Forecasting of Egypt Wheat Imports Using Multivariate Fuzzy Time Series Model Based on Fuzzy Clustering

Ashraf K. Abd-Elaal, Member, IAENG, Hesham A. Hefny, and Ashraf H. Abd-Elwahab

Abstract—This paper presents Multivariate-Factors fuzzy time series model for improving forecasting accuracy. The proposed model is based on fuzzy clustering and it employs eight main procedures to build the multivariate-factors model. The model is evaluated by studying the Egypt Wheat imports as a forecasting problem. Forecasting Egypt wheat imports depend on three factors: population size, wheat area, and wheat production. This forecasting problem is considered to be a good benchmark for comparing different forecasting techniques since it exhibits highly nonlinearities over a long period of time and it provides important economical indicators needed for national future planning. Experimental results show that the proposed model provides higher forecasting accuracy than ARIMA model, Regression model and neural network model. Therefore, the proposed model can lead to satisfactory high performance for fuzzy time series.

Index Terms—ARIMA, Egypt wheat imports, Fuzzy clustering, Multivariate-Factors fuzzy time series.

I. INTRODUCTION

The uncertainties existed in historical data represent a real challenge for traditional time series forecasting models. The main motivation for using Fuzzy time series forecasting models is their abilities to handle such uncertainties in historical data of real-world forecasting problem. Song and Chissom presented the first concept of fuzzy time series model for forecasting the enrolments of the University of Alabama based on the fuzzy set theory [1]-[2]. Since then, many researchers have contributed to developing and improving fuzzy time series models.

Qiu et al. presented a new method to generalize the conventional models for forecasting process, where the data of the University of Alabama and Shanghai stock index are adopted to illustrate the processes [3]. Bahrepour et al. presented a novel approach for high-order fuzzy time series. Their model was based on two facets. First was the use of a self-organizing map (SOM), as a fast clustering technique, to partition the universe of discourse unequally. The second facet of their model was the adoption of three different agents; voting, statistical and emotional, for estimating the best order of the high-order fuzzy time series model [4]. Abd-Elaal et al. introduced a fuzzy time series model, which depended on fuzzy clustering for partitioning the universe of discourse to forecast the Gold Reserves of Egypt based on official data starting from the first quarter of 2002 up to the first quarter of 2010. The comparison result, with other fuzzy time series models as well as the traditional ARIMA model, showed that the proposed model provided a higher accuracy and an efficient performance [5]. Liu et al. presented an approach to improve the derivation of fuzzy relationships in the fuzzy time series model using rough sets. Their proposed model, not only required no prior knowledge or pre-review dataset to generate heuristic rules, but also, effectively reduced computational effort by decreasing the numbers of fuzzy sets of linguistic variables [6]. Duru presented a study that aimed to improve the fuzzy logical forecasting model by utilizing multivariate inference. The model also allows handling the partitioning problem for an exponentially distributed time series by using a multiplicative clustering approach [7]. Hassan et al. presented a hybrid fuzzy time series model, based on Interval Type-2 Fuzzy Inference System (IT2-FIS) and ARIMA model. The model improved the forecasting result by handling the measurement and parametric uncertainties of ARIMA model by using Fuzzy approach [8]. Egrioglu et al. presented a novel hybrid fuzzy time series approach in which fuzzy c-means (FCM) method and artificial neural networks were employed for fuzzification and defuzzification, respectively. The model has successfully been applied to the well known enrollment data for the University of Alabama [9]. Khiabani et al. proposed combination of the adaptive time-variant model (ATVF) with PSO algorithm to improve Alabama University enrollments forecasting. ATVF model automatically adapts the analysis window size of fuzzy time series based on the predictive accuracy in the training phase and uses heuristic rules to determine forecasting values [10].

In this paper, researchers introduce multivariate-factors fuzzy time series forecasting model based on fuzzy clustering to handle real-world multivariate forecasting problems. The proposed model is examined by the problem of forecasting Egypt Wheat imports based on the official data of Central Agency for Public Mobilization and Statistics (CAPMAS).

II. BACKGROUND

A. ARIMA Model

For Stationary time series, i.e. time series with constant mean, the ARMA (Autoregressive Moving Average) forecasting model is widely used. The ARMA(p,q) model consists of two models, Autoregressive (AR(p)) and moving average (MA(q)) models. The forecasting of a time series
using the ARMA model is given in (1).
\[ Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \xi_t + \theta_1 \xi_{t-1} + \theta_2 \xi_{t-2} + \ldots + \theta_q \xi_{t-q} \] (1)

Where \( \xi_t \) is the error, \( \phi_1, \theta_q \) are parameters for ARMA model. This model can be used when data are stationary. However, it can’t be used in case of non-stationary data, i.e. Stationary means that there is no growth or decline in the data. That means that both mean and variance remain constant over time, then the forecasting model will be ARIMA (autoregressive integrated moving average) mode, ARIMA(p,d,q) where p is the order of the autoregressive part, d is the degree of the difference and q is the order of the moving average part [11]. For instance, ARIMA(4,1,2) model has four autoregressive parameters, two moving-average parameters, computed after the series have a difference of order one. The ARIMA model is a widely used time series model that represents the basis of many fundamental ideas in time-series analysis.

B. Neural Networks

Neural network (NN) is a computational structure inspired by the study of biological neural processing. NN consists of many processing units, called neurons. Each neuron has multi-inputs and a single output. The inputs emulate the external signals received by the biological neurons. The output represents the neuron response according to the accumulated weighted input signals. The weights representing the force of the synaptic union: positive weight representing an excitatory effect, and negative weight which is an inhibitory effect. If the result of the sum of weighted inputs is higher than a certain threshold (normally +1); in the opposite case, the output presents a zero or a negative value (normally -1). Fig. 1 shows the schematic representation of an artificial neuron.

Fig. 1. Schematic representation of an artificial neuron

schematic representation of an artificial neuron. [12]. Due to their powerful learning capabilities, neural networks have been adopted by many researchers for solving several forecasting problems [13]-[17].

C. Fuzzy Clustering

Fuzzy clustering aims at partitioning a data set into homogenous fuzzy clusters [18]. Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to more than one cluster with different degrees of membership values. Fuzzy C-Mean Iterative (FCMI) is a famous implementation of fuzzy clustering algorithms [19]. Assuming a set of m patterns X=(x_1, x_2, ..., x_m) is distributed in d-dimensional pattern space and c fuzzy clusters, whose centers have initial values \( y_{10}, y_{20},...,y_{c0} \). At each iteration (K), the distance \( d_{ij} \) between each pattern \( x_i \) and each cluster \( c_j \) is computed by:
\[ d_{ij}^{(k)} = ||x_i - y_j^{(k)}|| \] (2)

Fuzzy clustering is carried out through an iterative optimization of the objective function \( d_{ij}^{(k)} \), with the update of membership \( u_{ij}^{(k)} \) and the cluster centers \( y_j \) by:
\[ y_j^{(k+1)} = \frac{\sum_{i=1}^{n} u_{ij}^{(k)} x_i}{\sum_{i=1}^{n} u_{ij}^{(k)}} \] (3)

Where
\[ u_{ij}^{(k)} = \left[ \sum_{j=1}^{c} \left( \frac{d_{ij}^{(0)}}{d_{ij}^{(k)}} \right)^{2 \beta} \right]^{-1} \]

The process terminates when the difference between two consecutive clusters centers do not exceed a given tolerance.
\[ \left( \sum_{i=1}^{m} ||y_j^{(k+1)} - y_j^{(k)}|| \right)^{1/2} < \epsilon \] (4)

D. Fuzzy Time Series

Song and Chissom presented the concept of fuzzy time series (FTS) for allowing solving forecasting problems of considerably short time series with uncertainties. They presented the time-invariant fuzzy time series model and the time-variant fuzzy time series model based on the fuzzy set theory for forecasting the enrollments of the University of Alabama. The following are definitions of basic concepts of the fuzzy time-series [1][5][20]-[24].

Definition 1. Assume \( Y(t) (t = 0, 1, 2, \ldots) \) is a subset of a real numbers. Let \( Y(t) \) is the universe of discourse defined by the fuzzy set \( f_1(t) \). If \( F(t) \) is a collection of \( f_1(t), f_2(t), \ldots \) then \( F(t) \) is defined as a FTS on \( Y(t) (t = 0, 1, 2, \ldots) \).

Definition 2. If there is a fuzzy logical relationship \( R(t-1, t) \), such that \( F(t) = F(t-1) \circ R(t-1, t) \), where “\( \circ \)” represents a max-min composition operation, then \( F(t) \) is induced by \( F(t-1) \). The fuzzy logical relationship (FLR) between \( F(t) \) and \( F(t-1) \) is denoted by \( F(t) \rightarrow F(t) \).

Definition 3. Suppose \( F(t-1) = A_i \) and \( F(t) = A_j \) where \( A_i \) and \( A_j \) are two fuzzy sets defined on \( Y(t) \). The relationship between two consecutive observations, \( F(t) \) and \( F(t-1) \), referred to as a FLR, can be denoted by \( A_i \rightarrow A_j \), where \( A_i \) is called the Left-Hand Side (LHS) and \( A_j \) is called the Right-Hand Side (RHS) of the FLR.

Definition 4. All fuzzy logical relationships (FLRs) in the training dataset can be grouped together into different fuzzy logical relationship groups according to the same LHS of the FLR. For example, if there are two fuzzy logical relationships with the same LHS \( A_i \) as: \( A_i \rightarrow A_3; A_i \rightarrow A_2 \), then these two FLRs can be grouped into a fuzzy logical relationship group (FLRG) as: \( A_i \rightarrow A_j, A_2 \).

Definition 5. IF \( F(t) \) is FTS that is caused by \( F(t-1), F(t-2), \ldots \) and \( F(t-n) \), then \( F(t) \) is called “\( n^{th} \) order FTS” with “\( n^{th} \) order fuzzy logical relation” represented as: \( F(t-n), \ldots, F(t-2), F(t-1) \rightarrow F(t) \).

Definition 6. Let \( F_i(t) \) and \( F_2(t) \) be two FTS. If \( F_1(t) \) is...
caused by \((F(t-1), F_2(t-2), \ldots, (F(t-n)), (F(t-n-1), F_2(t-n-2), \ldots, (F(t-2), F_2(t-3), \ldots, F_2(t-n)))\), then \(F(t)\) is called “two-factors \(n^\text{th}\) order FTS”, with “\(n^\text{th}\) order fuzzy logical relationships represented as” \((F(t-1), F_2(t-1), (F(t-2), F_2(t-2), \ldots, \ldots, \ldots, (F(t-n), F_2(t-n-2), \ldots, F_2(t-2), F_2(t-3), \ldots, F_2(t-n))) \rightarrow F(t)\).

### III. METHODOLOGY

**A. The Proposed Multivariate-Factors Model**

If \(F(t)\) is caused by \((F(t-1), F(t-2), \ldots, F(t-n+1)), (F(t-2), F(t-3), \ldots, F(t-2-n+1)), \ldots, (F(t-n), F(t-n-1), \ldots, F(t-n)), (F(t-n-1), F(t-n-2), \ldots, F(t-n-1)), \ldots, (F(t-2), F(t-3), \ldots, F(t-2)), F(t-1), F(t), F(t+1))\), then the FLR is represented by

\[
\begin{align*}
(F(t-1), F(t-2), \ldots, F(t-n)), & \quad \cdots, \\
(F(t-n), F(t-n-1), \ldots, F(t-n+1)), & \quad \cdots, \\
(F(t-n), F(t-n-1), \ldots, F(t-n+1)), & \quad \cdots, \\
(F(t), F_2(t), \ldots, F_2(t-n)), & \quad \cdots, \\
(F(t-1), F(t-2), \ldots, F(t-n)), & \quad \cdots,
\end{align*}
\]

This is called Multivariate-factors fuzzy time series model (MFFTS), where \((F(t-n), F(t-n)), \ldots, F(t-n))\) are called antecedents factors, and \(F(t)\) is called consequent factor.

1) The Eight Steps of the Proposed Model

The stepwise computational procedure of the proposed MFFTS model is explained in the following eight steps:

**Step 1. Cluster antecedents and consequent factors data into \(c\) clusters:** Clustering \((F(t-1), F(t-2), \ldots, F(t-n+1)), (F(t-2), F(t-3), \ldots, F(t-2-n+1)), \ldots, (F(t-n), F(t-n-1), \ldots, F(t-n)), (F(t-n-1), F(t-n-2), \ldots, F(t-n-1)), \ldots, (F(t-2), F(t-3), \ldots, F(t-2), F(t-1), F(t), F(t+1))\) with \(n\) observation into \(c\) \((2 \leq c \leq n)\) clusters.

At iteration \(k=0\), initialize \(Y_{ij} = y_{ij,0}, 1 \leq i \leq c_j\)

\[
Y_{ij,0} = D_{j,\text{min}} / (a * i)
\]

Where, \(D_{j,\text{min}}\) is the minimum value of \(j\) factor, \(c_j\) is the number of clusters of \(j\) factor and \(a\) is a positive integer.

**Step 2. Determine membership values for each cluster:** In this step, the proposed model selected the maximum membership grade of each value for each cluster for antecedents and consequent factors which it belongs to.

**Step 3. Define the Universe of Discourse:** In this step, the proposed model defines the universe of discourse \(V_j\) as

\[
V_j = [V_{j,\text{min}} - V_{j,\text{down}}] V_{j,\text{max}} + V_{j,\text{up}} \}
\]

Where, \(V_{j,\text{min}}\) is the minimum value of \(j\) factor, \(V_{j,\text{max}}\) is the maximum value of \(j\) factor, and \(V_{j,\text{down}}, V_{j,\text{up}}\) are the positive real numbers to divide \(V_j\) into \(c_j\) intervals.

**Step 4. Partition the universe of discourse:** According to this step, the proposed model partitions the universe of discourse of antecedents and consequents factors into \(c_j\) intervals.

**Step 5. Fuzzify the historical data:** Proposed model fuzzifies historical data, by determining the best fuzzy cluster to each actual data for antecedents and consequents factors.

**Step 6. Calculate the crisp value for each linguistic term:** The crisp value that represents each linguistic term is calculated by:

\[
\text{Crisp value} = \frac{\sum_{j=1}^{c_j} m_{g_{aj}} x_{k,j} + m_{g_{aj}} x_{k,j}}{\sum_{j=1}^{c_j} m_{g_{aj}} x_{k,j}}
\]

Where \(m_{g_{aj}}\) is the membership grade, and \(X_{k,j}\) is the actual value.

**Step 7. Re-fuzzify of historical data:** Linguistics with highly frequent of occurrence is selected for further partitioning. The interval corresponding to linguistic with highest frequency of occurrence is re-clustered into \(k\) clusters, where \(2 \leq k \leq n-k\). The interval corresponding to the next linguistic term with highest frequency of occurrence is re-clustered to \(k-1\) clusters, and so on. Thus, historical data can be fuzzified again using a larger number of linguistics. The crisp values, which were produced in this step after re-fuzzification of the historical data, are the required forecasting value for antecedents and consequent factors.

**Step 8. Determine the forecasted values of consequent factor:** The proposed model uses the fuzzy logical relationship between the forecasting values of the antecedent’s factors and the forecasting values of consequent factor to calculate the forecasting values.

\[
\text{Forecasting (A)} = \frac{\sum_{k=1}^{n-k} \text{Crisp - value}(X_{k, (i-r)}) \times \sum_{e=1}^{n-e} \text{Crisp - value}(X_{e, (i-r)})}{n}
\]

Where \(n\) is the order degree

**IV. EXPERIMENTAL RESULTS**

To examine the above proposed model, we use the problem of forecasting “Egypt Wheat imports” as the experimental test problem. Egypt Wheat imports depend on three factors: population size, wheat cultivated area, and wheat production. If population size increases, then wheat imports should be increased too. If wheat production increases or wheat area increases, then wheat imports should be decreased. Thus, there are many relationships among these factors that may be positive or negative relationship.

In our experimental study, we use the original data provided by the Central Agency for Public Mobilization and Statistics (CAPMAS), for “Egypt Wheat imports”, population size, wheat cultivated area and Wheat production during the period starting from year 1986 up to 2008.

The forecasting accuracy is compared by using the “Normalized Root Mean Squared Error” (NRMSE), and the “Normalized Root Mean Squared Error” (NMSE) performance indices.

NRMSE, in statistics, is the square root of the sum of the squared deviations between actual and predicted values divided by the sum of the square of actual values. The NRMSE is often expressed in units of percent. Smaller values indicate a better agreement between measured and calculated values.

\[
\text{NRMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\text{actual}_i - \text{predict}_i)^2}{\sum_{i=1}^{n} (\text{actual}_i)^2}}
\]
NMSE, in statistics, is the sum of the squared deviations between actual and predicted values divided by the sum of the squared deviations between actual and mean of the actual values.

$$\text{NMSE} = \frac{\sum (\text{actual} - \text{predict})^2}{\sum (\text{actual} - \text{mean actual})^2}$$ \hspace{1cm} (11)

A. Testing the Multivariate-Factors Proposed Model

According to Table I, the universe of discourse of wheat imports is found to be $U=[3011, 8173]$ with $D_{\text{up}}=72$, and $V_{1,\text{up}}=90$, then $U$ is partitioned into 9 intervals. The universe of discourse of population size is found to be $V_1=[47677, 57315]$ with $V_{1,\text{up}}=97$, and $V_{1,\text{down}}=90$, then $V_1$ is partitioned into 9 intervals. The universe of discourse of wheat cultivated area is $V_2=[1174,3129]$ with $V_{2,\text{down}}=32$, and $V_{2,\text{up}}=65$, then $V_2$ is partitioned into 9 intervals. Finally, the universe of discourse of wheat production is $V_3=[1832, 8425]$ with $V_{3,\text{down}}=97$, and $V_{3,\text{up}}=151$. Then, $V_3$ is partitioned into 9 intervals, with see Fig. 2.

<table>
<thead>
<tr>
<th>Years</th>
<th>Wheat Imports (000 tons)</th>
<th>Population (000 people)</th>
<th>Wheat Area (000 feddans)</th>
<th>Wheat Production (000 tons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>6811</td>
<td>47751</td>
<td>1206</td>
<td>1929</td>
</tr>
<tr>
<td>1987</td>
<td>6874</td>
<td>48816</td>
<td>1373</td>
<td>2722</td>
</tr>
<tr>
<td>1988</td>
<td>6890</td>
<td>49826</td>
<td>1421</td>
<td>2839</td>
</tr>
<tr>
<td>1989</td>
<td>7527</td>
<td>50858</td>
<td>1532</td>
<td>3183</td>
</tr>
<tr>
<td>1990</td>
<td>7712</td>
<td>51911</td>
<td>1955</td>
<td>4268</td>
</tr>
<tr>
<td>1991</td>
<td>5550</td>
<td>52985</td>
<td>2215</td>
<td>4483</td>
</tr>
<tr>
<td>1992</td>
<td>5496</td>
<td>54082</td>
<td>2092</td>
<td>4618</td>
</tr>
<tr>
<td>1993</td>
<td>6118</td>
<td>55201</td>
<td>2171</td>
<td>4786</td>
</tr>
<tr>
<td>1994</td>
<td>4509</td>
<td>56344</td>
<td>2111</td>
<td>4437</td>
</tr>
<tr>
<td>1995</td>
<td>6256</td>
<td>57642</td>
<td>2512</td>
<td>5722</td>
</tr>
<tr>
<td>1996</td>
<td>5868</td>
<td>58835</td>
<td>2421</td>
<td>5735</td>
</tr>
<tr>
<td>1997</td>
<td>5202</td>
<td>60053</td>
<td>2486</td>
<td>5849</td>
</tr>
<tr>
<td>1998</td>
<td>5817</td>
<td>61296</td>
<td>2421</td>
<td>6093</td>
</tr>
<tr>
<td>1999</td>
<td>4776</td>
<td>62564</td>
<td>2380</td>
<td>6347</td>
</tr>
<tr>
<td>2000</td>
<td>3930</td>
<td>63860</td>
<td>2463</td>
<td>6584</td>
</tr>
<tr>
<td>2001</td>
<td>4982</td>
<td>65182</td>
<td>2342</td>
<td>6255</td>
</tr>
<tr>
<td>2002</td>
<td>4819</td>
<td>66531</td>
<td>2450</td>
<td>6625</td>
</tr>
<tr>
<td>2003</td>
<td>5205</td>
<td>67908</td>
<td>2506</td>
<td>6845</td>
</tr>
<tr>
<td>2004</td>
<td>3083</td>
<td>69313</td>
<td>2605</td>
<td>7178</td>
</tr>
<tr>
<td>2005</td>
<td>6272</td>
<td>70748</td>
<td>2985</td>
<td>8141</td>
</tr>
<tr>
<td>2006</td>
<td>5372</td>
<td>72212</td>
<td>3064</td>
<td>8274</td>
</tr>
<tr>
<td>2007</td>
<td>8073</td>
<td>73608</td>
<td>2716</td>
<td>7379</td>
</tr>
<tr>
<td>2008</td>
<td>7381</td>
<td>75225</td>
<td>2920</td>
<td>7977</td>
</tr>
</tbody>
</table>

Hence, the intervals of Egypt Wheat imports are $u_1; u_2; u_3; u_4; u_5; u_6; u_7; u_8$ where:

$u_1=[3011.00, 3584.56]$ \hspace{1cm} $u_2=[3584.56, 4158.11]$ \hspace{1cm} $u_3=[4158.11, 4731.67]$ \hspace{1cm} $u_4=[4731.67, 5305.22]$ \hspace{1cm} $u_5=[5305.22, 5878.78]$ \hspace{1cm} $u_6=[5878.78, 6452.33]$ \hspace{1cm} $u_7=[6452.33, 7025.89]$ \hspace{1cm} $u_8=[7025.89, 7599.44], u_9=[7599.44, 8173.00]$.

A linguistic variable named "Wheat imports" ($Y$) is defined to have the following linguistic values: $A_1 =$ (very very few), $A_2 =$ (very very few), $A_3 =$ (very few), $A_4 =$ (few), $A_5 =$ (moderate), $A_6 =$ (many), $A_7 =$ (many many), $A_8 =$ (many many many), $A_9 =$ (too many). The proposed model fuzzifies the historical data by assigning the corresponding linguistic values as shown in Table VI. Table II shows the membership grade of wheat import actual values, which affect on linguistic values. The proposed model selects the maximum membership grade for each cluster. The representing crisp value for each cluster is calculated as follows:

$$\text{Crisp value}(A_1) = 0.7 \times (2000)/0.7 = 3930$$

$$\text{Crisp value}(A_2) = 1.0x(1994)/1.0 = 4509$$

$$\text{Crisp value}(A_3) = (1.0x(1987) + 1.0x(1988))/2.0 = 6882$$

The intervals of population number are $V_{1,\text{up}}; V_{1,\text{down}}; V_{2,\text{up}}; V_{2,\text{down}}$.
A linguistic variable named "population size" (X₁), is defined to have the following linguistic values which are defined as: B₁₁ = (very very very few), B₁₂ = (very very few), B₁₃ = (very few), B₁₄ = (few), B₁₅ = (moderate), B₁₆ = (many), B₁₇ = (many many), B₁₈ = (many many many), B₁₉ = (too many). The proposed model fuzzifies the historical data by assigning the corresponding linguistic values as shown in Table VI. Table III shows the membership grade of population actual values, which affect on linguistic values. The proposed model selects the maximum membership grade for each cluster. The representing crisp value for each cluster calculating as:

\[
\text{Crisp}_\text{value}(B_i) = \left( 0.9 \times (1988) + 0.9 \times (1989) \right) / 1.8 = 55936
\]

\[
\text{Crisp}_\text{value}(B_{12}) = 1.0 \times (1991) / 1.0 = 52985
\]

\[
\text{Crisp}_\text{value}(B_{19}) = 1.0 \times (2000) / 1.0 = 72212
\]

<table>
<thead>
<tr>
<th>Year</th>
<th>X₁</th>
<th>B₁₁</th>
<th>B₁₂</th>
<th>B₁₃</th>
<th>B₁₄</th>
<th>B₁₅</th>
<th>B₁₆</th>
<th>B₁₇</th>
<th>B₁₈</th>
<th>B₁₉</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>47751</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1987</td>
<td>48816</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1988</td>
<td>49826</td>
<td>0.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1989</td>
<td>50858</td>
<td>0.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1990</td>
<td>51911</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1991</td>
<td>52985</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1992</td>
<td>54082</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1993</td>
<td>55201</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1994</td>
<td>56344</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1995</td>
<td>57642</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1996</td>
<td>58835</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1997</td>
<td>60053</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1998</td>
<td>61296</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1999</td>
<td>62564</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>63860</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2001</td>
<td>65182</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2002</td>
<td>66531</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2003</td>
<td>67908</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2004</td>
<td>69313</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2005</td>
<td>70748</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2006</td>
<td>72212</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2007</td>
<td>73608</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2008</td>
<td>75225</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The intervals of Wheat production are v₁₃₁; v₁₃₂; v₁₃₃; v₁₃₄; v₁₃₅; v₁₃₆; v₁₃₇; v₁₃₈; v₁₃₉ where:

\[
v_{131} = [1832.00, 2564.56], \quad v_{132} = [2564.56, 3297.11], \quad v_{133} = [3297.11, 4029.67], \quad v_{134} = [4029.67, 4762.22],
\]
\[
v_{135} = [4762.22, 5494.78], \quad v_{136} = [5494.78, 6227.33], \quad v_{137} = [6227.33, 6959.89], \quad v_{138} = [6959.89, 7692.44], \quad v_{139} = [7692.44, 8425.00]
\]

A linguistic variable named "Wheat area" (X₂), is defined to have the following linguistic values which are defined as: B₂₁ = (very very very few), B₂₂ = (very very few), B₂₃ = (very very few), B₂₄ = (few), B₂₅ = (moderate), B₂₆ = (many), B₂₇ = (many many), B₂₈ = (many many many), B₂₉ = (too many).

\[
\text{Crisp}_\text{value}(B_{21}) = 1.0 \times (1987) / 1.0 = 1397
\]

\[
\text{Crisp}_\text{value}(B_{22}) = 0.7 \times (1990) / 0.7 = 1955
\]

\[
\text{Crisp}_\text{value}(B_{29}) = 1.0 \times (2008) / 1.0 = 2920
\]
assigning the corresponding linguistic values as shown in Table VI. Table V shows the membership grade of Wheat production actual values, which affect on linguistic values.

<table>
<thead>
<tr>
<th>Year</th>
<th>X₁</th>
<th>B₁</th>
<th>B₂</th>
<th>X₂</th>
<th>B₁</th>
<th>B₂</th>
<th>X₃</th>
<th>B₁</th>
<th>B₂</th>
<th>B₃</th>
<th>B₄</th>
<th>B₅</th>
<th>B₆</th>
<th>B₇</th>
<th>B₈</th>
<th>B₉</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>1929</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1987</td>
<td>2722</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1988</td>
<td>2839</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1989</td>
<td>3183</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1990</td>
<td>4268</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1991</td>
<td>4483</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1992</td>
<td>4618</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1993</td>
<td>4786</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1994</td>
<td>4437</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1995</td>
<td>5722</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1996</td>
<td>5735</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1997</td>
<td>5849</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1998</td>
<td>6093</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1999</td>
<td>6347</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>6564</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2001</td>
<td>6255</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2002</td>
<td>6625</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2003</td>
<td>6384</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2004</td>
<td>7178</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2005</td>
<td>8141</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2006</td>
<td>8274</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2007</td>
<td>7379</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2008</td>
<td>7977</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The proposed model selects the maximum membership grade for each cluster. The representing crisp value for each cluster calculating is:

\[
\text{Crisp}_{\text{value}}(e; B_{ij}) = \begin{cases} 
\frac{(1.0 \times (x_{1987}) + 1.0 \times (x_{1988}))}{2} & : 6882 \\
\text{Crisp}_{\text{value}}(e; B_{ij}) = 1.0 \times (1990) / 1.0 = 7712 \\
\vdots \\
\text{Crisp}_{\text{value}}(e; B_{ij}) = 1.0 \times (2008) / 1.0 = 7381 
\end{cases}
\]

The proposed model selects linguistics with highly frequency of occurrence then partitions its interval into k sub-partitions. For consequent factor “wheat imports”, the proposed model divides \(u_e=[4731.67, 5305.22] \) into three partitions and \(u_e=[5305.22, 5878.78] \) into two partitions. When proposed model divided \(u_e \) then linguistic \(A_2 \) will be converted to be linguistic \(A_2 \) and linguistic \(A_7 \) to be linguistic \(A_{10} \) and so on.

For antecedent’s factors: population linguistic remains the same, but for “Wheat area”, the proposed model divides \(v_{25}=[2260.11, 2477.33] \) into three partitions and \(v_{25}=[2042.89, 2260.11] \) into two partitions. When proposed model divided \(v_{25} \) then linguistic \(B_{26} \) will be converted to be linguistic \(B_{27} \) and linguistic \(B_{27} \) to be linguistic \(B_{28} \) and so on, and for Wheat production, the proposed model divides \(v_{37}=[6227.33, 6959.89] \) into three partitions and \(v_{38}=[5494.78, 6227.33] \) into two partitions. When proposed model divided \(v_{36} \) then linguistic \(B_{37} \) will be converted to be linguistic \(B_{38} \) and linguistic \(B_{38} \) to be linguistic \(B_{39} \) and so on. By using the fuzzy logical relationship between the forecasting values of the antecedents factors and the forecasting values for consequents factors:

\[
\text{Forecasting}(A_j) = \left[ \frac{\sum_{r=0}^{k} \text{Crisp}_{\text{value}}(X_k, (i-r)) \times \text{Crisp}_{\text{value}}(Y_f, (r-f))}{k} \right]^{1/3} 
\]

The forecasting value for year 1998 is found to be 5554 while the actual value was 5817 and the forecasting value for year 2003 is found to be 5204 while the actual value was 5205. Table VI shows linguistic terms and forecasting values deduced by proposed model.

**B. Testing Regression Model**

The regression equation is:

\[
\text{Imprts} = -2.542 + 0.212 \text{ Population} + 2.84 \text{ Area} - 1.96 \text{ Production} 
\]

The R-squared and adjusted R-squared values are estimates of the 'goodness of fit' of the line. They represent the variation percentage of the data explained by the fitted line; the closer the points to the line, the better the fit.

**(Advance online publication: 29 November 2013)**
Adjusted R-squared is not sensitive to the number of points within the data. R-squared is derived from:

\[ R\text{-}squared = 100 \times \frac{SSR}{SST} \]

(14)

Adjusted R-squared tends to optimistically estimate how well the models fits the real data, is derived from:

\[ \text{adjusted } R\text{-}squared = 100 \times \left(1 - \frac{MSE}{MST}\right) \]

(15)

R-Sq = 20.1%, indicate that this model is not good to fit the data. The R-S(adj)= 7.5%, indicate that this model can’t represent the four factors due to their logical relationships that are sometimes positive and other times negative.

C. Testing ARIMA Model

The time series has four seasons (Import, population, Area, Production) by year. The best model is ARIMA(0,0,0)(2,1,2) which has the following form:

\[ Y_t = 0.6181 Y_{t-4} + 0.3832 Y_{t-1} + 1.068 \xi_t + 0.1907 \xi_{t-2} \]

(16)

Where \( Y_t \) = \( Y_t - Y_{t-4} \), \( Y_t = \text{log}(Z_t) \) and \( Z_t \) is the data series.

From Fig. 3 and Fig. 4 note that: the error plots of ACF and PACF, it is clear that there are no lags out of ranges. So the model is the best time series model for representing the data.

D. Testing Neural Network Model

The proposed feed forward neural network, which we call PROPOSEDFF, is designed using the MATLAB neural network toolbox. The PROPOSEDFF creates a feed-forward back-propagation network. It requires three inputs: population, Wheat area, and Wheat production and returns one output: Wheat import. See Fig. 5.


The first argument is a matrix \([0 1]\) of minimum and maximum values for the input vector. The second argument is an array \([10 5 3 1]\) containing the sizes of each layer. The third argument is a cell array,\{‘tansig’ ‘logsig’ ‘tansig’ ‘purelin’\} containing the names of the transfer functions to be used in each layer. The fourth argument contains ‘trainrp’ the name of the training function to be used. The fifth argument ‘learngd’ is back-propagation weight/bias learning function. The sixth argument ‘mse’ is the performance functions. And the final input ‘nn’ the number of inputs.
V. RESULTS

The researchers build four models: regression model, which has R-Sq = 20.1%, that due to the non-linearity in the problem, second model is ARIMA(0,0,0)(2,1,2) t model, the third model is PROPOSEDF, which designed using the MATLAB neural network toolbox, and the fourth model is the proposed multi-factors fuzzy time series model (MFFTS). The result of the forecasting accuracy for the models: Regression, ARIMA, PROPOSEDF, and MFFTS model compared with the actual data by using NRMSE performance indices illustrates in Fig. 6.

It is clear that the multivariate fuzzy time series proposed model provides higher accuracy in forecasting than ARIMA Model, Regression and PROPOSEDF models, so the proposed model can lead to satisfactory high performance for fuzzy time series.

VI. CONCLUSION

This paper presents a new Multivariate-factors fuzzy time series model (MFFTS). The proposed model has been used successfully to forecast Egypt Wheat imports as consequent factor depending on: population size, Wheat area, and Wheat production as antecedent’s factors. The experimental test is based on the official data provided by the Central Agency for Public Mobilization and Statistics (CAPMAS) starting from 1986 up to 2008. The Multivariate-factors proposed model has been evaluated through a comparison with Regression, ARIMA, and neural network model. The result of the comparison ensures the superiority of the proposed model over the other three models in terms of lowest NRMSE.

REFERENCES