International Journal of Latest Research in Humanities and Social Science (IJLRHSS)
Volume 06 - Issue 01, 2023
www.ijlrhss.com || PP. 39-47

Azerbaijan Natural Gas Production 2022-2024 Forecast by ARIMA (0,1,2) Method

Taleh Mammadov
A Doctoral Student
Kütahya Dumlupınar University, Faculty of Economics, Turkey
talehsekili55@gmail.com

Abstract: The aim of the study is to estimate the Azerbaijani natural gas production data between 1950-2021 and the natural gas production between 2022-2024 with the ARIMA method. The ARIMA method is a method of prediction with the autoregressive integrated moving average method proposed by (Box-Jenkins, 1976). In the ARIMA method, stagnation is important in series. For this reason, the series with logarithmic transformation were subjected to unit root test and it was examined from which order they were static. It was seen that there was no stationary at the serial level and that it was stationary when the first difference was taken. In the next stage, the most suitable model for our series was determined to be ARIMA (0,1,2) and when the residue correlogram was drawn, the AC and PAC coefficients were found to be statistically insignificant. In the continuation of the study, a forecast for the year 2022-2024 was realized to produce natural gas in Azerbaijan.
Keywords: Natural gas production, Time series, ARIMA method, Azerbaijan

Introduction

Natural gas is mainly used as fuel and raw material for production. For domestic stoves, water heaters and cookers. As an industrial fuel in brick, cement, and tile kilns; in glassmaking; for generating steam in kettles; and as a source of clean heat for instrument sterilization and food processing. As a raw material for petrochemical manufacturing, natural gas is used to produce hydrogen, sulfur, soot and ammonia. Ammonia is used in a range of fertilizers and as a secondary raw material in the manufacture of other chemicals, including nitric acid and urea. Ethylene, an important petrochemical, is also made from natural gas. Natural gas is considered an environmentally friendly, clean fuel with significant environmental benefits over other fossil fuels.Its superior environmental quality is negligible emissions of sulfur dioxide, or lower emissions of nitrogen oxides and carbon dioxide compared to coal or crude oil. This helps reduce problems caused by acid rain, the ozone layer or greenhouse gases. Natural gas is also a very safe energy source when transported, stored and used . Historically, natural gas has been discovered during the exploration for crude oil. Natural gas is often an unwanted by-product as gas deposits are extracted during drilling and workers are forced to stop drilling and allow the gas to vent freely into the air. Now, especially after the oil shortages of the 1970s, natural gas is an important energy source in the world. Throughout the 19th century, natural gas was used almost exclusively as a light source, and its use remained localized due to the lack of transport structures, which made it difficult to transport large quantities of gas over long distances. A major change occurred with the invention of the leak-proof pipe joint in 1890, but it wasn't until the 1920s that delivering natural gas to customers over long distances became feasible due to advances in pipeline technology. Moreover, it was only after World War II that the use of natural gas grew rapidly because of the development of pipeline networks and storage systems (Mokhatab, Poe, & Speight, 2006).

When we look at Table 1, the figures in terms of natural gas production (trillion cubic meters) estimated in the world in 2020-2050 are included. When we look at it, there is an increasing trend in natural gas production in the world. The amount of natural gas produced in 2020 is estimated to be 141.7 trillion cubic meters, in 2025 this figure will be 152.4 trillion cubic meters, and in 2050, natural gas production will reach 186.0 trillion cubic meters. The average annual percentage change is 0.9%. When we look at the same table, the highest production takes place in the United States. The amount of natural gas produced in 2020 is 33.9, this figure is estimated to reach 43.0 in 2050. In second place, Russia is on the list of the countries that produce the most natural gas in the world. Natural gas production in 2020 is estimated to be 24.2 and 25.5 trillion cubic meters in 2050.

Table 1. World natural gas production by region, reference case

Region	2020	2025	2030	2035	2040	2045	2050	change, 2020–2050
United States	33.9	36.3	37.9	38.6	39.9	41.5	43	0.8
Canada	6.2	6.1	6.7	7.5	8.1	8.6	8.9	1.2
Mexico and other OECD Americas	0.9	1	1.2	1.5	1.7	1.9	1.8	2.4
OECD Europe	8.6	8.4	7.1	6.1	5.3	5	5	-1.8
OECD Asia	5.6	6.1	6.3	6.9	7.2	7.1	7.2	0.8
Japan	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
South Korea	0	0	0	0	0	0	0	0.1
Australia and New Zealand	5.5	6	6.2	6.8	7.1	7.1	7.1	0.8
Russia	24.2	27.5	29.3	30.5	32.2	34	35.5	1.3
Other Europe and Eurasia	7.3	8.2	8.2	8.9	9.8	10.4	10.6	1.3
Non-OECD Asia	18.7	20.4	22.2	22.4	22.3	23.5	26.2	1.1
China	6.6	7.4	8.5	8.6	8.3	8.9	10.2	1.5
India	1	1.1	1.6	1.9	2.2	2.9	4	4.7
OtherAsia	11.1	11.9	12.2	11.9	11.8	11.6	12	0.3
Middle East	23	24.4	26.5	26.6	27.9	28.4	28.8	0.8
Africa	7.2	8.1	8.8	9.8	11	11.4	12.1	1.7
Non-OECD Americas	6.1	6.2	6	6	6.1	6.2	6.7	0.3
Brazil	1	1.1	1	0.9	0.8	0.8	0.9	-0.3
OtherNon-OECD Americas	5	5	5	5.1	5.2	5.4	5.8	0.4
Total Non-OECD	86.4	94.6	101	104.2	109.3	114	120	1.1
Total World	141.7	152.4	160.2	164.9	171.4	178.1	186	0.9

Source: (U.S. Energy Information Administration, 2022)

In Table 2, natural gas consumption figures are given in the world between 2020-2050. While it is 141.7 trillion cubic meters in 2020, this figure increases to 186 trillion cubic meters in 2050. On a yearly basis, average consumption is projected to occur the most in India with 4.3% and the lowest in Japan with -1.1%.

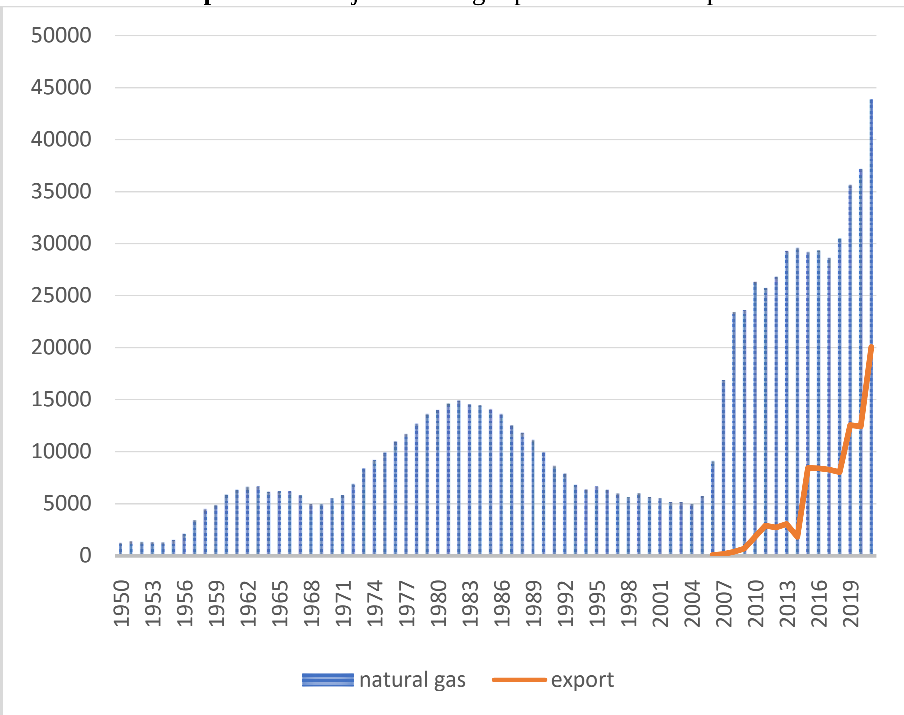
Table 2. World natural gas consumption by region, reference case

Region	2020	2025	2030	2035	2040	2045	2050	change, 2020-2050
United States	30.8	30.8	31	31.4	32.7	34.1	35.4	0.5
Canada	4.3	4.6	4.9	5.4	5.9	6.3	6.6	1.5
Mexico and other OECD Americas	3.4	4.2	4.5	5	5.2	5.4	5.4	1.5
OECD Europe	19	20	20.3	20.2	20.2	20.7	21.3	0.4
OECD Asia	7.8	7.7	7.7	7.3	7	6.7	6.7	-0.5
Japan	3.8	3.5	3.5	3.2	2.9	2.8	2.7	-1.1
South Korea	2.2	2.2	2.3	2.2	2.1	2	2	-0.3
Australia and New Zealand	1.8	1.9	2	2	2	1.9	2	0.3
Russia	17.6	18	18.9	19.4	20.1	20.7	21.2	0.6
Other Europe and Eurasia	6.8	7.2	7.4	7.5	7.6	7.8	8	0.6
Non-OECD Asia	23	27.7	31.5	33.8	36.6	39.4	42.5	2.1
China	11	14	16.2	17.4	18.6	19.9	21	2.2
India	2	2.7	3.5	4.3	5.1	6	7	4.3
OtherAsia	10.1	11	11.7	12.2	12.8	13.5	14.5	1.2
Middle East	18.3	20.8	21.7	21.7	22.5	23	23.3	0.8
Africa	5.4	5.8	6.5	7.2	7.6	7.8	8.7	1.6
Non-OECD Americas	5.4	5.7	5.7	5.9	6.1	6.3	6.8	0.8
Brazil	1.3	1.7	1.6	1.5	1.5	1.4	1.5	0.6
OtherNon-OECD Americas	4.2	4	4.1	4.4	4.6	4.9	5.3	0.8
Total Non-OECD	76.5	85.2	91.8	95.6	100.5	105	110.5	1.2
Total World	141.7	152.4	160.2	164.9	171.4	178.1	186	0.9

Source: (U.S. Energy Information Administration, 2022)

In the information published in 2021, BP points out that the volume of Azerbaijan's confirmed natural gas resources is 2.5 trillion cubic meters (Statistical Review of World Energy 2021, 2021).When we look at Chart 1, data on Azerbaijani natural gas production between 1950-2021 and natural gas exports in 2006-2021 are included. Looking at the chart, it is seen that Azerbaijani natural gas production is gaining an upward trend and the increase in production is higher with the export of natural gas especially after 2006. The pipeline, which was inaugurated in 2006 (Baku-Tbilisi-Erzurum Natural Gas Pipeline) and opened in 2018 (Trans Anatolian natural gas pipeline), which exports the natural gas in the Azerbaijani Shah Deniz field to Europe through Turkey, has an important place in Azerbaijan's natural gas exports. Looking at recent years, especially in 2021, compared to 2020, Azerbaijan increased its total gas exports by 61% and exported 20 billion cubic meters of gas, and about 1/3 of this was realized to European countries. The most important reason for this was the completion of the construction of the Trans-Adriatic pipeline on October 13, 2020. Prior to this, the pipeline had begun to operate in test mode, but all construction work was completed in October, and natural gas exports to Europe via the pipeline began from the beginning of 2021 (Azərbaycan Respublikasının Dövlət Statistika Komitəsi, 2022).

Graph 1: Azerbaijan natural gas production and export

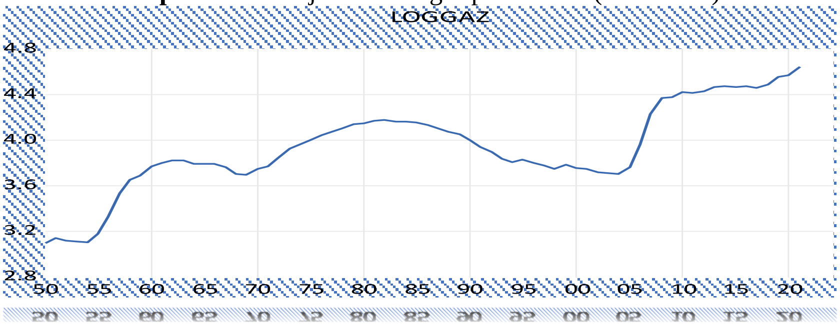


Source: (Azərbaycan Respublikasının Dövlət Statistika Komitəsi, 2022)

The war that has occurred in recent years, especially between Russia and Ukraine, and the embargo imposed on Russia by America, Britain, and many western states due to this war and Russia's export of Russian gas under various conditions by using it as a weapon, especially drag European states to search for alternative gas. For this reason, Azerbaijan and many Arab states are alternatives. The frequent meetings of Azerbaijani and European officials in recent times indicate that Azerbaijan's natural gas exports to Europe will increase in the future.

1. Data and Method

Graph 2: Azerbaijan natural gas production (LOGGAZ)



Our series (1950-2021) is obtained annually from the data sheet of **the State Statistical Committee of the Republic of Azerbaijan**. Then, the logarithmic value of the series was taken and loaded into the program. Since it is annual, no seasonality has been seen.

1.1. Augmented Dickey-Fuller test (ADF)

The concept of stagnation is important in time series, if work on non-stationary series takes place, then false regression is used, even if statistically significant answers are found(Mert & Çağlar, 2019).

$y_t = \varphi y_{t-1} + \varepsilon_t$(1.1)

It takes advantage of the first difference of the random gait process given in the standard Dickey-Fuller (DF) test (equation 1.1) (Dickey & Fuller, 1976). For the first difference of equation (1.2), the equation is y_{t-1} subtracted from both sides.

$y_t - y_{t-1} = \varphi y_{t-1} - y_{t-1} + \varepsilon_t$

$\Delta y_t = (\varphi - 1)y_{t-1} + \varepsilon_t$ and $\delta = \varphi - 1$ the following equation (1.2) is obtained.

$\Delta y_t = \delta y_{t-1} + \varepsilon_t$(1.2)

Model number (1.2) is the unfixed and trendless model. Fixed model (1.3), fixed and trended model (1.4) is given.

$$\Delta y_t = \mu + \delta y_{t-1} + \varepsilon_t \qquad (1.3)$$

$$\Delta y_t = \mu + \beta t + \delta y_{t-1} + \varepsilon_t \qquad (1.4)$$

The hypothesis for standard DF testing is given below.

$H_0: \delta = 0$ (has a unit root)

$H_s: \delta < 0$ (no unit root)

However, if there is a high degree of correlation in the series, this time ε_t will lose its clean series. To solve this problem, the Augmented Dickey-Fuller (ADF) test uses the AR(p) process rather than the AR (1) process, adding the terms p delayed difference to the equation (Dickey & Fuller, 1981). Equations (1.5), (1.6), (1.7)

$$\Delta y_t = \delta y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t \qquad (1.5)$$

$$\Delta y_t = \mu + \delta y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t \qquad (1.6)$$

$$\Delta y_t = \mu + \beta t + \delta y_{t-1} + \sum_{i-1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t \qquad (1.7)$$

Table 3: Augmented Dickey-Fuller (ADF)unit root test result

Null Hypothesis: LOGGAZ has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.593940	0.4803
Test critical values:	1% level	-3.527045	
	5% level	-2.903566	
	10% level	-2.589227	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: D(LOGGAZ) has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.398330	0.0143
Test critical values:	1% level	-3.527045	
	5% level	-2.903566	
	10% level	-2.589227	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: LOGGAZ has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.609793	0.2775
Test critical values:	1% level	-4.094550	
	5% level	-3.475305	
	10% level	-3.165046	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: D(LOGGAZ) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.355762	0.0659
Test critical values:	1% level	-4.094550	
	5% level	-3.475305	
	10% level	-3.165046	

*MacKinnon (1996) one-sided p-values.

Looking at the results of the increased AugmentedDickey-Fuller test, our absence hypothesis is established as the absence hypothesis (LOGGAZ contains the root of units) for our LOGGAZ series. The series was first analyzed at the level value (intercept model) and the optimal delay length was taken as 1 according to the SIC information criterion. $t_\delta = -1.593940$ and the absence hypothesis cannot be rejected because the calculated test statistic is higher than the (Mac Kinnon,1996) critical values. In summary, the serial level value contains the volume root. The same result applies to the trend and interceptsmodel.

Since the series are not static in their level values, the first difference is calculated by taking I (1). And since the statistical value calculated for the intercept model is less $t_\delta = -3.398330$ than the critical values of 5% and 10%, the absence hypothesis cannot be made. Therefore, it is seen that our series is a stationary series when the difference is taken. In the same way, when the difference between the trend and intercept model is taken, it is seen that it is a stationary series.

1.2. Ng-Perron unit root test

Ng-Perron (2001) proposed four different test statistics based on the train-free GLS series. These statistics are $Zy_t^{d}{}'_a$ andZt statistics as modified PP statistics, R_1 statistics and ERS test statistics (Mert & Çağlar, 2019). Modified test statistics equation (1.8), (1.9), (1.10), (1.11)

$$MZ_a^d = (\frac{T^{-1}(y_T^d)^2 - f_0}{2k} \qquad (1.8)$$

$$MZ_t^d = MZ_a * MSB \qquad (1.9)$$

$$MSB^d = (k/f_0)^{1/2} \qquad (1.10)$$

$$MP_T^d = \begin{cases} (\bar{c}^2 k - \frac{\bar{c}T^{-1}(y_T^d)^2)}{f_{0,}} x_t = \{1\} \\ (\bar{c}^2 k + (1 - \frac{\bar{c})T^{-1}(y_T^d)^2)}{f_0} x_t = \{1, t\} \end{cases} \qquad (1.11)$$

Table 4: Ng-Perron unit root test result

Null Hypothesis: LOGGAZ has a unit root
Exogenous: Constant
Lag length: 1 (Spectral GLS-detrended AR based on SIC, maxlag=11)
Sample: 1950 2021
Included observations: 72

		MZa	MZt	MSB	MPT
Ng-Perron test statistics		-1.36372	-0.51672	0.37890	11.2347
Asymptotic critical values*:	1%	-13.8000	-2.58000	0.17400	1.78000
	5%	-8.10000	-1.98000	0.23300	3.17000
	10%	-5.70000	-1.62000	0.27500	4.45000

*Ng-Perron (2001, Table 1)

Null Hypothesis: LOGGAZ has a unit root
Exogenous: Constant, Linear Trend
Lag length: 1 (Spectral GLS-detrended AR based on SIC, maxlag=11)
Sample: 1950 2021
Included observations: 72

		MZa	MZt	MSB	MPT
Ng-Perron test statistics		-12.4780	-2.49072	0.19961	7.34248
Asymptotic critical values*:	1%	-23.8000	-3.42000	0.14300	4.03000
	5%	-17.3000	-2.91000	0.16800	5.48000
	10%	-14.2000	-2.62000	0.18500	6.67000

*Ng-Perron (2001, Table 1)

Null Hypothesis: D(LOGGAZ) has a unit root
Exogenous: Constant
Lag length: 0 (Spectral GLS-detrended AR based on SIC, maxlag=11)
Sample (adjusted): 1951 2021
Included observations: 71 after adjustments

		MZa	MZt	MSB	MPT
Ng-Perron test statistics		-16.4460	-2.85581	0.17365	1.53405
Asymptotic critical values*:	1%	-13.8000	-2.58000	0.17400	1.78000
	5%	-8.10000	-1.98000	0.23300	3.17000
	10%	-5.70000	-1.62000	0.27500	4.45000

*Ng-Perron (2001, Table 1)

Null Hypothesis: D(LOGGAZ) has a unit root
Exogenous: Constant, Linear Trend
Lag length: 0 (Spectral GLS-detrended AR based on SIC, maxlag=11)
Sample (adjusted): 1951 2021
Included observations: 71 after adjustments

		MZa	MZt	MSB	MPT
Ng-Perron test statistics		-17.3136	-2.91389	0.16830	5.43646
Asymptotic critical values*:	1%	-23.8000	-3.42000	0.14300	4.03000
	5%	-17.3000	-2.91000	0.16800	5.48000
	10%	-14.2000	-2.62000	0.18500	6.67000

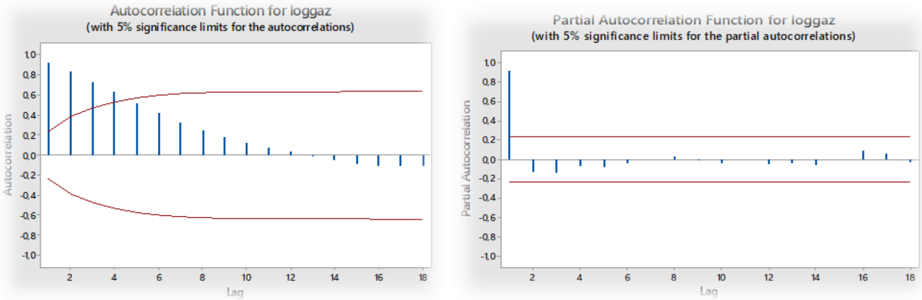
*Ng-Perron (2001, Table 1)

In the fixed model of the (Ng-Perron, 2001) unit root test, our absence hypothesis is that it contains serial unit roots. In the SIC information criterion, the maximum delay length was taken from 11 delays. MZ_a =-1.36372, MZ_t =-0.51672, MSB=0.37890, MPT=11.2347. The absence hypothesis cannot be rejected because each test statistic is greater than the given information criteria. At the serial fixed level value, has a unit root, the series is not static. The same table applies to the trend and intercepts model.

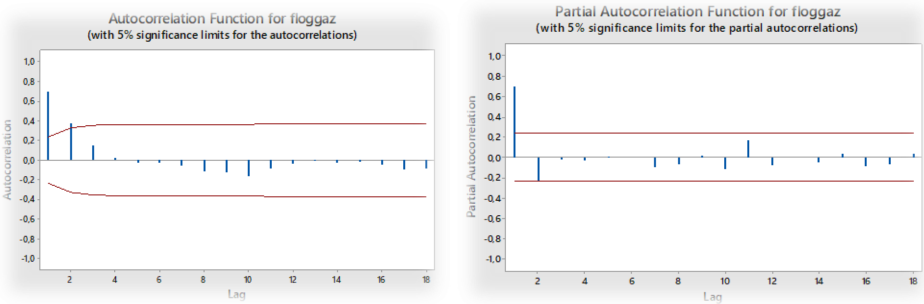
Taking the first difference in the series, the test statistic obtained in the fixed model was obtained as MZ_a =-16.4460, MZ_t =-2.85581, MSB= 0.17365, MPT= 1.53405. The series is static because the given test statistic is smaller than the given test statistic. The same table applies to the fixed and trendy models. In summary, the series is static when the first difference is taken.

An ACF chart is a bar chart that shows correlation coefficients between a time series and delayed values. In short, ACF describes how the present value of a particular time series relates to historical values. In the ACF drawing, the x-axis represents the correlation coefficient, and the y-axis represents the number of delays. PACF is a partial autocorrelation function that explains the partial correlation between series and their delays (Karcıoğlu, Tanışman&Bulut 2021). The level value of the series I (0) ACF was not static when looking at the PACF chart, the first difference is I (1) and the series is static.

Graph 3: Level value ACF, PACF distribution



Graph 4: ACF, PACF chart of series I (1)



The ARIMA method is the autoregressive integrated moving average method proposed by (Box-Jenkins, 1976). The ARIMA method is the method used by many researchers from the past to the present. The concept of stagnation is important in the model. If the series is stationary at the level value, the integration degree will be taken as zero, and if it is a series that becomes static by taking the first difference, this time integration degree will be taken as one.For this reason, the stability of our series has been checked and it has been determined that our series is static by taking the first difference I(1) and not I(0) at the level value.
ARIMA (p, d,q)
d- degree of integration
p- AR latency number
q- Shows the number of MA delays.

1.3. AR (1) model

If a series is affected by its own delayed values, it is an autoregressive AR series. AR (1) is given in the first-order autoregressive equation (Mert & Çağlar, 2019)(1.12).

$$(1-pL)(y_t-\mu)=\varepsilon_t \qquad (1.12)$$

L is the latency processor, p is the AR (1) parameter, the μ constant term, and the $\varepsilon_t \sim N(0,\sigma^2)$ clean array.
The clear version of equation (1.12) is given in (1.13).

$$y_t-\mu-pLy_t+p\mu=\varepsilon_t$$

$$y_t=\mu(1-p)+py_{t-1}+\varepsilon_t (1.13)$$

$c=\mu(1-p)$ the equation AR (1) is reached as in the open form (1.1 4) equation to be constant.

$$y_t=c+py_{t-1}+\varepsilon_t \qquad (1.14)$$

1.4. MA (1) model

The time series can be affected not by its own past values like the AR model, but by random shocks. These series are called moving average MA series. The equation MA (1) is given in (1.15) (Mert & Çağlar, 2019).

$$y_t-\mu=(1+\theta L)\varepsilon_t \qquad (1.15)$$

MA (1) opening equation of closed form (1.16)

$$y_t-\mu=\varepsilon_t+\theta L\varepsilon_t$$

$y_t = \mu + \theta\varepsilon_{t-1} + \varepsilon_t$ (1.16)

If our series is stable at the level, it will be ARMA (p,0, q) model, and if it is stationary when the first difference is taken, it will be ARIMA(p,1, q) model.

The closed equation of the ARIMA (p, d,q) model is given in (1.17).

$(1-L)^d(1-\sum_{j=1}^{p} p_j L^j)(y_t-\mu) = (1+\sum_{j-1}^{q} \theta_j L^j)\varepsilon_t$(1.17)

$(1-L)^d$the difference is the polynomial. If the series is stationary at the level, it will be equal to 1 since d = 0. When this time the model will turn into an ARMA(p,q) model.

Since our series is static in the first difference and not at the level, the closed form of the ARIMA (p,1, q) model (1.18) is shown in the equation.

$(1-L)(1-\sum_{j=1}^{p} p_j L^j)(y_t-\mu) = (1+\sum_{j-1}^{q} \theta_j L^j)\varepsilon_t$ (1.18)

The closed equation given in (1.18) will be opened as the equation given in (1.19).

$$(1-L)(y_t - \sum_{j=1}^{p} p_j L^j y_t - \mu + \sum_{j=1}^{p} p_j \mu) = \varepsilon_t + \sum_{j=1}^{q} \theta_j L^j \varepsilon_t$$

$$(1-L)(y_t - p_1 y_{t-1} - \cdots - p_p y_{t-p} - \mu + \sum_{j=1}^{p} p_j \mu) = \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q} + \varepsilon_t \; y_t - p_1 y_{t-1} - \cdots - p_p y_{t-p} - \mu + \sum_{j=1}^{p} p_j \mu$$

$$- y_{t-1} + p_1 y_{t-2} + \cdots + p_p y_{t-p-1} + \mu - \sum_{j=1}^{p} p_j \mu = \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q} + \varepsilon_t$$

$$\Delta y_t = p_1 y_{t-1} + \cdots + p_p y_{t-p} - p_1 y_{t-p} - p_1 y_{t-2} - \cdots - p_p y_{t-p-1} + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q} + \varepsilon_t$$
(1.19)

Table 5: Parameters of the ARIMA model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.023633	0.013502	1.750365	0.0846
MA(1)	0.859598	0.105204	8.170753	0.0000
MA(2)	0.341852	0.094742	3.608226	0.0006
SIGMASO	0.001830	0.000342	5.350960	0.0000

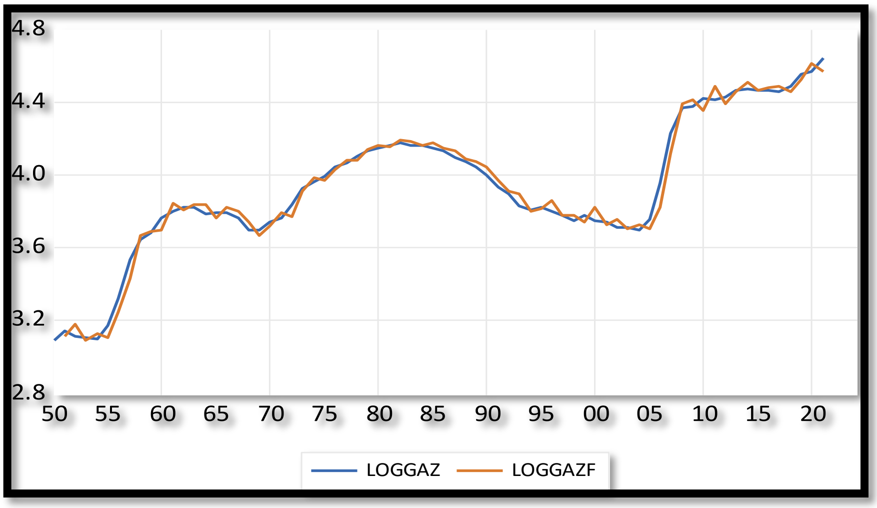
R^2 : 0.510212
Adjusted R^2 : 0.488281
Prob (F-statistic): 0.000000
DW value: 1.85 (a value close to 2 indicates that there is no autocorrelation in the remains)
C/constant 0.0846 <0. 10, MA (1) 0.0000 <0.01, MA(2) 0.0006 < 0.01 are seen to be less than0.01. For this reason, the ARIMA (0,1,2) model has been selected as the most suitable model for our series.
As another condition, the correlation graph was looked at to see if the remains were in a clean sequence.

Table 6: Residual correlation for ARIMA(0,1,2) model

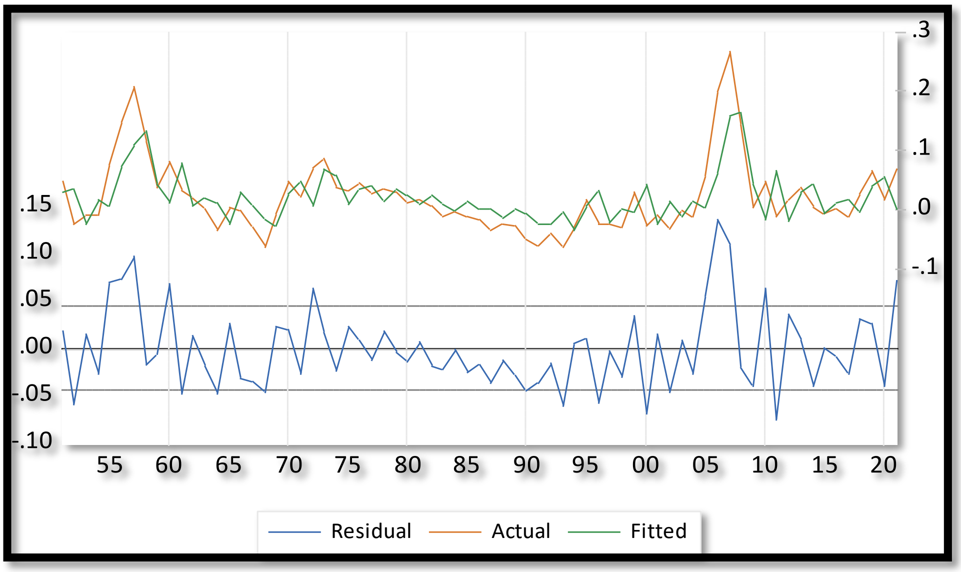
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.049	0.049	0.1754	0.675
		2	0.135	0.133	1.5402	0.463
		3	0.123	0.113	2.6918	0.442
		4	-0.044	-0.072	2.8402	0.585
		5	-0.040	-0.070	2.9679	0.705
		6	0.028	0.035	3.0304	0.805
		7	-0.018	0.010	3.0555	0.880
		8	-0.133	-0.137	4.5009	0.809
		9	0.072	0.073	4.9389	0.840
		10	-0.198	-0.174	8.2762	0.602
		11	0.001	0.034	8.2762	0.688
		12	0.005	0.021	8.2789	0.763
		13	0.015	0.054	8.2993	0.824
		14	-0.070	-0.098	8.7445	0.847
		15	0.081	0.066	9.3534	0.858
		16	0.037	0.044	9.4850	0.892
		17	-0.100	-0.090	10.452	0.884
		18	-0.005	-0.091	10.454	0.916
		19	-0.028	0.032	10.532	0.939
		20	-0.058	-0.051	10.876	0.949
		21	-0.058	-0.047	11.220	0.958
		22	0.043	0.035	11.416	0.968
		23	-0.051	0.016	11.693	0.975
		24	0.017	-0.024	11.724	0.983
		25	-0.036	-0.052	11.872	0.988
		26	0.002	0.043	11.872	0.992
		27	-0.005	-0.030	11.875	0.995
		28	-0.066	-0.118	12.407	0.995
		29	0.038	0.067	12.585	0.997
		30	-0.062	-0.042	13.066	0.997
		31	-0.049	-0.106	13.381	0.998

Looking at Table 6, it is understood that the remains of the model are clean series. The AC and PAC coefficients are statistically insignificant.

Graph 5: Graph of ARIMA (0,1,2) model forecast and actual values



Graph 6: Actual and estimated values of the LOGGAZ series with residues



2. Result

In the study, using the data of Azerbaijani natural gas production between 1950-2021, natural gas production between 2022-2024 was estimated by ARIMA method. Since the data is annual, no seasonality has been observed. In the second stage of the study, it was checked whether the series was static or not. Therefore, the series Dickey, Fuller (1981) and Ng-Perron (2001) were checked for unit root by unit root test. As a result of the test, it was concluded that our series contains unit root at the serial level because the test statistic calculated at the level value is greater than all critical values. However, when the first difference of the series is taken, it is seen that the series does not contain unit root because the calculated test statistic is smaller than all critical values. The Ng-Perron (2001) unit root test gives the same result as the ADF test. In the third stage, the ACF and PACF graphs were looked at and it was seen that they were static when the difference was taken. In the continuation of the study, it was determined that the most suitable model for our series was the ARIMA (0,1,2) model. Then, the residue correlogram of the ARIMA (0,1,2) model prediction was examined, and it was seen that it was a clean sequence. In Table 7, ARIMA model forecasts and predictions for LOGGAZ series are given. Looking at the projections for the year 2022-2024, it is seen that Azerbaijani natural gas production is increasing and increasing trend.

Table 7: ARIMA model prediction for LOGGAZ series

Coefficient	**ARIMA (0,1,2)**
C	0.023633 / 10% significant
MA (1)	0.859598 / 1% significant
MA(2)	0.341852 / 1% significant
DW	1.85
Residue	Clean array
Predictions / 3 terms	
2022	*4.79250729*
2023	*4.81613985*
2024	*4.83977241*

Bibliography

[1]. *Azərbaycan Respublikasının Dövlət Statistika Komitəsi*. (2022). https://www.stat.gov.az/: https://www.stat.gov.az/search/?q=T%C9%99bii+qaz%2C+qaz+hal%C4%B1nda%2C+min+kub+metr

[2]. Box, G.E.P. & Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control, Revised Edition, San Francisco: Holden- Day.

[3]. Dickey, D.A. & W.A. Fuller (1979), "Distribution of theestimatorsforautoregressive time serieswith a unitroot", Journal of theAmerican Statistical Association, 74, 427–431.

[4]. Dickey, D.A. &W.A. Fuller (1981), "Distribution of theestimatorsforautoregressive time serieswith a unitroot", Econometrica, 49, 1057-72.

[5]. NG, Serena& Pierre PERRON (2001), "LagLengthSelection and the Construction of UnitRootTestswithGood Size and Power", Econometrica, 69, 1519-1554

[6]. Mokhatab, S., Poe, W., & Speight, J. (2006). *Handbook Of Natural Gas Transmıssıon And Processıng.* Elsevier.

[7]. *Statistical Review of World Energy 2021*. (2021). 10 1, 2022 tarihinde www.bp.com: https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2021-full-report.pdf

[8]. *U.S. Energy Information Administration*. (2022). www.eia.gov: https://www.eia.gov/outlooks/ieo/tables_side_xls.php

[9]. Karcıoğlu, A. A. , Tanışman, S. & Bulut, H. (2021). Türkiye'de COVID-19 Bulaşısının ARIMA Modeli ve LSTM Ağı Kullanılarak Zaman Serisi Tahmini . Avrupa Bilim ve Teknoloji Dergisi ,Ejosat Özel Sayı 2021 (RDCONF) , 288-297 . DOI: 10.31590/ejosat.1039394

[10]. Mac Kinnon, J. G. (1996). "Numerical Distribution FunctionsforUnitRoot and CointegrationTests". Journal of AppliedEconometrics, 11(6), 601-618.

[11]. Mert, M., & Çağlar, A. (2019). *Eviews ve Gauss Uygulamalı Zaman Serisi Analizi.* Ankara: Detay Yayıncılık.