

# Forecasting the Number of Muslim Pilgrims Using Fuzzy Time Series

Mahmoud Elgamal

Email: mahmoud\_40@yahoo.com

**Abstract** – Pilgrimage of Muslims is considered the largest human gathering all over the world in which more than three millions move together through a very limited space in a short time period. The yearly number of pilgrims coming from outside Saudi Arabia, denoted by NPO for short, is more than two thirds of the total number of Pilgrims. Therefore forecasting the NPO is considered by Saudi Arabia as the most important indicator in determining the planning mechanism for future secure and comfortable pilgrimage seasons. The main objective of this article is to employ the fuzzy time series approach [1] to forecast the yearly series of NPO and to show that it gives better forecasts than Box-Jenkins Procedures. In order to achieve our objective, the fuzzy time series method is used to forecast the future five observations and the results are compared with the results given by Shaarawy et. al [2].

**Keywords** – Fuzzy Time Series; Fuzzy Sets; Linguistic Variable; Time Invariant Model.

## I. INTRODUCTION

Forecasting is important for any society to effectively control a plan which is often engaged by collecting a large number of time series data in order to be able to determine the trend of the development.

Pilgrimage (or hajj) of Muslims is one of the most important events all over the world. It is considered the largest human gathering in which pilgrims move together through a very limited space in a short period of time. This important event is repeated annually, and the number of pilgrims is increasing year after year. In addition, hajj is one of the main sources of gross national product (GNP) in Saudi Arabia. Therefore, it has received a great attention by the government of Saudi Arabia. Every year, the kingdom of Saudi Arabia spends a great deal of efforts and money to improve the hajj system, which includes security, economy, management of water and electrical resources, services and goods required by the vast number of pilgrims. However, without knowing the number of pilgrims in advance may make the process of improvement very difficult. Therefore, it is very important to have a mechanism for predicting and forecasting the number of pilgrims in order to determine the size and quality of expansions and maintenance needed in the two holy mosques in Makkah and Medina and to avoid any mistakes or disasters that may occur. One of the main components of the total number of pilgrims is the Number of Pilgrims coming from Outside the kingdom (NPO).

The first objective of this paper is to use the fuzzy time series approach to forecast the number of NPO data.

Fuzzy time series have been applied to a diversity of data such as stock exchange, temperature, enrollment, ...etc. In forecasting, statistical methods such as time series

models, Box-Jenkins, ...etc. are commonly used techniques. However, if the given data are in linguistic form or very little, the statistical methods will fail [3, 4, 5]. In order to handle such a problem, fuzzy time series models [3, 4, 5, 1, 6, 7] have been developed an applied in practice. The first time series model was introduced by Song-Chisson 1993. And since then, fuzzy time series have drawn attention of many researchers because of its good performance for some kind of forecasting problems.

In [1] they proposed a method for handling forecasting of university enrollments using fuzzy time series.

In this paper, we exploit the fore-mentioned method to forecast NPO.

The remainder of this paper is organized as follows: section (2) presents some basic characteristics of fuzzy time series. A complete description of the algorithm used to predict NPO in section (3). Section (4) is dedicated to evaluate the forecast performance of fuzzy time series procedures and compare their numerical results. Finally, the paper is concluded in section (5).

## II. FUZZY TIME SERIES

In this section we introduce a brief review of fuzzy time series. Let  $U$  be the universe of discourse,  $U = \{u_1, u_2, \dots, u_n\}$ , a fuzzy set [8]  $A$  of  $U$  is defined as

$$A = \mu_A(u_1)/u_1 + \mu_A(u_2)/u_2 + \dots + \mu_A(u_n)/u_n$$

where  $\mu_A$  is the membership function of  $A$ ,  $\mu_A : U \rightarrow [0, 1]$ ,  $\mu_A(u_i)$  denotes the grade of membership of  $u_i$  in  $A$ ,  $\mu_A(u_i) \in [0, 1]$ . Let  $Y(t) (t = \dots, 0, 1, 2, \dots)$ , be a subset of  $R^1$ , be the universe of discourse on which fuzzy sets  $\mu_i(t) (i = 1, 2, \dots)$  are defined and let  $F(t)$  be a collection  $\mu_i(t) (i = 1, 2, \dots)$ . Then,  $F(t)$  is called a fuzzy time series on  $Y(t) (t = \dots, 0, 1, 2, \dots)$ . It is obvious that  $F(t)$  can be regarded as a linguistic variable, and  $\mu_i(t)$  can be viewed as possible linguistic values of  $F(t)$ , where  $\mu_i(t) (i = 1, 2, \dots)$  are represented by fuzzy sets. Furthermore, we can also see that  $F(t)$  is a function of time  $t$ , i.e., the values of  $F(t)$  can be different at different times. If  $F(t)$  is caused by  $F(t-1)$  only, then this relationship is represented by

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

Let  $F(t)$  be a fuzzy time series. If for any time  $t$ ,  $F(t) = F(t-1)$  and  $F(t)$  only has finite elements then  $F(t)$  is called a time-invariant fuzzy time series, otherwise, called time-variant fuzzy time series [1].

Assume that the value of the time series at time  $t$  is  $T_t$  and assume that the value of the time series at time  $(t-1)$  is  $T_{t-1}$ , then the variation of the time series between time  $t$  and time  $(t-1)$  is equal to  $T_t - T_{t-1}$ . Here, we will introduce some heuristic rules that will be used in the algorithm [1].

**Rule 1:** The variation of the time series value between times  $t$  and  $(t - 1)$  is related to the variations of the time series values between time  $t$  and the past values, and the relationship of the time series between the current value and last value is closer than the one between this value and the other past values.

**Rule 2:** If the trend of the time series past values is increasing, then the current value of the time series is increasing. If the trend of the time series past values is decreasing, then the current value of the time series is decreasing.

**Rule 3:** Let the variation of last value be a criterion. Compute the fuzzy relationships between last value and the other past values based on data variations. From the derived fuzzy relationships, we can know the degrees of relationships between the variation of last value and the variations of other past values. The variation of current value can be obtained from the derived fuzzy relationships.

Based on these heuristic rules, firstly we can fuzzify the historical time series data. We use the fuzzified variation of the historical time series data and the linguistic values (big decrease), (decrease), (no change), (increase), (big increase), (too big increase) to forecast time series values.

The fuzzified variation of the historical time series between  $T_t$  and  $T_{t-1}$  can be described as follows:

$$F(t) = u_1/(\text{big decrease}) + u_2/(\text{decrease}) + \dots + u_i/L + \dots + u_m/(\text{too big increase}) \quad (1)$$

where  $F(t)$  denotes the fuzzified variation of the time series between  $t$  and  $(t-1)$ ,  $u_i$  is the grade of membership to the linguistic value  $L$ ,  $m$  is the number of the elements in the universe of discourse, and  $1 \leq i \leq m$ .

To forecast the time series value at  $t$ , we must decide how many values of the time series will be used, where the number of time series values to be used is called the *window basis*. Suppose we set a window basis  $w$  values, then the variation of last value is used to be a criterion and the other variations of  $w$  past values are used to form a matrix which is called the *operation matrix*. The *criterion matrix*  $C(t)$  and the operation matrix  $O^w(t)$  at  $t$  are expressed as follows [1]:

$$C(t) = F(t-1) = \begin{bmatrix} \text{(big decrease)} & \text{(decrease)} & \dots & \text{(too big increase)} \\ C_1 & C_2 & \dots & C_m \end{bmatrix} \quad (2)$$

$$O^w(t) = \begin{bmatrix} F(t-2) \\ F(t-3) \\ \vdots \\ F(t-w-1) \end{bmatrix} = \begin{bmatrix} \text{(big decrease)} & \text{(decrease)} & \dots & \text{(too big increase)} \\ O_{11} & O_{12} & \dots & O_{1m} \\ O_{21} & O_{22} & \dots & O_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ O_{w1} & O_{w2} & \dots & O_{wm} \end{bmatrix} \quad (3)$$

Table 1: Number of Pilgrims coming from outside Kingdom

| Year | NPO    | Year | NPO     | Year | NPO     | Year | NPO     |
|------|--------|------|---------|------|---------|------|---------|
| 1390 | 431270 | 1401 | 879368  | 1412 | 1012917 | 1423 | 1431012 |
| 1391 | 479339 | 1402 | 853555  | 1413 | 992813  | 1424 | 1419706 |
| 1392 | 645182 | 1403 | 1003911 | 1414 | 995611  | 1425 | 1534759 |
| 1393 | 607755 | 1404 | 919671  | 1415 | 1043274 | 1426 | 1557447 |
| 1394 | 918777 | 1405 | 851761  | 1416 | 1080465 | 1427 | 1654407 |
| 1395 | 894573 | 1406 | 856718  | 1417 | 1168591 | 1428 | 1707814 |
| 1396 | 719040 | 1407 | 960386  | 1418 | 1132344 | 1429 | 1729841 |
| 1397 | 739319 | 1408 | 762755  | 1419 | 1056730 | 1430 | 1613965 |
| 1398 | 830236 | 1409 | 774560  | 1420 | 1267555 | 1431 | 1799601 |
| 1399 | 862520 | 1410 | 828993  | 1421 | 1363992 | 1432 | 1828195 |
| 1400 | 812892 | 1411 | 720102  | 1422 | 1354184 | 1433 | 1752932 |

Table 2: Actual values and variation of NPO

|           |         |         |         |         |         |         |         |         |         |         |         |
|-----------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| Year      | 1390    | 1391    | 1392    | 1393    | 1394    | 1395    | 1396    | 1397    | 1398    | 1399    | 1400    |
| NPO       | 431270  | 479339  | 645182  | 607755  | 918777  | 894573  | 719040  | 739319  | 830236  | 862520  | 812892  |
| Variation | -       | 48069   | 165843  | -37427  | 311022  | -24204  | -175533 | 20279   | 90917   | 32284   | -49628  |
| Year      | 1401    | 1402    | 1403    | 1404    | 1405    | 1406    | 1407    | 1408    | 1409    | 1410    | 1411    |
| NPO       | 879368  | 853555  | 1003911 | 919671  | 851761  | 856718  | 960386  | 762755  | 774560  | 828993  | 720102  |
| Variation | 66476   | -25813  | 150356  | -84240  | -67910  | 4957    | 103668  | -197631 | 11805   | 54433   | -108891 |
| Year      | 1412    | 1413    | 1414    | 1415    | 1416    | 1417    | 1418    | 1419    | 1420    | 1421    | 1422    |
| NPO       | 1012917 | 992813  | 995611  | 1043274 | 1080465 | 1168591 | 1132344 | 1056730 | 1267555 | 1363992 | 1354184 |
| Variation | 292815  | -20104  | 2798    | 47663   | 37191   | 88126   | -36247  | -75614  | 210825  | 96437   | -9808   |
| Year      | 1423    | 1424    | 1425    | 1426    | 1427    | 1428    | 1429    | 1430    | 1431    | 1432    | 1433    |
| NPO       | 1431012 | 1419706 | 1534759 | 1557447 | 1654407 | 1707814 | 1729841 | 1613965 | 1799601 | 1828195 | 1752932 |
| Variation | 76828   | -11306  | 115053  | 22688   | 96960   | 53407   | 22027   | -115876 | 185636  | 28594   | -75263  |

The relation between the operation matrix  $O^w(t)$  and the criterion matrix  $C(t)$  can be calculated, and we can get a relation matrix  $R(t)_{w \times m}$  by performing  $R(t) = O^w(t) \otimes C(t)$ , where

$$R(t) = \begin{bmatrix} O_{11} \times C_1 & O_{12} \times C_2 & \cdots & O_{1m} \times C_m \\ O_{21} \times C_2 & O_{22} \times C_2 & \cdots & O_{2m} \times C_m \\ \cdots & \cdots & \ddots & \cdots \\ O_{w1} \times C_2 & O_{w2} \times C_2 & \cdots & O_{wm} \times C_m \end{bmatrix} \quad (4)$$

$$= \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ R_{21} & R_{22} & \cdots & R_{2m} \\ \cdots & \cdots & \ddots & \cdots \\ R_{w1} & R_{w2} & \cdots & R_{wm} \end{bmatrix}$$

where  $R_{ij} = O_{ij} \times C_j$ ,  $1 \leq i \leq w$ ,  $1 \leq j \leq m$ . From the relation matrix  $R(t)$ , we get the degree of relationships between the last value of time series and the other past values in time series data variations. Then, we can get the forecasting variation of the time series att, where

$$F(t) = [Max(R_{11}, R_{21}, \cdots, R_{w1}) \cdots Max(R_{12}, R_{22}, \cdots, R_{w2}) \cdots Max(R_{1m}, R_{2m}, \cdots, R_{wm})]$$

### III. PREDICTION ALGORITHM

In this section, we discuss in more detail the algorithm used to predict the time series of number of pilgrims coming from outside the kingdom of Saudi Arabia (NPO) consists of 44 observations (from year 1390AH # to year 1433 AH) table (1).

**Step 1:** From the time series data shown in table (1), compute the variations of the time series values between any two continuous values. The variation of time series current value is the time series current value minus the time series previous value as shown in table (2).

To define the universe of discourse  $U$ , let  $v_{min}$  be the minimum of variation and  $v_{max}$  be the maximum of variation. Then the universe of discourse  $U = [v_{min} - v_1, v_{max} + v_2]$ , where  $v_{min}$  and  $v_{max}$  are suitable positive numbers. Here,  $v_{min} = 369$ ,  $v_{max} = 478$  and  $U = [-198000, 311500]$

**Step 2:** Partition the universe of discourse  $U$  into several even length intervals  $u_1, u_2, \cdots, u_m$ . In this paper, we partition the universe of discourse  $U$  into six intervals:

$$\begin{aligned} u_1 &= [-198000, -113083], \\ u_2 &= [-113083, -28167], \\ u_3 &= [-28167, 56750], \\ u_4 &= [56750, 141667], \\ u_5 &= [141667, 226583], \\ u_6 &= [226583, 311500]. \end{aligned}$$

**Step 3:** Define fuzzy sets on the universe of discourse  $U$ . First, determine some linguistic values represented by fuzzy sets to describe the degree of variation between two continuous values. Here, we consider six fuzzy sets which are  $A_1 =$  (big decrease),  $A_2 =$  (decrease),  $A_3 =$  (no change),  $A_4 =$  (increase),  $A_5 =$  (big increase),  $A_6 =$  (too big increase). Then, define fuzzy sets  $A_1, A_2, \cdots, A_6$  on the universe of discourse  $U$  as follows:

$$\begin{aligned} A_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0/u_6 \\ A_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5 + 0/u_6 \\ A_3 &= 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + 0/u_5 + 0/u_6 \end{aligned} \quad (6)$$

$$\begin{aligned} A_4 &= 0/u_1 + 0/u_2 + 0.5/u_3 + 1/u_4 + 0.5/u_5 + 0/u_6 \\ A_5 &= 0/u_1 + 0/u_2 + 0/u_3 + 0.5/u_4 + 1/u_5 + 0.5/u_6 \end{aligned}$$

$$A_6 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6$$

**Step 4:** Fuzzify the values of time series historical data. If the number of variation of the time series at value  $i$  is  $p$ , where  $p \in u_j$ , and if there is a value represented by fuzzy set  $A_k$  in which the maximum membership value occurs at  $u_j$ , then  $p$  is translated to  $A_k$ . The fuzzified variations of the time series data are shown in table (3).

**Step 5:** Choose a suitable window basis  $w$ , and calculate the output from the operation matrix  $O^w(t)$  and the criterion matrix  $C(t)$ , where  $t$  is the year for which we want to forecast the time series value. The fuzzified forecasted variations for all years are given in table (4).

**Step 6:** Defuzzify the fuzzy forecasted variations derived in the former step as follow:

- If the grades of membership of the fuzzified forecasted variation are all 0, then we set the forecasted variation to 0.
- If the grades of membership of the fuzzified forecasted variation have only one maximum  $u_i$ , and the midpoint of  $u_i$  is  $m_i$ , then the forecasted variation is  $m_i$ . If the grades of membership of the fuzzified forecasted variation have more than one maximum  $u_1, u_2, \cdots, u_k$ , and their midpoints are  $m_1, m_2, \cdots, m_k$ , respectively, then the forecasted variation is  $(m_1 + m_2 + \cdots + m_k)/k$ .

**Step 7:** Calculate the forecasted time series value. The forecasted time series value is forecasted variation plus the true value of last year.

**Step 8:** Calculate the mean absolute percentage error (MAPE), which is a measure of prediction accuracy of a forecasting method in statistics. It usually expresses accuracy as a percentage, and is defined by formula (7):

$$MAPE = \frac{1}{m} \sum_{t=1}^m \left| \frac{\text{true value} - \text{forecast}}{\text{true value}} \right| \times 100 \quad (7)$$

Table (5) shows time series true values versus predicted values and the error estimate (MAPE) for all years for window base;  $w = 5$ . Figure (3.1) shows the graph of both true and predicted values.

#### 3.1. Accuracy of Prediction for Different Window Bases

From equation (3) it is clear that the operation matrix  $O^w(t)$  size depends on window base  $w$  and consequently the relation matrix  $R(t)$  which is the kernel of prediction. Table (6) shows prediction for different values of window base, e.g.,  $w = 5, 10, 15, 20, 25, 30$ . Also table (8) shows MAPE error for the different values of  $w$ .

Table 3: Fuzzified historical values of time series

| Year | Variations | Fuzzified Variations |
|------|------------|----------------------|
| 1390 |            |                      |
| 1391 | 48069      | $A_3$                |
| 1392 | 165843     | $A_5$                |
| 1393 | -37427     | $A_2$                |
| 1394 | 311022     | $A_6$                |
| 1395 | -24204     | $A_3$                |
| 1396 | -175533    | $A_1$                |
| 1397 | 20279      | $A_3$                |
| 1398 | 90917      | $A_4$                |
| 1399 | 32284      | $A_3$                |
| 1400 | -49628     | $A_2$                |
| 1401 | 66476      | $A_4$                |
| 1402 | -25813     | $A_3$                |
| 1403 | 150356     | $A_5$                |
| 1404 | -84240     | $A_2$                |
| 1405 | -67910     | $A_2$                |
| 1406 | 4957       | $A_3$                |
| 1407 | 103668     | $A_4$                |
| 1408 | -197631    | $A_1$                |
| 1409 | 11805      | $A_3$                |
| 1410 | 54433      | $A_3$                |
| 1411 | -108891    | $A_2$                |
| 1412 | 292815     | $A_6$                |
| 1413 | -20104     | $A_3$                |
| 1414 | 2798       | $A_3$                |
| 1415 | 47663      | $A_3$                |
| 1416 | 37191      | $A_3$                |
| 1417 | 88126      | $A_4$                |
| 1418 | -36247     | $A_2$                |
| 1419 | -75614     | $A_2$                |
| 1420 | 210825     | $A_5$                |
| 1421 | 96437      | $A_4$                |
| 1422 | -9808      | $A_3$                |
| 1423 | 76828      | $A_4$                |
| 1424 | -11306     | $A_3$                |
| 1425 | 115053     | $A_4$                |
| 1426 | 22688      | $A_3$                |
| 1427 | 96960      | $A_4$                |
| 1428 | 53407      | $A_3$                |
| 1429 | 22027      | $A_3$                |
| 1430 | -115876    | $A_1$                |
| 1431 | 185636     | $A_5$                |
| 1432 | 28594      | $A_3$                |
| 1433 | -75263     | $A_2$                |

Table 4: Forecasted variations with windows basis  $w = 5$

| Years | Membership functions of forecasted variations |       |       |       |       |       |
|-------|---|-------|-------|-------|-------|-------|
|       | $u_1$   | $u_2$ | $u_3$ | $u_4$ | $u_5$ | $u_6$ |
| 1396  | 0   | 0.5   | 1     | 0.25  | 0     | 0     |
| 1397  | 0.5   | 0.5   | 0     | 0     | 0     | 0     |
| 1398  | 0   | 0.5   | 1     | 0.25  | 0     | 0     |
| 1399  | 0   | 0     | 0.5   | 0.5   | 0.5   | 0     |
| 1400  | 0   | 0.5   | 1     | 0.5   | 0     | 0     |
| 1401  | 0.5   | 1     | 0.5   | 0     | 0     | 0     |
| 1402  | 0   | 0     | 0.5   | 1     | 0.5   | 0     |
| 1403  | 0   | 0.5   | 1     | 0.5   | 0     | 0     |
| 1404  | 0   | 0     | 0     | 0.5   | 1     | 0.5   |
| 1405  | 0.5   | 1     | 0.5   | 0     | 0     | 0     |
| 1406  | 0.5   | 1     | 0.5   | 0     | 0     | 0     |
| 1407  | 0   | 0.5   | 1     | 0.5   | 0     | 0     |
| 1408  | 0   | 0     | 0.5   | 1     | 0.5   | 0     |
| 1409  | 1   | 0.5   | 0     | 0     | 0     | 0     |
| 1410  | 0   | 0.5   | 1     | 0.5   | 0     | 0     |
| 1411  | 0   | 0.5   | 1     | 0.5   | 0     | 0     |
| 1412  | 0.5   | 1     | 0.5   | 0     | 0     | 0     |
| 1413  | 0   | 0     | 0     | 0     | 0.5   | 1     |
| 1414  | 0   | 0.5   | 1     | 0.5   | 0     | 0     |
| 1415  | 0   | 0.5   | 1     | 0.5   | 0     | 0     |
| 1416  | 0   | 0.5   | 1     | 0.5   | 0     | 0     |

|      |     |     |     |     |     |     |
|------|-----|-----|-----|-----|-----|-----|
| 1417 | 0   | 0.5 | 1   | 0.5 | 0   | 0   |
| 1418 | 0   | 0   | 0.5 | 1   | 0.5 | 0   |
| 1419 | 0.5 | 1   | 0.5 | 0   | 0   | 0   |
| 1420 | 0.5 | 1   | 0.5 | 0   | 0   | 0   |
| 1421 | 0   | 0   | 0   | 0.5 | 1   | 0.5 |
| 1422 | 0   | 0   | 0.5 | 1   | 0.5 | 0   |
| 1423 | 0   | 0.5 | 1   | 0.5 | 0   | 0   |
| 1424 | 0   | 0   | 0.5 | 1   | 0.5 | 0   |
| 1425 | 0   | 0.5 | 1   | 0.5 | 0   | 0   |
| 1426 | 0   | 0   | 0.5 | 1   | 0.5 | 0   |
| 1427 | 0   | 0.5 | 1   | 0.5 | 0   | 0   |
| 1428 | 0   | 0   | 0.5 | 1   | 0.5 | 0   |
| 1429 | 0   | 0.5 | 1   | 0.5 | 0   | 0   |
| 1430 | 0   | 0.5 | 1   | 0.5 | 0   | 0   |
| 1431 | 1   | 0.5 | 0   | 0   | 0   | 0   |
| 1432 | 0   | 0   | 0   | 0.5 | 1   | 0.5 |
| 1433 | 0   | 0.5 | 1   | 0.5 | 0   | 0   |

Table 5: True values versus forecasted values using the proposed fuzzy time-series method with the window basis  $w = 5$

| Years | Actual Value | Forecasted Value | Errors (%) |
|-------|--------------|------------------|------------|
| 1396  | 719040       | 908865           | 26.40      |
| 1397  | 739319       | 605957           | 18.04      |
| 1398  | 830236       | 753611           | 9.23       |
| 1399  | 862520       | 929444           | 7.76       |
| 1400  | 812892       | 876812           | 7.86       |
| 1401  | 879368       | 742267           | 15.59      |
| 1402  | 853555       | 978576           | 14.65      |
| 1403  | 1003911      | 867847           | 13.55      |
| 1404  | 919671       | 1188036          | 29.18      |
| 1405  | 851761       | 849046           | 0.32       |
| 1406  | 856718       | 781136           | 8.82       |
| 1407  | 960386       | 871010           | 9.31       |
| 1408  | 762755       | 1059594          | 38.92      |
| 1409  | 774560       | 649672           | 16.12      |
| 1410  | 828993       | 788852           | 4.84       |
| 1411  | 720102       | 843285           | 17.11      |
| 1412  | 1012917      | 649477           | 35.88      |
| 1413  | 992813       | 1239500          | 24.85      |
| 1414  | 995611       | 1007105          | 1.15       |
| 1415  | 1043274      | 1009903          | 3.20       |
| 1416  | 1080465      | 1057566          | 2.12       |
| 1417  | 1168591      | 1094757          | 6.32       |
| 1418  | 1132344      | 1267799          | 11.96      |
| 1419  | 1056730      | 1061719          | 0.47       |
| 1420  | 1267555      | 986105           | 22.20      |
| 1421  | 1363992      | 1451680          | 6.43       |
| 1422  | 1354184      | 1463200          | 8.05       |
| 1423  | 1431012      | 1368476          | 4.37       |
| 1424  | 1419706      | 1530220          | 7.78       |
| 1425  | 1534759      | 1433998          | 6.57       |
| 1426  | 1557447      | 1633967          | 4.91       |
| 1427  | 1654407      | 1571739          | 5.00       |
| 1428  | 1707814      | 1753615          | 2.68       |
| 1429  | 1729841      | 1722106          | 0.45       |
| 1430  | 1613965      | 1744133          | 8.07       |
| 1431  | 1799601      | 1500882          | 16.60      |
| 1432  | 1828195      | 1983726          | 8.51       |
| 1433  | 1752932      | 1842487          | 5.11       |

Table 6: Predicted values with different window Bases

| Years | <i>different window bases and the coressponding prediction</i> |            |            |            |            |            |            |
|-------|--|------------|------------|------------|------------|------------|------------|
|       | True values  | $w = 5$    | $w = 10$   | $w = 15$   | $w = 20$   | $w = 25$   | $w = 30$   |
| 1396  | 719040   | 908864.67  |            |            |            |            |            |
| 1397  | 739319   | 605956.67  |            |            |            |            |            |
| 1398  | 830236   | 753610.67  |            |            |            |            |            |
| 1399  | 862520   | 929444.33  |            |            |            |            |            |
| 1400  | 812892   | 876811.67  |            |            |            |            |            |
| 1401  | 879368   | 742267.00  | 742267.00  |            |            |            |            |
| 1402  | 853555   | 978576.33  | 978576.33  |            |            |            |            |
| 1403  | 1003911  | 867846.67  | 867846.67  |            |            |            |            |
| 1404  | 919671   | 1188036.00 | 1188036.00 |            |            |            |            |
| 1405  | 851761   | 849046.00  | 849046.00  |            |            |            |            |
| 1406  | 856718   | 781136.00  | 781136.00  | 781136.00  |            |            |            |
| 1407  | 960386   | 871009.67  | 871009.67  | 871009.67  |            |            |            |
| 1408  | 762755   | 1059594.33 | 1059594.33 | 1059594.33 |            |            |            |
| 1409  | 774560   | 649671.67  | 649671.67  | 649671.67  |            |            |            |
| 1410  | 828993   | 788851.67  | 788851.67  | 788851.67  |            |            |            |
| 1411  | 720102   | 843284.67  | 843284.67  | 843284.67  | 843284.67  |            |            |
| 1412  | 1012917  | 649477.00  | 649477.00  | 649477.00  | 649477.00  |            |            |
| 1413  | 992813   | 1239500.33 | 1239500.33 | 1239500.33 | 1239500.33 |            |            |
| 1414  | 995611   | 1007104.67 | 1007104.67 | 1007104.67 | 1007104.67 |            |            |
| 1415  | 1043274  | 1009902.67 | 1009902.67 | 1009902.67 | 1009902.67 |            |            |
| 1416  | 1080465  | 1057565.67 | 1057565.67 | 1057565.67 | 1057565.67 | 1057565.67 |            |
| 1417  | 1168591  | 1094756.67 | 1094756.67 | 1094756.67 | 1094756.67 | 1094756.67 |            |
| 1418  | 1132344  | 1267799.33 | 1267799.33 | 1267799.33 | 1267799.33 | 1267799.33 |            |
| 1419  | 1056730  | 1061719.00 | 1061719.00 | 1061719.00 | 1061719.00 | 1061719.00 |            |
| 1420  | 1267555  | 986105.00  | 986105.00  | 986105.00  | 986105.00  | 986105.00  |            |
| 1421  | 1363992  | 1451680.00 | 1451680.00 | 1451680.00 | 1451680.00 | 1451680.00 | 1451680.00 |
| 1422  | 1354184  | 1463200.33 | 1463200.33 | 1463200.33 | 1463200.33 | 1463200.33 | 1463200.33 |
| 1423  | 1431012  | 1368475.67 | 1368475.67 | 1368475.67 | 1368475.67 | 1368475.67 | 1368475.67 |
| 1424  | 1419706  | 1530220.33 | 1530220.33 | 1530220.33 | 1530220.33 | 1530220.33 | 1530220.33 |
| 1425  | 1534759  | 1433997.67 | 1433997.67 | 1433997.67 | 1433997.67 | 1433997.67 | 1433997.67 |
| 1426  | 1557447  | 1633967.33 | 1633967.33 | 1633967.33 | 1633967.33 | 1633967.33 | 1633967.33 |
| 1427  | 1654407  | 1571738.67 | 1571738.67 | 1571738.67 | 1571738.67 | 1571738.67 | 1571738.67 |
| 1428  | 1707814  | 1753615.33 | 1753615.33 | 1753615.33 | 1753615.33 | 1753615.33 | 1753615.33 |
| 1429  | 1729841  | 1722105.67 | 1722105.67 | 1722105.67 | 1722105.67 | 1722105.67 | 1722105.67 |
| 1430  | 1613965  | 1744132.67 | 1744132.67 | 1744132.67 | 1744132.67 | 1744132.67 | 1744132.67 |
| 1431  | 1799601  | 1500881.67 | 1500881.67 | 1500881.67 | 1500881.67 | 1500881.67 | 1500881.67 |
| 1432  | 1828195  | 1983726.00 | 1983726.00 | 1983726.00 | 1983726.00 | 1983726.00 | 1983726.00 |
| 1433  | 1752932  | 1842486.67 | 1842486.67 | 1842486.67 | 1842486.67 | 1842486.67 | 1842486.67 |

Table 7: Prediction errors with different window bases

| Window bases(w) | MAPE(%) |
|-----------------|---------|
| 5               | 9.78    |
| 10              | 8.21    |
| 15              | 6.54    |
| 20              | 4.77    |
| 25              | 2.90    |
| 30              | 1.92    |

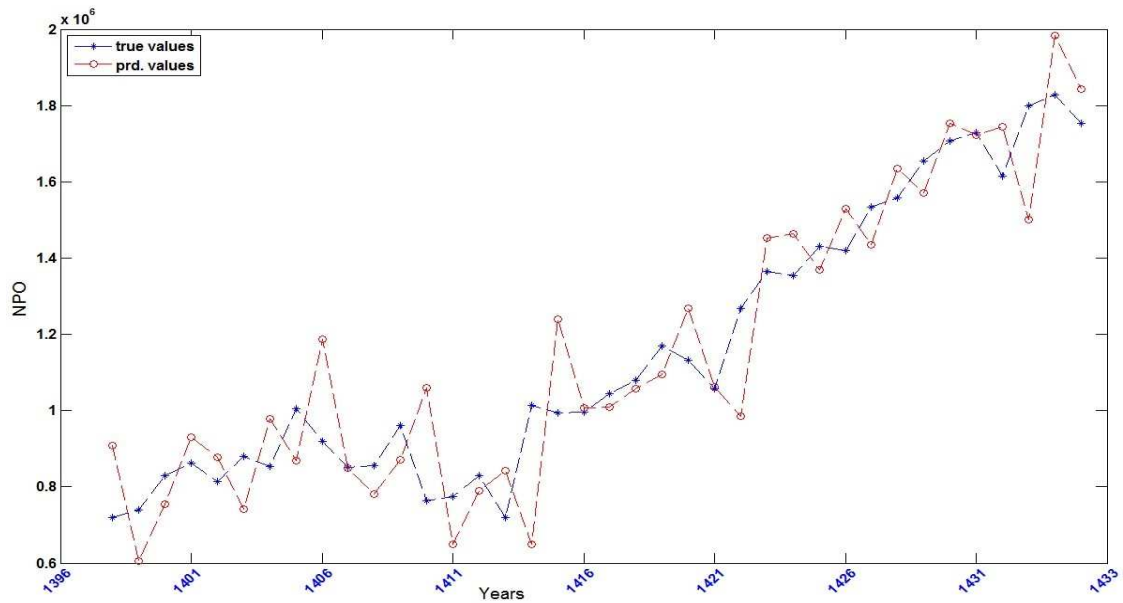


Fig.3.1: The true NPO and forecasts of observations.

#### IV. FORECASTING

In this experiment, fuzzy time series prediction algorithm was deployed to forecast the next five future observations. The point forecasts for these observations are given by table(8) and the result are shown in figure(4.2).

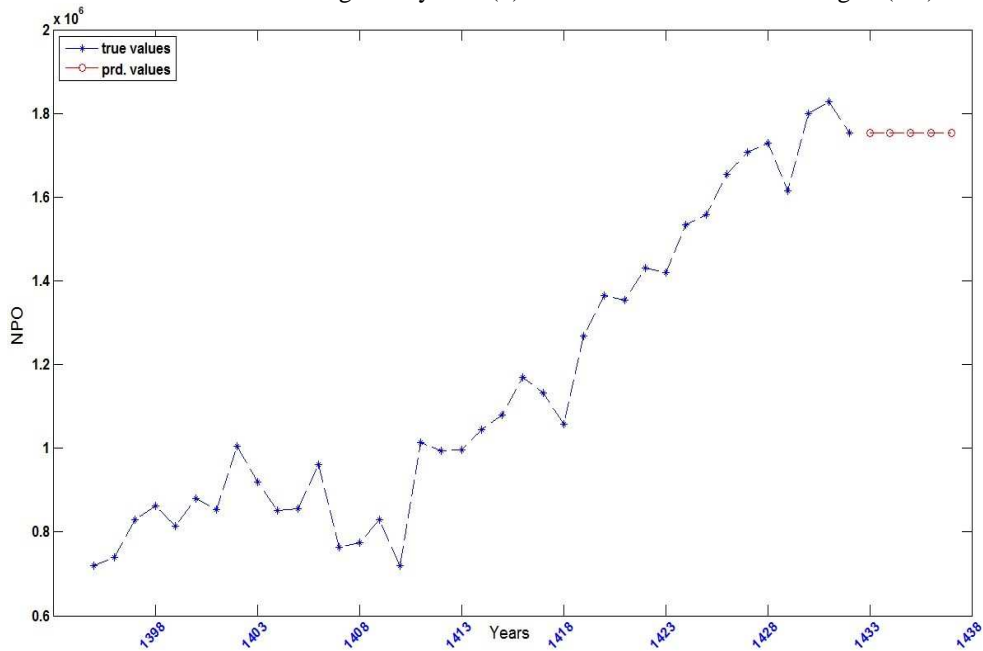


Fig. 4.2: The true NPO and future five observations.

Table 8: Predicted values for the next five years

| Year | Forecasted Value |
|------|------------------|
| 1434 | 1752930          |
| 1435 | 1752890          |
| 1436 | 1752900          |
| 1437 | 1752980          |
| 1438 | 1752950          |

#### V. CONCLUSION

The authors have proposed to use the fuzzy time series approach to forecast the series of number of Pilgrims coming from outside the Kingdom of Saudi Arabia from year 1390 AH to year 1433 AH. Point forecasts for the next five future years are provided by the authors using the proposed approach. The mean absolute percentage error

(MAPE) of the last five observation forecasts achieved by the approach is equal to 7.7%, compared with the results of Box-Jenkins approaches(9.6%)[2]. It has been shown that the proposed approach gives better forecasts than those achieved the traditional Box-Jenkins approaches.

## REFERENCES

- [1] J.R. Hwang, S.-M. Chen, C.-H. Lee, Handling forecasting problems using fuzzy time series, *Fuzzy Sets and Systems*, 100 217-228, (1998).
- [2] S. Shaarawy, E. Khan, M. Elgamal, Modeling and Forecasting the Number of Pilgrims Coming from Outside the Kingdom of Saudi Arabia Using Bayesian and Box-Jenkins Approaches, (IJACSA) *International Journal of Advanced Computer Science and Applications*, Vol.5, No. 4, 2014.
- [3] Q. Song and B.S. Chissom, Fuzzy time series and its models, *Fuzzy Sets and Systems* 54, 269-277, (1993).
- [4] Q. Song and B.S. Chissom, Forecasting enrollments with fuzzy time series{Part I, *Fuzzy Sets and Systems*54, 1-9, (1993).
- [5] Q. Song and B.S. Chissom, Forecasting enrollments with fuzzy time series{Part II, *Fuzzy Sets and Systems*62, 1-8, (1994).
- [6] F.M. Tseng, G.H. Tzeng, H.C. Yu and B.J.C. Yuan, Fuzzy ARIMA model for forecasting the foreign exchange market, *Fuzzy Sets and Systems* 118, 9-19, (2001).
- [7] W. Wu, Fuzzy reasoning and fuzzy relational equations, *Fuzzy Sets and Systems* 20, 67-78, (1986).
- [8] L.A. Zadeh, Fuzzy sets, *Fuzzy sets, Inform. and Control* 8 (1965) 338 353.

# The years are written using Lunar Calendar (AH) from 1390(1971) ~1433(2012); one lunar year is shorter than Gregorian year by about 11 days (see [http://en.wikipedia.org/wiki/Islamic\\_Calendar](http://en.wikipedia.org/wiki/Islamic_Calendar)).

## AUTHOR'S PROFILE



**Prof Dr. Mahmoud Elgamal** is presently working at The Custodian of the Two Holy Mosques Institute for Hajj and Omra Research, Umm Al -Qura University, Makkah, Saudi Arabia.

He is having more than 10 years of teaching experience with expertise in Image Processing, Artificial Intelligence, and Object Oriented Programming.

Email: [mahmoud\\_40@yahoo.com](mailto:mahmoud_40@yahoo.com)