İktisat Politikası Araştırmaları Dergisi -
Journal of Economic Policy Researches
Cilt/Volume: 9, Sayı/Issue: 2, 2022
E-ISSN: 2148-3876

İSTANBUL UNIVERSITY PRESS

RESEARCH ARTICLE / ARAŞTIRMA MAKALESİ

Stock Market Price Forecasting Using the Arima Model: an Application to Istanbul, Turkiye

Borsa İstanbul Fiyatlarının Arima Modeli İle Tahmin Edilmesi

Tamerlan MASHADIHASANLI[1]

ABSTRACT

Because of its critical position in open economies and its extremely high volatility, the stock market price index has been a popular subject of market research. In modern financial markets, traders and practitioners have had trouble predicting the stock market price index. In order to solve this problem, some methods have been researched by researchers and suitable methods have been found. To analyze and forecast monthly stock market price index, a variety of statistical and econometric models are extensively used. Thus, this study aims to investigate the application of autoregressive integrated moving averages (ARIMA) for forecasting monthly stock market price index in Istanbul for the period from 2009-M01 to 2021-M03. As compared to all other tentative models, the research showed that the ARIMA *(3,1,5)* model is the best fit model for predicting the stock market price index. Forecasting is conducted by using the developed model ARIMA *(3,1,5)* and the results indicated that the forecasted values are very similar to the actual ones, reducing forecast errors. In general, the stock market price index in Istanbul; showed a downwards trend over the forecasted period. The results of the study can set an example for researchers and practitioners working in the stock market and can be a guide for economic decision units and investors in the stock market.

Keywords: ARIMA, forecasting, stock market price index, time series, Turkiye

Jel Code: E47, G17, E37

ÖZ

Açık ekonomilerdeki kritik konumu ve son derece yüksek oynaklığı nedeniyle borsa fiyat endeksi, piyasa araştırmalarının popüler bir konusu olmuştur. Modern finans piyasalarında, tüccarlar ve uygulayıcılar borsa

DOI: 10.26650/JEPR1056771

[1]Istanbul University, Institute of Social Sciences/ Department of Economics (English), Istanbul, Turkiye

ORCID: T.M. 0000-0002-8186-8420

Corresponding author/Sorumlu yazar:
Tamerlan MASHADIHASANLI,
Istanbul University, Institute of Social Sciences/ Department of Economics (English),
Istanbul, Turkiye
E-mail/E-posta:
tamerlan.mashadihasanli@ogr.iu.edu.tr

Submitted/Başvuru: 12.01.2022
Revision Requested/Revizyon Talebi: 07.05.2022
Last Revision Received/Son Revizyon: 26.05.2022
Accepted/Kabul: 15.07.2022

Citation/Atıf: Mashadihasanli, T. (2022). Stock market price forecasting using the arima model: an application to Istanbul, Turkiye. *İktisat Politikası Araştırmaları Dergisi - Journal of Economic Policy Researches, 9*(2), 439-454. https://doi.org/10.26650/JEPR1056771

fiyat endeksini tahmin etmekte zorlanıyorlar. Bu soruna çözüm getirmek için araştırmacılar tarafından bazı yöntemler araştırılmış ve uygun yöntemler bulunmuştur. Aylık borsa fiyat endeksini analiz etmek ve tahmin etmek için çeşitli istatistiksel ve ekonometrik modeller yaygın olarak kullanılmaktadır. Bu nedenle, bu çalışma, 2009-M01 ile 2021-M03 arasındaki dönem için İstanbul'da aylık borsa fiyat endeksini tahmin etmek için otoregresif entegre hareketli ortalamalar (ARIMA) uygulamasını araştırmayı amaçlamaktadır. Araştırma, diğer tüm geçici modellerle karşılaştırıldığında, ARIMA (3,1,5) modelinin borsa fiyat endeksini tahmin etmek için en uygun model olduğunu göstermiştir. Tahmin, geliştirilen ARIMA (3,1,5) modeli kullanılarak yapılmıştır ve sonuçlar, tahmin edilen değerlerin gerçek değerlere çok benzer olduğunu ve tahmin hatalarını azalttığını göstermiştir. Genel olarak İstanbul'da borsa fiyat endeksi; tahmin edilen dönemde aşağı yönlü bir eğilim göstermiştir. Çalışmanın sonuçları borsada çalışan araştırmacı ve uygulayıcılara örnek teşkil edebileceği gibi borsada ekonomik karar birimlerine ve yatırımcılara yol gösterici olabilir.

Anahtar Kelimeler: ARIMA, tahminleme, borsa fiyat endeksi, zaman serisi, Türkiye
Jel Code: E47, G17, E37

1. Introduction

Prediction will remain to be an enthralling field of study, with domain researchers constantly looking for ways to develop current predictive models. The primary reason is that institutions and individuals now have the authority to make investment decisions as well as have ability to plan and improve successful strategies for their daily and future endeavors. For a long time, researchers have been focused on predicting stock market prices. Investors often seek to maximize their trading profits; hence, the need for higher level of accuracy in the prediction of future prices. However, achieving accurate predictions remains a challenge (Subing & Kusumah, 2017; Nandakumar, Uttamraj, Vishal, & Lokeswari, 2018; Shah, Campbell, & Zulkernine, 2018). Many investors desire to get their hands on any forecasting system that promises easy profits and reduces stock market risk. This continues to inspire academics to improve and create new predictive models (Atsalakis, Dimitrakakis & Zopounidis, 2011).

As a result, in recent time, various ways have been suggested. and adopted to predict the prices of equities of various corporations and stock indexes. Time series and machine learning techniques, fundamental analysis, technical analysis are some of the most common approaches for forecasting stock market prices (Kihoro & Okango, 2014). Artificial neural networks (ANNs) are one of them, and they are very popular because they can deduce answers from unknown data and learn patterns from it. There are a few associated works that used an ANNs model to forecast stock market prices (Mitra, 2009; Atsalakis & Valavanis, 2009; Mostafa, 2010). Another popular and effective method is Autoregressive Integrated Moving Averages (ARIMA). In financial time series forecasting, ARIMA models have been found to be more resilient and efficient, despite the most widely used ANNs strategies in short-term prediction technique (Yoo, 2007; Merh, Saxena, & Pardasani, 2010; Sterba & Hilovska, 2010).

As it can be understood from the academic studies mentioned above, the stock market price index is a popular subject of market research. In modern financial markets, traders and practitioners have had trouble predicting the stock market price index. It is thought that the ARIMA model can provide more accurate predictions by removing these difficulties. Thus, this study aims to investigate the application of autoregressive integrated moving averages (ARIMA) for forecasting monthly stock market price index in Istanbul for the period from 2009-M01 to 2021-M03. It is anticipated that this study will contribute to the literature in terms of forecasting and will encourage future academics to use the ARIMA model as a forecasting method.

The paper is structured as follows. Section two provides a brief overview of ARIMA model followed by the data and methodology section. Section four discusses the analysis and the research findings obtained. Finally, the conclusion presents a brief summary and critique of the findings.

2. Literature review

When we look at the literature, we can easily see that ARIMA models are applied to analyse and forecast time series data in much empirical research. To forecast the next day's electricity prices for the Spanish and California electricity markets, Contreras, Espinola, Nogales, and Conejo (2003) applied ARIMA models and they developed two ARIMA models to forecast hourly prices. To anticipate future prices, the Spanish model requires 5 hours., while the Californian model needs just 2 hours. In Brunei Darussalam, Kumar, Yadav, Singh, Hassan, and Jain (2004) employed the ARIMA model to estimate daily maximum surface ozone concentrations. They demonstrated that ARIMA (1, 0, 1) was acceptable for surface O3 data obtained at Brunei Darussalam's airport. Takahashi, Tamada, and Nagasaka (1998) suggested a neural network incorporating a multiple line-segments regression technique to forecast stock prices. The results demonstrated that the proposed technique was effective at predicting stock prices. The ARIMA model was used by Tsitsika, Maravelias, and Haralabous (2007) to predict pelagic fish production. During the estimations, it was revealed that ARIMA (1, 0, 1) and ARIMA (0, 1, 1) were best models to forecast data. The ARIMA model was used by Liu, Liu, Jiang, and Yang (2011) to forecast the occurrence of hemorrhagic fever associated with renal syndrome in China The goodness of fit test of the best ARIMA (0, 3, 1) model revealed non-significant autocorrelation in the model's residuals.

Yoon and Swales (1991) proposed a four-layered neural network for predicting US stock prices. According to the findings, MDA (multiple discriminant analysis) method is outperformed by the proposed method. In order to forecast inflation in the Bangladesh economy. Datta (2011) applied the ARIMA model. He demonstrated that the ARIMA (1, 0, 1) model satisfactorily matches the inflation data of Bangladesh. Al-Zeaud (2011) also applied the ARIMA model to model and predict volatility. According to the results, the ARIMA (2, 0, 2) model is the best at 95 percent confidence interval for the banking sector. Uko and Nkoro (2012) examined the ARIMA, VAR, and ECM models in predicting Nigerian inflation. according to the findings, among the other models, ARIMA is a better to forecast inflation in Nigeria and can be used as a benchmark model for forecasting inflation. Meyler, Kenny, and Quinn (1998) used ARIMA models to anticipate inflation in Ireland using quarterly data from 1976 to 1998, illustrating some practical challenges with ARIMA time series predicting. Kock and Teräsvirta (2013) used Artificial Neural Network (ANN) models to estimate consumer price inflation in Finland from March 1960 to December 2009, and found that Direct forecasts outperform recursive forecasts. Kharimah, Usman, Widiarti, and Elfaki (2015) used ARIMA models to evaluate CPI data from January 2009 to December 2013 and found that the ARIMA (1, 1, 0) was the best model for forecasting CPI in Malaysia. To predict Japanese stock markets, Baba and Kozaki (1992) applied a back-propagation neural network paired with a random optimization technique. The simulation results showed that the proposed approach did aid in stock price predictions.

Nyoni (2018) used ARIMA and GARCH models to simulate inflation in Kenya, utilizing annual time series data in 1960 - 2017. He discovered that there are 3 models that are best ones to be used for predicting inflation, and those are ARIMA (2, 2, 1), ARIMA (1, 2, 0), and the AR (1) - GARCH (1, 1) models. Most recently, Nyoni and Nathaniel (2018) explored inflation using data on in 1960-2016 for Nigeria and stated that the ARMA (1, 0, 2) is the best to predict inflation rate. To forecast stock prices in Tokyo, Kamijo and Tanigawa (1990) developed an interesting method-pattern recognition method. A new method for evaluating recurrent networks in order to reduce mismatching patterns has been proposed.

In his study, Zhang (2003) compared the results by applying the ARIMA, Artificial Neural Networks and ARIMA-Artificial Neural Networks hybrid method to the UK's 1980-1993 weekly exchange rate series and concluded that the performance of the ARIMA-Artificial Neural Networks method was superior to the others. In their study, Kumar and Thenmozhi (2014) used ARIMA, Artificial Neural Networks, Support Vector Regression, Random Forest methods as well as ARIMA-Artificial Neural Networks, ARIMA-Support Vector Regression, ARIMA to India's daily stock index data for the period 2003:01-2009:12. Applied Random Forest hybrid methods and found that the prediction success of ARIMA-Support Vector Regression method is superior to other methods. Çevik (2002) found that the most convenient model for the series was the ARIMA (1,2,1) model, using the ARMA method, with the monthly data of the 1986-2002 period in order to model the BIST index. Etuk, Uchendu, and Udo (2012) studied the Nigerian stock market with the Box-Jenkins approach using monthly data for the period 1987-2006. As a result, it was seen that the most suitable model was ARMA (2,1) and ARIMA (2,1,3), respectively. Sekreter and Gürsoy (2014) tried to predict the BIST-100 stock market with the daily data set for the period 2006-2012 with ARIMA and GARCH models and they revealed that the ARIMA model gave the best estimation result.

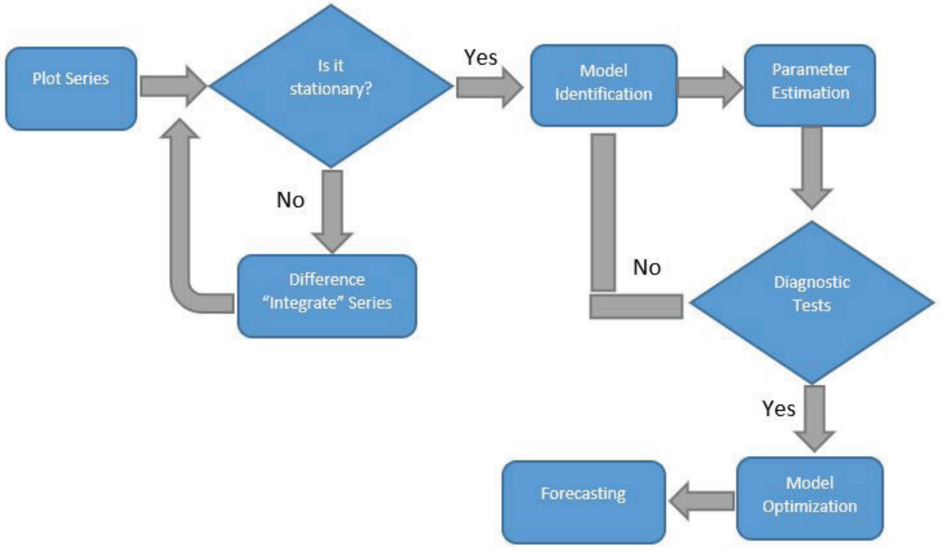
Bircan and Karagöz (2003) studied the most convenient estimation model for the monthly exchange rate series covering the period 1991-2002 using the Box Jenkins method. As a result of the estimation, the best model for the exchange rate series was determined as ARIMA (2,1,1). For the suitability of the model, the Q statistics were calculated, and it was decided that the estimation errors were randomly distributed and that the model was suitable for the exchange rate estimation at the 5 percent significance level. Gupta and Kashyap (2015) tried to estimate the fluctuations in the US dollar, Yen, Euro and GBP exchange rates with the ARIMA model, using monthly data to predict the changes in the exchange rate for the period 1999-2014 in India. Yaziz, Ahmad, Nian, and Muhammad (2011) tried to predict oil prices with Box-Jenkins and GARCH models using daily crude oil prices for the 1986-2009 period. They revealed that the most convenient estimation model is the GARCH (1,1) model. Saibu (2015) estimated Nigerian crude oil prices with the Box-Jenkins model with the monthly data set for the period 2000-2012.

As it can be understood from the academic studies mentioned above, the stock market price index is a popular subject of market research. In modern financial markets, traders and practitioners have had trouble predicting the stock market price index. It is thought that the ARIMA model can provide more accurate predictions by removing these difficulties. Generally, it is clear from the preceding studies that ARIMA can be used to forecast. The current study aims to find out the best ARIMA model for predicting the stock market price index in Istanbul.

3. Methodology

To predict stock market price index in Istanbul, this paper applies the ARIMA model. In 1970, The ARIMA model was developed by Box and Jenkins. It is also known as the Box-Jenkins methodology which consists of some major steps as identifying, estimating, and diagnosing.

Figure 1. Steps of Box-Jenkins Approach



In financial forecasting, the model is one of the most widely used approaches (Pai & Lin, 2005; Merh et al., 2010). ARIMA models have demonstrated their efficient ability to produce short-term predictions. In terms of short-term prediction, it consistently outperformed complicated structural models (Meyler et al., 1998). The ARIMA model consists of several steps such as identification, estimation and diagnostic (Tabachnick, Fidell, & Ullman, 2007). Figure 1 depicts the ARIMA modeling and forecasting procedure flow chart.

The ARIMA model is based on *AR* and *MA* models. While the *AR* model is used to show that the current observation is dependent on previous observations, the *MA* model is used to show that the current and previous residuals compose a linear function. (Chang, Sriboonchitta, & Wiboonpongse, 2009). General statement for these models is ARIMA (*p,d,q*) where *p* denotes the degree of *AR* model, *d* denotes the degree of different order and *q* denotes the degree of *MA* model. The ARIMA (p, d, q) model takes the following form:

$$\Delta dY_t = c + \varphi_p \Delta dY_{t-1} + \ldots + \varphi_p \Delta dY_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \qquad (1)$$

Where ΔdY_t indicates a differenced dependent variable at time t, ΔdY_{t-1}, ΔdY_{t-p} indicate the differenced lagged dependent variables, c is a constant, ϕ_1, ϕ_p, θ_1, θ_q indicate model parameters, ε_t is the residual term and ε_{t-1}, ε_{t-q} are the previous values of the residual.

To select the best ARIMA model among various experiments, the following criteria were employed in this analysis for stock market price index:

4. Data and Findings

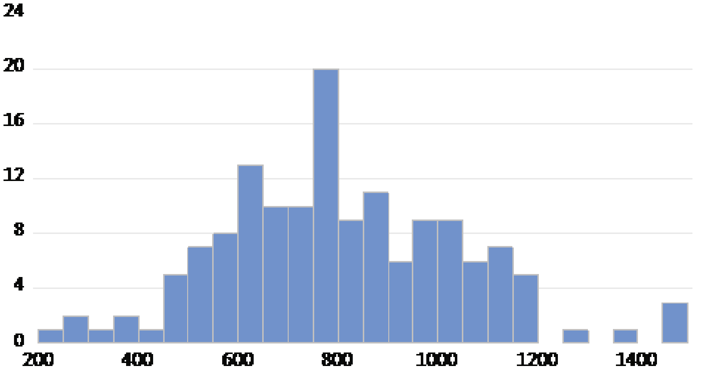
4.1. Data Description

Official monthly data of stock market price index of Istanbul between 2009-M01 and 2021-M03 has a total number of 147 observations which are used to estimate and forecast the model. It must be noted that the data is divided into two parts: first part is the in-sample data which covers the period from 2009-M01 to 2020-M12 that includes 144 observations and is used to estimate the model, second part is out-of-sample data which covers the period from 2021-M01 to 2021-M03 and is used for forecasting. The data used in the study is provided by Istanbul University. Figure 2 displays the descriptive statistics of the monthly stock market price index for the study selected period. It shows positive skewness (0.289393) which refers to the degree to which the data are asymmetric. Furthermore, it has a high positive kurtosis (3.280419), indicating that the distribution has larger tails than the normal distribution. And, according to the *Jarque-Bera statistics*, the stock market price index is normal at the confidence interval of 99% since probability is 0.281750 which is more than 0.01.

Table 1: Descriptive Statistics

Observations	147
Mean	805.5192
Median	784.8901
Maximum	1476.720
Minimum	240.2659
Std. Dev.	239.9296
Skewness	0.289393
Kurtosis	3.280419
Jarque-Bera	2.533470
Probability	0.281750

Figure 2. Distribution of the Monthly Stock Market Price Inde

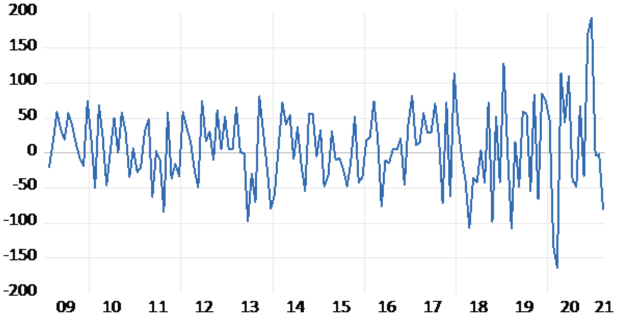


4.2. Stationarity test

According to the stock market price index series plot *(SPI)* between 2009-M01 and 2021-M03 in Figure 3, shows that the stock market price index series are non-stationary at level. As a result, the non-stationary series is used to transform stationary series using the lag differencing technique. The plot of the stock market price index and the differenced stock market price index have been illustrated in Figure 3 and Figure 4. Figure 3 clearly shows that there is a trend in that series. And Figure 4 shows that the data are stationary at first differenced.

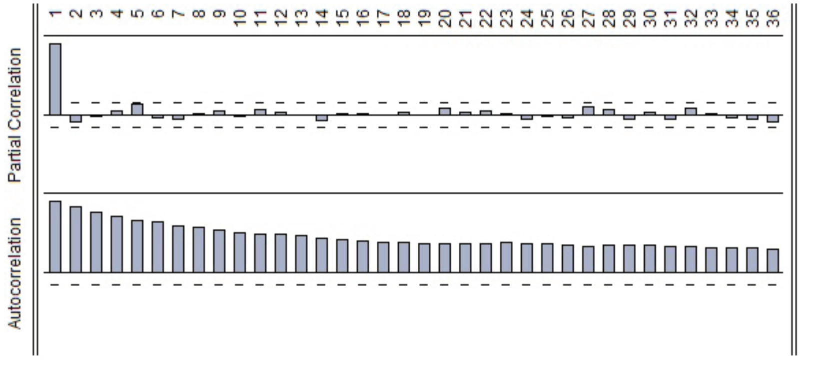
Figure 3. The Stock Market Price Index (SPI) Series Plot Between 2009-M01 and 2021-M03

Figure 4. The Differenced Stock Market Price Index D(SPI) Series Plot Between 2009-M01 and 2021-M03



The autocorrelation and partial autocorrelation function graphs of the *SPI* series have been illustrated in Figure 5. Figure 5 shows that there is partial autocorrelation in the 1st lag. In other lags, there is no autocorrelation since the values are between the significant lines. Beside that at lag 5 and 6, the bar still far from zero and the lags decline very quickly, so we can conclude that the *SPI* is not stationary.

Figure 5. The Autocorrelation and Partial Autocorrelation Function Graphs of the SPI Series



At next step, formal tests-standard unit root test at level was done and trend (because at Figure 3 it is seen that there is a trend) and intercept was included for unit root test. The results of unit root test have been illustrated in Table 1. When we look at the *p* value at level, it is seen that the value is 0.0591 (bigger than 0.05). It means the Null Hypothesis cannot be rejected. We accept that the variable has a unit root. So, we have a non-stationary variable and we are going to work with AR(I)MA Model (*p,d,q*). We apply first difference unit root test because of non-stationary variable. As we see from Table 1, *p* value of the unit root test at first difference is smaller than 0.05. In this case, first difference is going to be enough for identification of "possible models".

Table 2: The Result of Unit Root Test

Variable	**ADF**			
	Level		***1st Difference***	
SPI	*t-Statistic*	*Probability*	*t-Statistic*	*Probability*
	-3.372644	0.0591	-11.86094***	0.0000

Notes: *** 1 percent level

4.3. Model Identification

The second step is model identification. We check the correlogram to determine *p* for *AR* component and *q* for *MA* component of AR(I)MA Model. To determine *p* and *q* values, we are going to use the autocorrelation and partial autocorrelation functions. *ACF* and *PACF* may suggest diverse "possible models". We check the correlogram in first difference. The autocorrelation and partial autocorrelation function graphs of the differenced series *D(SPI)* have been illustrated in Figure 6. Figure 6 shows that the *SPI* data, at the *ACF* bar indicates non-significant at lag 3, as a result, it's safe to believe the data came from MA *(3)*. According to the PACF graph, the bar at lag 5 non-significant, as a result, it's safe to believe the data came from AR *(5)*, and we get the initial ARIMA (5,1,3) model with the differenced series equal to 1. For the next step, we will compare ARIMA (5, 1, 3) model with other ARIMA models such as ARIMA *(3,1,5),* ARIMA *(4,1,5)* and ARIMA *(5,1,5)*.

Figure 6. The Autocorrelation and Partial Autocorrelation Function Graphs in First Difference

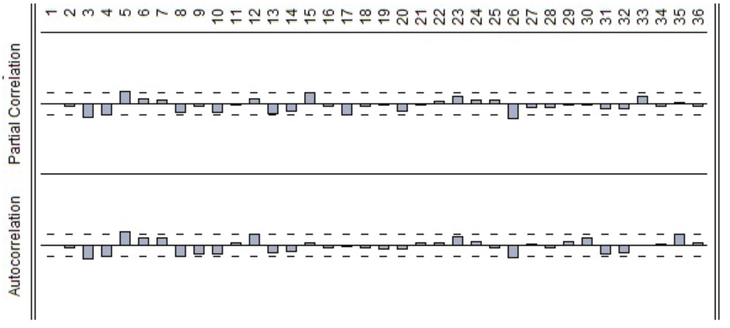


Table 2 shows the tentative ARIMA (*p,d,q*) test results for various parameters. *Adjusted R-squared, AIC*, *SC*, *HQC* values and the parameter significance are all crucial criteria for selecting models. In general, the larger the coefficient of determination and adjusted *R-squared*, and the smaller the *AIC, HQC,* and *SC* values, better ARIMA (*p,d,q*) model. So, the "possible models" are going to be following:

Table 3: Statistical Results of the Tentative ARIMA models.

D(SPI)	*ARIMA(3,1,5)*	ARIMA(4,1,5)	ARIMA(5,1,3)	ARIMA(5,1,5)
Adj. R2	***0.099854***	0.090075	0.069838	0.096732
AIC	***10.83624***	10.84681	10.86582	10.84135
SBC	***10.89754***	10.90812	10.92712	10.90266
HQC	***10.86115***	10.87172	10.89073	10.86626

Although the appropriate ARIMA model is usually chosen using the aforementioned criteria, other tests, such as residual randomness, s*erial correlation LM test, White test for heteroskedasticity, Ramsey RESET test for stability* and *normality,* are performed and checked for all tentative models. If the model passes the test, it is considered the optimal model; if it fails, the second model with the lowest *AIC* and *SC* value is chosen, and the relevant diagnostic tests are run until the appropriate model is found. Based on these criteria, ARIMA (*3,1,5*) is the optimal model.

4.4. Model Selection and Diagnostic Tests

Tables 3 and 4 indicate the estimated results for the chosen ARIMA (*3,1,5*) model and diagnostic tests, respectively.

Table 4: Estimation Results of the ARIMA (3,1,5) Model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.551335	5.195932	1.645775	0.1020
AR(3)	-0.159669	0.083095	-1.921518	0.0567
MA(5)	0.359797	0.082110	4.381889	0.0000
R-squared	0.099854	*Mean dependent var*		7.756071
Adjusted R-squared	0.087265	*S.D. dependent var*		56.37517
S.E. of regression	53.85924	*Akaike info criterion*		10.83624
Sum squared resid	414817.0	*Schwarz criterion*		10.89754
Log likelihood	-788.0453	*Hannan-Quinn criter.*		10.86115
F-statistic	7.931573	*Durbin-Watson stat*		2.049177
Prob(F-statistic)	0.000541			

C has a coefficient value of 8.551335, and *t-Statistic* is equal to 1.645775 with *p-value* 0.1020. AR (3) coefficient is estimated to be -0.159669and *t-Statistic* is equal to -1.921518 with *p-value* 0.0567. On the other hand, MA *(5)* has a coefficient value of 0.359797 and *t-Statistic* is equal to 4.381889 with *p-value* 0.0000. ARIMA (3,1,5) model estimation is:

$$D(SPI) = 8.551335 - 0.159669D(SPI)_{t-1} + 0.359797\varepsilon_{t-1} + \varepsilon_t \quad (2)$$

Table 5: Diagnostic Tests Result of the ARIMA (3,1,5) Model

Diagnostic Tests			
Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.119958	*Prob. F(2,141*	0.8870
*Obs*R-squared*	0.248001	*Prob. Chi-Square(2)*	0.8834
Heteroskedasticity Test: White			
F-statistic	0.563024	*Prob. F(9,136)*	0.8253
*Obs*R-squared*	5.244405	*Prob. Chi-Square(9)*	0.8125
Scaled explained SS	6.411169	*Prob. Chi-Square(9)*	0.6982
Normality Test			
Jarque Bera	1.945047	*Probability*	0.378128
Ramsey RESET Test			
	Value	*df*	*Probability*
t-statistic	0.777283	142	0.4383
F-statistic	0.604170	(1, 142)	0.4383
Likelihood ratio	0.462410	1	0.4965

The diagnostic tests in Table 4 show that there is no heteroskedasticity where *p-value* (0.6982) is greater than 5%. Moreover, *LM Test* (*p-values* 0.7304) reveals that the model has no serial correlation. Finally, the *Ramsey RESET test* confirms the stability of the chosen model because the *p-value* (0.4965) is greater than the *threshold* of 5%.

At the next step, the autocorrelation and partial autocorrelation function graphs of the residual series and squared residuals were checked. The graphs have been illustrated in Figure 7 and Figure 8. On the graph of series' residuals indicates that the bar at lag 0 to lag 37 at the graph of white noise process is located below the significant line. According to the graph, the *p-value* for lag 0 to lag 37 are greater than 0.05. So, it means we cannot reject *Null Hypothesis* (Residuals are white noise). These results imply that the residuals are white noise, which indicates that the model is valid.

Figure 7. The Autocorrelation and Partial Autocorrelation Function Graphs for D(LOGSPI) Series' Residuals

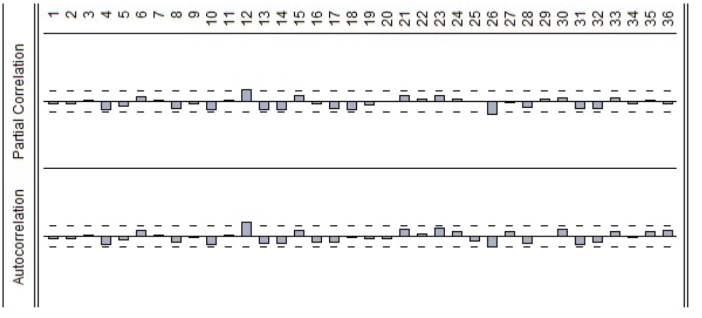
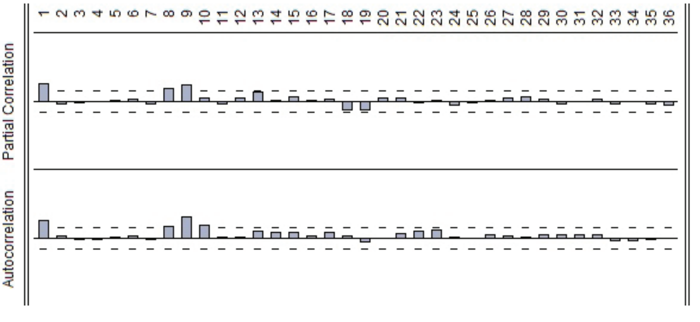
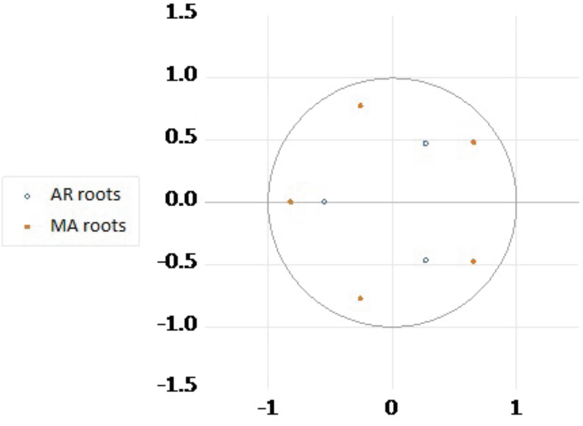


Figure 8. The Autocorrelation and Partial Autocorrelation Function Graphs for D(LOGSPI) Squared Residuals



The next step is to check if estimated ARIMA process is (covariance) stationary or not and check ARIMA process is invertible or not. The results of ARIMA process have been illustrated in Figure 9. As seen from Figure 8, *AR* and *MA* roots are located within the unit circle. So, it means that ARIMA (*3,1,5*) process is stationary and invertible.

Figure 9. The Results of ARIMA (3,1,5) Process



4.5. Data Forecasting

After ensuring that the residuals are white noise and ARIMA process is (covariance) stationary and invertible, so we can forecast with this ARIMA (*3,1,5*) model. ARIMA (*3,1,5*) model is used to forecast the stock market price index from 2021-M01 to 2021-M03. The static forecast has been chosen because of better performance than the dynamic one. Table 5 shows the ARIMA (*3,1,5*) static forecast statistical performance measures showing that the statistic forecast has lower *RMSE*, *MAR*, and *MAPE* values. Additionally, since ARIMA (*3,1,5*) is the only model with significant coefficients and passed all diagnostic tests, no other models were considered.

Table 6: The Statistical Performance Measures of the ARIMA (3,1,5) Model

Forecast Sample: 2020M01-21M03	*ARIMA (3,1,5)*	
	Static Forecast	***Dynamic Forecast***
Root Mean Squared Error	*88.05508*	*148.9128*
Mean Absolute Error	*65.88420*	*134.6992*
Mean Abs. Percent Error	*5.733617*	*11.68781*
Theil Inequality Coefficient	*0.036609*	*0.060916*

Figure 10 shows that the real values of the stock market price index closely follow the forecasted value, indicating that the developed model can accurately predict the stock market price index.

Figure 10. Static Forecasting- Actual and Fitted at Level and First Difference

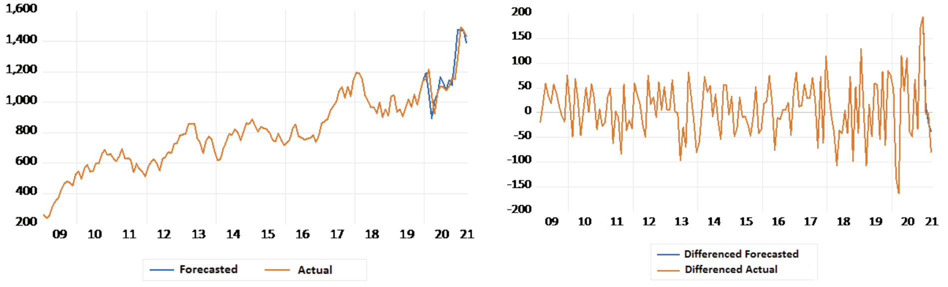


Table 6 displays the predicted values of the best model, ARIMA (3, 1, 5) for Istanbul stock market price index. The data appear to be very similar. So, it means that using the ARIMA (*3, 1, 5*) model is a correct decision and the results are very close to real figures.

Table 7: Sample of Empirical Results of ARIMA (3, 1, 5) of the Stock Market Price Index in Istanbul

Sample Period	Actual Values	Predicted Values
2021-M01	1473.45	1492.04
2021-M02	1471.39	1471.45
2021-M03	1391.73	1433.83

5. Conclusion

The study aims to investigate the application of autoregressive integrated moving averages (ARIMA) for forecasting the stock market price index of Istanbul for the period from 2009-M01 to 2021-M03. After applying the *Box-Jenkins analysis*, the findings revealed that the stock market price index of Istanbul can be determined using ARIMA approach. As compared to all other tentative models, the research shows that ARIMA (*3,1,5*) model is the best fit model for predicting the stock market price index. Forecasting is conducted by using the developed model ARIMA (*3,1,5*), and the results indicated that the forecasted values are very similar to the actual ones, reducing forecast errors. In general, the stock market price index in Istanbul; showed a downwards trend over the forecasted period. These results are almost similar to the studies of Baba and Kozaki (1992), Kamijo and Tanigawa (1990), Kihoro and Okango (2014), Nandakumar, Uttamraj, Vishal, and Lokeswari (2018), Pai and Lin (2005), Subing and Kusumah (2017), Yoo (2007) and Yoon and Swales (1991).

The results of the study can set an example for researchers and practitioners working in the stock market and can be a guide for economic decision units and investors in the stock market. The number of indicators in the obtained data set can be increased and it is predicted that in future studies, hourly, weekly and monthly data can be added to increase the amount of data and to obtain results with higher accuracy.

Peer-review: Externally peer-reviewed.
Conflict of Interest: The author has no conflict of interest to declare.
Grant Support: The author declared that this study has received no financial support.

References

Al-Zeaud, H. A. (2011). Modelling and forecasting volatility using ARIMA Model. *European Journal of Economics, Finance & Administrative Science*, *35*, 109–125.

Atsalakis, G. S., & Valavanis, K. P. (2009). Forecasting stock market short-term trends using a neuro-fuzzy based methodology. *Expert systems with Applications*, *36*(7), 10696–10707.

Atsalakis, G. S., Dimitrakakis, E. M., & Zopounidis, C. D. (2011). Elliott wave theory and neuro-fuzzy systems, in stock market prediction: The WASP system. *Expert Systems with Applications*, *38*(8), 9196–9206.

Baba, N., & Kozaki, M. (1992). An intelligent forecasting system of stock price using neural networks. In *[Proceedings 1992] IJCNN International Joint Conference on Neural Networks* (Vol. 1, pp. 371–377).

Bircan, H., & Karagöz, Y. (2003). Box-Jenkins modelleri ile aylık döviz kuru tahmini üzerine bir uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Dergisi,* (6), 49–62.

Çevik, O (2002). İMKB endeksinin Box-Jenkins yöntemi ile modellenmesi. *Afyon Kocatepe Üniversitesi İİBF Dergisi*, (C.IV, S.1), 17–31.

Chang, C. L., Sriboonchitta, S., & Wiboonpongse, A. (2009). Modelling and forecasting tourism from East Asia to Thailand under temporal and spatial aggregation. *Mathematics and computers in simulation*, *79*(5), 1730–1744.

Contreras, J., Espinola, R., Nogales, F. J., & Conejo, A. J. (2003). ARIMA models to predict next-day electricity prices. *IEEE transactions on power systems*, *18*(3), 1014–1020.

Datta, K. (2011). ARIMA forecasting of inflation in the Bangladesh Economy. *IUP Journal of Bank Management*, *10*(4), 7-15.

Etuk, H. E., Uchendu, B., & Udo, E. O. (2012). Box-Jenkins modeling of Nigerian stock prices data. *Greener Journal of Science Engineering and Technological Research, 2*(2), 32–38.

Gupta, S., & Kashyap, S. (2015). Box Jenkins approach to forecast exchange rate in India. *Prestige International Journal of Management and Research, 8*(1), 1–11.

Kamijo, K. I., & Tanigawa, T. (1990). Stock price pattern recognition-a recurrent neural network approach. In *1990 IJCNN international joint conference on neural networks* (pp. 215–221).

Kharimah, F., Usman, M., Widiarti, W., & Elfaki, F. A. (2015). Time series modeling and forecasting of the consumer price index Bandar Lampung. *Science International*, *27*(5 (B)), 4619–4624.

Kihoro, J. M., & Okango, E. L. (2014). Stock market price prediction using artificial neural network: an application to the Kenyan equity bank share prices. *Journal of Agriculture, Science and Technology*, *16*(1), 160–171.

Kock, A. B., & Teräsvirta, T. (2013). Forecasting the Finnish consumer price inflation using artificial neural network models and three automated model selection techniques. *Finnish Economic Papers*, *26*(1), 13–24.

Kumar, K., Yadav, A. K., Singh, M. P., Hassan, H., & Jain, V. K. (2004). Forecasting Daily Maximum Surface Ozone. *Journal of the Air & Waste Management Association, 54*(7)*,* 80–814

Kumar, M., & Thenmozhi, M. (2014). Forecasting stock index returns using ARIMA-SVM, ARIMA-ANN, and ARIMA-random forest hybrid models*. International Journal of Banking, Accounting and Finance, 5*(3), 284–308.

Liu, Q., Liu, X., Jiang, B., & Yang, W. (2011). Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. *BMC infectious diseases*, *11*(1), 1–7.

Merh, N., Saxena, V. P., & Pardasani, K. R. (2010). A comparison between hybrid approaches of ANN and ARIMA for Indian stock trend forecasting. *Business Intelligence Journal*, *3*(2), 23–43.

Meyler, A., Kenny, G., & Quinn, T. (1998). Forecasting Irish inflation using ARIMA models.

Mitra, S. K. (2009). Optimal combination of trading rules using neural networks. *International business research*, *2*(1), 86–99.

Mostafa, M. M. (2010). Forecasting stock exchange movements using neural networks: Empirical evidence from Kuwait. *Expert systems with applications*, *37*(9), 6302–6309.

Nandakumar, R., Uttamraj, K. R., Vishal, R., & Lokeswari, Y. V. (2018). Stock price prediction using long short term memory. *International Research Journal of Engineering and Technology*, *5*(03), 1–9.

Nyoni, T. (2018). Modeling and forecasting inflation in Kenya: Recent insights from ARIMA and GARCH analysis. *Dimorian Review*, *5*(6), 16–40.

Nyoni, T., & Nathaniel, S. P. (2018). Modeling rates of inflation in Nigeria: An application of ARMA, ARIMA and GARCH models. *MPRA Paper No. 91351,* 1–29.

Pai, P. F., & Lin, C. S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega*, *33*(6), 497–505.

Saibu, O. (2015). Determining optimal crude oil price benchmark in Nigeria: An empirical approach. *Romanian Economic Journal Year XVIII no, 58.*

Sekreter, A., & Gursoy, A. (2014). Combining forecasting method vs. individual forecasting methods: Evidence from Istanbul Stock Exchange National 100 Index. *The Empirical Economics Letters, 13*(7), 735–743.

Shah, D., Campbell, W., & Zulkernine, F. H. (2018). A comparative study of LSTM and DNN for stock market forecasting. In *2018 IEEE International Conference on Big Data (Big Data)* (pp. 4148–4155).

Sterba, J., & Hilovska, K. (2010). The implementation of hybrid ARIMA neural network prediction model for aggregate water consumption prediction. *Aplimat—Journal of Applied Mathematics*, *3*(3), 123–131.

Subing, H. J. T., & Kusumah, R. W. R. (2017). An empirical analysis of internal and external factors of stock pricing: Evidence from Indonesia. *Problems and Perspectives in Management*, *15*(4), 178–87.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics.* Boston, MA: Pearson.

Takahashi, T., Tamada, R., & Nagasaka, K. (1998). Multiple line-segments regression for stock prices and long-range forecasting system by neural network. In *Proceedings of the 37th SICE Annual Conference. International Session Papers* (pp. 1127–1132).

Tsitsika, E. V., Maravelias, C. D., & Haralabous, J. (2007). Modeling and forecasting pelagic fish production using univariate and multivariate ARIMA models. *Fisheries science*, *73*(5), 979–988.

Uko, A. K., & Nkoro, E. (2012). Inflation forecasts with ARIMA, vector autoregressive and error correction models in Nigeria. *European Journal of Economics, Finance & Administrative Science*, *50*, 71–87.

Yaziz, S. R., Ahmad, M. H., Nian, L. C., & Muhammad, N. (2011). A comparative study on Box-Jenkins and Garch models in forecasting crude oil prices. *Journal of applied sciences, 11*(7), 1129–1135.

Yoo, S. (2007). Neural Network Model vs. SARIMA Model in Forecasting Korean Stock Price Index (KOSPI). *Issues in Information Systems*, *8*(3), 372–378.

Yoon, Y., & Swales, G. (1991). Predicting stock price performance: A neural network approach. In *Proceedings of the twenty-fourth annual Hawaii international conference on system sciences* (Vol. 4, pp. 156–162).

Zhang, G. P. (2003). Time Series Forecasting using A Hybrid ARIMA and Neural Network Model. *Neurocomputing*, *50,* 159–175.