Machine Learning Model of Seasonal Autoregressive Integrated Moving Average with Exogenous Variables for Logistics

Sergei Ivanov, Tatiana Zudilova, Nikita Osipov, Irina Osetrova, Anantchenko Igor, Lubov Ivanova
ITMO National Research University (ITMO University)
Saint Petersburg, Russian Federation
serg_ie@mail.ru, zudilova@ifmo.spb.ru, nikita@ifmo.spb.ru, irina@ifmo.spb.ru, igor@anantchenko.ru, 45is@mail.ru

Abstract — **The article considers the problem of forecasting the volume of seasonal logistics transportation of fruits and vegetables using time series based on the seasonal autoregressive integrated moving average with exogenous variables (SARIMAX) machine learning model. To analyze and build forecasts for a data set, the volumes of logistics transportation of fruits and vegetables are analyzed taking into account seasonality. The purpose of this article is to develop and study a model for forecasting the volume of seasonal logistics transportation using SARIMAX. The problem of time series analysis for determining seasonal fluctuations using the SARIMAX machine learning model is solved. The study shows the effectiveness of the constructed SARIMAX model for forecasting seasonal logistics transportation in comparison with other models, such as autoregressive integrated moving average (ARIMA), random forest model, exponential smoothing, Holt-Winters model, artificial neural network (ANN), multilayer perceptron (MLP), GARCH model, support vector model. The evaluation of the modeling results shows the effectiveness of the SARIMAX model with an R-square of 0.99 and a low mean absolute error (MAE) of 0.045. Python programs were developed to solve the problem of modeling and forecasting logistics transportation taking into account seasonality. The results show the accuracy of the forecasting model of about 96% and its effectiveness in applying to a wide range of forecasting problems.**

I. INTRODUCTION

Modeling and optimization methods are widely used in modern logistics. The use of modern machine learning methods increases the transparency of the supply chain and the level of service for customers. Mathematical modeling, which includes building representations of real operations, is widely used in supply chain analysis. Event-based modeling helps to identify constraints and improve logistics workflows. Modeling also allows you to determine the interaction between various components in the supply chain, such as transport networks, warehouses and production sites. System dynamics methods are used to study long-term changes in the supply chain.

Optimization models in logistics operations help find optimal solutions given certain constraints. Inventory and route optimization are implemented to improve the performance of logistics networks. Several approaches are used for effective inventory management: ABC analysis, Just-in-Time. Route optimization allows you to create effective delivery routes taking into account important factors of time and cost. Mathematical modeling, supply chain optimization are modern tools for logistics.

Forecasting and forward planning allow you to effectively manage both freight and passenger transportation. Using various forecasting approaches allows logistics companies to anticipate future needs. Strategic planning helps companies to adapt to changing market conditions in a timely manner. Forecasting allows logistics companies to effectively allocate resources. Let's consider the main methods and strategies for forecasting and planning in logistics. Qualitative approaches are used for assessment - expert opinions in the transport sector. The Delphi method is used, in which a group of experts anonymously develops estimates. Quantitative methods, including statistics, regression analysis, and time series analysis, are also widely used. Modern logistics companies increasingly rely on machine learning methods to analyze large datasets and identify complex patterns. A review of existing research shows the use of several forecasting methods and models, including “autoregressive integrated moving average” (ARIMA), random forest model, exponential smoothing, Holt-Winters model, artificial neural network (ANN) model, multilayer perceptron (MLP), GARCH model, and support vector model.

The study [1] investigated the use of SARIMAX model to solve sales trend analysis problems for large retailers using large datasets. The study analyzed time series data using various machine-learning models.

In [2], SARIMAX modeling was used to forecast the consumer price index and estimate inflation. The process covering data collection, processing, stationarity testing, model building, and accuracy evaluation is described in detail. This work demonstrated the superior performance of SARIMAX in forecasting grain and food prices compared to alternatives such as ARIMA, support vector machines, random forest, Holt-Winters, and exponential smoothing.

The study [3] solved the problem of forecasting daily hotel room demand using ARIMA, SARIMAX, neural networks, and sGARCH and GJR-GARCH models, with SARIMAX achieving the most accurate results.

In [4], SARIMAX modeling was used to accurately forecast rainfall in India. In this study, forecasting models such as

SARIMAX, Decision Tree, Support Vector Machine, ARIMA, and Exponential Smoothing were evaluated. The results demonstrated excellent accuracy of the SARIMAX model, achieving an R-2 value of 0.99 and a low mean absolute error of 0.044.

Another study [5] investigated time series forecasting using a decomposition method for the SARIMAX model. It applied decomposition into trend-cycle-irregularity and seasonality components using a multiplicative decomposition method as a pre-processing step. This pre-decomposition reduced the average error rate in forecasting.

Many studies have used the SARIMAX model to forecast various time series data. In [6], the model was used to forecast monthly rainfall, helping in the preparation for agricultural and seasonal disasters, demonstrating high accuracy (R2 = 0.91) and low error (RMSE = 54.5). A doctoral study [7] used SARIMAX to forecast electricity consumption in a university setting, including temperature and humidity, and identify weekly trends, achieving 96% accuracy.

A study [8] successfully predicted maximum temperature changes in Ahmedabad using SARIMAX, showing high accuracy valuable for urban planning and environmental surveillance.

The capabilities of the SARIMAX model are extended to the electricity sector in [9] to forecast demand, generation, peak load, and power. The results showed high accuracy compared to simpler moving average methods.

A comparison of forecasting methods in [10] showed that the hybrid SARIMAX-MLP model outperformed both the standard SARIMAX and LSTM models in forecasting the domestic transportation market (MAPE 4.81 vs. 6.14 and 6.52, respectively).

Forecasting corn production in [11] also benefited from the application of the SARIMAX model, showing high accuracy.

A hybrid approach WD-SARIMAX, which combines wavelet decomposition and SARIMAX, was used in [12] to forecast temperature in Delhi, achieving impressive results (MAE 1.13, MedAE 0.76, RMSE 1.67, MAPE 4.9, R2 0.91) by decomposing non-stationary data.

In [13], considering electricity supply and demand in a dynamic tariff system, the superiority of SARIMAX based on RMSE, MAE and R2 was found, although its limitations were noted compared to the BATS model.

In [14] a comparison of natural gas production and consumption forecasting in the US, the application of SARIMAX was found to be more effective than SARIMA, as indicated by RMSE and MAPE.

In addition, [15] applied SARIMAX to forecast domestic flights in Indonesia, taking into account airports, COVID-19 cases, calendar effects and restrictions, which provided a detailed understanding of air travel dynamics.

A hybrid SARIMAX-LSTM model was investigated in [16] for short-term energy load forecasting, which was found to be acceptable, but no direct comparison with ARIMA models was presented.

In [17], the usefulness of the SARIMAX model was demonstrated in forecasting cyanobacterial biomass in a lake, with exogenous variables such as chlorophyll and phosphorus being influential.

In [18], forecasts of monkeypox outbreaks using an improved SARIMAX model enhanced by the Ditperthroated Technique (DTACO) were detailed. This DTACO-SARIMAX framework was found to be useful for disease management and monitoring its progression.

In [19], a hybrid model incorporating multivariate decomposition together with a SARIMAX - Gated Recurrent Unit pair was used to forecast rice prices in West Java. This strategy partitions the data into trend, seasonal, and residual elements using SARIMAX for the seasonal forecasts and a combination of Gated Recurrent Unit for the trend and residual forecasts.

For grain prices in Indonesia, [20] presents a comparison of the SARIMA and SARIMAX approaches. Including external factors such as government-mandated procurement costs in the SARIMAX model significantly improved the forecasting accuracy.

In [21] investigates food and beverage consumption patterns in Indonesia by incorporating weather information using the SARIMAX analysis method. The results show that integrating weather data significantly improves the model performance.

Finally, [22] demonstrates the application of the SARIMAX model to advanced solar radiation forecasting in Muscat, Oman, implemented using standard Python libraries. The results of the study indicate superior performance compared to alternative models, achieving an MSE of 0.075, an MAE of 0.209, and an R2 value of 0.92.

Based on the analysis of the articles, it can be concluded that SARIMAX is the most effective model for forecasting taking into account seasonality and exogenous variables.

II. Machine learning models for forecasting

Forecasting in logistics transportation is an important tool for optimizing logistics processes and resource planning.

The purpose of this work is to create a model that will allow forecasting future transportation volumes based on historical data and factors affecting demand.

To solve the problem, the authors prepared a data set of seasonal fruit and vegetable transportation volumes by month since 2018.

Fig. 1 shows a graph of the dynamics of transportation volumes.

The ARIMA model represents a common strategy for analyzing data, effectively capturing trends and recurring seasonal patterns.

It's frequently employed to predict logistics supply volumes, while also accounting for predictable seasonal fluctuations.

The model's architecture rests on three core components: autoregression, integration, and moving average. Autoregression establishes the link between a supply volume's present value and its past values. Integration transforms the time series into a stable form. The moving average element manages random fluctuations and background noise within the data.

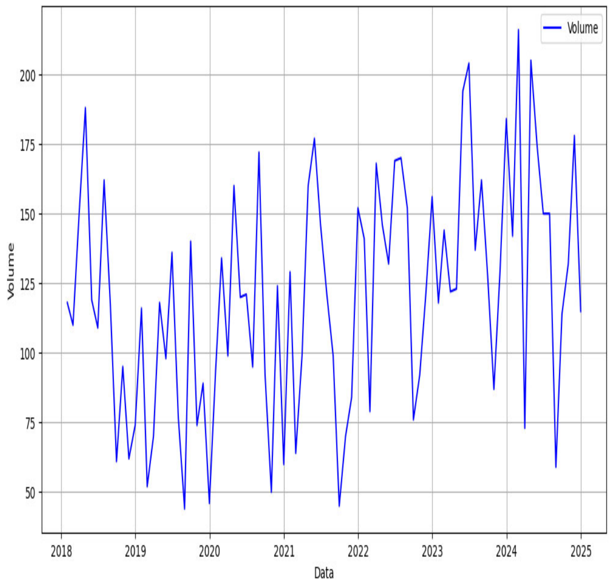

Fig. 1. Initial dataset of transportation volumes of seasonal fruits and vegetables.

Before constructing the model, the Dickey-Fuller test is carried out. If the series proves non-stationary, integration is then applied.

Subsequently, after parameter selection and model construction, a quality assessment is conducted using the "Akaike Information Criterion" (AIC) and "Bayesian Information Criterion" (BIC).

Once the model is validated, it's utilized to forecast future supply volumes.

ARIMA is a broadly recognized tool for time series forecasting, distinguished by its minimal assumptions regarding data distribution.

In contrast, Random Forest, another machine learning technique, addresses both classification and regression challenges. To enhance forecast precision, it constructs numerous decision trees and merges their outputs.

Random Forest achieves improved outcomes by aggregating forecasts from multiple models. It utilizes a random selection of training data (Bootstrap method) to generate diverse trees and employs random feature selection to minimize correlation between them. The final forecast is calculated by averaging predictions across all trees, making the method less prone to overfitting.

Exponential smoothing, an alternative forecasting approach, uses a weighted average of past observations. Each prior value is multiplied by a factor that diminishes exponentially over time, emphasizing recent data. The smoothing factor, ranging from 0 to 1, governs the degree of smoothing; values near 1 produce forecasts heavily influenced by recent observations. This method performs well with time series exhibiting clear patterns, but struggles with abrupt shifts in the data and inherently overlooks seasonal patterns.

The exponential smoothing is appropriate for stationary time series lacking trends or seasonality.

The Holt-Winters model takes into account trends and seasonal fluctuations and includes the components: the level of the time series, the trend in the level change, and the seasonal nature of the repeating fluctuations at intervals.

The additive Holt-Winters model is appropriate when seasonal variations remain constant. In this model the components are summed.

Conversely, the multiplicative Holt-Winters model is used when seasonal variations vary proportionally. In this model the components are multiplied.

However, the Holt-Winters model can be unreliable when dealing with substantial variations in the datasets.

Artificial neural networks (ANNs) find frequent application in predictive tasks. These systems operate with nodes that accept data, process it internally, and then deliver results. The strength of these connections, represented by weights, dictates how input translates to output. ANNs are organized into successive layers; the initial input layer takes in raw data. Following layers then refine this data, culminating in an output layer that produces forecasts. In supervised learning, ANNs learn from datasets containing labeled examples, pairing inputs with expected outcomes. Data flows forward through the network, generating an initial output. Subsequently, backpropagation refines the node weights, minimizing errors and reducing loss. Activation functions shape each node's output – sigmoid is often used for binary decisions, while Rectified Linear Unit and Softmax are common for multiple categories. To prevent overfitting, regularization techniques are employed; Dropout randomly deactivates nodes during training, and L2 regularization discourages excessively large weights within the loss function. After training, the network can predict outcomes for unseen data by feeding in inputs and repeating forward propagation. Nevertheless, utilizing ANNs demands considerable training time and a substantial dataset.

The GARCH model is a valuable method for time series analysis. This model is designed to forecast volatility fluctuations. Predicting volatility using the GARCH method entails determining model coefficients from historical data through maximum likelihood calculations; then, calculating the conditional variance to anticipate future volatility fluctuations;

and lastly, producing forecasts of volatility for the future, informed by these coefficients and present returns. The success of the GARCH model hinges on the initial assumptions and the specific parameters chosen.

The multi-layer perceptron (MLP) is a flexible neural network architecture suitable for tackling a wide range of tasks, including categorization and prediction problems. Data enters the MLP through an input layer, where each neuron corresponds to a unique characteristic. Hidden layers follow, refining the data through intricate mathematical operations; the quantity of these layers and neurons is adjusted to meet the demands of the particular task.

The training of MLP is characterized by a dual-phase methodology. Initially, in forward propagation, data traverses the neural network. Every node performs a computation, generating an output derived from incoming data and assigned weight parameters. Nonlinearity is introduced via activation functions integrated within each node.

Following this, the backpropagation phase commences. Here, a loss function quantifies the discrepancy between predicted and actual values. An optimization algorithm then fine-tunes the network's weights, striving to reduce the error. MLP leverage various activation functions tailored to specific tasks. Sigmoid functions are suited to binary classification, RLU are applied within the hidden layers, and Softmax functions for multi-class classification scenarios. To combat overfitting, techniques such as Dropout and L2 regularization are integrated into the training procedure.

The MLP model excels at modeling difficult relationships and has shown considerable effectiveness when dealing with extensive datasets; however, optimal performance typically necessitates prolonged training durations.

The "Support Vector Machine" (SVM) also functions effectively for classification and regression, proving especially beneficial for high-dimensional data and predictive tasks. SVM establishes a hyperplane that maximizes the separation of classes within the feature space. In regression scenarios, SVM seeks to predict continuous values by identifying a function whose deviation from actual values remains below a specified threshold (epsilon).

The method's success hinges on selecting the appropriate kernel and fine-tuning its parameters, requires data standardization, and can demand substantial training. Traditional machine learning approaches for forecasting often falter, frequently overlooking seasonal patterns and external variables – a limitation that the SARIMAX was designed to overcome.

The SARIMAX performs optimally when applied to stationary dataset, necessitating preliminary corrections. For this study, we initially assessed stationarity using the test of Dickey-Fuller. Techniques like logarithms, and root transformations can be implemented to achieve stationarity. Averages are computed to smooth the data and diminish random noise.

The Dickey-Fuller test statistic for our dataset validated the stationarity of the series.

III. MODELING RESULTS USING SARIMAX

To account for seasonal components, the SARIMAX ("seasonal autoregressive integrated moving average with exogenous regressors") method is used, which extends the ARIMA by taking into account external regressors.

The main components of the SARIMAX model are the following elements: autoregressive, integrated, moving average, seasonal and exogenous regressors.

In the study and computational experiment, prices and returns are chosen as external regressors for the SARIMAX model. The seasonal component defines periodic patterns, such as harvest time.

For the SARIMAX model, the current value of the time series is calculated using the formula:

$$D_t = c + f_1 D_{t-1} + f_2 D_{t-2} + \dots + A_1 e_{t-1} + A_2 e_{t-2} + S + R + e_t \tag{1}$$

where:

D_t is the current value of time series, c is a constant, f is the autoregressive coefficients, A is the moving average coefficients, e_t is the error at time t, S is the seasonal component, R is the component of the external regressors.

The core steps for SARIMAX involve:

1) The dataset acquisition and preprocessing, including dataset cleansing.
2) The dataset exploration, verifying time series stationarity and identifying seasonality.
3) The parameter determination for model.
4) The construction of SARIMAX model and subsequent evaluation.
5) The forecasting and assessment of forecast accuracy.

The SARIMAX model can also be impacted by atypical outliers in the dataset, potentially reducing forecast accuracy, highlighting the importance of dataset preparation. Mitigation strategies include Z-score analysis, interquartile range (IQR) estimation, and visual inspection. Smoothing techniques such as moving averages and exponential smoothing can also lessen the influence of outliers and random fluctuations, with outlier replacement using the median proving valuable. The "Akaike Information Criterion" (AIC) and "Bayesian Information Criterion" (BIC) were utilized to evaluate the model's performance, confirming its suitability.

The model was used to forecast future shipping volumes for 2026.

Fig. 2 shows the forecast results for shipping volumes up to 2026.

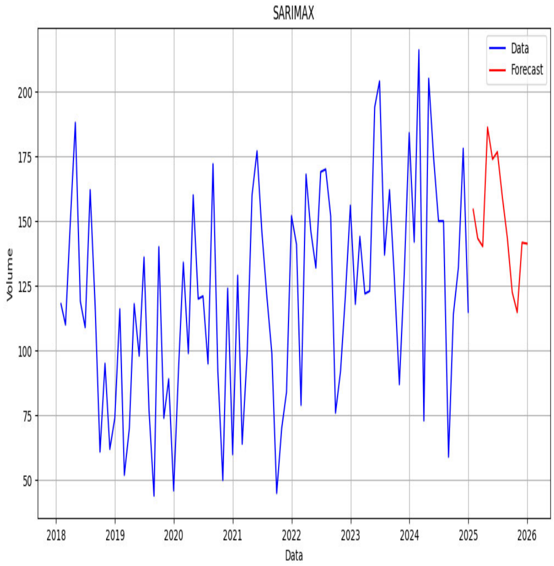

Fig. 2. Forecasting future shipments of seasonal fruits and vegetables

The SARIMAX model showed high forecasting accuracy of about 96% for shipping volumes. Compared to the basic ARIMA model, the SARIMAX model provides a 10% improvement in forecasting accuracy by taking into account seasonal fluctuations.

The SARIMAX model can be adapted to various types of dataset and also takes into account seasonal fluctuations, which improves forecasting accuracy.

The study shows the effectiveness of the constructed SARIMAX model for forecasting seasonal logistics transportation compared to other models such as autoregressive integrated moving average (ARIMA), random forest model, exponential smoothing, Holt-Winters model, artificial neural network (ANN) model, multilayer perceptron (MLP), GARCH model, support vector model. The evaluation of the modeling results shows the effectiveness of the SARIMAX model with an R-squared value of 0.99 and a low mean absolute error (MAE) of 0.045.

In modern logistics companies, the use of machine learning methods and models for forecasting provides significant advantages in the market and allows optimizing the company's logistics processes. In addition, digital tools - programs developed by the authors of the article allow companies to respond more flexibly to market changes and take into account various restrictions. It is important to note the high accuracy of the SARIMAX forecasting model used in the programs developed by the authors.

The programs developed by the authors allow you to clearly visualize the modeling results and save forecast graphs in various formats. The programs also allow you to download user datasets for processing large volumes via files. In addition, the user receives the results of model evaluation and stationarity testing.

The results of testing the execution time of the forecasting program showed a direct dependence of time on the volume of the initial dataset, as well as some slowdown in the execution of modeling when it is necessary to bring the initial series to a stationary one, which is also explained by the need to perform differentiation. In general, the developed software tool showed the best performance on various stationary data sets.

Data-frames were used to process the loaded data set in the program. If testing of the input data showed non-stationarity, differentiation is performed in the program to transform it to a stationary series. Then, the hyperparameters of the model are selected and the best ones are used in it. The forecast period in the program is set by the user, but the accuracy of the forecast depends significantly on the volume and period of the input historical data. Finally, the program evaluates the accuracy of the model.

The developed programs are documented in detail for ease of use and the project is hosted on github. The programs were used to forecast logistics shipments for a small retail chain and showed high accuracy for constructing weekly forecasts.

The program developed by the authors includes the following important sections: loading a data set from a user file, preliminary data processing, creating data frames, checking for data stationarity using the Dickey-Fuller test, differentiating when the time series is not stationary, defining variables, building a SARIMAX model and selecting the best hyperparameters, forecasting for a user-defined period of time, evaluating the model using metrics and visualizing the results with the construction of graphs and the ability to save the results to files.

The use of the developed digital tool will allow logistics companies to forecast transportation volumes taking into account various factors and restrictions. This will increase profitability and reduce costs in all components of the logistics chain. In addition, the tool developed by the authors will allow fine-tuning to solve various logistics problems.

Based on the developed programs, it is planned to further develop hybrid forecasting methods and explore complex integrated machine learning models. The authors plan to explore hybrid models based on a combination of SARIMAX and ANN, MLP, LSTM, DTACO and wavelet decomposition.

IV. Conclusion

The paper solves the problem of forecasting the volume of seasonal fruit and vegetable logistics transportation using the SARIMAX time series model. When analyzing and constructing forecasts for a dataset, the paper considers the volume of fruit and vegetable logistics transportation taking into account seasonality.

The purpose of the work is to develop and study a seasonally adjusted model for forecasting the volume of logistics transportation using SARIMAX methods.

The problem of time series analysis for determining seasonal fluctuations using the SARIMAX model is solved and its accuracy is assessed.

The study shows the effectiveness of the constructed SARIMAX model for forecasting seasonal logistics transportation in comparison with other models, such as autoregressive integrated moving average (ARIMA), random forest model, exponential smoothing, Holt-Winters model, artificial neural network (ANN), multilayer perceptron (MLP), GARCH model, support vector model.

The evaluation of the modeling results shows the effectiveness of the SARIMAX model with an R-square of 0.99 and a low mean absolute error (MAE) of 0.045.

Python programs were developed to solve the problem of modeling and forecasting logistics transportation with seasonality.

The results show the accuracy of the forecasting model of about 96% and its effectiveness in applying to a wide range of forecasting problems.

REFERENCES

[1] Özmen, Emre S., "Time Series Performance and Limitations with SARIMAX: An Application with Retail Store Data.", *Electronic Turkish Studies* 16, no.5, 2021.

[2] Vivek, K., A. R. Deepti, and Farzeen Basith., "Forecasting the Consumer Price Index using SARIMAX Modeling.", *2024 Second International Conference on Inventive Computing and Informatics (ICICI)*. IEEE, 2024.

[3] Ampountolas, Apostolos., "Modeling and forecasting daily hotel demand: A comparison based on sarimax, neural networks, and garch models.", *Forecasting*, 3, no.3, 2021, pp. 580-595.

[4] Kumar, Dinesh, and Pankaj Richariya., "Leveraging SARIMAX for Accurate Rainfall Forecasting in India.", *International Journal of Innovations in Science, Engineering And Management* , 2024, pp. 37-47.

[5] Nontapa, Chalermrat, et al., "A new time series forecasting using decomposition method with SARIMAX model.", *Neural Information Processing: 27th International Conference, ICONIP 2020*, Bangkok, Thailand, 2020, Proceedings, Part V 27. Springer International Publishing.

[6] Mulla, Shahenaz, Chaitanya B. Pande, and Sudhir K. Singh., "Times Series Forecasting of Monthly Rainfall using Seasonal Auto Regressive Integrated Moving Average with EXogenous Variables (SARIMAX) Model.", *Water Resources Management* 38, no.6, 2024, pp. 1825-1846.

[7] Shadkam, Arash., "Using SARIMAX to forecast electricity demand and consumption in university buildings.", *Diss. University of British Columbia*, 2020.

[8] Shah, Vyom, et al., "Forecasting Maximum Temperature Trends with SARIMAX: A Case Study from Ahmedabad, India.", *Sustainability* , 16, no.16, 2024, p.7183.

[9] Alharbi, Fahad Radhi, and Denes Csala., "A seasonal autoregressive integrated moving average with exogenous factors (SARIMAX) forecasting model-based time series approach.", *Inventions* , 7, no.4, 2022, p.94.

[10] Fiskin, Cemile Solak, et al., "Time series forecasting of domestic shipping market: comparison of SARIMAX, ANN-based models and SARIMAX-ANN hybrid model.", *International Journal of Shipping and Transport Logistics* , 14, no.3, 2022, pp. 193-221.

[11] Airlangga, Gregorius., "Advanced Forecasting of Maize Production using SARIMAX Models: An Analytical Approach.", *Journal media informatika Budidarma* 8, no.1, 2024, pp. 361-370.

[12] Elshewey, Ahmed M., et al., "A Novel WD-SARIMAX model for temperature forecasting using daily delhi climate dataset.", *Sustainability* , 15, no.1, 2022, p.757.

[13] Karki, Aayush Maan, et al., "Electric load prediction using SARIMAX.", *Australian Journal of Electrical and Electronics Engineering* , 2024, pp. 1-14.

[14] Manigandan, Palanisamy, et al., "Forecasting natural gas production and consumption in United States-evidence from SARIMA and SARIMAX models.", *Energies* , 14, no.19, 2021, p.6021.

[15] Zeinawaqi, Zahrah, and Abdullah Ahmad Dzikrullah., "Factors Affecting The Number Of Domestic Flights In Indonesia During Covid-19 Pandemic Using SARIMAX Method.", *Jurnal Matematika, Statistika dan Komputasi*, 21, no.1, 2024, pp. 1-9.

[16] Sheng, Feng, and Li Jia., "Short-term load forecasting based on SARIMAX-LSTM.", *2020 5th International Conference on Power and Renewable Energy (ICPRE)*. IEEE, 2020.

[17] Guo, Yan, Xijun Lai, and Min Gan., "Cyanobacterial biomass prediction in a shallow lake using the time series SARIMAX models.", *Ecological Informatics* , 78, 2023, p.102292.

[18] Elshabrawy, Mahmoud, et al., "Forecasting of Monkeypox Cases Using Optimized SARIMAX Based Model.", *2023 3rd International Conference on Electronic Engineering (ICEEM)*. IEEE, 2023.

[19] Gde Artadana Mahaputra Wardhiana, I. Nyoman, et al., "Price Forecasting of West Java Rice using Multivariate Decomposition SARIMAX-Gated Recurrent Unit Combination.", *International Journal of Intelligent Engineering & Systems*, 18, no.1, 2025.

[20] Yulianti, Riska, et al., "Comparison of SARIMA and SARIMAX methods for forecasting harvested dry grain prices in Indonesia.", *BAREKENG: Jurnal Ilmu Matematika dan Terapan* , 19, no.1, 2025, pp. 319-330.

[21] Alzami, Farrikh, et al., "Demand Prediction for Food and Beverage SMEs Using SARIMAX and Weather Data.", *Ingénierie des Systèmes d'Information* , 29.1, 2024.

[22] Baloch, Mazhar, et al., "An Intelligent SARIMAX-Based Machine Learning Framework for Long-Term Solar Irradiance Forecasting at Muscat, Oman.", *Energies*, 17, no. 23, 2024, p.6118.