Statistical analysis of ARIMA models for natural gas consumption in Romania

1 Introduction

The project contains multiple ARIMA models for forecasting natural gas consumption in Romania. The dataset was collected from Eurostat and it contains monthly information about the volume of natural gas measured in TJ/GCV, starting Jan. 2008 and ending in Apr. 2024, having 196 rows.[1]

2 Dataset

The time series presents clear seasonality, the natural gas consumption being the highest in winter (Dec., Jan., Feb.), decreases until reaching a minimum point (June, July) and increases back starting with autumn months (Sep., Oct.). Before modelling the time series, it was split in train and test (last 20 values) in order to analyse better the efficiency of the models.

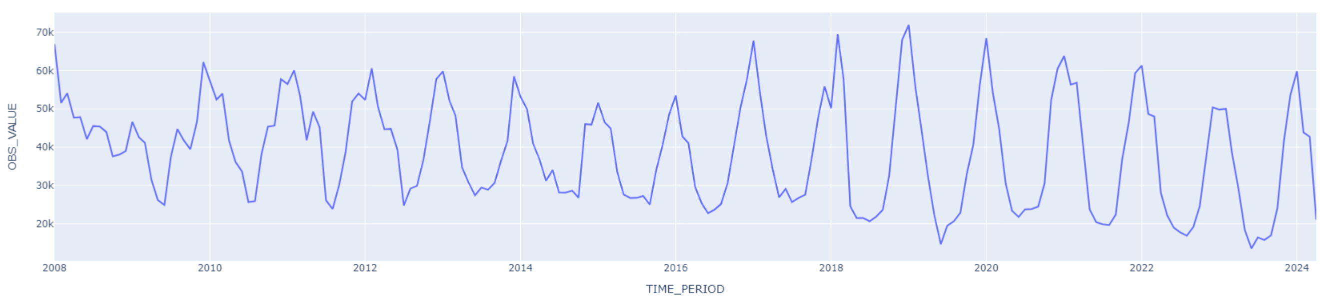

Figure 1: Plot of the consumption of natural gas in Romania, Jan. 2008 - Feb. 2024

In the case of a not stationary time series, is is important to also inspect the plot of the differenced time series, either by 1 lag or 2, as it is a method very effective in transforming the data into an appropiate stationary time series ready for modelling. The plot below contains the time series differenced by 1 lag.

Figure 2: Plot of the time series differenced by 1 lag

It was also plotted the time series differenced seasonally (by lag=12) and it is visible how much it resembles a stationary process, looking more similar to the white noise, which indicates the direction of the analysis, suggesting that the SARIMA model would be a better fit for this dataset.

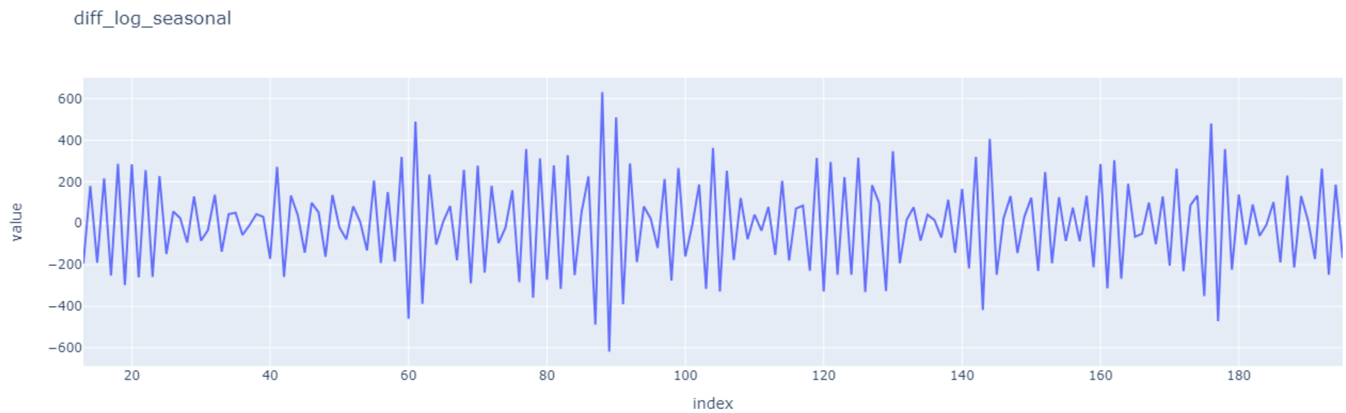

Figure 3: Plot of the time series differenced seasonally

For a better understanding of the components of the time series, there was also used a seasonality decomposition function, based on the additive model. The plots show a downward trend and yearly seasonality.

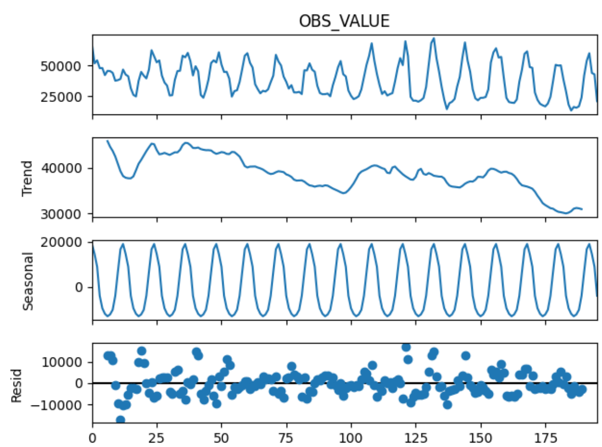

Figure 4: Plot of time series components

3 Stationarity

Before testing the model, the time series needs to be checked if it verifies the stationarity condition.

3.1 ACF and PACF

In order to find the right model, the Autocorrelation Function and the Partial Autocorrelation Function computed on the original dataset may indicate if the dataset is stationary or it needs differencing, and they can also show the right lags to use further in the ARIMA models. [2]

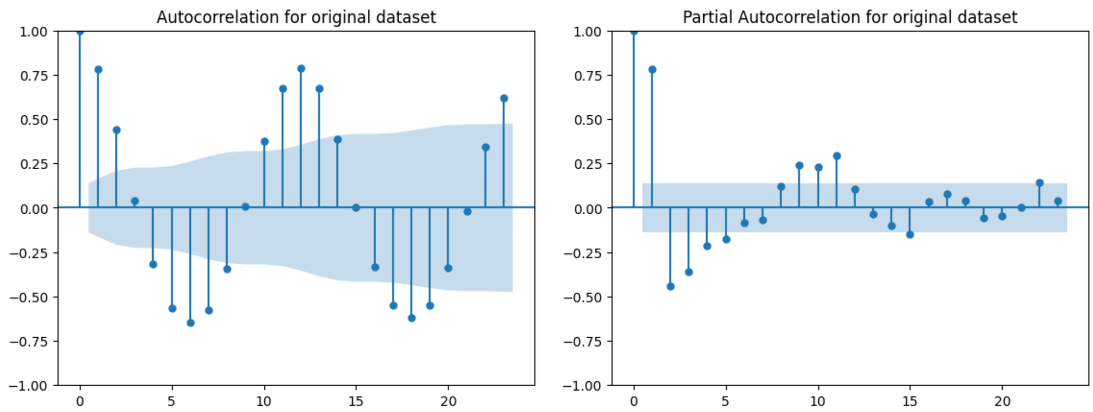

Figure 5: ACF and PACF plots for original time series

In the plots above it is noticeable that the time series is not stationary, there is high autocorrelation between the lags (high values for lags 1-6 for PACF) and seasonality visible in the ACF plot (the highest values correspond to lag 1 and lag 12, which indicate yearly seasonality).

Because the original time series is not stationary, it was also examined the differenced one to see if it is enough to proceed to the ARIMA models. In this case, the plots ACF and PACF for the differenced time series, showed below, still present significant lags, which probably mean the time series is still not stationary.

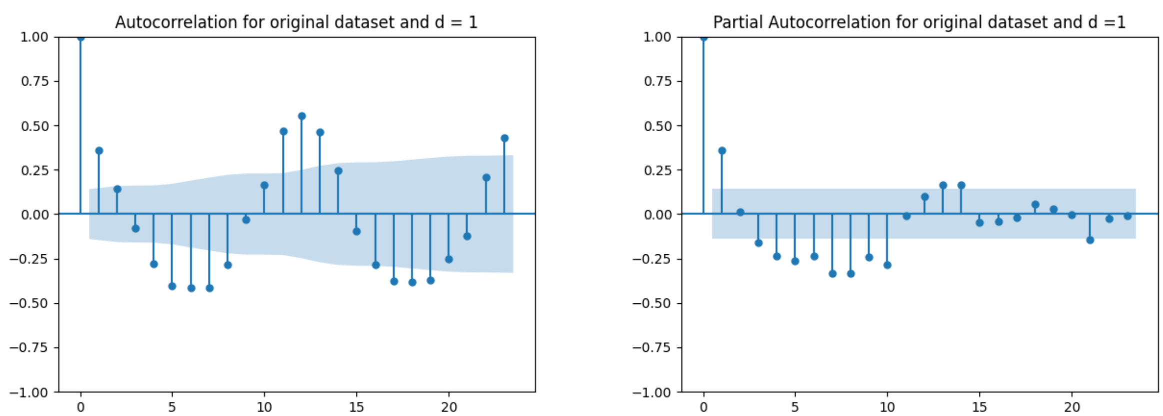

Figure 6: ACF and PACF plots for differenced time series

The time series was also seasonally differenced (lag=12) and checked if it is closer to being stationary. The ACF and PACF plots below show almost no lag correlation, except for the lag 12. This may indicate the time series is stationary and they also give ideas for models to test based on the significant lags: an ARIMA model composed of MA(2), AR(2), SMA(1), SAR(1).

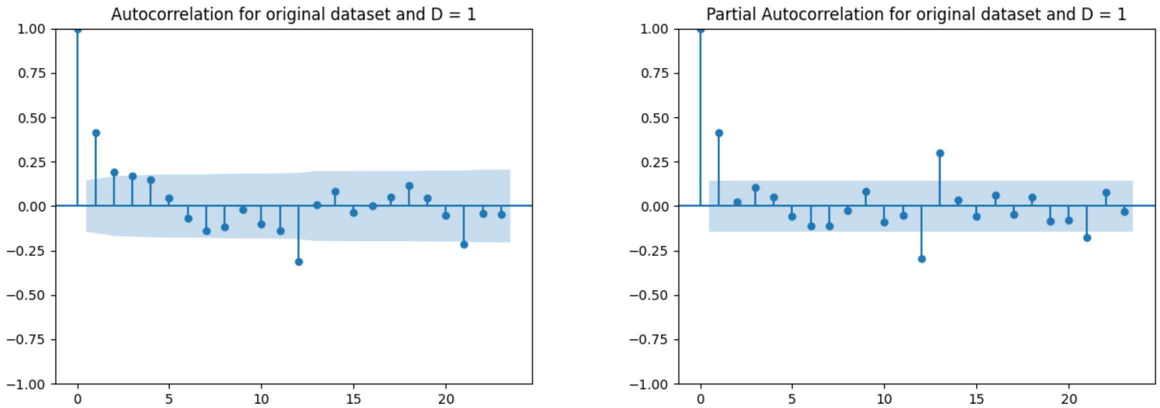

Figure 7: ACF and PACF plots for seasonal differenced time series

3.2 Unit Root Tests

To confirm which time series are stationary, there are two main unit root tests that verify this information: Augmented Dickey-Fuller test (ADF) and Kwiatkowski–Phillips–Schmidt–Shin test (KPSS). For ADF, the H_0 hypothesis states that the unit root is present, which means the time series is not stationary. [3][4] Instead, in the case of KPSS test, the H_0 claims the time series is trend-stationary. [5]

Below is presented a table with all the configured time series tested for stationarity:

Data	**Differenced Data**	**Seasonally Differenced Data**	**p-value ADF**	**p-value KPSS**
original	0	0	0.58	0.04
original	1	0	< 0.05	> 0.05
original	2	0	< 0.05	> 0.05
original	0	1	< 0.05	> 0.05
original	1	1	< 0.05	> 0.05
log	0	0	0.87	0.01
log	1	0	1.76	> 0.05
log	0	1	< 0.05	> 0.05

Table 1: Time Series Data Transformations and Stationarity Tests

Based on the tests performed, the stationary time series that can be modelled with ARIMA should be differenced or seasonally differenced.

4 Models

The ARIMA model algorithm used and most of the evaluation metrics were part of the statsmodels Python library [6].

4.1 Metrics

The ARIMA fitted models can be evaluated by several metrics:

- Log Likelihood represents the probability that the observed data came from estimated model and an increased value indicates a better model
- AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), HQIC (Hannan-Quinn information criterion) are indicators of the quality of the model; lower values are related to better models
- The coefficients of the components and their statistical significance explained through p-values; if there are parts of the model not significant statistically, they should be removed from the model

- Mean Absolute Error (MAE) and RMSE (Root Mean Squared Error) compute the overall error of the fitted model, lower values indicate a better model

ARIMA	Data type	Log Likelihood	MAE	RMSE	AIC	BIC	All P-Values signif.
(2,1,2)	orig.	-1792.68	5391.59	8361.42	3595.36	3611.18	Yes
$(2,0,2) \times (1,0,1)_{12}$	orig.	-1777.64	4482.18	6031.59	3571.28	3596.65	No
$(2,0,2) \times (1,1,1)_{12}$	orig.	-1672.18	6621.13	10674.65	3358.37	3380.06	No
$(2,0,2) \times (0,1,0)_{12}$	orig.	-1673.97	7577.78	12996.60	3357.94	3373.44	Yes
$(1,1,1) \times (1,1,0)_{12}$	orig.	-1673.74	5465.66	9059.69	3355.49	3367.86	Yes
$(0,1,0) \times (1,0,1)_{12}$	orig.	-1796.57	5435.11	8579.98	3599.15	3608.64	Yes
$(1,0,1) \times (0,0,1)_{12}$	orig.	-1810.88	5485.74	7181.75	3631.76	3647.61	Yes
$(0,1,0) \times (1,0,1)_{12}$	log.	69.21	0.18	41828.43	-132.43	-122.93	Yes
$(0,1,0) \times (0,1,1)_{12}$	log.	64.25	0.21	41828.40	-124.51	-118.32	Yes
$(2,0,2) \times (1,1,1)_{12}$	log.	83.10	0.83	41829.14	-152.20	-130.51	No
$(1,0,0) \times (0,1,1)_{12}$	log.	82.01	0.83	41829.14	-158.02	-148.72	Yes
$(1,0,0) \times (0,1,0)_{12}$	log.	58.65	0.85	41829.14	-113.31	-107.11	Yes

Table 2: ARIMA Models and main goodness of fit criterias

Based on the main metrics, there were chosen two ARIMA models with the best scores: $SARIMA(1,1,1) \times (1,1,0)_{12}$ fitted on the original time series and $SARIMA(0,1,0) \times (0,1,1)_{12}$ on the log data. Throughout the remaining analysis, those will be the principal models examined in the upcoming tests.

4.2 Residuals

Another important way to evaluate the accuracy of the ARIMA models are the residuals diagnostics methods consisting of: the plot of standardized residuals, histogram of residuals, QQ plot, and correlogram of residuals.

Those plots show if the residuals are close to white noise or if there are any patterns not captured by the model. The standardized residuals plot should look very similar to white noise, the histogram should resemble a normal distribution and in the correlogram all lags should be lower than the confidence intervals. [7]

In both models, the plot of standardized residuals appear to be similar to white noise, the correlogram of residuals for $SARIMA(1,1,1) \times (1,1,0)_{12}$ shows better results, almost all lags are within the range, and the histograms display a normal distribution.

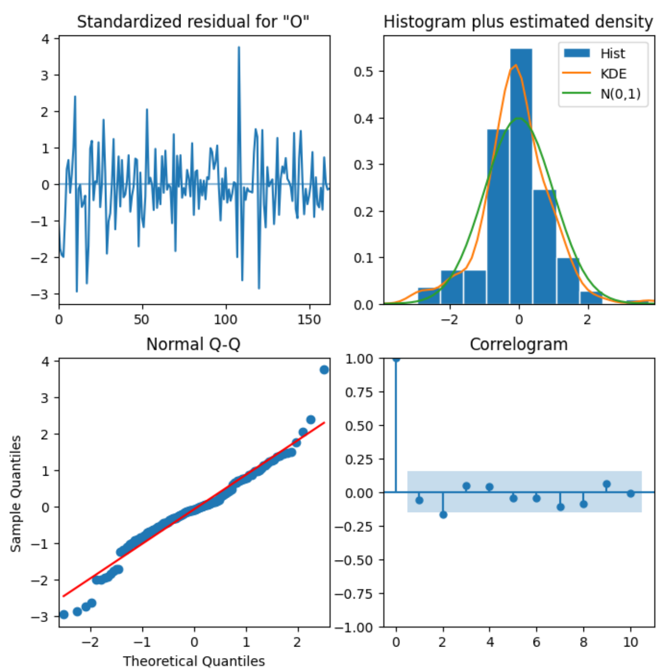

Figure 8: Diagnostics plots for residuals of $SARIMA(1,1,1) \times (1,1,0)_{12}$ on original time series

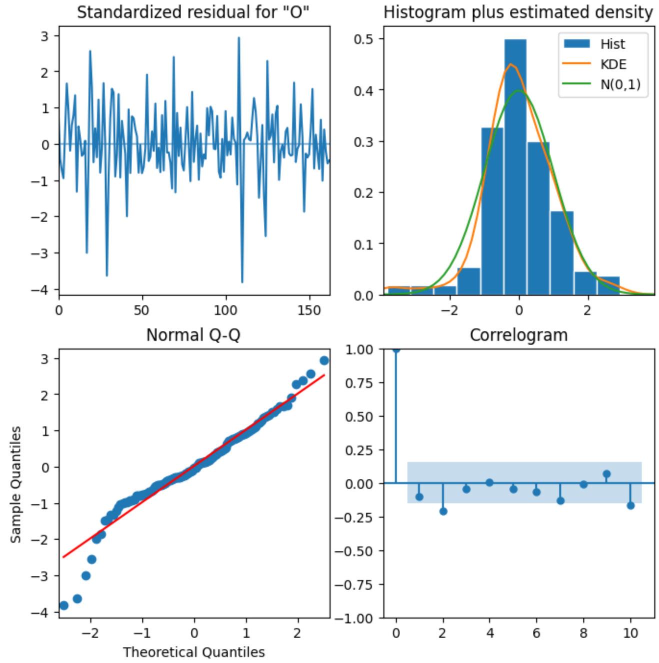

Figure 9: Diagnostics plots for residuals of $SARIMA(0, 1, 0) \times (0, 1, 1)_{12}$ on log of time series

The Ljung-Box was also performed on the residuals of the models, as it is a test that shows if the residuals are independent or correlated.

In the selected models, the Ljung-Box test revealed that all lags for residuals were not correlated, all p-values were higher than 0.05 threshold.

5 Forecasts

The two models were used for testing the predictions on known data (last 20 values were removed from the dataset) and also forecasting 20 more values. In both cases, the models succeeded in identifying the seasonality, the spikes of consumption in the coldest months. The $SARIMA(1, 1, 1) \times (1, 1, 0)_{12}$ also captured better the downward trend of the last 3 years.

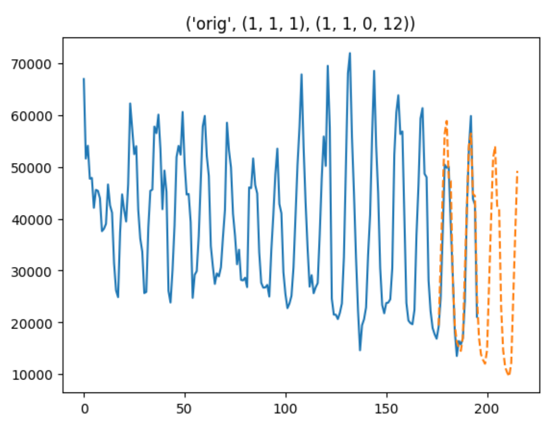

Figure 10: Forecasts of $SARIMA(1, 1, 1) \times (1, 1, 0)_{12}$ on original time series

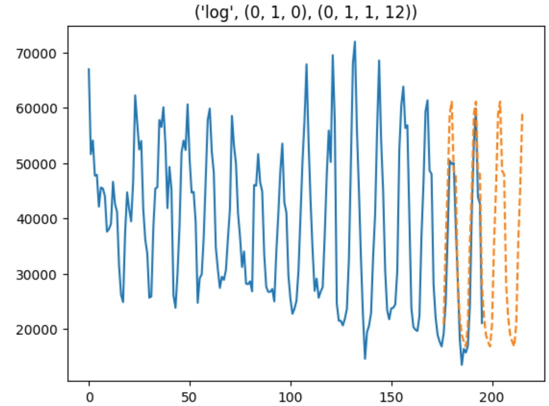

Figure 11: Diagnostics plots for residuals of $SARIMA(0,1,0) \times (0,1,1)_{12}$ on log time series

6 Conclusion

Based on multiple criteria, the best model is the $SARIMA(1,1,1) \times (1,1,0)_{12}$, due to the smallest values for AIC, BIC indicators, the highest log likelihood, and almost the best values for RMSE and MAE. The predicitions of the model are closer to the test data and the forecast values seem to preserve the trend and seasonality better.

References

[1] Eurostat, *Total consumption of natural gas, tj (gcv)*, 2022. DOI: 10.2908/EI_ISEN_M. [Online]. Available: https://ec.europa.eu/eurostat/databrowser/product/page/EI_ISEN_M (visited on 06/02/2024).

[2] *8.1 Stationarity and differencing — Forecasting: Principles and Practice (2nd ed)*. [Online]. Available: https://otexts.com/fpp2/stationarity.html (visited on 06/02/2024).

[3] *Augmented Dickey–Fuller test*, en, Page Version ID: 1225407051, May 2024. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Augmented_Dickey%E2%80%93Fuller_test&oldid=1225407051 (visited on 06/02/2024).

[4] *ARIMA modelling in Python*, en, Jun. 2017. [Online]. Available: https://www.geo.fu-berlin.de/en/v/soga-py/Advanced-statistics/time-series-analysis/ARIMA-modelling-in-Python/index.html (visited on 06/02/2024).

[5] *KPSS test*, en, Page Version ID: 1179751435, Oct. 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=KPSS_test&oldid=1179751435 (visited on 06/02/2024).

[6] "Statsmodels/statsmodels: Release 0.14.2," en, DOI: 10.5281/zenodo.10984387. [Online]. Available: https://zenodo.org/records/10984387 (visited on 06/02/2024).

[7] D. Andrés, *Step-by-Step Guide to Time Series Forecasting with SARIMA Models - ML Pills*, en-US, Jun. 2023. [Online]. Available: https://mlpills.dev/time-series/how-to-train-a-sarima-model-step-by-step/ (visited on 06/02/2024).