# Comparison Of Fuzzy Time Series And ARIMA Model

K. Senthamarai Kannan, M. SulaigaBeevi, S. Syed Ali Fathima

**Abstract**—Crude Oil price, deregulated commodity, which plays a vital criterion in the global economy. Government of India give permission to Oil Companies to revise the price of fuel daily based on the change of international crude oil price and Dollar currency exchange rate. Forecasting is the one of the essential tool to predict the future environment of the fuel price. This paper collates the applications of two Forecasting models such as Auto Regressive Moving Average (ARIMA) model and Fuzzy time series model, on petrol price prediction. The error values, Root Mean Square Error (RMSE), Mean Square Error (MSE) and Mean Absolute Error (MAE) are calculated numerically and graphically for the forecasted values.

**Index Terms** ACF, ARIMA, Fuzzy time series, MAE, MSE,PACF, petrol price

## 1 INTRODUCTION

India is the fourth largest consumer of oil and also imports nearly two-thirds of crude oil requirements. Fluctuation of oil prices will definitely affect the economy. Lower oil prices will reduce the import bill by Rs. 4000 crore per every decrease in dollar, save foreign exchange, keep the inflation low and also reduce the fuel subsidies. It is estimated that a $ 10 fall in crude could reduce the current account and the fiscal deficit by approximately 0.5% and 0.1% respectively of GDP. Hence, the government allow the Oil Companies to revise the fuel price monthly twice based on the changes in the international crude oil price and Dollar currency exchange rate since 2002. The system is dropped in June 15, 2017 and it leads the changes in the price of petrol on daily basis.

Planning is essential for the country, business firm and every individual. Time series is a set of observation taken at specified times, at equal intervals. Time series relation between "Cause" and "Effects". One variable "Time", which is independent variable and the "Data", the dependent variable.


- *Dr. K. Senthamarai Kannan, Senior Professor, Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli-627012. senkannan2002@gmail.com*
- *M. SulaigaBeevi, Research Scholar, Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli-627012. ssulaiga@gmail.com*
- *Dr. S. Syed Ali Fathima, Assistant Professor, Department of Mathematics, Sadakathullah Appa College, Tirunelveli - 627011. syedalifathima2014@gmail.com*


## 2. LITERATURE SURVEY

Fuzzy logic was introduced by L.A.Zadeh in 1965 defined as set of elements with varying degree of membership function μn: U→[0,1].In 1993 and1996 many researchers used max-min composition in their forecasting model. Song and Chissom[4]proposedmamadani's method to compute fuzzy relation and compare the difference between the traditional time series and fuzzy time series. Song and Chissom[5]divide universe of discourse in to 7 same lengths of intervals and used max-min composition as defuzzification of the output. Further the researcher in [1]Chen used new model of max-min composition technique with new fewer complexes then used max-min composition. Most of the researcher used fixed length of intervals and Chen's approach in between year 1996 to 2000 after that various hybrid model was used. Wang [6]discussed a comparison method of forecasting using fuzzy time series and ARIMA model. For this purpose, they have used the data related to the Taiwan export data. Yun – Sheng Hsu et al., [7] discussed traditional methods of forecasting and heuristics model of forecasting and then compared the two methods. Dani and Sarma[3] forecasted using first order time variant method for this purpose total monsoon rainfall in 18 yrs of one of the region of Chhattisgarh.Sarma and Chouhan[2]analyzed a traditional method ARIMA comparing Neural network fuzzy time series model.

## 3.MATERIALS AND METHODS

### 3.1 DataSource

The Daily price changes of Petrol price data for the period of 16 June 2017 to 4 December 2017 obtained from the Indian Oil Corporation Ltd.

The basic statistics for all the data are given in Table 1 as follows

TABLE1
DESCRIPTIVE STATISTICS

| Statistics | Petrol Price |
|---|---|
| Mean | 70.21 |
| S.D | 2.48764 |
| Maximum | 73.48 |
| Minimum | 65.44 |
| Skewness | -0.654 |
| Kurtosis | -1.068 |

### 3.2 Time Series Forecasting Models:

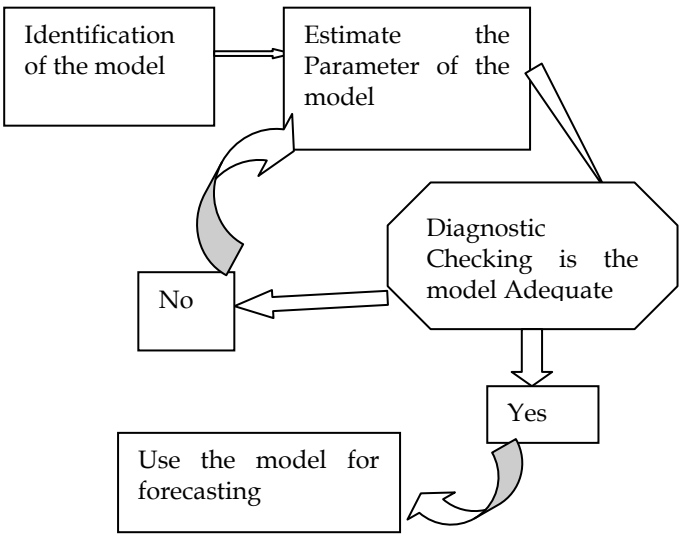
Forecasting is a significant part of Time series. In our paper we are going to discuss the two methods of forecasting which have become generally popular of time series.

- Auto Regressive (AR) Integrated (I) Moving Average (MA) ARIMA generally known as the Box-Jenkins methodology.
- Fuzzy Time series.

### 3.2.1ARIMA Model

Time Series ARIMA model was planned by Box-Jenkins in 1970, The past data of economic fluctuations are observed by the model which examines every variable using auto regression, AR(p) and MA(q).

**The algorithm is to be had as follows:**




**Step1 (Identification of the model):**

To determine, whether the variable is stationary or not. This can be done with the correlogram. If it is not stationary it needs to be First Differenced.

**Step2 (Estimation):**

(i) To determine the (p) and (q) in the ARIMA (p, d, q) model.

(ii) How many times the data needs to be differenced to produced a stationary series.

**Step3 (Diagnostic Checking):**

In diagnostic checking, the assumption of the model was tested so that any areas in the model is insufficient for this purpose, whether the parameter achieves statistical significance or multicollinearity and whether the residual term is white noise or not which was included in the statistical identification and repeat step 3 until a best model is identified.

**Step4 (Forecasting):**

(i)Model is stationary

(ii) Identify the (p, d, q)

(iii) Identify the sufficient (p, d,q). Then go to forecasting.

### 3.2.2 Fuzzy Time Series Model

The main feature of time series analysis, which distinguishes it from other statistical analysis, is the importance of the order in other problems the study is statistically independent but they are dependent in time series consecutive observation.

Time series analysis comprises methods for analyzing time series data in order to extract statistics and characteristics of the data which have natural temporal ordering. This model is to predict future values based on previously observed values.

In time series analysis the main objective is to understand and change in the aspect in the hope of more properly expecting the course of future events. A good compassionate of the mechanism generating the series may also helps us to control the aspect involved in it. In this manner generating mechanism control the future behavior of the process.

The time-variant and time-invariant fuzzy time series definitions are given below.

**Definition 1:**

Let Y(t), a subset of real numbers, be the universe of discourse on which fuzzy sets f(t) are defined. If F(t) is a collection of f1, f2,…. then F(t) is called a fuzzy time series defined on Y(t).

**Definition 2:**

Suppose F(t) is implied by F(t-1) only, that is F(t-1) F(t). Then this relation can be expressed as

F(t) = F(t-1) * R(t-1), where R(t-1) is the fuzzy relationship between F(t-1) and F(t) and is called the first order model of F(t). Here, "*" is maximum–minimum composition operator. The relation R is called first-order model of F(t).

**Definition3:**

Suppose F(t) is calculated by F(t-1) only, and F(t) = F(t-1) R (t-1, t). For any t, if R (t-1, t) is independent of t, then F(t) is considered a time-invariant fuzzy time series. Otherwise, F(t) is time-variant.

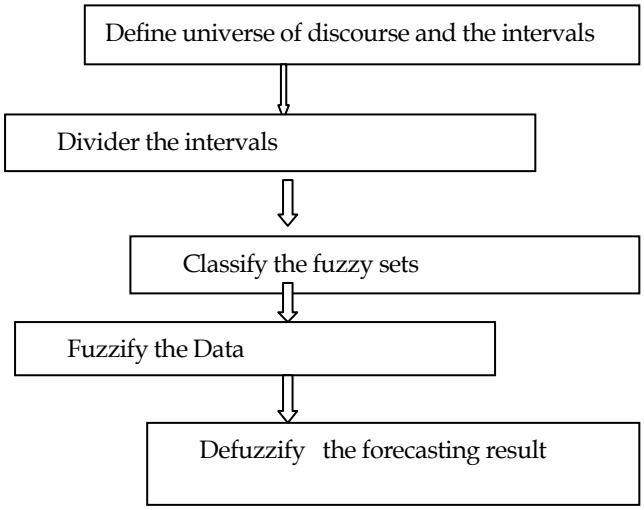
[3]

Fuzzy Time series is assumed to be a fuzzy variable along with associated member ship function Song and Chissom (1993) have proposed a procedure for solving fuzzy time series model described as follows.

**The analysis of time series of observation consists of the followings**:

Membership function μ, we can define a fuzzy set F on a universe of discourse U as μF(x):U→ [0,1]which is nothing but a mapping from the universe of discourse U into the unit interval [0,1] and μF(x) represent the scope to which x containing F. The notion of membership functions allows any element within the universe of discourse to have partial membership to a specific fuzzy set and also to have partial membership to other fuzzy sets Song and Chissom applied time invariant and time variant models to forecast the enrollment at the University of Alabama.

**The time variant and invariant model includes the following step**s:




**Step1 (Define universe of Discourse and the intervals):**

Universe of Discourse U, based on ranges available data in the historical time series data, by the following rule: U= [Dmax-D1,Dmin-D2] where D1 and D2 are two suitable positive numbers.

**Step2 (Divider the Intervals):**

Divider the universe of Discourse into equal length of intervals u1, u2……un.the number of intervals will be in the unity with the number of linguistic variable (Fuzzy sets)A1, A2,……An to be consider. If we notice the relation the fact that predict with fuzzy time series which have the least average error. We have found the middle point of the intervals.

**Step3 (Classify the fuzzy set):**

Classify the fuzzy set into linguistic variablesA1=very poor, A2=poor, A3=good, A4=very good and so on. Every linguistic value corresponds a fuzzy variable which is assigned beside a correlating fuzzy set determinative the significance of this variable. If the value of variable U in method is accepted as the middle point of the corresponding interval of fuzzy set defined as,

Ai= (μAi (ui)/ui) ui∈U μAi(ui) ∈ [0,1] is a fuzzy set.

**Step4 (Fuzzify the data):**

Fuzzify the historical data and establish the fuzzy logical relationships by the following rule: If Ai is the fuzzy price of the day n and Aj the fuzzy price of the day n + 1, then the fuzzy logical relation is denoted as Ai→ Aj, here Ai is called current state and Aj is called next state.

**Step5 (Defuzzify the forecasting results):**

- First finding mean of the data
- Find the sum of the data values and divide the sum by the number of data values
- Find the absolute value of the difference each data value and the mean
- Find the sum of absolute value of the difference

## 4. RESULTS AND DISCUSSIONS

### 4.1 Fuzzy Forecasting Analysis

The implementation of the above algorithm for the daily petrol revision on 172 data were collected on 16 Jun 2017 to 4 Dec 2017.

**Step 1 (Define universe of Discourse and the intervals):**

Let U= [65-74]is the universe of discourse of the petrol price and the length of the equal intervals is 9.

**Step 2(Divider the interval):**

Partition the universe of discourse in to nine equal length intervals u1, u2. . . u9, where

u1=[65-66],u2=[66-67],u3=[67-68],u4 = [68-69],u5=[69-70],u6=[70-71],u7=[71-72],

u8= [72-73], u9= [73-74].

It is essential to find the middle point of the intervals that forecasts with fuzzy time series which exhibits the least average error.

Step 3(Classify the fuzzy set):

Fuzzy set are the universal set U. In this case "the difference in total petrol price" is a linguistic variable that assumes the subsequent linguistic values: A1=(To much low level price), A2 =(To low level price), A3 =( moderate low level price), A4=( low level), A5=(medium level), A6= (high level ), A7= (moderate high level ), A8=(To high level ), A9=( To much high level ), It is essential to find the middle point of the intervals that forecasts with fuzzy time series which exhibits the least average error. For each linguistic value it correspond a fuzzy variable when it is assigned against a corresponding fuzzy set determinative the meaning of the variables.

The value of variable U in formula is accepted as the middle point than the corresponding interval of fuzzy set is as follows:

Ai(i=1...9) will be defined as follows: Ai=(μAi (ui )/ui)ui ƐU μAi(ui)Ɛ [0,1] is a fuzzy set.

Step 4(Fuzzify the data):

If Ai is the fuzzy price of the day n and Aj the fuzzy price of the day n + 1, then the fuzzy logical relation is denoted as Ai→ Aj, here Ai is called current state and Aj is called next state.

Step5 (Defuzzify the forecasting results):

Results (fuzzy forecasted variations) of the previous step are summarized to obtain crisp integer value (forecasted petrol and diesel price) for each day under consideration. This process is known as defuzzification.

### 4.2. ARIMA Forecasting Analysis

Box-Jenkins Analysis (1970) refers to systemic method of identifying fitting, checking and using ARIMA Time series model the method of appropriate for time series of medium to long length.LetYtis a discrete time series variable which takes different values over a period of time. The corresponding AR(p)model of Yt series, which is the generalizations of autoregressive model can be expressed

$$Y_t = c + b_0 + b_1 Y_{t-1} + b_2 Y_{t-2} + \cdots + b_p Y_{t-p} + e_t \quad (1)$$

where c is a constant and et is white noise. This is like a multiple regression but with lagged values of Yt as predictors. We refer to this as an AR(p) model. Autoregressive models are remarkably flexible at handling a wide range of different time series patterns. Rather than use past values of the forecast variable in a regression, a moving average model uses past forecast errors in a regression-like model.

$$Y_t = c + e_t - \varphi_1 e_{t-1} - \varphi_2 e_{t-2} - \cdots - \varphi_q e_{t-q} \quad (2)$$

where, cis the constant mean of the seriesφ1, φ2, . . . φq is the coefficients of the estimated error term; $e_t$is the error term. Combining both the model is called as ARIMA models, which has general form as:

$$Y_t = c + b_1 Y_{t-1} + b_2 Y_{t-2} \cdots + b_p Y_{t-p} + e_t - \varphi_1 e_{t-1} - \varphi_2 e_{t-2} - \cdots - \varphi_q e_{t-q} \quad (3)$$

If Yt is stationary at level or d (0) or at first differenced (1) determines the order of integration, which is called asARIMA model.

P, d,q are the level for each of AR, I and MA parts.

**Step1 (Identification of the model):**

To determine if the variable is stationary. Our data has non stationary. This can be done with the correlogram Figure (1) and (2) tells that so it is not stationary it needs to be First Differenced.

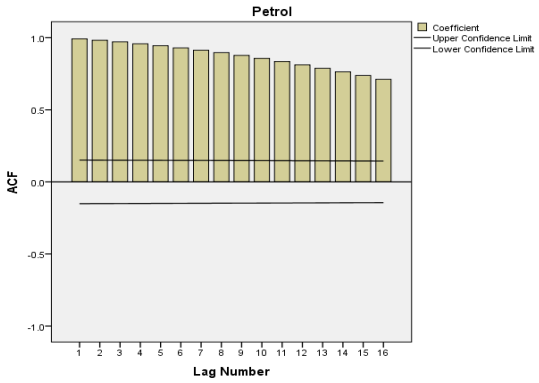



Figure 1: ACF for Petrol Price

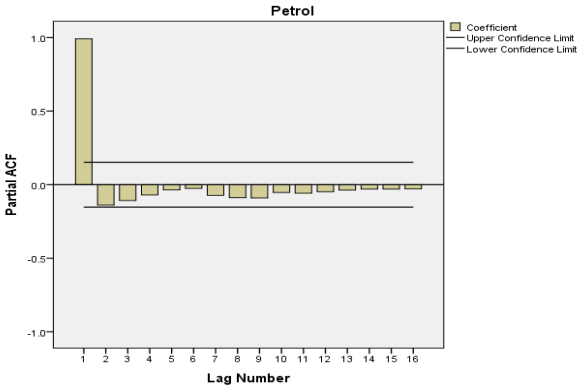



Figure 2: PACF for Petrol Price

**Step2 (Estimation):**

To determine the (p) and (q) in the ARIMA (P, I, Q) model: Our data needs to be differenced to produce a stationary series. First order produced stationary series.so our model is (0,1,0).

**Step3 (Diagnostic Checking):**

The statistical identification method includes whether the parameter achieves statistical significance. (0, 1, 1) is the parameter of achieving significance.

**Step4 (Forecasting):**

The accuracy of the fitted model is measure by MAE, MSE, RMSE, which is a measure of accuracy of the fitted model.

Based on the above two models the predicted values are compared numerically and graphically. For all measures, smaller values generally indicate a better fitting model for fore casting errors shows in table (2).

$$MAE = \sum_{t=1}^{n} \frac{(Forecasted\ \ Value - Actual\ \ Value)}{n} \quad (4)$$

$$MSE = \sum_{t=1}^{n} \frac{(Forecasted\ \ Value - Actual\ \ Value)^2}{n} \quad (5)$$

$$RMSE = \sqrt{MSE} \quad (6)$$

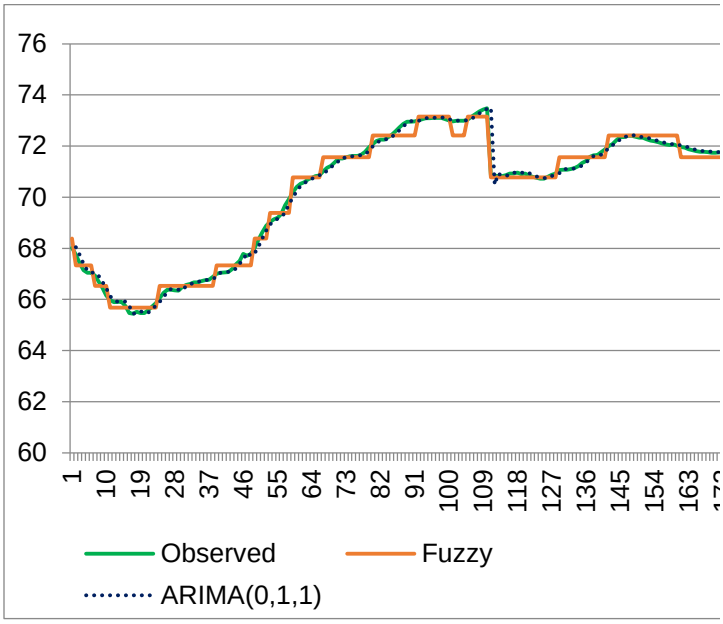



Figure 3: Comparison of Petrol Price

## 5. CONCLUSIONS

Parameters are estimated for ARIMA (0, 1, 1) and error values are obtained both numerically and graphically using the ARIMA and Fuzzy time series methods. The methodology used in time series in Box Jenkins methodology it can forecast the average error. We can find the average error using the formula in fuzzy time series .Fig (3) shows that comparison of actual model, ARIMA Model and Fuzzy Time series models. It was observed that error values of fuzzy time series is less than the ARIMA time series error values, which is shown in Table 2. So it is conclude that fuzzy time series is better result than other time series model and also its appropriate for predict fuel price in India.

## REFERENCES


- S. M. Chen, "Forecasting enrollments based on fuzzy time series," *Fuzzy Sets and Systems*, vol. 81, no. 3, pp. 311 – 319, 1996..
- S. Sharma and M. Chouhan,"A Review: Fuzzy Time Series Model for Forecasting,"*International Journal of Advances in Science and Technology*, vol. 2, no 3, pp. 32-35, sep2014..
- S. Dani and S. Sarma ,"Forecasting Rainfall of a Region by using Fuzzy Time Series," *Asian Journal of Mathematics and Applications*, pp. 1-10,2013 ..
- Q. Song, and B. S. Chissom,"Fuzzy time series and its models," *Fuzzy Sets andSystems*, vol. 54,no 3, pp. 269-277 , mar(1993)..
- Q. Song, and B. S. Chissom ,"Forecasting enrollments with fuzzy time series- Part II, " *Fuzzy Sets and Systems*, vol. 62, no l, pp. 1-8, 1994.
- C. C. Wang, "A Comparison study between fuzzy time series model and ARIMA Model for Forecasting Taiwan Export,"*Expert Systems with Applications*, vol. 38, pp. 9296–9304 (2011a).
- Yun-sheng Hsu et.al, "A Comparison of ARIMA fore casting and heuristic modeling, "Applied *Financial Economics*, vol. 21, pp1095-1102(2011b.).