*Global Journal of Energy Technology Research Updates, 2022, 87-96*

Published by Avanti Publishers

**Global Journal of Energy Technology Research Updates**

ISSN (online): 2409-5818


# Grid Search for SARIMAX Parameters for Photovoltaic Time Series Modeling

Todizara Andrianajaina[1,*], David Tsivalalaina Razafimahefa[1], Raonirivo Rakotoarijaina[1] and Cristian Goyozo Haba[2]

[1]*Ecole Supérieure Polytechnique Antsiranana, University of Antsiranana, Antsiranana, Madagascar*

[2]*Faculty of Electrical Engineering, Gheorghe Asachi Technical University of Iaşi, Iaşi, Romania*

## ARTICLE INFO

*Article Type*: Research Article

*Keywords*:
Photovoltaic
SARIMAX grid
Time series modeling
Mean absolute error (MAE)
Root mean squared error (RMSE)

*Timeline*:
Received: November 11, 2022
Accepted: December 15, 2022
Published: December 23, 2022

*Citation:* Andrianajaina T, Razafimahefa DT, Rakotoarijaina R, Haba CG. Grid search for SARIMAX parameters for photovoltaic time series modeling. Glob J Energ Technol Res Updat. 2022; 9: 87-96.

*DOI*: https://doi.org/10.15377/2409-5818.2022.09.7

*Corresponding Author
Email: todizara.andrianajaina@gmail.com
Tel: +(261) 32 72 807 74

## ABSTRACT

The SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous regressors) model is a time series model that can be used to forecast future values of a time series, given its past values. It is beneficial for modeling time series data that exhibits seasonality and incorporating additional exogenous variables (variables that are not part of the time series itself but may affect it).

One way to optimize the performance of a SARIMAX model is to use a grid search approach to find the best combination of hyperparameters for the model. A grid search involves specifying a set of possible values for each hyperparameter and then training and evaluating the model using all possible combinations of these values. The combination of hyperparameters that results in the best model performance can then be chosen as the final model. To perform a grid search for a SARIMAX model, you must define the grid of hyperparameters you want to search over. This will typically include the values of the autoregressive (AR) and moving average (MA) terms and the values of any exogenous variables you want to include in the model. We will also need to define a metric to evaluate the model's performance, such as mean absolute or root mean squared error.

Once we have defined the grid of hyperparameters and the evaluation metric, you can use a grid search algorithm (such as a brute force search or a more efficient method such as random search or Bayesian optimization) to evaluate the performance of the model using all possible combinations of hyperparameters. The combination of hyperparameters that results in the best model performance can then be chosen as the final model.

In this article, we will explore the potential of SARIMAX for PV time series modeling. The objective is to find the optimal set of hyperparameters. Grid Search passes all hyperparameter combinations through the model individually and checks the results. Overall, it returns the collection of hyperparameters that yield the most outstanding results after running the model. One of the most optimal *SARIMAX (p,d,q) x (P, D, Q,s)* combinations is *SARIMAX (0,0,1) x (0,0,0,4)*.

# 1. Introduction

Photovoltaic (PV) time series energy modeling involves predicting the amount of electricity that a PV system will generate over time, usually in the form of a time series of data points. This can be useful for various purposes, such as estimating the energy production of a PV system for a particular location, optimizing the design of a PV system, or predicting the performance of a PV system over its lifetime [1,2].

Several approaches to PV time series energy modeling can be classified into two main categories: statistical and physical models. Statistical models are based on statistical techniques, such as regression or time series analysis, and are typically used to forecast PV energy production using past data as input. On the other hand, physical models are based on physical principles, such as the laws of thermodynamics. They are used to predict PV energy production based on factors such as the PV system's location, orientation, and surface properties [3-6].

Grid search is a hyperparameter optimization method in which a range of values for each hyperparameter is specified. The model is trained and evaluated for every combination of these values. It can be used to find the optimal set of hyperparameters for a SARIMAX model, which stands for Seasonal AutoRegressive Integrated Moving Average with eXogenous regressors [7, 8].

The rise in solar energy usage has resulted in the need for efficient and accurate methods for Photovoltaic (PV) time series modeling [9]. SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eXogenous variables) is an increasingly used technique in this domain. In this article, we will explore the potential of SARIMAX for PV time series modeling, its advantages, best practices for grid search optimization, and examples of successful implementations.

# 2. Introduction to SARIMAX

SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous regressors) is a statistical model that can be used to analyze and forecast time series data that exhibits autocorrelation and seasonality. It is an extension of the ARIMA (Autoregressive Integrated Moving Average) model, used for time series data that exhibits autocorrelation but not seasonality [10, 11].

The SARIMAX model adds the ability to include seasonal terms in the model and exogenous variables (also known as predictor or explanatory variables) in the model. This allows the model to capture the influence of these variables on the time series data and account for the periodic nature of the data.

The SARIMAX model is specified by three main components: the autoregressive (AR) term, the moving average (MA) term, and the seasonal terms. The AR term captures the autocorrelation in the data by including lagged values of the time series as predictors. The MA term captures the impact of random shocks or errors in the data by including lagged errors as predictors. The seasonal terms capture the periodic nature of the data by including lagged values of the time series at different seasonal intervals (e.g., monthly, quarterly, etc.) as predictors.

In order to fit a SARIMAX model, we will need to specify the values for the various parameters of the model, such as the order of the AR and MA terms, the order of the seasonal terms, and the values of the exogenous variables to include in the model. In SARIMAX, the model's parameters are determined by fitting the model to the historical data. This is done by finding the parameters that minimize the error between the actual data and the model's predictions. This process, called grid search optimization, is used to determine the best set of parameters for the model.

Once the SARIMAX model has been fit to the data, it can be used to make forecasts of future values of the time series. It can also be used to analyze the impact of the exogenous variables on the time series data and to identify any underlying trends or patterns.

The parameters for the SARIMAX model are:

- Seasonality is denoted by the letter S. This indicates that the statistics are seasonal. Examples include the seasons of the year, which would affect how quickly the temperature changes in a particular region depending on the season, often being warmer in the summer and cooler in the winter. The "S" would be 12 if our data included monthly averages. Symbolized in our model as "s."
- AR is the abbreviation for Autoregressive. This indicates how similar the data are to any preceding time interval or period. It represents recurring patterns in the data, to put it simply. The term "auto-regressive" refers to a method that determines the points where the regression is greatest, indicating a pattern in the data, after regressing the data with a specific lag period of prior data, referred to in our model as "p."
- I is an acronym for integrated. "I" means that the difference between the new and old values has been used to replace the data values, referred to in our model as "d."
- The moving Average is referred to as MA. This phrase calculates the moving average over a specified number of periods. It is used to smooth out or minimize the noise in a model. The noise would be more smoothed out the longer the moving average period was, referred to in our model as "q."
- X is an acronym for exogenous. This considers a recognized outside element. In our model, this is not a parameter but an optional argument. Determining ideal model parameters is optional.

So, in our model, our parameters look like this:

SARIMAX (p,d,q) x (P,D,Q,s).

## 3. Photovoltaic Time Series Modeling

Photovoltaic (PV) time series modeling refers to the process of using statistical models to analyze and forecast time series data related to photovoltaic systems. Photovoltaic systems generate electricity from sunlight, and the time series data may include variables such as solar radiation, temperature, and electricity generation [1, 12, 13].

PV time series modeling is helpful for various purposes, such as predicting the electricity generation of a PV system at a future point in time, identifying trends and patterns in the data, and understanding the factors that influence electricity generation.

Several types of statistical models can be used for PV time series modeling, including autoregressive integrated moving average (ARIMA) models, seasonal decomposition models, and machine learning models. The choice of model will depend on the data's characteristics and the analysis's goals.

In order to fit a PV time series model to data, you will typically need to pre-process the data by removing any missing values and transforming the data as needed (e.g., taking differences or logarithms of the data to stabilize the variance). You will then need to specify the model and select appropriate model parameters. This can be done using maximum likelihood estimation or grid search techniques.

Time series modeling with PV data requires accurate and efficient models that can capture the seasonal and daily variations in the PV system's output. SARIMAX is one such model used successfully in PV time series modeling.

## 4. Advantages of Using SARIMAX for Photovoltaic Time Series Modeling

There are several advantages to using a SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous regressors) model for photovoltaic (PV) time series modeling [14-16]:

- SARIMAX models can capture both autocorrelation and seasonality in the data: The SARIMAX model is an extension of the ARIMA model, which is used for time series data that exhibits autocorrelation but not seasonality. The SARIMAX model adds the ability to include seasonal terms in the model, which allows it to capture the periodic nature of the data. This is particularly useful for PV time series data, which often exhibits autocorrelation and seasonality.

- SARIMAX models can include exogenous variables: The SARIMAX model allows you to include exogenous variables (also known as predictor or explanatory variables) in the model. This can be useful for PV time series modeling. Consider including variables such as solar radiation or temperature in the model to understand better the factors influencing electricity generation.
- SARIMAX models can be fitted using a variety of techniques. Several different techniques can be used to fit a SARIMAX model, such as maximum likelihood estimation or grid search. This gives you flexibility in how you fit the model to the data.
- SARIMAX models are widely used and well-studied: the SARIMAX model is a widely used and well-studied statistical model with a large body of literature and software implementations available. This means that a wealth of resources is available to help you understand and use the model effectively.

Overall, the SARIMAX model is a powerful tool for PV time series modeling, as it allows you to capture both autocorrelation and seasonality in the data and include exogenous variables in the model.

# 5. Grid Search Optimization of SARIMAX Models

Grid search optimization is a method for finding the best combination of parameters for a statistical model, such as a SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous regressors) model. It involves defining a grid of parameter values to test, fit and evaluate the model for each combination of parameters and selecting the combination that performs the best [17, 18].

To perform grid search optimization of a SARIMAX model, you must define the parameter grid you want to test. This will typically include the values for the autoregressive (AR), moving average (MA), and seasonal terms of the model, as well as any exogenous variables you want to include. You will also need to define a metric to use for evaluating the model's performance, such as root mean squared error (RMSE) or mean absolute error (MAE) [19-24]

Once the parameter grid and evaluation metrics have been defined, you can use a loop or a library such as sci-kit-learn's GridSearchCV to iterate over the parameter grid, fitting and evaluating the SARIMAX model for each combination of parameters. After the grid search is complete, you can select the combination of parameters that resulted in the best model performance as the optimal set of parameters for the model [25, 26].

It is also a good idea to visualize the grid search results by plotting the error for each combination of parameters to understand better how the different parameters affect the model's performance.

Overall, grid search optimization is a valuable method for finding the optimal set of parameters for a SARIMAX model. It allows you to systematically explore the parameter space and select the combination of parameters that results in the best model performance.

It should be noted that the data used in this article are from the Miroslava, Iasi, Romania site. Miroslava Photovoltaic Park is a good example of an investment in the local community's future. This 5 hectares solar farm, funded by the European Union, supplies power for the Miroslava commune - Iasi City, lowering administrative expenses and harmful emissions. The solar power plant has a capacity of 1MWp and is located in the municipality of Miroslava (Ciurbeşti) [27]. The efficiency of the installation is unique to this solar farm: electrical trackers position the panels to maximize exposure to sunshine. The photovoltaic park ensures the energy to public power lightning for all the commune and the local public institutions.

The solar plant is made up of the following parts:

- Mono-crystalline photovoltaic panels - 4,200 units producing electricity with a unitary installed power of 240 Wp/each unit, including their interconnection equipment, called concentrating boxes;

- Support elements for photovoltaic panels, consisting of 97 independent trackers with the possibility of tracking the sun on 2 axes, possibility of inclination 120 degrees - for increased efficiency;
- Three phase inverters - 9 units with a power of 12 kw each;
- Low-voltage power lines in the cable convert the DC electricity produced into AC electricity through inverters;
- Technical control room;
- A 45 kw diesel power group; HYUNDAI DHY45KSE; Maximum power 44 kVA / 35 kW (10.5 Kw in 230 V); Nominal power 40 kVA/32 kW; Regulation type AVR.
- 20 kV underground electric lines connect the power plant to the national power grid.

As mentioned in the previous section, the SARIMAX model is defined by three primary parameters: *(p,d,q) x (P, D, Q,s).*

- The first set of parameters *(p,d,q)* refers to the autoregressive (AR), the difference (I), and the moving average (MA) components of the model. These parameters are used to model the dependencies between the current value of the time series and its past values, as well as any trends or seasonality.
- The second set of parameters *(P, D, Q,s)* refers to the seasonal components of the model. These parameters are used to model the dependencies between the current value of the time series and past values of the same season, as well as the presence of any seasonality. "s" is the period of the seasonality, which is the number of time steps between repeated seasons. For example, if s is 12, the model assumes a monthly seasonality with 12 months in a year.

The combinations *(p,d,q) x (P, D,q,s)* that we have tested to find the optimum grid are the right ones:

(p,d,q) = [(0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1), (1, 0, 0), (1, 0, 1), (1, 1, 0), (1, 1, 1)]

*(P, D, Q,s)* = [(0, 0, 0, s), (0, 0, 1, s), (0, 1, 0, s), (0, 1, 1, s), (1, 0, 0, s), (1, 0, 1, s), (1, 1, 0, s), (1, 1, 1, 1, s)]

With

s = [3,4,12], which means seasonality 3, 4, and 12.

For the dataset we have treated in this paper, the results for the optimal grids are represented in the following Fig. (**1**).

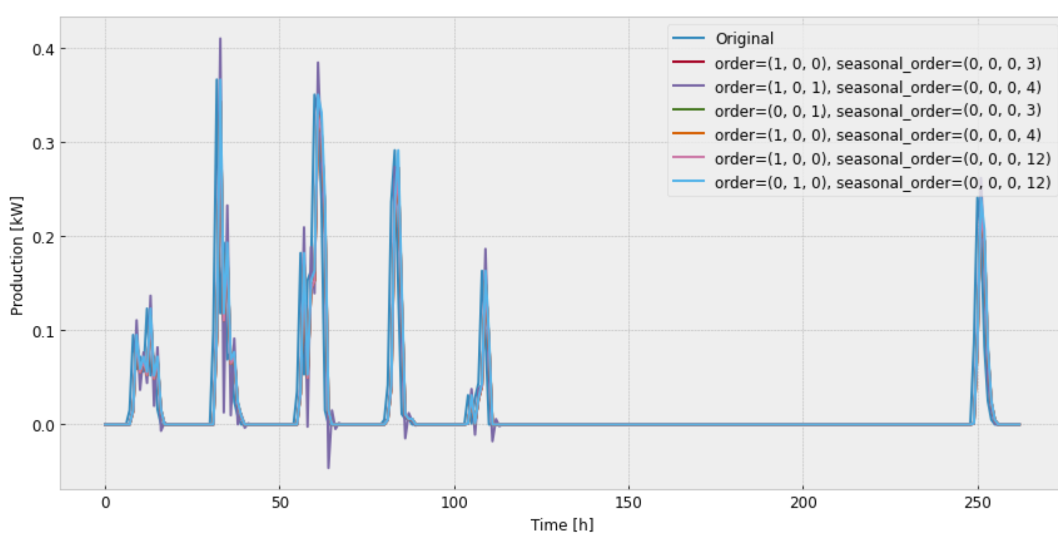


**Figure 1:** SARIMAX Forecasting.

We used the following metrics to rate the accuracy of each algorithm's prediction:

- **Mean Absolute Error (MAE):** Without considering their direction, MAE calculates the average size of mistakes in a set of forecasts. All individual differences are equally weighted in the test sample's average of the absolute disparities between prediction and observation.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

- **Root mean squared error (RMSE)** is a scoring formula for quadratic equations that also calculates the average error magnitude. It is the average of the discrepancies between predicted results and actual observations.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2}$$

- **Mean absolute percentage error (MAPE)** is often referred to as the mean absolute percentage deviation (MAPD), a metric for forecasting technique accuracy. The accuracy is often expressed as a ratio determined by the following formula:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{3}$$

- **Coefficient of determination R²**: The percentage of variance in the dependent variable that can be predicted from the independent variables is known as the coefficient of determination, abbreviated $R^2$ and pronounced: "R squared." It is a statistic applied to statistical models whose main objective is to forecast future events or test hypotheses using data from other relevant sources. Based on the percentage of the overall variance in outcomes that the model accounts for, it gives a gauge of how the model duplicates well-observed results.

A data set has n labeled values $y_1, \ldots$, each with a predetermined value $\hat{y}_1, \ldots, \hat{y}_n$.

Residuals are stated as $e_i = y_i - \hat{y}_i$

Considering that $\bar{y}$ is the observed data's mean:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{4}$$

then the variability of the data set can be measured with two Mean squared error formulas:

- The residual sum of squares (also known as the mean square error)

$$SS_{res} = \sum_i (y_i - \hat{y}_i)^2 = \sum_i e_i^2 \tag{5}$$

- The total sum of squares:

$$SS_{tot} = \sum_i (y_i - \bar{y}_i)^2 \tag{6}$$

The most general definition of the coefficient of determination is

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{7}$$

In the ideal scenario, the predicted and observed values coincide precisely, leading to $SS_{res}=0$ and $R^2=1$. A baseline model, which always predicts $\bar{y}$, will have $R^2=0$. Models that have worse predictions than this baseline will have a negative $R^2$.

The following Fig. (**2**) shows the metrics for each grid.

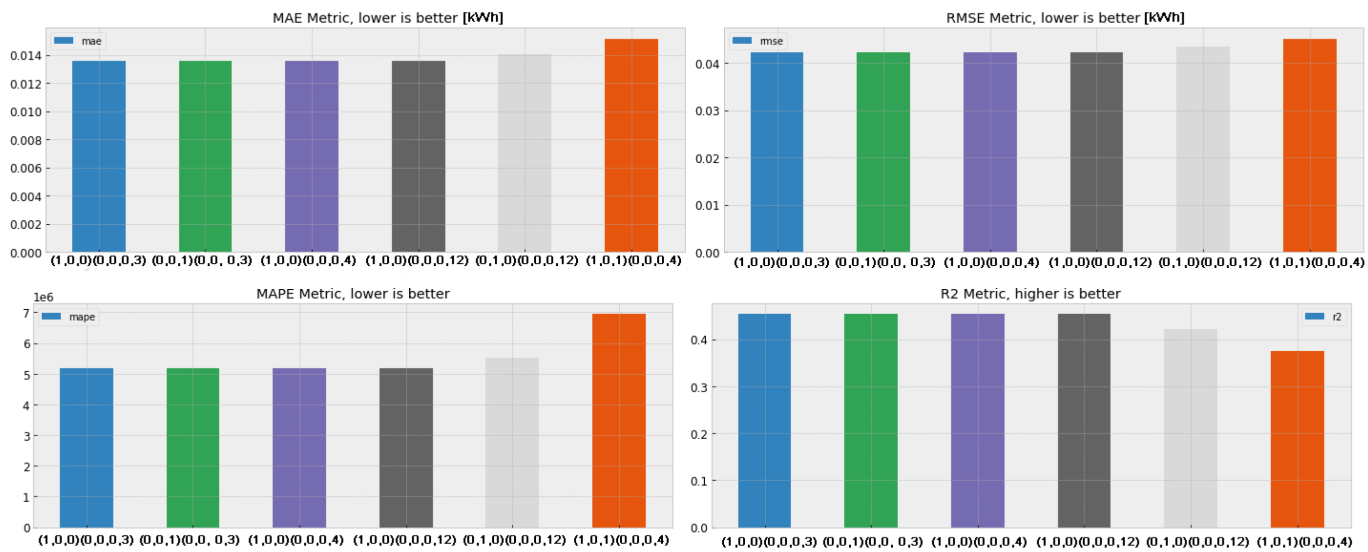


**Figure 2:** Metrics for each grid.

# 6. Implementing SARIMAX for Photovoltaic Time Series Modeling

To implement SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous regressors) for photovoltaic (PV) time series modeling, you can follow these steps [28, 29].

- Load and pre-process the data: Load the PV time series data and perform any necessary pre-processing, such as removing missing values and transforming the data as needed (e.g., taking differences or logarithms of the data to stabilize the variance).
- Define the exogenous variables: Determine which ones you want to include in the model and extract them from the data. Exogenous variables are additional predictors or explanatory variables that may influence the time series data.
- Specify the SARIMAX model: Use the SARIMAX function from the statsmodels library to specify the SARIMAX model, including the values for the AR, MA, and seasonal terms, as well as the exogenous variables.
- Fit the SARIMAX model: Use the appropriate method of the SARIMAX model object to fit the model to the data. You may use a technique such as maximum likelihood estimation or grid search to find the optimal set of parameters for the model.
- Evaluate the model fit: Use a metric such as root mean squared error (RMSE) or mean absolute error (MAE) to evaluate the model's fit to the data. You may also want to visualize the model's fit to the data using the plot_diagnostics method of the model object.
- Make forecasts: Use the prediction method of the model object to make forecasts of future values of the time series. You can also use the plot_predict method to visualize the forecasts made by the model.
- Analyze the results: Use the model fit and forecasts to analyze the variables' relationships and identify any underlying trends or patterns in the data.

Overall, implementing SARIMAX for PV time series modeling involves fitting the model to the data, evaluating the model fit, making forecasts, and analyzing the results to understand the factors that influence electricity generation and predict future values of the time series.

## 7. Challenges in Using SARIMAX for Photovoltaic Time Series Modeling

There are several challenges that you may encounter when using a SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous regressors) model for photovoltaic (PV) time series modeling:

- Selecting the optimal set of parameters: Finding the optimal set of parameters for the SARIMAX model can be challenging, as the model has several parameters (e.g., the orders of the AR, MA, and seasonal terms) that need to be specified. Use a technique such as grid search or maximum likelihood estimation to find the optimal set of parameters for the model.
- Handling missing values and outliers: Missing values and outliers in the data can affect the SARIMAX model's fit and the forecasts' accuracy. You may need to apply imputation or outlier detection and removal techniques to handle these issues properly.
- Dealing with non-stationarity and seasonality: Time series data that exhibit non-stationarity or seasonality can be challenging to model accurately. You may need to apply techniques such as differencing or log transformation to stabilize the data's variance and make it more suitable for modeling.
- Incorporating exogenous variables: Incorporating exogenous variables (additional predictor or explanatory variables) in the SARIMAX model can be challenging. You may need to carefully select and pre-process the variables to get good results.

Overall, using SARIMAX for PV time series modeling requires careful consideration of the characteristics of the data and the goals of the analysis, as well as a thorough understanding of the model and the techniques used to fit and evaluate it.

## 8. Best Practices for SARIMAX Grid Search Optimization

When using the grid search function in the SARIMAX class to optimize the model's parameters, it is important to follow certain best practices. Firstly, choosing an appropriate range of parameters for the grid search is important. This range should be wide enough to ensure optimal parameters are found but not so vast that the grid search takes too long.

Secondly, it is important to choose an appropriate error metric for evaluating the model's performance. The most common error metric for PV time series modeling is the Root Mean Squared Error (RMSE), which measures the average difference between actual and predicted values.

Finally, it is important to use a validation set when evaluating the model's performance. A validation set is a set of data used to evaluate the model's performance after it has been trained on the training data. The validation set should be chosen carefully to ensure that it represents the data the model will be used to predict.

## 9. Conclusion

SARIMAX is a powerful PV time series modeling tool and has been used successfully in various applications. It can easily incorporate seasonal components in the model and can be optimized using the grid search method.

When using SARIMAX for PV time series modeling, it is important to:

- Define the parameter grid carefully: Make sure to consider the range of values for each parameter you want to test based on the characteristics of the data and any prior knowledge or assumptions you have about the data. It can be helpful to visualize the data (e.g., by plotting the autocorrelation and partial autocorrelation plots) in order to inform your choice of parameter values.

- Split the data into training and testing sets: It is important to evaluate the model on unseen data, so make sure to split the data into a training set and a testing set. You can use the training set to fit the models and the testing set to evaluate their performance.
- Use a cross-validation procedure: In addition to splitting the data into training and testing sets, it can be helpful to use a cross-validation procedure such as k-fold cross-validation to further evaluate the model's performance. This can reduce the risk of overfitting the model to the data.
- Use a robust evaluation metric: Choose an evaluation metric appropriate for the data's characteristics and the analysis's goals. For example, root means squared error (RMSE) or mean absolute error (MAE) is often used for time series data. Make sure to use the same evaluation metric for all of the grid search models to make meaningful comparisons between them.
- Visualize the results: It can be helpful to visualize the grid search results, such as by plotting the error for each combination of parameters, to understand better how the different parameters affect the model's performance. This can help you to identify patterns and trends in the data and to choose the best-performing model.

In this paper, we survey the possible grid combinations for SARIMAX and evaluate the metrics of each combination in order to choose the most suitable one to model our dataset. The data used to test and evaluate the model's prediction are from the Miroslava site and two years of data from 2020 and 2021. Six combinations best represent the model of our dataset. The results of the production prediction made by the model using these combinations have been presented, as well as the metrics of the latter. In short, the optimal grid search method allows to optimization of the SARIMAX-based model. This article has demonstrated its usefulness and advantages.


## Acknowledgements

The Eugen Ionescu Scientific Mobility Grant, provided by the Ministry of Foreign Affairs of Romania and administered by the Francophone University Agency (AUF), allowed this research to be completed. Without identifying them, we want to express our gratitude to everyone who contributed to the grant-awarding process.

Additionally, we would like to express our gratitude to the Gheorghe Asachi Technical University of Iaşi (TUIASI) for allowing us to perform research there, particularly in the Faculty of Electrical Engineering, Energetics and Applied Informatics (EEEIA). We also want to thank the EEEIA academic staff and researchers for sharing their pearls of wisdom with us and providing insight and expertise that greatly assisted the research.

Of course, this research would not have been possible without the contribution of Miroslava’s team, who made their data accessible, for which we also thank them.


## References


[1] Huang X, Li Q, Tai Y, Chen Z, Liu J, Shi J, *et al*. Time series forecasting for hourly photovoltaic power using conditional generative adversarial network and Bi-LSTM. Energy. 2022; 246: 123403. https://doi.org/10.1016/j.energy.2022.123403

[2] Vogt S, Schreiber J, Sick B. Synthetic photovoltaic and wind power forecasting data. ArXiv 2022 ; arXiv:2204.00411v1 [cs.LG].

[3] Markovics D, Mayer MJ. Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction. Renew Sustain Energy Rev. 2022; 161: 112364. https://doi.org/10.1016/j.rser.2022.112364

[4] Carneiro TC, de Carvalho PCM, Alves dos Santos H, Lima MAFB, Braga AP de S. Review on photovoltaic power and solar resource forecasting: current status and trends. J Sol Energy Eng. 2022; 144: 010801. https://doi.org/10.1115/1.4051652

[5] Zhou Y, Wang J, Li Z, Lu H. Short-term photovoltaic power forecasting based on signal decomposition and machine learning optimization. Energy Convers Manag. 2022; 267: 115944. https://doi.org/10.1016/j.enconman.2022.115944

[6] Fan GF, Wei HZ, Chen MY, Hong WC. Photovoltaic power generation forecasting based on the ARIMA-BPNN-SVR model. Glob J Energy Technol Res Updat. 2022; 9: 18-38. https://doi.org/10.15377/2409-5818.2022.09.2

[7] Sah S, Surendiran B, Dhanalakshmi R, Yamin M. Covid-19 cases prediction using SARIMAX Model by tuning hyperparameter through grid search cross-validation approach. Expert Syst. 2022; e13086. https://doi.org/10.1111/exsy.13086

[8] Zhao Z, Zhai M, Li G, Gao X, Song W, Wang X, *et al*. Study on the prediction effect of a combined model of SARIMA and LSTM based on SSA for influenza in Shanxi Province, China. Res Sq. 2023; 23: Article number 71. https://doi.org/10.21203/RS.3.RS-2081379/V1

[9] Ismail BI. Maximizing power generation of a solar PV system for a potential application at musselwhite gold mine site in northwestern ontario, canada. Glob J Energy Technol Resh Updat. 2020; 7: 1-11. https://doi.org/10.15377/2409-5818.2020.07.1

[10] Harat ZA, Asadi Zarch MA. Comparison of SARIMA and SARIMAX for long-term drought prediction. Desert Management 2022; 10: 1-16.

[11] Kumar SS, Kumar A, Agarwal S, Syafrullah M, Adiyarta K. Forecasting indoor temperature for smart buildings with ARIMA, SARIMAX, and LSTM: A fusion approach. 2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), IEEE; 2022, p. 186-92. https://doi.org/10.23919/EECSI56542.2022.9946498

[12] Sharadga H, Hajimirza S, Balog RS. Time series forecasting of solar power generation for large-scale photovoltaic plants. Renew Energy. 2020; 150: 797-807. https://doi.org/10.1016/j.renene.2019.12.131.

[13] Zheng J, Zhang H, Dai Y, Wang B, Zheng T, Liao Q, *et al*. Time series prediction for output of multi-region solar power plants. Appl Energy. 2020; 257: 114001. https://doi.org/10.1016/j.apenergy.2019.114001

[14] Zhou H, Zhang Y, Yang L, Liu Q, Yan K, Du Y. Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism. IEEE Access 2019; 7: 78063-74. https://doi.org/10.1109/ACCESS.2019.2923006

[15] Khazaei H, Moslemi R, Sharma R. Stochastic decision-making model for aggregation of residential units with PV-systems and storages. 2020 IEEE Power & Energy Society General Meeting (PESGM), IEEE; 2020, p. 1-5. https://doi.org/10.1109/PESGM41954.2020.9281448

[16] Rafi A, Lee T, Wu W. A comparative study of methods to forecast domestic energy consumption aggregated with photovoltaic and heat pumps system. 2021 26th International Conference on Automation and Computing (ICAC), IEEE; 2021, p. 1-6. https://doi.org/10.23919/ICAC50006.2021.9594149

[17] Steenbergen W. Adjusted grid search to find hyper-parameters in SARIMAX models: efficiently filling the shelves in kruidvat stores. Student Res Conf. 2018; 4: 1-4.

[18] Omar MS, Kawamukai H. Comparison between the holt-winters and SARIMA models in the prediction of NDVI in an arid region in Kenya using pixel-wise NDVI time series. Acad J Res Sci Publish. 2021; 2: 1-15. https://doi.org/10.52132/Ajrsp/en.2231

[19] Singh RK, Rani M, Bhagavathula AS, Sah R, Rodriguez-Morales AJ, Kalita H, *et al*. Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated moving average (ARIMA) model », JMIR public health and surveillance. JMIR Public Health Surveill. 2020; 6(2): e19115. https://doi.org/10.2196/19115

[20] Chicco D, Warrens MJ, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput Sci. 2021; 7. https://doi.org/10.7717/peerj-cs.623

[21] Wang W, Lu Y. Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model. IOP Conf Ser Mater Sci Eng 2018; 324: 012049. https://doi.org/10.1088/1757-899X/324/1/012049

[22] Karunasingha DSK. Root mean square error or mean absolute error? Use their ratio as well. Inf Sci (N Y). 2022; 585: 609-29. https://doi.org/10.1016/j.ins.2021.11.036

[23] Priyadarshini I, Mohanty P, Kumar R, Taniar D. Monkeypox outbreak analysis: an extensive study using machine learning models and time series analysis. Computers. 2023;12: 36. https://doi.org/10.3390/computers12020036

[24] Barajas MA, Murphy MP, Lasseter LC, Sunny GI, Mazumdar H, Gohel HA, *et al*. Seasonal trend assessment for groundwater contamination detection and monitoring using ARIMA model. 2023 IEEE 2nd International Conference on AI in Cybersecurity (ICAIC), févr; 2023, p. 1-7. https://doi.org/10.1109/ICAIC57335.2023.10044126

[25] Alhakeem ZM, Jebur YM, Henedy SN, Imran H, Bernardo LFA, Hussein HM. Prediction of ecofriendly concrete compressive strength using gradient boosting regression tree combined with GridSearchCV hyperparameter-optimization techniques. Materials. 2022; 15: 7432. https://doi.org/10.3390/ma15217432

[26] Priya SM. Hyper tuning using gridsearchcv on machine learning models for prognosticating dementia. Research Square. 2022; p. 1-15.

[27] B. Jovičić. Mecasolar supplies 1 MW in 2-Axis trackers. Balkan Green Energy News, May 27, 2015.

[28] Kim E, Akhtar MS, Yang O-B. Designing solar power generation output forecasting methods using time series algorithms. Electr Power Syst Res. 2023; 216: 109073. https://doi.org/10.1016/j.epsr.2022.109073

[29] Idman E, Idman E, Yildirim O. Estimating solar power plant data using time series analysis methods. 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey: IEEE; 2020, p. 1-6. https://doi.org/10.1109/HORA49412.2020.9152839