Punjab University Journal of Mathematics
(ISSN 1016-2526)
Vol. 52(1)(2020) PP 1-15

Improved Fuzzy Forecasting Model for Stock Exchange Market

Saima Mustafa, Sophia Siddique
Department of Mathematics and Statistics,
PMAS-Arid Agriculture University Rawalpindi, Pakistan
Email: saimamustafa28@gmail.com, Email: sophiasiddique92@gmail.com

Cem Kadilar
Department of Statistics,
Hacettepe University, Beytepe, Ankara, Turkey
Email: kadilar@hacettepe.edu.tr
Received: 30 May, 2019 / Accepted: 25 October, 2019 / Published online: 01January, 2020

Abstract. Stock price forecasting, which is an important topic in finance and economics, has prodded the enthusiasm of specialists throughout the years to develop models for better forecasts. Stock market is a key factor of monetary markets and signs of economic growth. Therefore, differ-ent forecasting models have been explored in literature for the stock data prediction. In this article, classical ARIMA model is mentioned and then Fuzzy Auto Regressive Integrated Moving Average (FARIMA) model is proposed as a new model. These two models are used for forecasting the stock exchange market of Attock cement Pakistan limited. Due to vague natures of stock data and parameter, fuzzy least square method is used in proposed FARIMA model. FARIMA model is based on the possibil-ity of success. These possibilities are defined by linguistic term, such as very low, low, average, high, and very high. This model makes it possible for decision makers to forecast the best based on fewer observations than the ARIMA model. Finally, comparison between the proposed model and ARIMA shows that the proposed model has better performance than ARIMA by using different criteria, such as mean squared error, mean ab-solute percentage error, and mean absolute deviation.

Key Words: ARIMA model, FARIMA model, Stock exchange market.

1. Introduction

Forecasting is a technique for the estimation or prediction of the future trend using the past and present information. Forecasting gives upcoming pattern about future occasions and their consequences to the organization so that they must prepare themselves or make the strategies according to those results.

In international economics forecasting, the conduct of ostensible trade rates and stock venture has been a central subject in business analysts' work while the Stock business sector contributes a genuine and testing fiscal movement. Stock trade is the field of the capital business division. The economy of a country depends upon capital business division, forecasting the stock costs and their patterns in accomplishing huge increase in money related markets. The stock cost has profound impact in the financial event of the country and the large-scale economic approach. Stock value variances depict the states economy change to a specific degree influencing the national macroeconomic arrangement. However, predicting the stock market prices and movements is not an easy task because of the critical impact of full scale financial variables, including political occasions, general monetary conditions, speculators' desires, financial specialists, decisions, sudden circumstances in security exchanges, and speculators brain research.

Enormous profits can be earned when a good forecasting model gives precise expectations, but violent fluctuations sometimes caused by economic proxy factors, such as fiscal levy, price level, war disaster, make forecasting challengeable in stock market activity. It is essential for building a more exact forecasting model or giving a superior model to business sector that can empower the speculators to foresee the costs ahead of time; it would also help the financial specialists in keeping independent national economy.

In the previous era, various models and techniques have been created for stock value forecast. Among them quantitative methodologies are utilized for estimating the investigation of chronicled information by utilizing factual standards and ideas. Quantitative methodology consists of Time Series Strategies,such as Autoregressive Integrated Moving Average, Autoregressive Conditional Heteroscedasticity (ARCH), and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) using the previous information of the variable for estimating its future qualities.

In view of the time series methodology, BoxJenkins [3] presented an AutoRegressive Integrated Moving Average (ARIMA) model that depicts homogeneous non-stationary process. It is also stated as Box-Jenkins methodology made out of set of events for identifying the values, evaluating the unknown parameter and diagnostic testing whether the model is fitting the information well. This model is utmost important in economic technique for forecasting the future trends in the most important fields.

Box-Jenkins models have the benefit of precise forecasting in a small period; it likewise has the impediment that not less than 50 and ideally 100 observations ought to be utilized. Due to this, Autoregressive Integrated Moving Average has limitation that large amount of data is required to run the model.

In modern era, lots of developments have occurred due to the modern technology and uncertain factors in environment. Owing to this uncertain situation, we generally should estimate future circumstances utilizing little information in a short period. In this scenario, autoregressive integration moving average model becomes less applicable and cannot work for short term forecasting. In addition to this, we sometimes have to deal with uncertain observations that cannot be handled by simple quantitative concepts so in this context, fuzzy is used to deal with this impreciseness and uncertain environment.

Tanaka et al [5] used the fuzzy regression to deal with fuzzy uncertain environment and elude the error of model. Song and Chissom [8] proposed the method related to fuzzy time series for demonstrating the forecasting. They also examined the robustness of the method.

This article is grounded upon the works of quantitative models of Box Jenkins ARIMA models - and proposed fuzzy model. We have joined the points of interest of two strategies to build up the proposed fuzzy ARIMA model .To compare the worth of proposed model fuzzy ARIMA with ARIMA model, we conduct a delineation for forecasting the stock exchange of Attock cement Industry Pakistan limited. In the outcomes, we found that the proposed model fuzzy autoregressive integrated moving average is appropriate and better than autoregressive integrated moving model in forecasting.

This article is structured into the following sections: Concepts of ARIMA are mentioned in Section 2. In Section 3, the proposed FARIMA model is introduced and examined. The ARIMA and Proposed FARIMA model are applied to forecast the stock exchange of

Attock Cement Pakistan in Section 4. Finally, the comparison of both models and results are discussed.

2. REVIEW OF ARIMA MODEL

The Autoregressive Integrated Moving Average model of Box Jenkins methodology is represented by

$$(1-\sum_{r=1}^{\rho}\tau_r B^r)(1-B)^d z_t = (1-\sum_{r=1}^{q} {}_r B^r)\varepsilon_t \, , \tag{2. 1}$$

where $(1-\sum_{r=1}^{\rho}\tau_r B^r)$ is the autoregressive process consists of $\tau_1, \tau_2, \tau_3, ..., \tau_\rho$ known as autocorrelation parameters, $(1-\sum_{r=1}^{q} {}_r B^r)\varepsilon_t$ is the moving average process consists of co-efficients $({}_1, {}_2, {}_3, ..., {}_q)$ are known as partial autocorrelation function with order q. Both processes contain B polynomial represented as a back shift operator. The term of $(1-B)^d z_t$ represents the part of the stationary series with the order of d ;which means the number of times the series is integrated to make it stationary.
The Box Jenkins Methodology consists of the following phases:

i. Phase 1
   Phase 1 is based on the model identification to select the appropriate model.
ii. Phase 2
   In the phase 2, parameters of the selected model are estimated.
iii. Phase 3
   Phase 3 comprises of diagnosis testing of the selected model.
iv. Phase 4
   In this phase, forecasts are obtained from the selected model.

It is assumed that ε_t is the error term as a white noise series.

3. PROPOSED FUZZY AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL

Proposed Fuzzy Auto-Regressive Integrated Moving Average Model is given with fuzzy parameters as follows:

$$\tilde{z}_t = \tilde{\kappa}_1 z_{t-1} + \tilde{\kappa}_2 z_{t-2} + \Lambda + \tilde{\kappa}_\rho z_{t-\rho} + \varepsilon_t - \tilde{\kappa}_{\rho+1}\varepsilon_{t-1} - ... - \tilde{\kappa}_{\rho+\hbar}\varepsilon_{t-\hbar} \tag{3.1}$$

In (3.1), $\tilde{z}_t$ is the estimated fuzzy variable as an output variable,$(\tilde{\kappa}_1, \tilde{\kappa}_2, \tilde{\kappa}_\rho, \tilde{\kappa}_{\rho+1}, ..., \tilde{\kappa}_{\rho+\hbar})$ are the parameters with ρ terms known as fuzzy Autocorrelation parameters and fuzzy parameters with $\hbar$ terms known as fuzzy partial autocorrelations. By explaining the time series data, they have the cause and effect associations among the observed variables. Therefore, it is difficult to measure the degree of dependency that involved different factors. Due to this, the impreciseness and conciseness have been occurred in parameters. This impreciseness has been tackled by connecting parameters with ρ and $\hbar$ orders into fuzzy parameters.

3.1. **Construction of the proposed model.** In order to estimate the fuzzy parameters of the proposed FARIMA model, fuzzy least square method is used. Fuzzy least square method is used to determine the distance between the fuzzy value which is estimated from the model and observed data. Fuzzy least square approach is represented as:

$$d(\beta_0, \beta_1) = \left[\int_0^1 f(\xi) d^2 \Big\{ \{\beta_0, \beta_1\} \xi \Big\} d\xi \right]^{\frac{1}{2}} , \tag{3.2}$$

where β_0 and β_1 are two fuzzy numbers. Here,β_0 is the trapezoidal fuzzy number with four points such that $\beta_0 = \{b_{0m}, b_{0u}, b_{0l}, b_{0v}\}$ and β_1 is another trapezoidal fuzzy number with four points such that $\beta_1 = \{b_{1m}, b_{1u}, b_{1l}, b_{1v}\}$.For both fuzzy numbers β_0 and β_1 the parameter of b_m represents the left fuzzy point, b_u represents the right center fuzzy point, b_l represents the left center fuzzy point and b_v represents the right fuzzy point of trapezoidal membership function.

In (3.2), $f(\xi)$ is used as the weighting function for determining the square of distance between two fuzzy numbers that prompt putting more significance on higher degree of the membership function.

3.2. **Parameter estimation.** The parameters of the model (3.1) are estimated by using the least square method to gain the unique solution, whereas the least square method is defined by as the sum of squared errors distance between observed values denoted as $V_t(\xi)$ and estimated output denoted as $S_t(\xi)$. Mathematically represented as

$$SSE = \sum_{i=1}^{\rho} d\big[V_t(\xi), S_t(\xi)\big] , \tag{3.3}$$

where index t denotes the non-fuzzy time series data used in $V_t(\xi)$ and $S_t(\xi)$. Here, $S_t(\xi) = \{f(b_m), f(b_u), f(b_l), f(b_v)\}$.

From (3.2), each fuzzy parameter can be converted to the autoregressive integrated moving average in the form of functions shown as:

$$f(b_m) = \varepsilon_t + \sum_{i=1}^{\rho} b_m z_{t-i} - \sum_{j=1}^{\hbar} b_m \varepsilon_{t-j} , \tag{3.4}$$

$$f(b_u) = \varepsilon_t + \sum_{i=1}^{\rho} b_u z_{t-i} - \sum_{j=1}^{\hbar} b_u \varepsilon_{t-j} , \tag{3.5}$$

$$f(b_l) = \varepsilon_t + \sum_{i=1}^{\rho} b_l z_{t-i} - \sum_{j=1}^{\hbar} b_l \varepsilon_{t-j} , \tag{3.6}$$

$$f(b_v) = \varepsilon_t + \sum_{i=1}^{\rho} b_v z_{t-i} - \sum_{j=1}^{\hbar} b_v \varepsilon_{t-j} , \tag{3.7}$$

Now, $S_t(\xi)$ with zeta-cut interval for trapezoidal number can be represented as:

$$S_t(\xi) = \big[\{f(b_u) - f(b_m)\}\xi + f(b_m), f(b_v) - \xi\{f(b_v) - f(b_l)\}\big] . \tag{3.8}$$

In the same way, observed value is $V(\xi)$ expressed on $S_t = [W_1, W_2]$ such as

$$W_1 = \{f(b_u) - f(b_m)\}\xi + f(b_m),\ W_2 = f(b_v) - \xi\{f(b_v) - f(b_l)\}$$

where W_1 and W_2 represent the lower bound and upper bound, respectively. Note that observed value is equal to $V_t(\xi) = [W_1, W_2]$.

Consider (3.8) and weighting functions in sum of squared error equation(SSE) given by

$$SSE=\sum_{i=1}^{\rho}\int_{0}^{1}f(\xi)\Big[W_1-\big\{\big[f(b_u)-f(b_m)\big]\xi+f(b_m)\big\}\Big]^2+\Big[W_2-\big\{f(b_v)+\xi\big[f(b_v)-f(b_l)\big]\big\}\Big]^2d\xi\ . \tag{3.9}$$

Using (3.9)for finding the partial derivation with respect to b_m, b_u, b_l and b_v, we obtain the simplified form of equations by putting the condition $f(\xi)=\xi$, $\big|z_{t-i}-\varepsilon_{t-j}\big|=y_{ij}$ as follows:

$$\sum_{i=1}^{\rho}2\int_{0}^{1}\xi(\xi-1)y_{ij}\big[W_1-\big\{f(b_u)-f(b_m)\big\}\xi+f(b_m)\big]d\xi=0\ , \tag{3.10}$$

$$\sum_{i=1}^{\rho}2\int_{0}^{1}\xi^2y_{ij}\big[-W_1+\big\{f(b_u)-f(b_m)\big\}\xi-f(b_m)\big]d\xi=0\ , \tag{3.11}$$

$$\sum_{i=1}^{\rho}2\int_{0}^{1}\xi y_{ij}\big[-W_2+f(b_v)-\xi\big\{f(b_v)-f(b_l)\big\}\big]d\xi=0\ , \tag{3.12}$$

$$\sum_{i=1}^{\rho}2\int_{0}^{1}\xi(\xi-1)y_{ij}\big[W_2-f(b_v)+\xi\big\{f(b_v)-f(b_l)\big\}\big]d\xi=0\ , \tag{3.13}$$

Solving the integrals and replacing the values in (3.10) - (3.13), we obtain the following equations:

$$b_{m_0}\sum_{i=1}^{\rho}y_{i0}y_{ij}+b_{m_1}\sum_{i=1}^{\rho}y_{i1}y_{ij}+...+b_{m_\hbar}\sum_{i=1}^{\rho}y_{i\hbar}y_{ij}=\sum_{i=1}^{\rho}p_iy_{ij}\ . \tag{3.14}$$

$$b_{u_0}\sum_{i=1}^{\rho}y_{i0}y_{ij}+b_{u_1}\sum_{i=1}^{\rho}y_{i1}y_{ij}+...+b_{u_\hbar}\sum_{i=1}^{\rho}y_{i\hbar}y_{ij}=\sum_{i=1}^{\rho}q_iy_{ij}\ . \tag{3.15}$$

$$b_{l_0}\sum_{i=1}^{\rho}y_{i0}y_{ij}+b_{l_1}\sum_{i=1}^{\rho}y_{i1}y_{ij}+...+b_{l_\hbar}\sum_{i=1}^{\rho}y_{i\hbar}y_{ij}=\sum_{i=1}^{\rho}r_iy_{ij}\ . \tag{3.16}$$

$$b_{v_0}\sum_{i=1}^{\rho}y_{i0}y_{ij}+b_{v_1}\sum_{i=1}^{\rho}y_{i1}y_{ij}+...+b_{v_\hbar}\sum_{i=1}^{\rho}y_{i\hbar}y_{ij}=\sum_{i=1}^{\rho}k_iy_{ij}\ . \tag{3.17}$$

where $y_{i0}=1$ and $j=0,1,2,...,\hbar$. Representing these equations in the matrix form:

$$Bm=P,\ Bu=Q,\ Bl=R,\ Bv=K\ ,$$

where

$$m=\big\{b_{m_0},b_{m_1},...,b_{m_n}\big\}^T,\ P=(\sum_{i=1}^{\rho}P_{t-1}y_{t-1}+...)^T$$

$$u=\big\{b_{u_0},b_{u_1},...,b_{u_n}\big\}^T,\ Q=(\sum_{i=1}^{\rho}Q_{t-1}y_{t-1}+...)^T$$

$$l=\big\{b_{l_0},b_{l_1},...,b_{l_n}\big\}^T,\ R=(\sum_{i=1}^{\rho}R_{t-1}y_{t-1}+...)^T$$

$$v = \left\{b_{v_0}, b_{v_1}, ..., b_{v_n}\right\}^T,\ K = (\sum_{i=1}^{\rho} K_{t-1}y_{t-1} + ...)^T$$

Here, $B = Y^TY$ and $\begin{bmatrix} 1 & \cdots & y_{1_\hbar} \\ \vdots & \ddots & \vdots \\ 1 & \cdots & y_{p_\hbar} \end{bmatrix}$ and matrix B is the positive definite with the rank of $(n+1)$. Note that if matrix $B = Y^TY$, then inverse of matrix B, which can be easily determined, consists of unique solution. Then, problem can be written as follows:

$$m = B^{-1}P,\ u = B^{-1}Q,\ l = B^{-1}R,\ v = B^{-1}K$$

3.3. **Construction of fuzzy membership function.** Constructing the membership function plays an important role in the fuzzy method.Various techniques are used to create the fuzzy membership functions such as observations investigations and functions. Fuzzy membership function was first presented by Zadeh [7] based on fuzzy number. In this study, trapezoidal function is used for creating the fuzzy membership function.

Trapezoidal fuzzy membership function consists of four parameters. These four parameters determine the figure of membership function giving the four corners point of z coordinate. Trapezoidal fuzzy membership function is described by the mathematical formula given by

$$Trapezoidal(z; m, u, l, v) = \begin{cases} 1 - \frac{m-z}{u-m}, & m \leq z \leq u \\ 1, & u \leq z \leq l \\ 1 - \frac{z-v}{v-l}, & l \leq z \leq v \end{cases}$$

where m represents the left point parameter, u represents the left center point parameter, l represents the right center point parameter, v represents the right point parameter.

3.4. **Construction of linguistic categories of Zeta interval.** In this study, five categories of linguistic term are constructed. These categories consist of following steps:

Step 1: Arrange the observations in data set in the ascending order.

Step 2: Calculating the average distance and standard deviation of the arranged data. The average distance is given by

$$AverageDistance(AD) = \frac{1}{t-1}\sum_{r=1}^{t-1}\left|y_{p(r)} - y_{p(r+1)}\right|,$$

where $y_{p(r)} \leq y_{p(r+1)}$. Here, p is permutation. The standard deviation of average distance is calculated as follows:

$$\sigma_{AD} = \sqrt{\frac{1}{t}\sum_{r=1}^{t-1}(y_r - AD)^2}\,.$$

Step 3: Outliers are eliminated by using the following condition:

$$AD - \sigma_{AD} \leq y_r \leq AD + \sigma_{AD}\ ,$$

where the observations, which are greater and equal to the left side condition and less and equal to the right side condition, are included; whereas, the observations, which do not satisfy the conditions, are outliers and outliers are not included in dataset.

Step 4: Revised average distance is computed from the remaining sorted dataset.

Step 5: The universe of discourse is defined as follows:

$$Universe of discourse = \left[lower\ bound\ , upper\ bound\right]\ ,$$

where $Lower\ bound = \left[y_{min} - AD_R\right]$ and $Upper\ bound = \left[y_{max} - AD_R\right]$, y_{min} and y_{max} are the minimum and maximum values in the dataset.

Step 6: Construct the fuzzification interval approach as follows:

$$u_B = \begin{cases} \frac{y-b_1}{b_2-b_1}, & b_1 \le y \le b_2 \\ 1, & b_2 \le y \le b_3 \\ \frac{b_4-y}{b_4-b_3}, & b_3 \le y \le b_4 \\ 0, & otherwise\ . \end{cases}$$

where b_1 is equal to AD_R, b_2 is equal to $(b_1 + AD_R)$, b_3 is equal to $(b_2 + AD_R)$ and b_4 is equal to $(b_3 + AD_R)$,

$u_\xi = \left[b_1 - AD_R(1-\xi),\ b_4 + AD_R(1-\xi)\right]$. In interval formulation, first trapezoidal consists of b_1, b_2, b_3, b_4. They are used in (3.10) - (3.13) to estimate the lower and upper bounds.

4. Application

In this section, ARIMA and Proposed FARIMA are performed to forecast the Attock Cement Pakistan Limited Stock Series.

4.1. **ARIMA model.**

Attock Cement Pakistan Limited Stock Data used in this article covers the period of the years from 2005 to 2012. Figure1 depicts the original pattern of the series to have general overview of the movement of the time series during the time period.

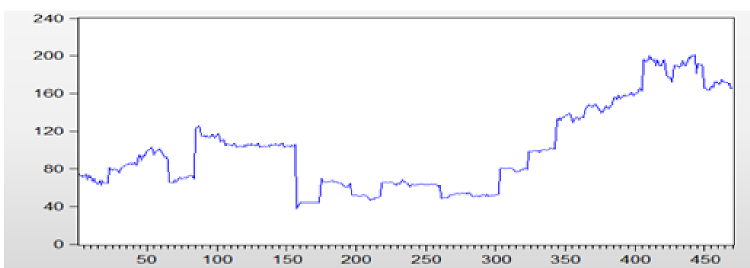


FIGURE 1. Time series plot of Attock Cement stock index (2005-2012)

Figure 1 is the plot of the data against the time period. The graph shows irregular patterns of upward and downward trends at all levels indicating that the series does not scatter horizontally around the constant mean. Therefore, the Augmented Dickey Fuller (ADF) Unit Root Test is carried out to test if the stock price is stationary.

TABLE 1. ADF Unit Root test of original stock index series

Series	Series	Prob.	First Differenced Series	Lag	Prob.
Stock	0	0.3090	ΔStock	2	0.0000

From Table 1, we observe that Stock series is non-stationary at level; whereas, after the first differenced, we see that the differenced series is stationary. Therefore, Stock series is a $I(1)$ series. Note that the error terms of the models are white noise series and lags are determined using Schwarz Bayesian Information Criterion.

After the first differenced of the series, we get the ACF and partial ACF graphs in Figure 2.

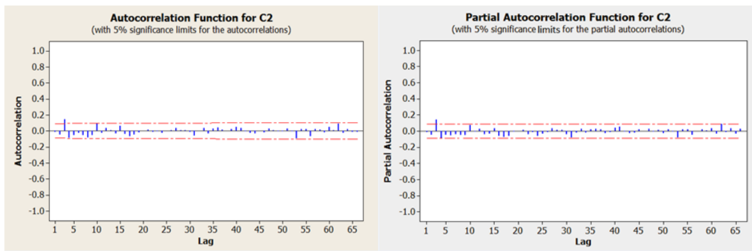


FIGURE 2. ACF and PACF graphs of first differenced stock index series

From Figure 2, we think the period is 3 or 4. Therefore, we try some seasonal ARIMA models such as ARIMA (0,1,0) $(1,0,0)_3$, ARIMA (0,1,0) $(2,0,0)_3$, ARIMA (0,1,0) $(0,0,1)_3$, ARIMA (0,1,0) $(0,0,2)_3$ models. We also consider the same models for period 4. However, for all of these models, we cannot obtain the error term as a white noise series. Therefore, we decide to use different criteria, such as Akaike Information Criterion (AIC), Schwarz Bayesian Information Criterion (BIC) and Hannan-Quinn Criterion (HQ), to find the best model for the data. In Table 2, we see that ARIMA (4,1,3) is the best model according toAIC and HQ criteria as the smallest values of these criteria are in this model. BIC criterion selects ARIMA (0,1,0) for the best model. As it is only a differenced model, we do not take care of BIC for this data.

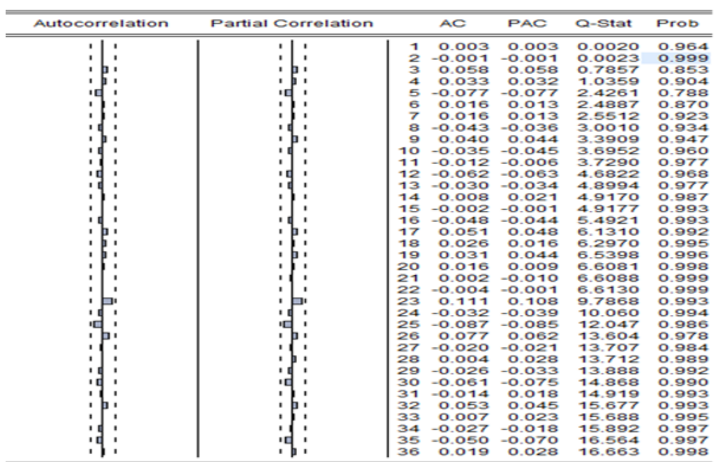
TABLE 2. Model selection based on various criteria

Models	AIC	BIC	HQ
(4,1,3)	7.178597	7.258246	7.209936
(4,1,4)	7.182774	7.271273	7.217595
(4,1,1)	7.186098	7.248047	7.210472
(1,1,4)	7.189021	7.250970	7.213395
(4,1,2)	7.190354	7.261153	7.218211
(2,1,4)	7.193254	7.264054	7.221111
(3,1,4)	7.197518	7.277167	7.228857
(0,1,4)	7.202806	7.255905	7.223698
(4,1,0)	7.203440	7.256540	7.224333
(3,1,3)	7.203795	7.274594	7.231651
(0,1,3)	7.204640	7.248889	7.222050
(3,1,1)	7.205572	7.258671	7.226464
(1,1,3)	7.205572	7.258671	7.226464
(2,1,3)	7.206024	7.267974	7.230399
(3,1,0)	7.206972	7.251221	7.224382

TABLE 2. Model selection based on various criteria

Models	AIC	BIC	HQ
(3,1,2)	7.209798	7.271747	7.234172
(2,1,1)	7.211605	7.255855	7.229016
(2,1,2)	7.212325	7.265425	7.233218
(1,1,2)	7.214051	7.258301	7.231461
(0,1,0)	7.220807	7.238507	7.227771
(0,1,1)	7.224769	7.251318	7.235215
(1,1,0)	7.224797	7.25147	7.235243
(0,1,2)	7.225925	7.261324	7.239853
(2,1,0)	7.226429	7.261828	7.240357

Next step is to diagnose the selected model through testing the errors at each lagif the errors have the white noise series properties. For this reason, Figure 3 is given.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.003	0.003	0.0020	0.964
		2	-0.001	-0.001	0.0023	0.999
		3	0.058	0.058	0.7857	0.853
		4	0.033	0.032	1.0359	0.904
		5	-0.077	-0.077	2.4261	0.788
		6	0.016	0.013	2.4887	0.870
		7	0.016	0.013	2.5512	0.923
		8	-0.043	-0.036	3.0010	0.934
		9	0.040	0.044	3.3909	0.947
		10	-0.035	-0.045	3.6952	0.960
		11	-0.012	-0.006	3.7290	0.977
		12	-0.062	-0.063	4.6822	0.968
		13	-0.030	-0.034	4.8994	0.977
		14	0.008	0.021	4.9170	0.987
		15	-0.002	-0.001	4.9177	0.993
		16	-0.048	-0.044	5.4921	0.993
		17	0.051	0.048	6.1310	0.992
		18	0.026	0.016	6.2970	0.995
		19	0.031	0.044	6.5398	0.996
		20	0.016	0.009	6.6081	0.998
		21	0.002	-0.010	6.6088	0.999
		22	-0.004	-0.001	6.6130	0.999
		23	0.111	0.108	9.7868	0.993
		24	-0.032	-0.039	10.060	0.994
		25	-0.087	-0.085	12.047	0.986
		26	0.077	0.062	13.604	0.978
		27	-0.020	-0.021	13.707	0.984
		28	0.004	0.028	13.712	0.989
		29	-0.026	-0.033	13.888	0.992
		30	-0.061	-0.075	14.868	0.990
		31	-0.014	0.018	14.919	0.993
		32	0.053	0.045	15.677	0.993
		33	0.007	0.023	15.688	0.995
		34	-0.027	-0.018	15.892	0.997
		35	-0.050	-0.070	16.564	0.997
		36	0.019	0.028	16.663	0.998

FIGURE 3. Correlogram of residuals for selected model

From Figure 3, it is clearly seen that all probability values for all lags are bigger than 0.05 which means that the errors of the selected model are white noise.

4.2. **FARIMA model.** The optimal Proposed Fuzzy Auto-Regressive Integrated Moving Average Model with fuzzy parametersis given as follows:

$$\tilde{z}_t = \tilde{\kappa}_1 z_{t-1} + \tilde{\kappa}_2 z_{t-2} + \Lambda + \tilde{\kappa}_\rho z_{t-\rho} + \varepsilon_t - \tilde{\kappa}_{\rho+1}\varepsilon_{t-1} - \dots - \tilde{\kappa}_{\rho+\hbar}\varepsilon_{t-\hbar}$$

In this model, possibilities of success are categories in to five linguistic terms; each linguistic term is represented by the degree of trapezoidal fuzzy number as

$$\tilde{V}_{verylow_k} = \left[14.22 - 16.22(1-\xi)51.60 + 16.22(1-\xi)\right] \quad (4.1)$$

$$\tilde{V}_{low_k} = \left[41.44 - 16.22(1-\xi)81.17 + 16.22(1-\xi)\right] \quad (4.2)$$

$$\tilde{V}_{average_k} = \left[64.8 - 16.22(1-\xi)113.54 + 16.22(1-\xi)\right] \quad (4.3)$$

$$\tilde{V}_{high_k} = \left[97.36 - 16.22(1-\xi)145.98 + 16.22(1-\xi)\right] \quad (4.4)$$

$$\tilde{V}_{veryhigh_k} = \left[129.76 - 16.22(1-\xi)178.42 + 16.22(1-\xi)\right] \quad (4.5)$$

We solve the integrals of equations (3.10) - (3.13) and put the intervals performed in Matlab with the coefficient values in the model. Then,trapezoidals are computed along with the defuzzification values obtained from SOM method in Matlab. These values are given in Table 3.

TABLE 3. Trapozoidal fuzzy numbers and defuzzification by using SOM method

Fuzzy Set	m	u	l	v	Smallest value of maximum method
B_1	-3.6641	-2.2175	34.4170	137.6678	0.0359
B_2	-3.6536	-2.2070	34.2681	137.0724	0.0464
B_3	-3.3501	-1.9540	30.3398	121.3594	0.0499
B_4	-3.5027	-2.2975	35.6587	142.6348	-0.0021
B_5	-3.3587	-1.9613	30.4413	121.7654	0.0413
B_6	-3.4793	-2.0625	32.0113	128.0452	0.0207
B_7	-3.8814	-2.3987	37.2287	148.914	0.0186
B_8	-0.1006	1.9613	30.4413	121.7654	1.9994
B_9	0.3225	5.5368	33.3946	133.5787	5.455
B_{10}	0.3211	5.815	33.2502	133.0009	5.055
B_{11}	0.2843	4.2638	29.4390	117.7546	4.2843
B_{12}	0.3341	6.0113	34.5894	138.398	6.0341
B_{13}	0.2852	4.2781	29.5375	118.1486	4.2852
B_{14}	0.2999	4.4987	31.0606	124.2419	4.4999
B_{15}	0.3488	7.319	36.1225	144.4914	7.4588
B_{16}	0.2852	7.2781	29.5275	118.1486	7.1652
B_{17}	11.1106	33.3308	61.3880	122.7760	33.4106
B_{18}	11.0625	13.1866	61.1225	122.2450	13.2625
B_{19}	9.8637	29.3822	54.1158	108.2317	29.4637
B_{20}	11.5122	16.5331	63.6029	127.2057	16.567
B_{21}	9.8249	21.3807	54.2964	108.5938	21.431
B_{22}	10.3326	31.0011	57.0971	114.1943	31.0326
B_{23}	12.0915	36.0535	66.403	132.8062	36.0915
B_{24}	9.8249	24.5805	54.2969	108.5938	24.405
B_{25}	28.5243	42.7865	52.6686	105.3372	42.8243
B_{26}	28.4009	42.6014	52.4408	104.8815	42.7009
B_{27}	25.1453	37.7179	46.4293	92.8587	37.7453
B_{28}	29.5535	44.3302	54.5688	109.1377	44.4000
B_{29}	25.2294	37.8441	46.5847	93.1693	37.9294
B_{30}	26.5304	39.7958	48.9872	97.9743	39.8304
B_{31}	30.8546	46.2819	56.9713	113.9427	46.3546
B_{32}	25.2294	37.8441	46.5847	93.1693	37.9294
B_{33}	51.6005	77.4008	98.9176	177.8352	77.5005
B_{34}	51.3773	77.0660	98.4897	196.9795	77.0773
B_{35}	45.4878	68.2317	87.1996	174.3992	68.2870
B_{36}	53.4622	80.1937	102.4865	204.9730	80.2622
B_{37}	46.6400	68.4600	87.4913	174.9827	68.5400

TABLE 3. Trapozoidal fuzzy numbers and defuzzification by using SOM method

Fuzzy Set	m	u	l	v	Smallest value of maximum method
B_{38}	47.9938	71.9906	92.0035	184.0070	71.9938
B_{39}	55.8160	83.7240	106.9987	213.9973	83.8160
B_{40}	45.6400	68.4600	87.4913	174.9827	68.5400
B_{41}	55.6815	64.9614	96.8700	193.5040	64.9815
B_{42}	55.4070	64.6804	96.4510	192.6670	64.7070
B_{43}	49.0854	57.2660	85.3945	170.5810	57.2854
B_{44}	57.6905	67.3052	100.3650	200.4855	67.3905
B_{45}	49.2496	57.4575	85.6802	171.1517	57.5496
B_{46}	51.7895	60.4208	90.0990	179.9785	60.4895
B_{47}	60.2304	70.2085	104.7838	209.3123	70.2304
B_{48}	49.2496	57.4575	85.6802	171.1517	57.5496
B_{49}	66.4842	91.4666	99.7263	182.9332	91.5000
B_{50}	66.1966	91.0710	99.2949	182.1420	97.1000
B_{51}	58.6083	80.6313	87.9125	161.2625	80.7000
B_{52}	68.8829	83.3244	94.7667	189.5334	83.4000
B_{53}	58.8044	80.9010	88.2066	161.8020	80.9044
B_{54}	61.8371	85.0733	92.7556	170.466	90.7000
B_{55}	71.9156	94.8734	98.9390	197.8780	94.9156
B_{56}	58.8044	78.2066	80.9010	161.802	78.3000
B_{57}	83.9393	92.9110	125.9090	165.4781	92.9393
B_{58}	83.5763	92.5092	125.3644	164.7623	92.5763
B_{59}	73.9957	81.9046	110.9935	145.852	81.9957
B_{60}	86.9678	96.2632	130.4517	171.4485	96.2678
B_{61}	74.2432	82.178	111.3649	146.3632	82.2432
B_{62}	78.0722	86.4168	117.1083	153.9115	86.4722
B_{63}	90.7968	100.5014	136.1951	178.9968	100.5968
B_{64}	74.2432	82.1786	111.3649	146.3632	82.2432
B_{65}	99.1160	148.6740	172.5324	263.6888	148.793
B_{66}	98.6872	148.0308	171.7861	262.5481	148.1419
B_{67}	87.3744	131.0617	152.0938	232.4515	131.172
B_{68}	102.6920	154.0380	178.7572	273.2025	154.149
B_{69}	87.6668	131.5002	152.6026	233.2292	131.611
B_{70}	92.1880	138.2820	160.4274	259.5894	138.394
B_{71}	107.2132	160.8198	186.6274	258.4666	160.921
B_{72}	87.6668	131.5002	152.6026	228.8378	131.611
B_{73}	103.2051	154.8077	170.1576	268.9553	154.918
B_{74}	102.7587	154.1380	179.4254	229.6034	154.249
B_{75}	90.9792	136.4688	160.1489	228.8378	136.579
B_{76}	106.9236	160.3930	186.2505	268.9553	160.410
B_{77}	91.2836	136.9253	180.6448	229.6034	136.103
B_{78}	95.9913	143.9870	168.3495	241.4447	143.108
B_{79}	111.6364	167.4547	193.9552	280.7966	167.665
B_{80}	91.2836	136.9253	160.6448	229.6034	136.103

TABLE 3. Trapozoidal fuzzy numbers and defuzzification by using SOM method

Fuzzy Set	m	u	l	v	Smallest value of maximum method
B_{81}	114.0078	171.0117	175.1517	248.8885	171.123
B_{82}	113.5146	170.2719	174.3940	247.8119	170.382
B_{83}	100.5021	150.7532	154.4028	219.4046	150.802
B_{84}	118.1211	177.1817	181.4710	257.8683	177.219
B_{85}	100.8384	157.2576	154.9193	220.1386	157.338
B_{86}	106.0389	159.0583	163.9090	237.4917	159.168
B_{87}	123.3216	184.9824	189.4607	269.2214	184.102
B_{88}	100.8384	151.2576	154.9193	220.1388	151.368
B_{89}	131.4629	197.2351	217.3832	231.3316	197.362
B_{90}	130.8943	196.3819	216.4435	230.3310	196.494
B_{91}	115.8895	173.8701	191.1093	203.9275	173.989
B_{92}	136.2060	204.3512	225.2208	239.6779	204.4060
B_{93}	116.2772	174.4518	192.2900	204.6098	174.577
B_{94}	122.2740	183.4488	202.1992	215.2302	183.574
B_{95}	142.2028	213.3482	234.5900	250.2302	213.4028
B_{96}	116.2772	174.4518	192.2900	204.6098	174.587
B_{97}	165.3560	248.0340	270.6423	329.6745	248.145
B_{98}	164.6408	246.9612	269.4716	328.2485	246.103
B_{99}	145.7676	218.6513	238.5814	290.6205	218.7676
B_{100}	171.3220	256.9830	280.4070	341.5690	256.104
B_{101}	146.2552	219.3828	239.3776	291.5928	219.4155
B_{102}	153.7980	254.1678	251.7250	306.6310	254.2980
B_{103}	178.8648	253.0683	292.7524	356.6072	253.1648
B_{104}	146.2552	224.0584	239.3796	291.5928	224.1552
B_{105}	169.4452	263.3380	264.5127	325.5854	263.4452
B_{106}	168.7122	224.8080	263.3380	324.1770	224.919
B_{107}	149.3723	236.4020	233.1779	287.0157	236.5723
B_{108}	175.5587	274.9320	274.0562	337.3323	274.1058
B_{109}	149.8720	224.8080	233.9580	287.9760	224.9190
B_{110}	157.6013	236.4020	246.0238	302.8277	236.5013
B_{111}	183.2880	274.9320	286.1220	352.7840	274.1088
B_{112}	149.8720	224.8080	233.9580	287.9760	224.9720
B_{113}	180.2478	270.3718	246.0238	314.7827	270.4847
B_{114}	179.4682	269.2022	286.1220	313.4277	269.3682
B_{115}	158.8952	238.3429	233.9580	277.4928	238.479
B_{116}	186.511	280.1266	279.4554	326.1399	280.2000
B_{117}	159.4268	239.1403	278.2465	278.4212	239.2600
B_{118}	167.6489	251.4734	246.3504	292.7801	251.5489
B_{119}	194.9732	292.4597	302.2854	340.4988	292.5732
B_{120}	159.4268	239.1403	247.1746	278.4212	239.2600
B_{121}	197.7030	296.5545	352.1038	399.0476	296.6430
B_{122}	196.8478	295.2717	350.5807	397.3214	295.665
B_{123}	174.2826	261.4239	310.3927	354.7754	261.534

TABLE 3. Trapozoidal fuzzy numbers and defuzzification by using SOM method

Fuzzy Set	m	u	l	v	Smallest value of maximum method
B_{124}	204.8360	307.2540	364.0875	413.4450	307.365
B_{125}	174.8657	262.2985	311.4312	352.9523	262.310
B_{126}	183.8840	275.8260	327.4925	371.1550	275.914
B_{127}	213.8543	320.7815	380.8688	431.6477	320.854
B_{128}	174.8657	262.2985	311.4312	352.9523	262.310
B_{129}	231.2787	346.9181	370.7674	35210.38	346.1065
B_{130}	230.2787	345.4175	350.5807	369.1636	345.5300
B_{131}	203.8809	305.8213	310.3927	326.8454	305.980
B_{132}	239.6231	359.4347	364.8075	384.1444	359.5231
B_{133}	204.5630	306.8445	327.4925	327.9389	306.9630
B_{134}	215.1129	322.6693	380.8688	344.8516	322.778
B_{135}	250.1730	375.2595	380.8688	401.0571	375.3600
B_{136}	204.5630	306.8445	311.4312	327.9389	306.946
B_{137}	235.3597	353.0396	375.8351	391.5813	353.1
B_{138}	234.3417	351.5125	374.2093	389.8875	351.6500
B_{139}	207.4784	311.2177	331.3127	345.1936	311.34
B_{140}	243.8514	365.7771	389.3950	405.7093	365.8000
B_{141}	208.1726	312.2589	332.4212	346.3485	312.36900
B_{142}	218.9086	328.3629	349.5650	364.2107	328.4000
B_{143}	254.5874	381.8811	406.5388	423.5715	381.9600
B_{144}	208.1726	312.2589	332.4212	346.3485	312.3900
B_{145}	246.1648	369.2473	383.6573	390.9506	369.3100
B_{146}	245.1000	367.6500	381.9978	389.2596	367.7000
B_{147}	217.0036	325.5053	338.2083	344.6377	325.6300
B_{148}	255.0036	382.5695	397.4995	405.0559	382.6000
B_{149}	217.7296	326.5943	339.3398	345.7907	326.6300
B_{150}	228.9585	343.4377	356.8405	363.6241	343.548
B_{151}	266.2752	399.4729	415.0002	422.8893	399.5752
B_{152}	217.7296	326.5943	339.3398	345.7907	326.6300
B_{153}	262.6176	395.4263	400.3699	406.2086	395.5917
B_{154}	262.4772	393.7159	398.6381	404.4516	393.8677
B_{155}	232.3888	348.5832	352.9411	358.0882	348.6288
B_{156}	273.1287	404.6931	414.8150	420.8644	404.728
B_{157}	233.1663	349.7494	354.1219	359.2862	349.8000
B_{158}	245.1913	367.7869	372.3850	377.8156	367.8000
B_{159}	285.1537	427.7306	433.0781	439.3938	427.8537
B_{160}	233.1663	349.7494	354.1219	359.2862	349.8400

In order to find the fuzzy relationship, dataset is arranged according to the years order with fuzzy set value.

TABLE 4. Actual prices of Attock Cement Industry with fuzzy set

Year	Prices	Fuzzy Set
2005	5.55	B_9
2006	6.055	B_{12}
2007	7.16	B_{15}
2008	7.45	B_{16}
2009	13.71	B_{18}
2010	16.8	B_{20}
2011	21.41	B_{21}
2012	167.75	B_{79}

Now fuzzy relationship is designed from fuzzified dataset given in Table 4. If the time series variable $F(t-1)$ is fuzzified as B_9 in 2005 and $F(t)$ as B_{12} in 2006, then B_9 is related to B_{12}.In the same way, for all of the years, fuzzy relationship is formed as in Table 5. Note that there is no fuzzy relationship that consists of more than one set that can be merged into one group.

TABLE 5. Fuzzy relationships

Year Relationship	Fuzzy Relationship
2005 $\longrightarrow$ 2006	$B_9 \rightarrow B_{12}$
2006 $\longrightarrow$ 2007	$B_{12} \rightarrow B_{15}$
2007 $\longrightarrow$ 2008	$B_{15} \rightarrow B_{16}$
2008 $\longrightarrow$ 2009	$B_{16} \rightarrow B_{18}$
2009 $\longrightarrow$ 2010	$B_{18} \rightarrow B_{20}$
2010 $\longrightarrow$ 2011	$B_{20} \rightarrow B_{21}$
2011 $\longrightarrow$ 2012	$B_{21} \rightarrow B_{79}$

The forecast value for 2006 is obtained using the fuzzified interval midpoint value of 2005.The fuzzy relationship group of 2005 is $B_9 \rightarrow B_{12}$. According to this case, the highest degree of B_{12} interval is $u_{12} = \left[6.0113\ 34.5894\right]$ so the forecast value of 2006 is the midpoint of u_{12} which is 20.30035. In the same way, all the forecasted prices are obtained in Figure 4.

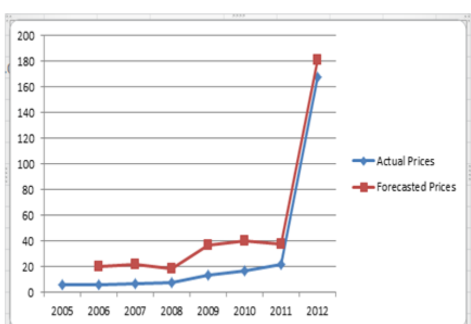


FIGURE 4. Forecasted prices

To compare the proposed model with ARIMA model, values of several statistics for the models are given in Table 6. From Table 6, it is clear that the proposed model fits better than ARIMA model to Stock Index in Pakistan.

TABLE 6. Comparison of models

Evaluation Criteria	ARIMA	FARIMA
Mean Square Error (MSE)	6978.4964	294.8119
Mean Absolute Percentage Error (MAPE)	114.9294	1.3994
Root Mean Square Error (RMSE)	83.5373	17.1700
Mean Absolute Deviation (MAD)	83.0082	16.5507

5. CONCLUSION

In application, we see that the mean squared error, mean absolute percentage error, mean absolute deviation, root mean squared error of the proposed fuzzy autoregressive integrated moving average are quite small as compared to the autoregressive integrated moving average model. This indicates that proposed FARIMA forecasts better and performs better than ARIMA. In addition to application, there are some more advantages of FARIMA with respect to ARIMA as follows:

- To perform ARIMA, there should be at least 50 observations, whereas FARIMA can also work well with small number of observations.
- ARIMA provides confidence intervals, whereas FARIMA gives possibility intervals of parameters.
- ARIMA uses information of previous function, whereas FARIMA uses information of fuzzy functions.

REFERENCES

[1] A.K. Azad, M. Mahsin, *Exchange Rates of Bangladesh using ANN & ARIMA models: A comparative study*, International Journal of Advanced Engineering Science & Technologies **10**, (2011) 31-36.

[2] G. E. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day Inc., San Francisco, CA,1976.

[3] E. Saarika, Hadavandi , I. B. Turksen, *A Hybrid Fuzzy Intelligent Agent Based System for Stock Price Prediction*, International Journal of Intelligent Systems **27**(11), (2014) 947-969.

[4] H. Saima, J. Jaafar, S. Belhaouari,*ARIMA based Interval Type-2 Fuzzy Model for forecasting*, International Journal of Computer Applications **28**(3), (2011) 975-8887.

[5] H. Tanaka, S. Uejima, K. Asai, *Linear regression analysis with fuzzy model*, IEEE Trans. Systems Man Cybernet **12**(6), (1980) 903-907.

[6] H. K.Tiffany, K. H. Huarng, *A neural network-based fuzzy time series model to improve forecasting*, Expert Systems with Applications **37**(4), (2010) 3366-3372.

[7] L.A .Zadeh,*Fuzzy sets, information and Control* **8**, (1965) 338-353.

[8] Q. Song, B.S. Chissom, *Forecasting enrollments with fuzzy time series Part I*, Fuzzy Sets and Systems **54**(1), (1993) 1-9.