Conference Name: International Conference on Business, Economics, Law, Language & Psychology, 11-12 January 2024, Paris
Conference Dates: 11-12 January 2024
Conference Venue: Pullman Paris Centre – Bercy Address: 1 Rue de Libourne, 75012 Paris, France
Appears in: PEOPLE: International Journal of Social Sciences (ISSN 2454-5899)
Publication year: 2024

Xhoxhi, 2024

Volume 2024, pp. 24-31

DOI- https://doi.org/10.20319/icssh.2024.2431

This paper can be cited as: Xhoxhi. O (2024). Analysis of the Profitability of Albanian Bank Using Time Series Models. International Conference on Business, Economics, Law, Language & Psychology, 11-12 January 2024, Paris. Proceedings of Social Science and Humanities Research Association (SSHRA), 2024, 24-31.

ANALYSIS OF THE PROFITABILITY OF ALBANIAN BANK USING TIME SERIES MODELS

Olsi Xhoxhi
Faculty of Economy, University "Ismail Qemali" of Vlora, Vlorë, Albania
olsixhoxhi0@gmail.com

Abstract

This research focuses on predicting the profitability of the Bank of Albania, specifically measured by Return on Assets (ROA), by employing advanced time series analysis techniques. ROA is a fundamental financial metric reflecting the bank's ability to efficiently utilize its assets to generate profit, and it holds crucial implications for the stability and sustainability of the central bank. In this research, we have obtained a monthly dataset regarding the profitability of Albanian banks (Roa), covering the period from January 2016 to March 2023. Following an extensive data analysis, we have determined that the SARIMAX(0, 0, 1)x(2, 0, [1, 2], 12) model is the most suitable option for modeling our dataset. The SARIMAX(0, 0, 1) element signifies a non-seasonal moving average component, addressing short-term fluctuations and irregularities in the data. The seasonal component, (2, 0, [1, 2], 12), takes into account both annual and semi-annual patterns, aligning with the observed seasonal trends in the data. The results are expected to be instrumental for policymakers, financial analysts, and stakeholders concerned with the Bank of Albania's financial health. By applying time series analysis to ROA prediction, this research not only aims to enhance the central bank's financial decision-making capabilities but also contributes to the broader understanding of financial stability within the context of Albania's economic landscape.

Keywords

SARIMA Model, Bank of Albania, Profitability, Time Series Forecasting Methods

1. Introduction

Forecasting a bank's profitability, measured by Return on Assets (ROA), is critically important in financial analysis. It's vital for understanding a bank's financial health, efficient resource allocation, and risk management within the banking industry. Banks, serving as financial intermediaries, have a significant impact on economic growth and stability through their ability to generate profits and influence credit availability. Predicting a bank's ROA is closely tied to its contribution to economic development. Furthermore, in the global financial system, bank profitability ripples through international markets, affecting investor confidence, cross-border finances, and capital allocation. Accurate ROA predictions empower regulators, investors, and policymakers to respond effectively to challenges and opportunities. Given the dynamic banking landscape, predictive modeling of ROA becomes even more crucial for adaptability, resilience, and maintaining the banking sector's stability and competitiveness. This study focuses on predictive modeling of bank profitability, with a particular emphasis on ROA, to provide insights for strategic decision-making, enhance financial stability, and fortify the banking industry's enduring strength.

2. Literature Review

Forecasting is a widely used technique in various fields of research and industry to predict future trends, behaviors, or events based on historical data and analysis. Authors from around the world utilize forecasting methods and models in their research to make informed decisions, develop strategies, and gain insights into various domains.

Mashadihasanli (2022) study aimed to investigate the application of autoregressive integrated moving averages (ARIMA) for forecasting the monthly stock market price index in Istanbul from January 2009 to March 2021. When compared to all other tentative models, their research demonstrated that the ARIMA (3,1,5) model was the best-fit model for predicting the stock market price index.

Jayatilake (2019) gathered secondary data on ABC bank's gross NPLs from the fourth quarter of 2008 to the fourth quarter of 2018 for analysis and utilized the ARIMA technique as the primary analytical tool. Additionally, she used Mean Absolute Percentage Error (MAPE) and Mean Absolute Deviation (MAD) to assess the accuracy of the fitted model. Subsequently, sheidentified the ARIMA (0,2,1) model as the best-fitting model for forecasting future NPLs at ABC bank, after evaluating accuracy using both training and validation datasets.

Amin *et al.* (2022) research aimed to determine the optimal model for forecasting the profitability of Islamic banks in Malaysia based on the Interbank Rate, consequently predicting the interbank rate's profitability from April 2021 to December 2021. The data was divided into two sections: estimation and evaluation. In the estimation phase, they sought the most fitting forecast model by minimizing the Akaike Information Criterion (AIC), Mean Average Error (MAE), and Mean Average Percentage Error (MAPE), while the evaluation phase involved utilizing this model for forecasting. Their findings revealed that ARMA(1,1) was the most suitable model, as it exhibited the lowest AIC, MAE, and MAPE among the other models, establishing its superiority for data forecasting. Subsequently, this model was applied to the evaluation section to predict the interbank rate values from April 2021 to December 2021, indicating a gradual month-by-month increase in the Islamic interbank rate, indicative of the Islamic Bank's rising profitability in Malaysia.

Based on Syarif (2020) research findings regarding forecasting using the Auto-Regressive Integrated Moving Average (ARIMA) model for Islamic bank development, it can be summarized that the ARIMA (0,1,1) model produced the most accurate forecasting results. These results indicate that the growth of Islamic banks in 2020 is projected to rise by 7.4% in assets and is expected to increase by 7.3% by the conclusion of 2022, resulting in a total asset value of IDR 437 trillion. Additionally, third-party funds are forecasted to grow by 8% by the end of 2022, reaching a total of IDR 361 trillion.

The aim of Aryani, Aidi, & Syafitri (2018) research was to forecast the impact of predictor variables on Islamic bank profitability using two distinct approaches, namely ARIMAX and ARIMA errors. Both approaches represent a fusion of the regression and ARIMA models. The comparison between the two models was based on criteria such as the smallest Schwarz Bayesian Criterion (SBC) value, significant parameter estimations, and the lowest Mean Absolute Percentage Error (MAPE) value to determine the superior model. Consequently, the ARIMA (1, 1, 1) error model with the exchange rate variable was identified as the best model, exhibiting a MAPE value of 32.013%, while the ARIMAX (0, 1, 2) model with the exchange rate variable yielded a MAPE value of 375.78%. Thus, the most effective model for predicting the impact of predictor factors on Islamic banking profitability was found to be the ARIMAX model (0,1,2) with the exchange rate variable.

3. Data and Methodology

In this study, we acquired the monthly dataset on the profitability of Albanian bank (Roa) spanning from January 2016 to March 2023, a total of 87 data points. This dataset was sourced from the official website of the Bank of Albania.

This study aims to apply the Seasonal Autoregressive Integrated Moving Average (SARIMA) model to forecast Albanian bank (Roa). It combines autoregressive (AR), differencing (I), and moving average (MA) components, along with seasonal variations, to make predictions about future data points in a time series. SARIMA is a powerful tool for modeling and forecasting time-dependent data, particularly when there are seasonal patterns or trends in the data.

4. Data analysis of Albanian bank profitability

The descriptive statistics provided in the table below offer valuable insights into the characteristics of the time series dataset spanning 87 observations. The mean, which stands at 1.224828, serves as the central measure of the series, indicating that the typical value over this time period is approximately 1.22. The standard deviation, at 0.358670, highlights the degree of variability within the series, suggesting that data points tend to cluster around the mean with relatively limited dispersion. The minimum value of 0.44 reflects the lowest point observed during this time frame, while the maximum value of 2.26 represents the highest point, demonstrating the range of values the time series covers. The quartiles (25%, 50%, and 75%) provide further insights into the distribution, with the median (50%) at 1.29 indicating the middle point of the series. The interquartile range (IQR) between the 25th and 75th percentiles (0.91 to 1.5) encapsulates the central 50% of the data, emphasizing that a substantial portion of the series falls within this range.

Table 1: *Descriptive Statistics*

count	87.000000	**25%**	0.910000
mean	1.224828	**50%**	1.290000
std	0.358670	**75%**	1.500000
min	0.440000	**max**	2.260000

The graph, shown below, displays the time series data over a specific time period, typically on the x-axis, and the variable of interest (profitabilityof albanian bank) on the y-axis. Each data point on the graph corresponds to a specific time point, allowing us to discern how the variable evolves over time.

Figure 1: *Bank of Albania profitability from 2016 to 2023*

Checking the stationarity is a fundamental step in time series analysis, so we will apply an Augmented Dickey-Fuller (ADF) test. Stationarity is a fundamental property in time series modeling as it ensures that statistical characteristics like mean and variance remain constant over time.

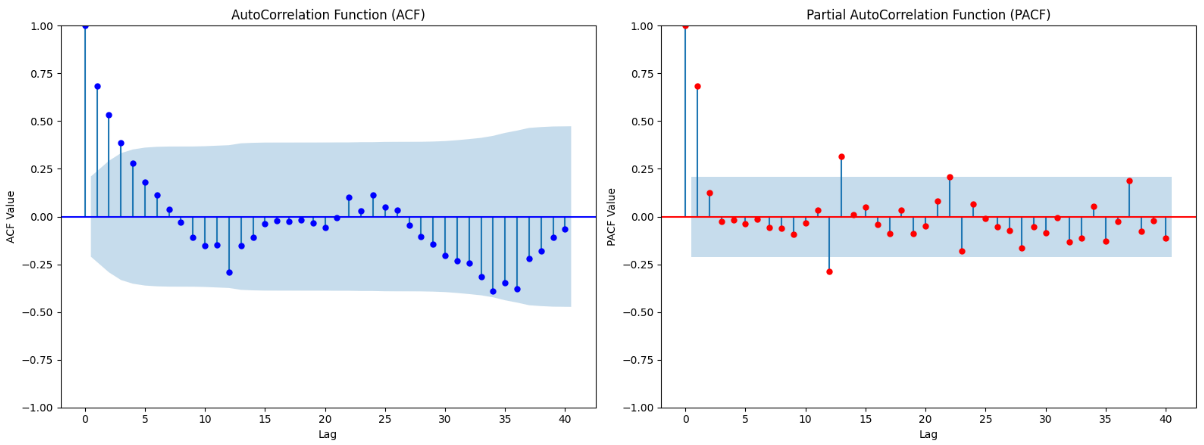
Table 2: *Results of Dickey-Fuller Test*

	Original Series
Test Statistic	-3.828032
p-value	0.002633
Critical Value (1%)	-3.508783
Critical Value (5%)	-2.895784
Critical Value (10%)	-2.585038

In the above results, the test statistic of -3.828032 is lower than the critical values at the 1%, 5%, and 10% significance levels (-3.508783, -2.895784, and -2.585038, respectively). Additionally, the p-value of 0.002633 is significantly less than the typical significance level of 0.05. These results indicate that we can reject the null hypothesis that the time series is non-stationary. In other words, the data shows strong evidence of stationarity, suggesting that the underlying process may be stable over time.

To find the optimal SARIMA parameters, rather than relying on manual analysis of ACF and PACF plots that are below we can enhance the efficiency of this process by developing a Python function. This function will automate the parameter selection for SARIMA, making the model building process more systematic and less dependent on visual inspections.

Figure 2: *The Autocorrelation and Partial Autocorrelation Function*



The results obtained from a Python function SARIMA time series model are as below:

Table 3: *Model Results*

Model	SARIMAX(0,0,1)x(2,0,[1, 2],12)
Mean Absolute Error(MAE)	0.2959
Mean Squared Error(MSE)	0.1604
Mean Absolute Percentage Error(MAPE)	26.87%
Akaike Information Criterion(AIC)	78.9070
Bayesian Information Criterion(BIC)	93.7025

The SARIMAX(0,0,1)x(2,0,[1,2],12) model is considered the most suitable time series model for the given data. It's been chosen because it effectively captures both seasonal and non-seasonal patterns in the data, making it the best option for making accurate forecasts and predictions. The specific numbers and parameters in the model notation have been fine-tuned to match the data's characteristics and improve its performance. The model's performance can be described as fairly good. It has a low Mean Absolute Error (MAE) and Mean Squared Error (MSE), indicating that it is making accurate predictions with relatively small errors. The Mean Absolute Percentage Error (MAPE) is reasonably low at 26.87%, suggesting that the model's predictions are, on average, within 26.87% of the actual values. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are also at acceptable levels, indicating that the model fits the data well without being overly complex.

Table 4: *SARIMA model*

Ljung-Box (L1) (Q):	3.11	**Jarque-Bera (JB):**	4.56
Prob(Q):	0.08	**Prob(JB):**	0.10
Heteroskedasticity (H):	0.49	**Skew:**	0.24
Prob(H) (two-sided):	0.06	**Kurtosis:**	4.01

The statistical measures and test results that are in the table 4 are as below:

- **Ljung-Box (L1) (Q):** The value of 3.11 is accompanied by a probability of 0.08 (Prob(Q)). With a p-value above the conventional significance level of 0.05, this suggests that there is not strong evidence of autocorrelation in the data, meaning the autocorrelation is not statistically significant.
- **Jarque-Bera (JB):** The JB statistic yields a value of 4.56 with a p-value of 0.10 (Prob(JB)). Again, with a p-value higher than 0.05, this indicates that there is no strong evidence to reject the null hypothesis of normality. In other words, the residuals might reasonably follow a normal distribution.
- **Heteroskedasticity (H):** The H value is 0.49, and the two-sided p-value (Prob(H)) is 0.06. In this case, the p-value is slightly above 0.05, suggesting a lack of strong evidence for heteroskedasticity. This means that the assumption of constant variance across the data may hold.
- **Skew (0.24):** Skewness indicates the symmetry of the data distribution. A value of 0.24 suggests a slight asymmetry, implying that the data is somewhat skewed, but not severely so. If it were perfectly symmetric, the skewness would be zero.
- **Kurtosis (4.01):** Kurtosis measures the tail heaviness of the data distribution. A value of 4.01 indicates that the data has slightly heavier tails compared to a normal distribution. This implies there may be some outliers or extreme values in the dataset.

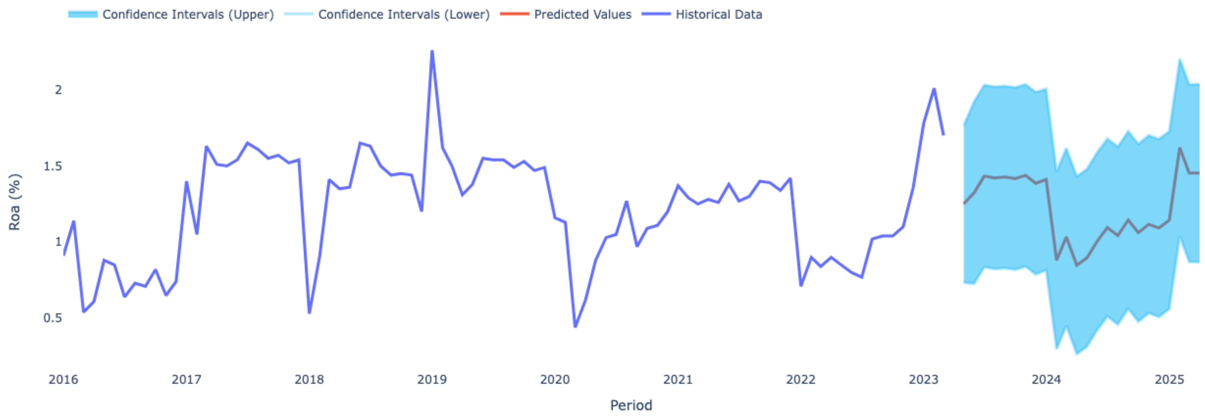
Below are the coefficients that provide information about the model's parameters. These coefficients help us understand how different factors contribute to the model's performance and predictions. Each coefficient has associated statistics like standard error, z-score, and p-value, which indicate the significance and influence of that particular factor on the model's results. These coefficients are essential for interpreting and fine-tuning the model.

Table 5: *SARIMA coefficients*

	coef	std err	z	P>|z|	[0.025	0.975]
ma.L1	0.5854	0.134	4.384	0.000	0.324	0.847
ar.S.L12	0.0045	0.037	0.122	0.903	-0.068	0.077
ar.S.L24	0.9953	0.040	24.972	0.000	0.917	1.073
ma.S.L12	-0.0381	0.305	-0.125	0.901	-0.636	0.560
ma.S.L24	-0.8829	0.714	-1.237	0.216	-2.282	0.516
sigma2	0.0592	0.039	1.524	0.127	-0.017	0.135

In the figure below, we have generated forecasts for 24 future periods using a SARIMA model. These predictions are based on historical data and the model's ability to capture underlying patterns and seasonality, offering valuable insights into the expected trends and values over the specified time frame. This visual representation allows us to explore and analyze the forecasted data points, providing a glimpse into future trends and aiding in decision-making and planning.

Figure 3: *Forecasting future periods*



5. Conclusions

After a comprehensive analysis of our data, we have arrived at the conclusion that the SARIMAX (0, 0, 1) x (2, 0, [1, 2], 12) model is the most appropriate choice for modeling our dataset. This conclusion stems from a rigorous examination of the data's characteristics, which included assessing its temporal patterns and seasonality. The SARIMAX (0, 0, 1) component indicates a non-seasonal moving average component, which accounts for the short-term fluctuations and irregularities present in the data. This is important to capture any residual patterns that might not be explained by the seasonal components. The seasonal component, (2, 0, [1, 2], 12), incorporates both yearly and biyearly patterns, which align with the observed seasonality in the data. The choice of a 12-month seasonal cycle is logical for many time series data, as it corresponds to the yearly calendar cycle.

However, it is important to note that this choice is not set in stone, and we should continue to monitor the model's performance over time. Regular model re-evaluation and potential re-fitting may be necessary as the underlying data patterns evolve. Additionally, we may explore more advanced techniques and incorporate exogenous variables if our data complexity or external factors require it.

REFERENCES

Amin, H. A., Izani, M. A., Azam, A. A., Nordin, M. H., & Shafie, N. A. (2022). FORECAST THE PROFITABILITY OF ISLAMIC BANKS IN MALAYSIA BASED ON ISLAMIC INTERBANK RATE. Mathematics in Applied Research.

Aryani, S., Aidi, M. N., & Syafitri, U. D. (2018). Analysis of The Profitability of Islamic Banking Using Arimax Model and Regression with Arima Errors Model. International Journal of Scientific Research in Science, Engineering and Technology, 49-53.

Jayatilake , A. C. (2019). MODELLING NON PERFORMING LOANS (NPLs) IN ABC BANK SRI LANKA. 4th Interdisciplinary Conference of Management Researchers.

Mashadihasanli, T. (2022). Stock Market Price Forecasting Using the Arima Model: an Application to Istanbul, Turkiye. Journal of Economic Policy Researches, 439-454.

Syarif, A. (2020). Forecasting the Development of Islamic Bank in Indonesia: Adopting ARIMA Model. JTAM (Jurnal Teori dan Aplikasi Matematika), 190-203.