Forecasting Tourist Arrivals Based on Fuzzy Approach with Average Length and New Base Mapping

Siti Musleha Ab Mutalib
Faculty of Computer & Mathematical Sciences
Universiti Teknologi MARA
Malaysia
musleha78@gmail.com

Nazirah Ramli, Farah Natasha Zainoordin & Zulkifli Ab Ghani Hilmi
Faculty of Computer & Mathematical Sciences
Universiti Teknologi MARA Cawangan Pahang
Malaysia

***Abstract*—Fuzzy time series model (FTSM) has been developed since years ago to handle data in linguistic values. However, most of the FTSM defined data in terms of discrete fuzzy sets which provide the forecasted output as a single point. Various interval length methods have been proposed to obtain higher forecasting accuracy but some of them cannot deal with data in large value of range. This paper proposes an improved fuzzy forecasting model based on trapezoidal fuzzy numbers with new base mapping. The proposed FTSM involves four phases which are data collection on tourist arrival, testing data for seasonality and trend pattern, model development and verifying the forecasting accuracy. The result shows that the data has seasonal and trend pattern; and 13 linguistic values in terms of trapezoidal fuzzy numbers were obtained. Various forecasted intervals under different degree of confidence can be produced compared to a single point forecasted value offered by the previous FTSM. In addition, a wide range of data can be catered based on the new base mapping table.**

Keywords—Average length, Base mapping, Fuzzy, Time series, Trapezoidal fuzzy

I. INTRODUCTION

Forecasting is a process of predicting the future performance based on the existing historical data. Predicting the future is important in managing and is a key component of planning and decision making. Various traditional forecasting models have been proposed to predict the future performance and solving the problem such as Grey Markov model [1], Box Jenkins [2] and exponential smoothing model [3]. However, the traditional forecasting cannot handle the vagueness and incompleteness in data. Moreover, the traditional model requires more historical data, must obey normal distribution and cannot be used for forecasting problems pertaining linguistic values. Thus, [4] proposed fuzzy time series model (FTSM) to deal with these issues by using the enrollment at University of Alabama as a sample set.

Based on that method, [5] concluded that Song and Chissom's method required large amount of computation to derive fuzzy relations and max – min composition operations. Thus, [5] used simple arithmetic operations for time series forecasting. Thereafter, various related research works have been made that follow their framework such as altering the length interval to improve forecasting accuracy. In [4, 6], Song and Chissom partitioned the universe of discourse into seven length of intervals, However, the reason to use seven intervals has not been discussed. [7] mentioned that the length of interval will affect the forecaster accuracy. Thus, [7] proposed average based and distribution based length method with smaller error compared to the previous one. Since that, many improvements and altering in defining the length of interval have been proposed. In 2004, [8] proposed natural partitioning and [9] proposed frequency density based method. [10] used ratio based length method and [11] proposed mean based discretization method to partition the universe of discourse. However, the aforementioned methods are complicated as it needs an extra task during the calculation process.

The average based length is significantly used in determining the universe of discourse [12, 13, 14, 15]. However, the base mapping table proposed for this method cannot cater data in large value of range. Therefore, this paper proposes forecasting tourist arrivals with FTS approach with an average based length method and construct the new based mapping table whereby it can cater for data in large value of range.

The paper is organized as follows. Section 2 introduces some basic definition of FTS and TrFNs. Section 3 presents the new base mapping table. In Section 4, we proposed FTS forecasting method with average based length and the new base mapping table. Section 5, presents the forecasting of tourist arrivals to illustrate the proposed method. The discussion is presented in Section 6. Lastly, the paper is concluded in Section 7.

II. PRELIMINARIES

In this section, some basic concepts of fuzzy time series and trapezoidal fuzzy numbers (TrFNs) are presented. The concept of FTS was first presented and defined by [4, 6] while TrFNs was first used in [12]. The definitions are given as follows:

Definition 1. Let $Y(t)(t=...,0,1,2,...)$ be a subset of R and $Y(t)$ be the universe of discourse defined by fuzzy set $\mu_i(t)(i=1,2,...)$. If $F(t)$ consists of $\mu_i(t)(i=1,2,...)$, then $F(t)$ is called FTS on $Y(t)(t=...,0,1,2,...)$.

ISBN 978-967-14020-0-9 ©BMCRC2016

Definition 2. Let $F(t)$ is a FTS. $F(t)$ is caused by $F(t-1)$ if there exists a fuzzy relationship $R(t-1,t)$, such that $F(t)=F(t-1)\otimes R(t-1,t)$ where $\otimes$ represent as fuzzy operator. The relationship can be denoted as $F(t-1)\rightarrow F(t)$ or $\tilde{A}_i \rightarrow \tilde{A}_j$ if $F(t-1)=\tilde{A}_i$ and $F(t)=\tilde{A}_j$.

Definition 3. A TrFNs $\tilde{A}$ denoted as $\tilde{A}=(a,b,c,d)$ is defined as

$$\mu_{\tilde{A}}(x)=\begin{cases}0, & x<a\\ \frac{x-a}{b-a}, & a\le x\le b\\ 1, & b\le x\le c\\ \frac{d-x}{d-c}, & c\le x\le d\\ 0, & x>d\end{cases} \quad \ldots\ldots(1)$$

III. THE NEW BASE MAPPING

TABLE I. BASE MAPPING TABLE [7]

Range	Base
0.1 – 1.0	0.1
1.1 – 10.0	1.0
11.0 – 100.0	10.0
101.0 – 1000.0	100.0

Table I shows the base mapping table for average length method proposed by [7]. However, the base mapping table cannot cater for large of data. Therefore, a new base mapping table is construct based on the pattern of the existing base mapping table. From the pattern of the last two rows in Table I, we can construct the new base mapping table as follows:

Assume for *n* = 1, for range 11.0 – 100.0
n = 2, for range 101.0 – 1000.0
when *n* = 1, the range will be $1+10^1$ to 10^2 with base of 10^1,
n = 2, the range will be $1+10^2$ to 10^3 with base of 10^2.
Therefore, based on the pattern, when $n=k$, the range will be $1+10^k$ to 10^{k+1} with base of 10^k. The new base mapping table is simplified in Table II.

TABLE II. NEW BASE MAPPING TABLE

n	Range	Base
	0.1 – 1.0	0.1
	1.1 – 10.0	1.0
1	11.0 – 100.0	10.0
2	101.0 – 1000.0	100.0
3	1001.0 – 10000.0	1000.0
4	10001.0 – 100000.0	10000.0
↓	↓	↓
k	$1+10^k$ - 10^{k+1}	10^k

IV. PROPOSED FORECASTING MODEL

The proposed forecasting model with the new base mapping involved four phases whereby phase 2 is taken from [15]. The procedure of the proposed approach is given as follows.

- Phase 1: Collecting the data.
- Phase *2*: Investigation for seasonality and trend pattern.

Line chart method for determining seasonality and trend pattern was used. For data with seasonal variation, deseasonalize the data using ratio to moving average [15].

- Phase 3: Model Development.
  a. Determine the maximum $D_{\max}$ and the minimum $D_{\min}$ of the deseasonalize data D_t.
  b. Define the universe of discourse $U=[D_{\min}-D_1, D_{\max}+D_2]$ where D_1 and D_2 are proper integer.
  c. Determine the appropriate length of intervals l and partition the universe of discourse by using the average based length methods as follows:
     i. Calculate the absolute difference between the next deseasonalized data (D_{t+1}) and deseasonalized data (D_t) whereby $t=1 \ldots n-1$ and calculate the average of the difference between data.
     ii. Take half of the average (Avg_{new}) as the length.
     iii. Determine the range and base for the length in (ii) and rounded the length by using the new base mapping table in Table II.
     iv. Calculate the number of interval using formula $m=(D_{\max}+D_2-D_{\min}+D_1)/l$ where l is the length after rounded.
     v. Partition the universe of discourse using the length interval obtained in (iv). Assume there are m intervals which are $u_1=[d_1,d_2]$, $u_2=[d_2,d_3]$, …, $u_{m-1}=[d_{m-1},d_m]$ and $u_m=[d_m,d_{m+1}]$. Remove the interval that does not cover the data.
  d. Develop new TrFNs to represent the linguistic terms based on the interval obtained in (v) as follows:
     $\tilde{A}_1=(d_1,d_1,d_2,d_3)$,
     $\tilde{A}_2=(d_1,d_2,d_3,d_4)$,
     $\vdots \quad \vdots$
     $\tilde{A}_{m-1}=(d_{m-2},d_{m-1},d_m,d_{m+1})$,
     $\tilde{A}_m=(d_{m-1},d_m,d_{m+1},d_{m+1})$.
  e. Fuzzify the D_t. If the D_t is located in the range u_m, so $u_m \rightarrow \tilde{A}_m$.
  f. Create fuzzy logical relationship (FLR) based on *Definition 2.* Then construct FLR group.
  g. Determine the forecasted values $\tilde{O}_t$, by using the heuristic rules as follows:
     - Rule 1: if the FLR group of $\tilde{A}_i$ is empty; $\tilde{A}_i \rightarrow \phi$, then $\tilde{O}_t=\tilde{A}_i$.
     - Rule 2: if the FLR group of $\tilde{A}_i$ is one to one; $\tilde{A}_i \rightarrow \tilde{A}_j$, then $\tilde{O}_t=\tilde{A}_j$.
     - Rule 3: if the FLR group of $\tilde{A}_i$ is one to many; $\tilde{A}_i \rightarrow \tilde{A}_{j1}$, $\tilde{A}_i \rightarrow \tilde{A}_{j2}$, …, $\tilde{A}_i \rightarrow \tilde{A}_{jp}$ then

$$\tilde{O}_t=\frac{\tilde{A}_{j1}+\tilde{A}_{j2}+\ldots+\tilde{A}_{jp}}{p} \quad \ldots\ldots(2)$$

  h. Calculate the final forecasted values $\tilde{F}_t$ based on [15] as follows:

i. If the original historical data only have seasonal pattern, then $\tilde{F}_t = Si_t \bullet \tilde{O}_t$.

$$\tilde{F}_t = \left(Si_t a_t, Si_t b_t, Si_t c_t, Si_t d_t\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(3)}$$

Where $\tilde{O}_t = \left(a_t, b_t, c_t, d_t\right)$ and Si_t = seasonal index.

ii. If the original historical data only have a trend pattern, then $\tilde{F}_t = R_{t-1} \oplus \tilde{O}_t$.

$$\tilde{F}_t = \left(R_{t-1}+a_t, R_{t-1}+b_t, R_{t-1}+c_t, R_{t-1}+d_t\right) \ldots\ldots\text{(4)}$$

Where R_t = actual value

iii. If the original historical data have both seasonal and trend patterns, then $\tilde{F}_t = \frac{Si_t}{Si_{t-1}} R_{t-1} \oplus Si_t \bullet \tilde{O}_t$

$$\tilde{F}_t = \left(\frac{Si_t}{Si_{t-1}}R_{t-1}+Si_t a_t, \frac{Si_t}{Si_{t-1}}R_{t-1}+Si_t b_t, \frac{Si_t}{Si_{t-1}}R_{t-1}+Si_t c_t, \frac{Si_t}{Si_{t-1}}R_{t-1}+Si_t d_t\right) \ldots\text{(5)}$$

iv. If the original historical data do not have either seasonal or trend pattern, then $\tilde{F}_t = \tilde{O}_t$

$$\tilde{F}_t = \left(a_t, b_t, c_t, d_t\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(6)}$$

i. Defuzzify the forecasted value centroid method as follows:
This method is applied to translate the TrFNs into crisp value.

$$\tilde{F}_t = \frac{\int x \bullet \mu(x)\,dx}{\int \mu(x)\,dx} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(7)}$$

- Phase 4: Verifying the forecasting accuracy

Calculate the root mean square error (RMSE) to verify the performance of proposed method. The formulation is;

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(\tilde{F}_t - \tilde{R}_t\right)^2}{n}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(8)}$$

Where $\tilde{R}_t$ = actual value, $\tilde{F}_t$ = forecasted value after defuzzify and n = number of data.

V. FORECASTING TOURIST ARRIVALS

In this section, the proposed FTSM with average based length and new base mapping table are used to predict the upcoming tourist to Malaysia.

- Phase 1: Fig 1 show the FTS data for monthly tourist arrival to Malaysia from 2008 until 2010. The data is collected from Johor Tourism Department.

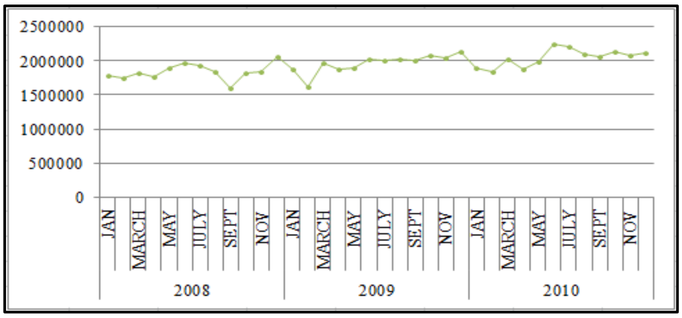

Fig. 1. Time series data of tourist arrivals to Malaysia from January 2008 to December 2010

- Phase 2: The seasonality of the data is determined based on the graph in Fig 1. It shows that there is a decreasing pattern in the upcoming tourist arrival to Malaysia from January to February, and also increasing pattern of tourist arrival from May to June. Thus, the historical data is seasonal since it has similar trend for three consecutive years.

Table III shows the deseasonalized value obtained using the ratio to moving average method and Fig 2 shows the trend pattern by using line chart.

TABLE III. DESEASONALIZED VALUE FOR TOURIST ARRIVAL TO MALAYSIA

Year	Month	Deseasonalized value	Year	Month	Deseasonalized value
2008	Jan	1826269	2009	July	1930060
	Feb	1970871		Aug	1993471
	Mar	1791907		Sept	2121450
	Apr	1849216		Oct	2041321
	May	1944118		Nov	2016863
	June	1836784		Dec	1960060
	July	1857199	2010	Jan	1946080
	Aug	1805839		Feb	2072478
	Sept	1698636		Mar	1991710
	Oct	1785792		Apr	1972763
	Nov	1817056		May	2039452
	Dec	1884638		June	2103429
2009	Jan	1919591		July	2132694
	Feb	1824781		Aug	2061364
	Mar	1945611		Sept	2180787
	Apr	1979002		Oct	2099512
	May	1938908		Nov	2049114
	June	1890138		Dec	1943536

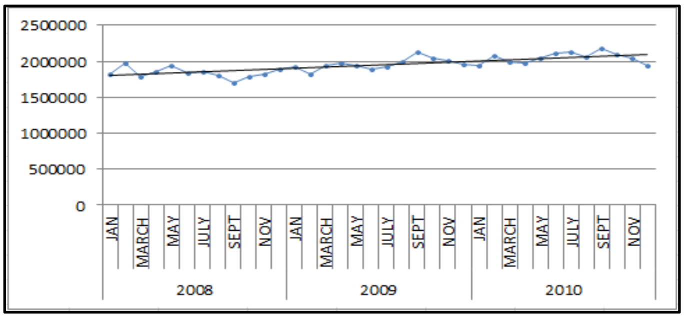

Fig. 2. Deseasonalizex data for trend pattern

- Phase 3: The average based length with new base mapping table is implemented to determine the appropriate length of interval which is 40000. The D_{min} and D_{max} deseasonalized value (from Table III) are 1698636 and 2180787 respectively. By choosing D_1 = 98636 and D_2 = 219213, the universe of discourse U is defined as [1600000, 2400000]. Thus, the number of appropriate interval is 20. However, the intervals which are not covered by the historical data have been removed. So that, the new numbers of intervals are 13 which are $u_1 = [1680000, 1720000]$, $u_2 = [1720000, 1760000]$, …, $u_{12} = [2120000, 2160000]$ and $u_{13} = [2160000, 2200000]$.

The new TrFNs can be defined as follows:

$$\tilde{A}_1 = (1680000, 1680000, 1720000, 1760000),$$
$$\tilde{A}_2 = (1680000, 1720000, 1760000, 1800000),$$
$$\vdots \qquad \vdots$$
$$\tilde{A}_{12} = (2080000, 2120000, 2160000, 2200000),$$
$$\tilde{A}_{13} = (2120000, 2160000, 2200000, 2200000).$$

Fuzzify the D_t. As an example, D_t for January 2008 is 1826269 and located at range $u_4 = [1800000, 1840000]$. Thus, the corresponding fuzzy number is assigned as $\tilde{A}_4$. Table IV shows some of the corresponding fuzzy number for the tourist arrivals in year 2008.

TABLE IV. CORRESPONDING FUZZY NUMBERS OF THE TOURIST ARRIVALS TO MALAYSIA FOR YEAR 2008

Month	**Dt**	**Fuzzy number**	**Month**	**Dt**	**Fuzzy number**
Jan	1826269	$\tilde{A}_4$	July	1857199	$\tilde{A}_5$
Feb	1970871	$\tilde{A}_8$	Aug	1805839	$\tilde{A}_4$
Mar	1791907	$\tilde{A}_3$	Sept	1698636	$\tilde{A}_1$
Apr	1849216	$\tilde{A}_5$	Oct	1785792	$\tilde{A}_3$
May	1944118	$\tilde{A}_7$	Nov	1817056	$\tilde{A}_4$
June	1836784	$\tilde{A}_4$	Dec	1884638	$\tilde{A}_6$

After fuzzifying the D_t, the FLR is created and further, the FLR group is produced as shown in Table V.

TABLE V. FUZZY LOGICAL RELATIONSHIP GROUPS

Group	**Fuzzy Logical Relationship**
1	$\tilde{A}_1 \rightarrow \tilde{A}_3$
2	$\tilde{A}_3 \rightarrow \tilde{A}_4$, $\tilde{A}_3 \rightarrow \tilde{A}_5$
3	$\tilde{A}_4 \rightarrow \tilde{A}_1$, $\tilde{A}_4 \rightarrow \tilde{A}_5$, $\tilde{A}_4 \rightarrow \tilde{A}_6$, $\tilde{A}_4 \rightarrow \tilde{A}_7$, $\tilde{A}_4 \rightarrow \tilde{A}_8$
4	$\tilde{A}_5 \rightarrow \tilde{A}_4$, $\tilde{A}_5 \rightarrow \tilde{A}_7$
5	$\tilde{A}_6 \rightarrow \tilde{A}_4$, $\tilde{A}_6 \rightarrow \tilde{A}_6$, $\tilde{A}_6 \rightarrow \tilde{A}_7$
6	$\tilde{A}_7 \rightarrow \tilde{A}_4$, $\tilde{A}_7 \rightarrow \tilde{A}_6$, $\tilde{A}_7 \rightarrow \tilde{A}_8$, $\tilde{A}_7 \rightarrow \tilde{A}_{10}$
7	$\tilde{A}_8 \rightarrow \tilde{A}_3$, $\tilde{A}_8 \rightarrow \tilde{A}_7$, $\tilde{A}_8 \rightarrow \tilde{A}_8$, $\tilde{A}_8 \rightarrow \tilde{A}_9$, $\tilde{A}_8 \rightarrow \tilde{A}_{12}$
8	$\tilde{A}_9 \rightarrow \tilde{A}_8$, $\tilde{A}_9 \rightarrow \tilde{A}_{11}$
9	$\tilde{A}_{10} \rightarrow \tilde{A}_7$, $\tilde{A}_{10} \rightarrow \tilde{A}_8$, $\tilde{A}_{10} \rightarrow \tilde{A}_9$, $\tilde{A}_{10} \rightarrow \tilde{A}_{13}$
10	$\tilde{A}_{11} \rightarrow \tilde{A}_{10}$, $\tilde{A}_{11} \rightarrow \tilde{A}_{12}$
11	$\tilde{A}_{12} \rightarrow \tilde{A}_{10}$
12	$\tilde{A}_{13} \rightarrow \tilde{A}_{11}$

The forecasted output $\tilde{O}_t$ is calculated using heuristic rules. Then, the final forecasted value $\tilde{F}_t$ is generated by using equation (5) since the D_t have both seasonal and trend pattern. Table VI shows the final forecasted value of tourist arrivals for year 2008 to 2010 with their defuzzified values.

TABLE VI. FORECASTED OUTPUT OF THE TOURIST ARRIVAL TO MALAYSIA IN YEAR 2008 TO 2010

Month	**Forecasted output**	**Defuzzified value**
Jan 2008		
Feb 2008	(1820000, 1860000, 1900000, 1940000)	1880000
Mar 2008	(3227242, 3255534, 3290898, 3326263)	3275263
Apr 2008	(3943072, 3983692, 4024312, 4064932)	4004002
May 2008	(3400208, 3438285, 3476362, 3514440)	3457324
June 2008	(3584343, 3623418, 3662493, 3701567)	3642955
July 2008	(4083471, 4126184, 4168897, 4211609)	4147540
Aug 2008	(3800504, 3833726, 3875252, 3916779)	3856893
Sept 2008	(3745201, 3785941, 3826681, 3867421)	3806311
Oct 2008	(3417819, 3447950, 3485613, 3523277)	3468962
Nov 2008	(3480875, 3521603, 3562331, 3603059)	3541967
Dec 2008	(3621895, 3662524, 3703153, 3743783)	3682838
Jan 2009	(3977306, 4012262, 4055956, 4099650)	4036638
Feb 2009	(3617551, 3656541, 3695530, 3734520)	3676036
Mar 2009	(3312108, 3347472, 3382836, 3418201)	3365154
Apr 2009	(3705353, 3737850, 3778470, 3819090)	3760511
May 2009	(3641717, 3679794, 3717871, 3755948)	3698833
June 2009	(3800999, 3840073, 3879148, 3918223)	3859611
July 2009	(4077908, 4120621, 4163334, 4206047)	4141978
Aug 2009	(3858664, 3900191, 3941717, 3983244)	3920954
Sept 2009	(3880520, 3921259, 3961999, 4002739)	3941629
Oct 2009	(3677351, 3715015, 3752678, 3790342)	3733847
Nov 2009	(4196484, 4237212, 4277940, 4318668)	4257576
Dec 2009	(4074433, 4115063, 4155692, 4186164)	4132443
Jan 2010	(4365971, 4409665, 4453359, 4497053)	4431512
Feb 2010	(3774245, 3813234, 3852224, 3891213)	3832729
Mar 2010	(3382678, 3418043, 3453407, 3488772)	3435725
Apr 2010	(4105154, 4145774, 4186394, 4216859)	4163150
May 2010	(3716062, 3754139, 3792216, 3830293)	3773176
June 2010	(3794904, 3833979, 3873054, 3912128)	3853516
July 2010	(4292053, 4334766, 4377478, 4420191)	4356122
Aug 2010	(4301570, 4343097, 4384624, 4426150)	4363860
Sept 2010	(4209120, 4249860, 4290600, 4331339)	4270230
Oct 2010	(3795890, 3833554, 3871217, 3899465)	3849665
Nov 2010	(4297629,4338357,4379086,4419814)	4358722
Dec 2010	(4204641, 4245270, 4285899, 4326529)	4265585
Jan 2011	(4390278, 4433972, 4477666, 4510436)	4452663

VI. DISCUSSION

In Table VI, the forecasted tourist arrival for January 2011 is (4390278, 4433972, 4477666, 4510436) and the possible forecasted intervals under different degree of confidence (DDoC) can be obtained. The decision analyst used TrFNs to find the possible forecasted interval by using the α - cut. It can be obtained under a degree of confidence (α) from interval $(0 \le \alpha \le 1)$. Here, we took June and calculate the degree of confidence because of the existence of a seasonal factor in these three consecutive years. The forecasted intervals are presented as in Table VII, VIII and IX below.

TABLE VII. FORECASTED INTERVALS FOR JUNE 2008 (FROM $\alpha = 0$ TO 1)

Degree of confidence, α	**Interval estimate**
$\alpha = 0$	[4083471, 4211609]
$\alpha = 0.1$	[4087742, 4207338]
$\alpha = 0.2$	[4092014, 4203067]
$\alpha = 0.3$	[4096285, 4198795]
$\alpha = 0.4$	[4100556, 4194524]
$\alpha = 0.5$	[4104828, 4190253]
$\alpha = 0.6$	[4109099, 4185982]
$\alpha = 0.7$	[4113370, 4181711]
$\alpha = 0.8$	[4117641, 4177439]
$\alpha = 0.9$	[4121913, 4173168]
$\alpha = 1.0$	[4126184, 4168897]

TABLE VIII. FORECASTED INTERVALS FOR JUNE 2009 (FROM α = 0 TO 1)

Degree of confidence, α	Interval estimate
α = 0	[4077908, 4206047]
α = 0.1	[4082179, 4201776]
α = 0.2	[4086451, 4197504]
α = 0.3	[4090722, 4193233]
α = 0.4	[4094993, 4188962]
α = 0.5	[4099265, 4184691]
α = 0.6	[4103536, 4180419]
α = 0.7	[4107807, 4176148]
α = 0.8	[4112078, 4171877]
α = 0.9	[4116350, 4167605]
α = 1.0	[4120621, 4163334]

TABLE IX. FORECASTED INTERVALS FOR JUNE 2010 (FROM α = 0 TO 1)

Degree of confidence, α	Interval estimate
α = 0	[4292045, 4420191]
α = 0.1	[4296324, 4415920]
α = 0.2	[4300596, 4411684]
α = 0.3	[4304867, 4407377]
α = 0.4	[4309138, 4403106]
α = 0.5	[4313410, 4398835]
α = 0.6	[4317681, 4394563]
α = 0.7	[4321952, 4390292]
α = 0.8	[4326223, 4386201]
α = 0.9	[4330495, 4381749]
α = 1.0	[4334766, 4377478]

The interval with higher value of α will give smaller interval of estimation. For example, for June 2010, when α = 0 the interval is [4292045, 4420119] while α = 1.0 the interval is [4334766, 4377478]. The differences between intervals are 128146 and 42712 respectively. This indicates that the range for higher α is more precise in making prediction. In addition, this estimation can be evaluated by the decision analyst. Furthermore, the previous FTSM produces the forecasted values in terms of single point value whereby some information will be dissipated.

To evaluate the performance of the forecasting model, the forecasting accuracy (RMSE) is calculated. The forecasted values need to be defuzzified using centroid method and the RMSE obtained is 322909.

VII. CONCLUSION

In this study, the fuzzy forecasting approach using an average based length with new base mapping is proposed. The new base mapping table can cater a large value of range. The propose forecasting model can cater for data up to 10^k where $k > 2$. Thus, this particular model allows unlimited range of data compared to the previous base mapping table that only caters up to 10^3 of data. These enhance the forecasting model to be used in more wide application of fuzzy forecasting problem. Other than that, the proposed method also produces the forecasted output in term of TrFNs, which allows forecasters to analyze various forecasted intervals under different degree of confidence instead of single point value offered by the existing FTSM. Thus, it can provide more information on the forecasted values.

ACKNOWLEDGMENT

This research is supported by the Ministry of Education Malaysia (MOE) and Universiti Teknologi MARA Malaysia (UiTM) under the Research Grant No. 600-RMI/RAGS 5/3 (149/2014).

REFERENCES

[1] X. Li and W. Chen, “A Grey-Markov predication for unemployment rate of graduates in China,” in *Grey Systems and Intelligent Services*, 2009, pp. 619–624.

[2] K. Mahipan, N. Chutiman, and B. Kumphon, “A forecasting model for Thailand ’ s unemployment rate,” *Mod. Appl. Sci.*, vol. 7, no. 7, pp. 10–16, 2013.

[3] M. Bratu, “Econometric models or smoothing exponential techniques to predict macroeconomic indicators in Romania,” *Zagreb Int. Rev. Econ. Bus.*, vol. 15, no. 2, pp. 87–100, 2012.

[4] Q. Song and B. S. Chissom, “Forecasting enrollments with fuzzy time series — Part I,” *Fuzzy Sets Syst.*, vol. 54, pp. 1–9, 1993.

[5] S.-M. Chen, “Forecasting enrollments based on fuzzy time series,” *Fuzzy sets Syst.*, vol. 81, pp. 311–319, 1996.

[6] Q. Song and B. S. Chissom, “Forecasting enrollments with fuzzy time series — part II,” *Fuzzy Sets Syst.*, vol. 62, pp. 1–8, 1994.

[7] K. Huarng, “Effective lengths of intervals to improve forecasting in fuzzy time series,” *Fuzzy sets Syst.*, vol. 123, pp. 387–394, 2001.

[8] S. Li and Y. Chen, “Natural partitioning-based forecasting model for fuzzy time series,” in *IEEE International Conference on Fuzzy System*, 2004, pp. 1355–1359.

[9] S. Chen and C. Hsu, “A new method to forecast enrollments using fuzzy time series,” *Int. J. Appl. Sci. Eng.*, vol. 2, pp. 234–244, 2004.

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

[11] P. Singh and B. Borah, “An efficient time series forecasting model based on fuzzy time series,” *Eng. Appl. Artif. Intell.*, vol. 26, no. 10, pp. 2443–2457, Nov. 2013.

[12] H. Liu, “An improved fuzzy time series forecasting method using trapezoidal fuzzy numbers,” *Fuzzy Optim Decis Mak.*, vol. 6, no. 1, pp. 63–80, 2007.

[13] H. T. Jasim, A. G. J. Salim, and K. I. Ibraheem, “A Novel Algorithm to Forecast Enrollment Based on Fuzzy Time Series,” *Appl. Appl. Math.*, vol. 7, no. 1, pp. 385–397, 2012.

[14] H.-T. Liu and M.-L. Wei, “An improved fuzzy forecasting method for seasonal time series,” *Expert Syst. Appl.*, vol. 37, no. 9, pp. 6310–6318, Sep. 2010.

[15] H.-T. Liu, “An integrated fuzzy time series forecasting system,” *Expert Syst. Appl.*, vol. 36, no. 6, pp. 10045–10053, Aug. 2009.