 15623.16 | 15719 | 15152.45 |
| **15163** | 15446.1 | 15567.8 | 15066.86 | 15567.8 | 15096.65 | 15603.75 | 14874.89 | 15446 | 15221.75 |
| **15984** | 15460.1 | 15567.8 | 16064.02 | 15567.8 | 16510.1 | 15612.75 | 15243.03 | 15460 | 16433.75 |
| **16859** | 16098.7 | 16196 | 17521.2 | 16824.2 | 17766.5 | 16718.75 | 16356.42 | 16599 | 17193.75 |
| **18150** | 16679.5 | 17452.4 | 18708.8 | 18708.8 | 18708.8 | 17226.38 | 17618.52 | 17678 | 18517.37 |
| **18970** | 17825 | 18708.8 | 18452.65 | 18708.8 | 19022.9 | 17871.88 | 19084.45 | 17825 | 18891.87 |
| **19328** | 18235 | 18708.8 | 18878.45 | 18708.8 | 18394.7 | 18970.00 | 19695.81 | 18735 | 19080.91 |
| **19337** | 18414 | 18708.8 | 18923.21 | 18708.8 | 18394.7 | 19328.00 | 19866.68 | 18414 | 19162.96 |
| **18876** | 18418.5 | 18708.8 | 18902.41 | 18708.8 | 18394.7 | 19337.00 | 19510.40 | 17919 | 18870.39 |
| | | | | | **Statistical Criterion** | | | | |
| **MAPE** | 3.2492 | 2.876 | 1.8846 | 2.3873 | 2.5151 | 2.494 | 2.9528 | 2.026 | **0.9556** |
| **RMSE** | 668.96 | 535.8 | 408.479 | 507.20 | 519.98 | 539.28 | 545.42 | 415.9 | **221.9** |
| **U Stat.** | 0.895 | 0.8603 | 0.63776 | 0.6667 | 0.6475 | 0.670 | 0.61745 | - | **0.431** |

networks (ANN) [21], Wavelet ARIMA (WARIMA) [55] and Wavelet artificial neural networks (WANN) [50]. These four models have been implemented using the API data for a comparison of the proposed model, as the results found are shown in Table 11. It is observed from Table 11 that the proposed model outperforms the existing models as the proposed model produces the smallest error values of the statistical criteria. In addition, the mean absolute error (MAE) and the coefficient of determination $R^2$ for the proposed model based on the actual and predicted values are calculated. A small value of MAE and a high value of $R^2$ indicate that the model has been adequately fitted. The coefficient of $R^2$ and the MAE can be computed respectively as follows:

$$R^2 = \frac{\sum_{i=1}^{N} \left(\hat{y}_i - \bar{y}\right)^2}{\sum_{i=1}^{N} \left(\hat{y}_i - \bar{y}\right)^2 + \sum_{i=1}^{N} \left(y_i - \hat{y}_i\right)^2} \tag{12}$$

$$\mathrm{MAE} = \sqrt{\frac{\sum_{i=1}^{N} \left(y_i - \hat{y}_i\right)^2}{N}} \tag{13}$$

where $y_i$ is observed data, $\hat{y}_i$ is the predicted data, $\bar{y}$ is the mean observed data and $N$ is the total number of data.

It can be seen from Table 12 that the value of $R^2$ is very high and the value of MAE is very small, indicates that the proposed model has been fitted very well. Table 13 shows the paired t-test for actual and predicted values in terms of the level of significance. It is found from this table that the null hypothesis is rejected, which indicates that the actual and predicted values are close enough to conclude that the model is well fitted, where the mean difference between the paired observations is not significant and does not differ from each other. Table 13 shows that there is no significance in the difference between the actual and predicted values, indicating that actual and predicted are very similar. This is implied

**TABLE 16. Fitting the best ARIMA model.**

| **Training dataset** | | **Testing dataset** | |
|---|---|---|---|
| **ARIMA(*p*,*r*,*q*)** | **AIC** | **ARIMA(*p*,*r*,*q*)** | **AIC** |
| **ARIMA(2,1,2)** | 151515.5 | ARIMA(1,1,0) | 19939.53 |
| **ARIMA(0,1,0)** | 153034.1 | ARIMA(0,1,1) | 19938.53 |
| **ARIMA(1,1,0)** | 153019.5 | ARIMA(0,1,0 | 19934.57 |
| **ARIMA(0,1,1)** | 153017.4 | ARIMA(1,1,2) | 19788.46 |
| **ARIMA(0,1,0)** | 153032.1 | ARIMA(2,1,1) | 19783.09 |
| **ARIMA(1,1,2** | 151732.4 | ARIMA(3,1,2) | 19761.71 |
| **ARIMA(2,1,1)** | 151672.7 | ARIMA(3,1,1) | 19759.88 |
| **ARIMA(3,1,2)** | 151491.5 | ARIMA(3,1,0) | 19881.77 |
| **ARIMA(3,1,1)** | 151502.1 | ARIMA(4,1,1) | 19763.27 |
| **ARIMA(4,1,2)** | 151494.0 | ARIMA(2,1,0) | 19937.6 |
| **ARIMA(3,1,3)** | 151500.5 | ARIMA(4,1,0) | 19859.43 |
| **ARIMA(2,1,3)** | **151488.5** | ARIMA(4,1,2) | 19764.44 |
| **ARIMA(1,1,3)** | 151555.3 | **ARIMA(3,1,1)** | **19757.88** |
| **ARIMA(2,1,4)** | 151489.1 | ARIMA(2,1,1) | 19781.08 |
| **ARIMA(1,1,2)** | 151730.4 | ARIMA(3,1,0) | 19879.76 |
| **ARIMA(1,1,4)** | 151507.9 | ARIMA(4,1,1) | 19761.26 |
| **ARIMA(3,1,2)** | 151489.5 | ARIMA(3,1,2) | 19759.7 |
| **ARIMA(3,1,4)** | 151491.5 | ARIMA(2,1,0) | 19935.6 |
| **-** | - | ARIMA(2,1,2) | 19764.01 |
| **-** | - | ARIMA(4,1,2) | 19762.43 |

that the model has been fitted very well, and is adequate for predicting air pollution time-series data.

### C. ENROLLMENT FORECASTING

Following the same algorithm of the proposed model is being implemented on the second dataset, which is the time-series data of enrollments at the University of Alabama for further validation of the proposed model. A description of the Tree Partition Method (TPM) processes for enrollments at the University of Alabama dataset is demonstrated in Table 14. As could be seen from the Table that the optimal number of partition is six partitions with unequal lengths of intervals. Figure 8 and Table 15 show the performance of the proposed model using the enrollment data of Alabama University. Table 15 shows a comparison of the proposed model and eight existing FTS models. It is observed from Table 15 and Fig. 8 that the proposed model outperformed the existing FTS model, where it gave better forecasts as compared to existing

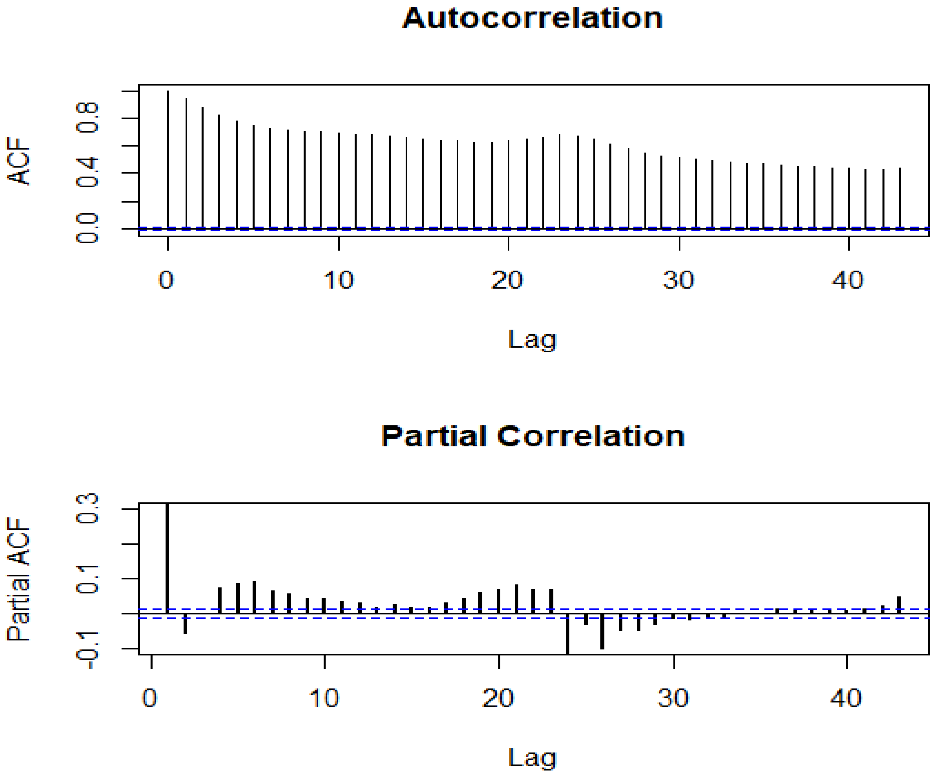


**FIGURE 8.** ACF and partial ACF plot for the training dataset.

models and produced the smallest values of the statistical criteria applied. This implied that the proposed model is a better option for predicting any type of random time-series data.

Based on the results, it is obvious that the proposed model is considered better than the other models with the smallest forecasting error according to MAPE; thus, the proposed model is the most accurate of the approaches used. Therefore, FTSMC-TPM can be used to establish the forecasting model with relative ease and accurate forecasting performance. Nevertheless, the method proposed might be not adequate for datasets that the value at time $t$ got affected by some values in the past with arbitrary time delay. Besides, in our proposed model, if the collected data set is too limited, we might not derive the transition probability matrix and may inadequately fit the model.

## V. CONCLUSION

This study proposed a novel fuzzy time series forecasting model based on the tree partition method (TPM), which provides an optimal partition of the universe of discourse. In this study, the algorithm of the proposed partition method offers a simple computational method compared to the clustering methods. It searches for a suitable defuzzification process and provides the forecasted values with a smaller error and better accuracy. In addition, the Markov transition probability Matrix (weights matrix) of the fuzzy logical relationships (FLRs) has been calculated. The model has been implemented for forecasting the air pollution in Malaysia, using the time series of real data of air pollution index (API) collected from Kuala Lumpur, for a period of three years. In addition, a comparative study of the proposed model and the existing classic and advanced time series models has been investigated. Further, the model has also been implemented on the historical time series data of enrollments of the University of Alabama. The forecasted values obtained by the model show its suitability in the fuzzy time series forecasting of air pollution without any prior knowledge of the production governing parameters. In forecasting the air pollution index, it shows that the proposed method produces a superior forecasting accuracy compared to the conventional and advanced time series models proposed in the literature.

In conclusion, the proposed partition method represents a promising method to improve forecasting accuracy where it can minimize the negative effects of abnormal observations on the performance of forecasting. Thus, the proposed model demonstrates its ability to avoid the arbitrary selection of intervals and dealing with recurrent observations and arbitrary length of intervals, which greatly improves model accuracy. For enrollment and air pollution forecasting, the proposed model produced a superior forecasting accuracy as compared to some conventional and advanced time series methods. For future studies, the proposed model could be performed to obtain more effective partitions of the universe of discourse and more accurate forecasts. It can also be extended by employing the high order fuzzy time series with considering the residuals in the calculations in order to avoid the model specification error, which may influence the model accuracy.

## ACKNOWLEDGMENT

The authors would like to thank the Department of Environment Malaysia for providing the air pollution data. They would also like to thank the editor and the reviewers for their constructive suggestions that improve the quality of this article.

## APPENDIX

(See Table 16 and Fig 8.)

## REFERENCES

[1] S. Yusuf, A. Mohammad, and A. A. Hamisu, "A novel two—Factor high order fuzzy time series with applications to temperature and futures exchange forecasting," *Nigerian J. Technol.*, vol. 36, no. 4, pp. 1124–1134, 2017.

[2] V. R. Uslu, E. Bas, U. Yolcu, and E. Egrioglu, "A fuzzy time series approach based on weights determined by the number of recurrences of fuzzy relations," *Swarm Evol. Comput.*, vol. 15, pp. 19–26, Apr. 2014.

[3] M. Habermann, M. Billger, and M. Haeger-Eugensson, "Land use regression as method to model air pollution. Previous results for Gothenburg/Sweden," *Procedia Eng.*, vol. 115, pp. 21–28, Jan. 2015.

[4] D. Turgut and İ. Temiz, "Time series analysis and forecasting for air pollution in Ankara: A box-Jenkins approach," *Alphanumeric J.*, vol. 3, no. 2, pp. 131–138, Dec. 2015.

[5] A. Russo, P. G. Lind, F. Raischel, R. Trigo, and M. Mendes, "Neural network forecast of daily pollution concentration using optimal meteorological data at synoptic and local scales," *Atmos. Pollut. Res.*, vol. 6, no. 3, pp. 540–549, May 2015.

[6] X. Leng, J. Wang, H. Ji, Q. Wang, H. Li, X. Qian, F. Li, and M. Yang, "Prediction of size-fractionated airborne particle-bound metals using MLR, BP-ANN and SVM analyses," *Chemosphere*, vol. 180, pp. 513–522, Aug. 2017.

[7] P. Wang, H. Zhang, Z. Qin, and G. Zhang, "A novel hybrid-garch model based on ARIMA and SVM for $PM_{2.5}$ concentrations forecasting," *Atmos. Pollut. Res.*, vol. 8, no. 5, pp. 850–860, Sep. 2017.

[8] C. Zafra, Y. Ángel, and E. Torres, "ARIMA analysis of the effect of land surface coverage on $PM_{10}$ concentrations in a high-altitude megacity," *Atmos. Pollut. Res.*, vol. 8, no. 4, pp. 660–668, Jul. 2017.

[9] L. Zhang, J. Lin, R. Qiu, X. Hu, H. Zhang, Q. Chen, H. Tan, D. Lin, and J. Wang, "Trend analysis and forecast of $PM_{2.5}$ in Fuzhou, China using the ARIMA model," *Ecol. Indicators*, vol. 95, pp. 702–710, Dec. 2018.

[10] J. He, S. Gong, Y. Yu, L. Yu, L. Wu, H. Mao, C. Song, S. Zhao, H. Liu, X. Li, and R. Li, "Air pollution characteristics and their relation to meteorological conditions during 2014–2015 in major Chinese cities," *Environ. Pollut.*, vol. 223, pp. 484–496, Apr. 2017.

[11] P. J. G. Nieto, F. S. Lasheras, E. García-Gonzalo, and F. J. de Cos Juez, "$PM_{10}$ concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study," *Sci. Total Environ.*, vol. 621, pp. 753–761, Apr. 2018.

[12] J.-I. Park, D.-J. Lee, C.-K. Song, and M.-G. Chun, "TAIFEX and KOSPI 200 forecasting based on two-factors high-order fuzzy time series and particle swarm optimization," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 959–967, Mar. 2010.

[13] Y. Alyousifi, K. Ibrahim, W. Kang, and W. Z. W. Zin, "Markov chain modeling for air pollution index based on maximum a posteriori method," *Air Qual., Atmos. Health*, vol. 12, no. 12, pp. 1521–1531, 2019.

[14] Y. Alyousifi, N. Masseran, and K. Ibrahim, "Modeling the stochastic dependence of air pollution index data," *Stochastic Environ. Res. Risk Assessment*, vol. 32, no. 6, pp. 1603–1611, Jun. 2018.

[15] H. Maleki, A. Sorooshian, G. Goudarzi, Z. Baboli, Y. T. Birgani, and M. Rahmati, "Air pollution prediction by using an artificial neural network model," *Clean Technol. Environ. Policy*, vol. 21, no. 6, pp. 1341–1352, May 2019.

[16] N. G. Dincer and Ö. Akkuş, "A new fuzzy time series model based on robust clustering for forecasting of air pollution," *Ecol. Informat.*, vol. 43, pp. 157–164, Jan. 2018.

[17] Y. Alyousifi, M. Othman, I. Faye, R. Sokkalingam, and P. C. L. Silva, "Markov weighted fuzzy time-series model based on an optimum partition method for forecasting air pollution," *Int. J. Fuzzy Syst.*, vol. 22, no. 5, pp. 1468–1486, Jul. 2020.

[18] Y. Alyousifi, M. Othman, R. Sokkalingam, I. Faye, and P. C. L. Silva, "Predicting daily air pollution index based on fuzzy time series Markov chain model," *Symmetry*, vol. 12, no. 2, p. 293, Feb. 2020.

[19] C.-H. Cheng, S.-F. Huang, and H.-J. Teoh, "Predicting daily ozone concentration maxima using fuzzy time series based on a two-stage linguistic partition method," *Comput. Math. Appl.*, vol. 62, no. 4, pp. 2016–2028, Aug. 2011.

[20] O. Cagcag, U. Yolcu, E. Egrioglu, and C. H. Aladag, "A novel seasonal fuzzy time series method to the forecasting of air pollution data in Ankara," *Amer. J. Intell. Syst.*, vol. 3, no. 1, pp. 13–19, 2013.

[21] N. H. A. Rahman, M. H. Lee, Suhartono, and M. T. Latif, "Artificial neural networks and fuzzy time series forecasting: An application to air quality," *Qual. Quantity*, vol. 49, no. 6, pp. 2633–2647, Nov. 2015.

[22] R.-C. Tsaur, "A fuzzy time series-Markov chain model with an application to forecast the exchange rate between the Taiwan and US Dollar," *Int. J. Innov. Comput., Inf. Control*, vol. 8, no. 7, pp. 4931–4942, 2012.

[23] L. A. Zadeh, "Fuzzy sets," in *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A Zadeh*. Singapore: World Scientific, 1996, pp. 394–432.

[24] Q. Song and B. S. Chissom, "Fuzzy time series and its models," *Fuzzy Sets Syst.*, vol. 54, no. 3, pp. 269–277, Mar. 1993.

[25] Q. Song and B. S. Chissom, "Forecasting enrollments with fuzzy time series—Part I," *Fuzzy Sets Syst.*, vol. 54, no. 1, pp. 1–9, 1993.

[26] S.-M. Chen, "Forecasting enrollments based on fuzzy time series," *Fuzzy Sets Syst.*, vol. 81, no. 3, pp. 311–319, Aug. 1996.

[27] K. Huarng, "Effective lengths of intervals to improve forecasting in fuzzy time series," *Fuzzy Sets Syst.*, vol. 123, no. 3, pp. 387–394, Nov. 2001.

[28] K. Huarng and T. H.-K. Yu, "Ratio-based lengths of intervals to improve fuzzy time series forecasting," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 36, no. 2, pp. 328–340, Apr. 2006.

[29] S.-M. Chen and N.-Y. Chung, "Forecasting enrollments using high-order fuzzy time series and genetic algorithms," *Int. J. Intell. Syst.*, vol. 21, no. 5, pp. 485–501, 2006.

[30] L. Lee, L. Wang, and S. Chen, "Temperature prediction and TAIFEX forecasting based on fuzzy logical relationships and genetic algorithms," *Expert Syst. Appl.*, vol. 33, no. 3, pp. 539–550, Oct. 2007.

[31] I.-H. Kuo, S.-J. Horng, T.-W. Kao, T.-L. Lin, C.-L. Lee, and Y. Pan, "An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 6108–6117, Apr. 2009.

[32] S. Davari, M. H. F. Zarandi, and I. B. Turksen, "An improved fuzzy time series forecasting model based on particle swarm intervalization," in *Proc. Annu. Meeting North Amer. Fuzzy Inf. Process. Soc. (NAFIPS)*, Jun. 2009, pp. 1–5.

[33] L.-Y. Hsu, S.-J. Horng, T.-W. Kao, Y.-H. Chen, R.-S. Run, R.-J. Chen, J.-L. Lai, and I.-H. Kuo, "Temperature prediction and TAIFEX forecasting based on fuzzy relationships and MTPSO techniques," *Expert Syst. Appl.*, vol. 37, no. 4, pp. 2756–2770, Apr. 2010.

[34] A. A. Almohammedi and V. Shepelev, "Saturation throughput analysis of steganography in the IEEE 802.11p protocol in the presence of non-ideal transmission channel," *IEEE Access*, vol. 9, pp. 14459–14469, 2021.

[35] E. Egrioglu, C. H. Aladag, M. A. Basaran, U. Yolcu, and V. R. Uslu, "A new approach based on the optimization of the length of intervals in fuzzy time series," *J. Intell. Fuzzy Syst.*, vol. 22, no. 1, pp. 15–19, 2011.

[36] E. Egrioglu, C. H. Aladag, U. Yolcu, V. R. Uslu, and M. A. Basaran, "Finding an optimal interval length in high order fuzzy time series," *Expert Syst. Appl.*, vol. 37, no. 7, pp. 5052–5055, Jul. 2010.

[37] N. F. Rahim, M. Othman, R. Sokkalingam, and E. A. Kadir, "Forecasting crude palm oil prices using fuzzy rule-based time series method," *IEEE Access*, vol. 6, pp. 32216–32224, 2018.

[38] J.-A. Jiang, C.-H. Syue, C.-H. Wang, J.-C. Wang, and J.-S. Shieh, "An interval type-2 fuzzy logic system for stock index forecasting based on fuzzy time series and a fuzzy logical relationship map," *IEEE Access*, vol. 6, pp. 69107–69119, 2018.

[39] Y.-N. Wang, J. Wang, X. Fan, and Y. Song, "Network traffic anomaly detection algorithm based on intuitionistic fuzzy time series graph mining," *IEEE Access*, vol. 8, pp. 63381–63389, 2020.

[40] S. Mao and F. Xiao, "Time series forecasting based on complex network analysis," *IEEE Access*, vol. 7, pp. 40220–40229, 2019.

[41] X. Sang, Q. Zhao, H. Lu, and J. Lu, "Weighted fuzzy time series forecasting based on improved fuzzy C-means clustering algorithm," in *Proc. IEEE Int. Conf. Prog. Informat. Comput. (PIC)*, Dec. 2018, pp. 80–84.

[42] C. Cheng, G. Cheng, and J. Wang, "Multi-attribute fuzzy time series method based on fuzzy clustering," *Expert Syst. Appl.*, vol. 34, no. 2, pp. 1235–1242, Feb. 2008.

[43] S.-T. Li, Y.-C. Cheng, and S.-Y. Lin, "A FCM-based deterministic forecasting model for fuzzy time series," *Comput. Math. Appl.*, vol. 56, no. 12, pp. 3052–3063, Dec. 2008.

[44] E. Bulut, O. Duru, and S. Yoshida, "A fuzzy time series forecasting model for multi-variate forecasting analysis with fuzzy C-means clustering," *Int. J. Comput. Inf. Eng.*, vol. 6, no. 3, pp. 671–677, 2012.

[45] U. Yolcu, C. H. Aladag, E. Egrioglu, and V. R. Uslu, "Time-series forecasting with a novel fuzzy time-series approach: An example for istanbul stock market," *J. Stat. Comput. Simul.*, vol. 83, no. 4, pp. 599–612, Apr. 2013.

[46] E. Bas, U. Yolcu, E. Egrioglu, and C. H. Aladag, "A fuzzy time series forecasting method based on operation of union and feed forward artificial neural network," *Amer. J. Intell. Syst.*, vol. 5, no. 3, pp. 81–91, 2015.

[47] A. Abbasov and M. H. Mamedova, "Application of fuzzy time series to population forecasting," Vienna Univ. Technol., Vienna, Austria, Tech. Rep. 2, 2003, pp. 545–552, vol. 12.

[48] Y. Alyousifi, K. Ibrahim, W. Kang, and W. Z. W. Zin, "Robust empirical Bayes approach for Markov chain modeling of air pollution index," *J. Environ. Health Sci. Eng.*, vol. 39, pp. 1–14, Jan. 2021.

[49] R. Efendi, Z. Ismail, and M. M. Deris, "Improved weight fuzzy time series as used in the exchange rates forecasting of US Dollar to Ringgit Malaysia," *Int. J. Comput. Intell. Appl.*, vol. 12, no. 1, Mar. 2013, Art. no. 1350005.

[50] S. R. Singh, "A computational method of forecasting based on fuzzy time series," *Math. Comput. Simul.*, vol. 79, no. 3, pp. 539–554, Dec. 2008.

[51] S.-M. Chen and C.-C. Hsu, "A new method to forecast enrollments using fuzzy time series," *Int. J. Appl. Sci. Eng.*, vol. 2, no. 3, pp. 234–244, 2004.

[52] E. Bai, W. K. Wong, W. C. Chu, M. Xia, and F. Pan, "A heuristic time-invariant model for fuzzy time series forecasting," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 2701–2707, Mar. 2011.

[53] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. Hoboken, NJ, USA: Wiley, 2015.

[54] S. Yaziz, N. Azizan, R. Zakaria, and M. Ahmad, "The performance of hybrid ARIMA-GARCH modeling in forecasting gold price," in *Proc. 20th Int. Congr. Modelling Simulation*, Adelaide, SA, Australia, 2013, pp. 1–6.

[55] N. K. Yong and N. Awang, "Wavelet-based time series model to improve the forecast accuracy of $PM_{10}$ concentrations in Peninsular Malaysia," *Environ. Monitor. Assessment*, vol. 191, no. 2, p. 64, Feb. 2019.

[56] N. Van Tinh and N. T. P. Nhung, "A hybrid forecasting model based on automatic clustering algorithm and fuzzy time series," *J. Multidisciplinary Eng. Sci. Stud.*, vol. 2, no. 10, pp. 1–7, Oct. 2016,

[57] A. B. Dariane, S. Azimi, and A. Zakerinejad, "Artificial neural network coupled with wavelet transform for estimating snow water equivalent using passive microwave data," *J. Earth Syst. Sci.*, vol. 123, no. 7, pp. 1591–1601, Oct. 2014.

**YOUSIF ALYOUSIFI** received the B.Sc. degree (Hons.) in mathematics from Thamar University, Yemen, in 2007, and the M.Sc. and Ph.D. degrees in statistics from Universiti Kebangsaan Malaysia. He is currently an Academic Staff with Thamar University. His research interests include statistical modeling based on classic and advanced time series.

**MAHMOD OTHMAN** has been an Associate Professor with Universiti Teknologi PETRONAS, Seri Iskandar, Malaysia, since March 2016. Before that, he was an Associate Professor with the Universiti Teknologi MARA, Malaysia. He has a total 25 years of working experience in education industries as an academician where 21 years with Universiti Teknologi MARA and another three years with Universiti Teknologi PETRONAS until present. His research interests include fuzzy mathematics, artificial intelligent, optimization, and decision making.

**AKRAM A. ALMOHAMMEDI** (Member, IEEE) received the B.Sc. degree in electronics engineering from Infrastructure University Kuala Lumpur, Malaysia, in 2012, the M.Sc. degree in electrical and electronics engineering majoring in computer and communication system from Universiti Kebangsaan Malaysia, Malaysia, and the Ph.D. degree in wireless communications and networks engineering from Universiti Putra Malaysia (UPM), in 2019. He is currently a Senior Researcher with South Ural State University (SUSU), Russia. His research interests include wireless communication, Markovian analysis, and data analytics.