Implementation of Revised Heuristic Knowledge in Average-based Interval for Fuzzy Time Series Forecasting of Tuberculosis Cases in Sabah

Suriana Lasaraiya¹, Suzelawati Zenian², Risman Mat Hasim³, Azmirul Ashaari⁴
Preparatory Centre for Science and Technology, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia¹
Faculty of Science and Natural Resources, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia¹,²
Dept. of Mathematics and Statistics-Faculty of Science, Universiti Putra Malaysia, Serdang, Selangor, Malaysia³
Azman Hashim International Business School, Universiti of Technology Malaysia, Johor Bahru, Johor, Malaysia⁴

Abstract—Fuzzy time series forecasting is one method used to forecast in certain reality problems. The research on fuzzy time series forecasting has been increased due to its capability in dealing with vagueness and uncertainty. In this paper, we are dealing with implementation of revised heuristic knowledge to basic average-based interval and showing that these models forecast better than the basic one. We suggest three different lengths of interval, size 5, size 10 and size 20 to be used in comparing these models of average-based interval, average-based interval with implementation of heuristic knowledge and, average-based interval with implementation of revised heuristic knowledge. These models applied to forecast the number of tuberculosis cases reported monthly in Sabah starting from January 2012 until December 2019. A few numerical examples are shown as well. The performances of evaluations are shown by comparison on the values obtained by Mean Square error (MSE) and Root Mean Square Error (RMSE).

Keywords—Fuzzy time series; forecasting; length of interval; average-based interval; heuristic knowledge

I. INTRODUCTION

Time series analysis, at the present time has become a very important research object since it has been utilized in large number of fields such as finance, health care, education and environmental. In all these fields, obtaining an accurate picture of the future is an essential requirement and this can only be possible with an appropriate and competent prediction tool. Since the main concern of the time series analysis is to predict the future, it can be considered as an inference problem [1]. We identify that there are three main kinds of inference system (IS) which are utilized in time-series analysis as a prediction tool, i.e Statistical Inference System (SIS), Fuzzy Inference System (FIS) and computational inference system (CIS).

Fuzzy time series (FTS) methods are one of the prediction tool groups based on fuzzy set theory designed for time-series and they take into consideration the dependency structure of time series. In FTS model, time series with crisp observations, by using various fuzzification techniques, is transformed into time series with fuzzy observations called fuzzy time series.

The concept of linguistics variables and their application reasoning was initially introduced by Zadeh [2], [3], [4], [5]. Song and Chissom [6], [7] were the first researchers who implemented and developed models to forecast in fuzzy time series. Their model is initially used to forecast the number of students enrolled in the University of Alabama from the year 1971 until year 1992. At the same time, they forecast using linear regression technique and compared the results obtained from time series technique. However, their methods require a large amount of computational time since their proposed methods is based on complex matrix operations. Due to limitation on time constraints, Chens [8] simplified the Song and Chissom methods by using simpler arithmetic operations and proved that his methods forecast better than the previous. Chens [9] continues his research, applying the same methods in [8], but he is focusing on high order fuzzy time series, which is more than one fuzzy logical relationship.

Huang [10] has been concerned on dealing with the suitable length of interval. A key point in choosing effective length of intervals is that they should not be too large as there will be no fluctuations in the fuzzy time series or too small as the meaning of fuzzy time series will be diminished. In his research, he proposed distribution- and average-based length in determining the effective length of intervals and proved that average-based length increases forecast accuracy compared to distribution-based length. Meanwhile, Li and Chen [11] proposed a concept of 3–4–5 rules of natural partitions to determine the length of intervals to be used, in their forecasting models.


Ramli and Mohamad (2017) [21] applied the models from [11] in their forecasting models. They used 3–4–5 natural partition rules to find the best length of intervals in their models in forecasting the unemployment rate under different degrees of confidence. The study [22] shows a comparison of
the fuzzy time series methods of Chen, Cheng and Markov chain model in predicting rainfall in Medan. In their study, Chen’s model gives MAPE=8.002%, Markov chain’s model gives MAPE=30.12% and Cheng’s model gives MAPE=34.5%. According to [23] a model has a very good performance if the MAPE value is below 10% and has a good performance if the MAPE value is between 10% and 20%. Thus, Chen’s model forecasting performances was very Good.

Norhayati et al. (2019) [24] proposed that fuzzy time series method is way more accurate compared to geometric brownian motion in forecasting stock prices in Bursa Malaysia. They apply average-based length focusing on trapezoidal membership members in their model Meanwhile, Susilo et al. (2022) [25] applied average-based fuzzy time series markov chain with some modification on partitioning using frequency density in predicting Covid-19 in Central Java, the ideal interval. The average based FTS Markov chain approach with adjustments to the frequency density partition achieves an accuracy rate of 89.3%, according to the findings of this study. Lasaraiya et. al. (2022) [27] shows that determination of the length of interval by average-based interval gives nearest forecasting value to the actual values compared to forecasting value obtained from determination of length of interval by natural partitions as it shows the smallest percentage error.

This paper will merge Chen’s model which is more on determination of suitable length of interval and Huang’s model which is more on choosing the suitable fuzzy logical relationship group (FLRG) in the model. Then, the comparison between the results of Chen’s model with basic average-based interval, Huang’s model with average-based interval with implementation of heuristic knowledge and the proposed average-based interval model with implementation of revised heuristic knowledge will be shown. Section II will briefly explain the fuzzy time series background. Section III will be on the general step in the proposed methods. The Section IV presents results and discussion followed by future work and conclusions.

II. Fuzzy Time Series

Fuzzy time series concept was proposed by Song and Chissom [6], [7]. It can be used as prediction tools in a real-life problem or cases where historical data are formed in linguistic value. This means that the data in fuzzy time series is linguistic data while the actual data is real number. The definition of fuzzy time series is briefly reviewed below.

Let $U$ be the universe of discourse, where $U = u_1, u_2, ..., u_n$. A fuzzy set of $A_i$ of $U$ is defined by

$$A_i = \frac{f_{A_1}(u_1)}{u_1} + \frac{f_{A_2}(u_2)}{u_2} + ... + \frac{f_{A_n}(u_n)}{u_n} (1)$$

where $f_{A_i}$ is the membership function of the fuzzy set $A_i$, $f_{A_i}: U \rightarrow [0,1]$. $u_k$ is the element of fuzzy set $A_i$, and $f_{A_i}(u_k)$ is the degree of belongingness of $u_k$ to $A_i$ for $f_{A_i}(u_k) \in [0,1]$ and $1 \leq k \leq n$.

Definition 1: $Y(t)(t = ..., 0, 1, 2, ...)$, is a subset of $R$. Let $Y(t)$ be the universe of discourse defined by fuzzy set $f_{Y(t)}(t = ..., 0, 1, 2, ...)$. $F(t)$ is defined as a fuzzy time series on $Y(t)(t = ..., 0, 1, 2, ...)$.

According to Definition 1, $F(t)$ can be understood as a linguistics variable and $f_{Y(t)}(t = 1, 2, ...)$ as the possible linguistics values of $F(t)$.

Definition 2: If there exist a fuzzy relationship $R(t-1, t)$, such that $F(t) = F(t-1) \times R(t-1, t)$, where $\times$ is an operator, then $F(t)$ is said to be caused by $F(t-1)$. The relationship between $F(t)$ and $F(t-1)$ can be denoted by

$$F(t-1) \rightarrow F(t) (2)$$

Definition 3: Suppose $F(t-1) = A_i$ and $F(t) = A_j$, a fuzzy logical relationship (FLR) is defined as $A_i \rightarrow A_j$, where $A_i$ and $A_j$ is on the left- and right-hand side respectively, while the repeated fuzzy logical relationship is removed.

Definition 4: Fuzzy logical relationship can be group together in fuzzy logical relationship group (FLRG) according to the same left-hand sides FLR.

According to Chen’s model [8], the repeated fuzzy sets will be removed in the FLRG. For example,

$$A_{i_1} \rightarrow A_{j_1}, A_{i_2} \rightarrow A_{j_2}, ..., A_{i_n} \rightarrow A_{j_n} \rightarrow A_{j_1}, A_{j_2}, ..., A_{j_n} (3)$$

Heuristic knowledge is introduced by [26] and this is used to guide the selection of proper fuzzy set for fuzzy logical relationship group (FLRG) before forecasting step. The definition is as shown below.

Definition 5: In the heuristic models, heuristic functions take fuzzy logical relationship groups and relevant variables as parameters. From these fuzzy logical relationship groups, the heuristic functions use the variables to select proper fuzzy set to establish heuristic fuzzy logical relationship groups.

All fuzzy sets $A_1, A_2, ..., A_n$ are well ordered. In other word, $A_k \geq A_i$ if $k > i$. Suppose $F(t - 1) = A_j$ and the FLRG for $A_j$ is given by $A_j \rightarrow A_{q_1}, A_{r_1}, ...$. Proper fuzzy sets, $A_{p_1}, A_{p_2}, ..., A_{p_k}$ can be selected by heuristic function $h()$,

$$h(x_1, x_2, ..., A_{q_1}, A_{r_1}, ...) = A_{p_1}, A_{p_2}, ..., A_{p_k} (4)$$

where $x_1, x_2, ...$ are heuristic variables; $A_{p_1}, A_{p_2}, ..., A_{p_k}$ are selected from $A_{q_1}, A_{r_1}, ...$ by the heuristics function. A heuristic fuzzy logical relationship group is obtained below:

$$A_j \rightarrow A_{p_1}, A_{p_2}, ..., A_{p_k} (5)$$

The heuristic fuzzy logical relationship group is then used to forecast $F(t)$.

III. THE PROPOSED METHOD

Average-based intervals and Average-based intervals with implementation of heuristic knowledge were introduced a few years ago in forecasting the number of students’ enrollments in University of Alabama. These models were hugely applied in forecasting in other real-life problems. In this paper, we try to show the Average-based intervals with implementation of revised heuristic knowledge forecast better than the other two models. There are five steps involved in this proposed method. The data as shown in Table I, was obtained from the Hospital of Queen Elizabeth II.
Step 1: Define the universe of discourse, \( U = [U_{\text{min}} - D_1, U_{\text{max}} + D_2] \), where \( U_{\text{min}} \) and \( U_{\text{max}} \) are minimum and maximum values in raw data, \( D_1, D_2 \) are two real numbers. \( U \) will be divided into \( n \) equal length of intervals \( u_1, u_2, ..., u_n \). We applied Chen’s method \([8]\) with some modifications in finding the length of intervals. The algorithm in average based interval by \([10]\) to set the suitable length of interval is according to Table II.

Step 2: Fuzzy sets \( A_i \). The linguistics variable is the raw data, \( A_i \) as possible linguistics values of the raw data. Each is defined by the intervals \( u_1, u_2, ..., u_n \).

\[
A_i = \frac{1}{u_1} + \frac{1}{u_2} + \frac{1}{u_3} + \frac{1}{u_4} + \cdots + \frac{1}{u_{n-1}} + \frac{1}{u_n}
\]

Step 3: Defined fuzzy logical relationship FLR. Repeated FLR will be removed.

Step 4: Group FLR in fuzzy logical relationship group FLRG.

Step 5: Calculated the forecasted output. There are 3 rules involved in Chen’s Rule.

- Rule 1: If the fuzzified number of cases of day \( i \) is \( A_i \), and there is only one FLR in the FLRG obtained in STEP 4, that is \( A_i \rightarrow A_k \), whereby \( A_k \) occurs in the interval \( u_k \) and the midpoint of \( u_k \) is \( m_k \), then the forecast number of cases of day \( i + 1 \) is \( m_k \).

- Rule 2: If the fuzzified number of cases of day \( A_i \) and there is more than one FLR in the FLRG obtained in STEP 4, that is, \( A_i \rightarrow A_{k1}, A_i \rightarrow A_{k2}, A_i \rightarrow A_{k3}, ..., A_i \rightarrow A_{kq} \) whereby \( A_{k1}, A_{k2}, A_{k3}, ..., A_{kq} \) occurs in the intervals \( u_{k1}, u_{k2}, u_{k3}, ..., u_{kq} \) and their midpoint are \( m_{k1}, m_{k2}, m_{k3}, ..., m_{kq} \), the forecast number of cases of the day \( i + 1 \) is

\[
\frac{m_{k1} + m_{k2} + m_{k3} + \cdots + m_{kq}}{q}
\]

- If the fuzzified number of cases of day \( i \) is \( A_i \), and empty FLR in the FLRG obtained in STEP 4, that is \( A_i \rightarrow \emptyset \), whereby \( A_i \) occurs in the intervals \( u_i \) and the midpoint of \( u_i \) is \( m_i \), then the forecast number of cases of the day \( i + 1 \) is \( m_i \).
IV. RESULTS AND DISCUSSION

According to Section III, there are 5 steps counted. The data used in this study is the number of tuberculosis cases reported monthly in Sabah, period of January 2012 to December 2019 as you can see from Fig. 1. The time series graphic as in Fig. 1 demonstrates that there is no discernible pattern in the number of tuberculosis cases reported. This can be caused by any unexpected factors, such as gender, family background, knowledge, etc., and did not become observation in this study.

Fig. 1. Plot actual number of tuberculosis cases reported in sabah period January 2012 – December 2019.

A. Length of Interval

The first modification was done in determining the length of intervals. According to the data in Table I, $U_{\min} = 199$ and $U_{\max} = 601$ respectively. We choose $D_1 = 9$ and $D_2 = 9$ respectively. Thus, $U = [199 - 9, 601 + 9] = [190, 610]$. The modifications (b) and (c) are as follows:

- Based on Table II, the length of interval obtained is 10.
- The length of interval is then divided by 2, giving the length of interval is 5.
- The length of interval is then multiplied by 2, giving the length of interval is 20.

Then, the lengths of intervals considered are size 5, size 10 and size 20, and hence list of intervals is shown below. All three sizes applied to the average-based interval model, average-based interval model with implementation of heuristic knowledge and proposed average-based interval model with implementation of revised heuristic knowledge.

<table>
<thead>
<tr>
<th>Length of intervals</th>
<th>Number of intervals</th>
<th>List of intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size 5</td>
<td>84</td>
<td>$u_4 = [190 - 195], u_2 = [195 - 200], ..., u_{64} = [605 - 610]$</td>
</tr>
<tr>
<td>Size 10</td>
<td>42</td>
<td>$u_4 = [190 - 200], u_2 = [200 - 210], ..., u_{42} = [600 - 610]$</td>
</tr>
<tr>
<td>Size 20</td>
<td>21</td>
<td>$u_4 = [190 - 210], u_2 = [210 - 230], ..., u_{21} = [590 - 610]$</td>
</tr>
</tbody>
</table>

B. Fuzzy Logical Relationship (FLR)

Next, is to find FLR. The number of tuberculosis cases reported monthly is to be defined according to intervals as shown in Table III and IV. The same step is repeated for all three models: average-based interval, average-based interval with implementation of heuristic knowledge and proposed average-based interval with implementation of revised heuristic knowledge.

C. Fuzzy Logical Relationship Group (FLRG)

Group of FLR is defined according to same fuzzy set in left hand side of the FLR. As for example, numerical FLR in the model for size 5 is shown. FLRG for fuzzy set $A_{ij}$ is chosen.

<table>
<thead>
<tr>
<th>Length of intervals</th>
<th>Number of intervals</th>
<th>List of intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size 5</td>
<td>84</td>
<td>$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + ... + 0/u_{n-1} + 0/u_{n4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$A_2 = 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + ... + 0/u_{n-1} + 0/u_{n4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$A_3 = 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + ... + 0/u_{n-1} + 0/u_{n4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$...$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$A_{64} = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + ... + 0.5/u_{n-1} + 1/u_{n4}$</td>
</tr>
<tr>
<td>Size 10</td>
<td>42</td>
<td>$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + ... + 0/u_{n-1} + 0/u_{n4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$A_2 = 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + ... + 0/u_{n-1} + 0/u_{n4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$A_3 = 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + ... + 0/u_{n-1} + 0/u_{n4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$...$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$A_{42} = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + ... + 0.5/u_{n-1} + 1/u_{n4}$</td>
</tr>
<tr>
<td>Size 20</td>
<td>21</td>
<td>$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + ... + 0/u_{n-1} + 0/u_{n4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$A_2 = 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + ... + 0/u_{n-1} + 0/u_{n4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$A_3 = 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + ... + 0/u_{n-1} + 0/u_{n4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$...$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$A_{21} = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + ... + 0.5/u_{n-1} + 1/u_{n4}$</td>
</tr>
</tbody>
</table>

\*Jupyter Notebook tools under R kernel
1) Average-based interval: FLRG will be grouped according to Definition 3. Refer Fig. 2.
FLRG for item (12) is $A_{33} \rightarrow A_{31}, A_{33}, A_{38}, A_{77}$
FLRG for item (13) is $A_{33} \rightarrow A_{31}, A_{33}, A_{38}, A_{77}$
FLRG for item (27) is $A_{33} \rightarrow A_{31}, A_{33}, A_{38}, A_{77}$
FLRG for item (84) is $A_{33} \rightarrow A_{31}, A_{33}, A_{38}, A_{77}$

2) Average-based intervals with heuristic knowledge according to Definition 5. Let the month $t - 1$ be $F(t - 1)$ and the FLRG for $F(t - 1)$ be $A_l \rightarrow A_r, \ldots$ This study assumes that there is heuristic knowledge showing the increment or decrement in the number of cases reported. The heuristic function is set as $h(x; A_l, A_r, \ldots)$ where $x$ is an indicator for number of cases forecasting. Fig. 3 shows the example on fuzzy sets of $A_{33}$. The heuristic FLRG is shown below.
FLRG for item (12) is $A_{33} \rightarrow A_{31}$
FLRG for item (13) is $A_{33} \rightarrow A_{31}, A_{38}, A_{77}$
FLRG for item (27) is $A_{33} \rightarrow A_{31}$
FLRG for item (84) is $A_{33} \rightarrow A_{31}, A_{38}, A_{77}$

3) Average-based intervals with revised heuristic knowledge. The modification is that, for the FLR who have the same value on left and right side, the whole FLRG will be chosen as the heuristic FLRG instead of looking back at the raw data to look whether the value increase or decrease. Fig. 4 shows the example on fuzzy sets of $A_{33}$. The heuristic FLRG is shown below.
FLRG for item (12) is $A_{33} \rightarrow A_{31}, A_{33}, A_{38}, A_{77}$
FLRG for item (13) is $A_{33} \rightarrow A_{31}, A_{33}, A_{38}, A_{77}$
FLRG for item (27) is $A_{33} \rightarrow A_{31}$
FLRG for item (84) is $A_{33} \rightarrow A_{31}, A_{33}, A_{38}, A_{77}$

D. Forecasting

The output will be defined according to Step 5. The forecast number of tuberculosis cases with respect to length of intervals size 5, size 10 and size 20 for the three models; average based intervals, average-based intervals with implementation of heuristic knowledge and average-based intervals with implementation of revised heuristic knowledge is shown in Table V, Table VI, and Table VII.

<table>
<thead>
<tr>
<th>Month</th>
<th>Cases</th>
<th>Forecast cases (according to respected size of intervals)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Size 5</td>
</tr>
<tr>
<td>Feb 2012</td>
<td>487</td>
<td>487.7917</td>
</tr>
<tr>
<td>Mar 2012</td>
<td>356</td>
<td>358.1012</td>
</tr>
<tr>
<td>Apr 2012</td>
<td>364</td>
<td>374.7282</td>
</tr>
<tr>
<td>May 2012</td>
<td>419</td>
<td>398.0060</td>
</tr>
<tr>
<td>Jun 2012</td>
<td>349</td>
<td>335.6548</td>
</tr>
<tr>
<td>July 2012</td>
<td>341</td>
<td>368.0774</td>
</tr>
<tr>
<td>Aug 2012</td>
<td>431</td>
<td>412.9702</td>
</tr>
<tr>
<td>Sep 2012</td>
<td>500</td>
<td>373.0655</td>
</tr>
<tr>
<td>Oct 2012</td>
<td>442</td>
<td>442.8988</td>
</tr>
<tr>
<td>Nov 2012</td>
<td>440</td>
<td>425.4405</td>
</tr>
<tr>
<td>Dec 2012</td>
<td>524</td>
<td>388.6925</td>
</tr>
</tbody>
</table>

Fig. 2. The FLRG of $A_{33}$ in Average-Based Intervals for Size 5 is $A_{33} \rightarrow A_{31}, A_{33}, A_{38}, A_{77}$.

Fig. 3. The FLRG of $A_{33}$ in average-based intervals with heuristic knowledge for size 5 is $A_{33} \rightarrow A_{31}, A_{33}, A_{38}, A_{77}$.

Fig. 4. The FLRG of $A_{33}$ in average-based intervals with revised heuristic knowledge for size 5 is $A_{33} \rightarrow A_{31}, A_{33}, A_{38}, A_{77}$.
The smaller the MSE value, the more accurate the model is in prediction. Lower values of RMSE indicate a better fit. Based on Table VIII, the average-based intervals with implementation of revised heuristic knowledge for size 5 give the smallest MSE = 1309.11 and RMSE = 36.18168 respectively compared to average-based interval and average-based interval with implementation of heuristic knowledge. Based on this MSE and RMSE value, it is concluded that the average-based intervals with implementation of revised heuristic knowledge model increase the accuracy of the forecasting value.

TABLE VIII. THE COMPARISON ON MSE AND RMSE OF EACH SIZE

<table>
<thead>
<tr>
<th>Models</th>
<th>Length of intervals</th>
<th>MSE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average-based interval</td>
<td>Size 5</td>
<td>2618.065</td>
<td>51.15726</td>
</tr>
<tr>
<td></td>
<td>Size 10</td>
<td>3337.299</td>
<td>57.76936</td>
</tr>
<tr>
<td></td>
<td>Size 20</td>
<td>4392.890</td>
<td>66.2788</td>
</tr>
<tr>
<td>Average-based intervals with implementation of heuristic knowledge</td>
<td>Size 5</td>
<td>1436.139</td>
<td>37.89642</td>
</tr>
<tr>
<td></td>
<td>Size 10</td>
<td>1926.555</td>
<td>43.89595</td>
</tr>
<tr>
<td></td>
<td>Size 20</td>
<td>2624.269</td>
<td>51.22762</td>
</tr>
<tr>
<td>Average-based intervals with implementation of revised heuristic knowledge</td>
<td>Size 5</td>
<td>1309.114</td>
<td>36.18168</td>
</tr>
<tr>
<td></td>
<td>Size 10</td>
<td>1887.522</td>
<td>43.44562</td>
</tr>
<tr>
<td></td>
<td>Size 20</td>
<td>2289.208</td>
<td>47.84564</td>
</tr>
</tbody>
</table>

V. FUTURE WORK

Overall, this paper shows only the results on training phase of the comparing model and length of interval chosen size 5, size 10 and size 20. Training phase is done for the period of January 2012 until December 2019. Testing phase on the next year 2020 (January-December) can be done as a future work for length of size 5, 10 and 20, limited to list of FLRG obtained from the training phase. The forecasting results will be compared with the actual number of tuberculosis cases.

ACKNOWLEDGMENT

The author would like to thank the main supervisor, Dr. Suzelawati Zenian and co supervisor for their continuous support in this research. Besides, million thanks to Mdm. Nadirah Sulaiman from Clinical Research Centre (CRC), Hospital Queen Elizabeth (HQE) as well as Dr. Roddy Teo from Communicable Disease Unit, Sabah State Health Department for their help throughout this research.

REFERENCES