

Statistical Validation and Machine Learning Prediction of Solar Power Generation in a Developed Automated Multi-Source Power Generation System

¹Akinrinade, N. A., ^{2,3*}Onawumi, A. S., ^{2,3}Sangotayo, E. O., ⁴Oke, A. M., ¹Omoniyi, E. B., ⁵Ajibowu, S. B., ⁶Oni, A. A. and ⁶Ajayi, M. O.

¹Department of Mechanical Engineering, College of Engineering, Bells University of Technology, Ota, Ogun State, Nigeria.

²Department of Mechanical Engineering, Faculty of Engineering, Ladoke Akinola University of Technology, Ogbomoso. Nigeria.

³Department of Mechanical and Mechatronics Engineering, College of Engineering and Technology, Achievers University, Owo. Nigeria.

⁴Department of Agricultural and Bio-resources Engineering, Bells University of Technology, Ota, Ogun State, Nigeria.

⁵Department of Electrical, Electronics and Telecommunication Engineering, Bells University of Technology, Ota, Ogun State, Nigeria.

⁶Department of Computer Engineering, College of Engineering, Bells University of Technology, Ota, Ogun State, Nigeria.

*Corresponding author: asonawumi@lautech.edu.ng

Abstract

The rising integration of renewables into modern energy grids, optimising operational management and reliability in Automated Multi-Source Power Generation Systems (AMPGS), is critical for grid stability. Models of AMPGS currently available are mostly not validated with statistical testing and do not provide accurate prediction tools for operation. This paper represents statistical validation and machine learning predictive modelling of an AMPGS. Two data set consist of 1000 experimental and simulated observations were analysed using six (6) variables, which include Solar Power Current (SPC), Battery Level (BL), Temperature (Temp), Irradiance (Ir), Response Time (RT) and Solar Power Generation (SPG). Kolmogorov-Smirnov (KS) test, t-test, histogram comparison and correlation test were considered for statistical analysis. Eight (8) machine learning algorithms: Artificial Neural Network (ANN), Random Forest (RF), Extra Trees, Gradient Boosting (GB), AdaBoost, Decision Tree, K-Nearest Neighbour (kNN) and XGBoost were trained using 80:20 training-to-testing split. These machine learning algorithms were evaluated by prediction accuracy values such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2). The KS test results were given by statistics as 0.045 and a p-value of 0.1383. The p-value from the t-test was 0.3329, proving no significant differences between experimental and simulation data. The Extra Trees machine learning algorithm yielded the best results with R^2 of 0.993 for the experimental and simulation datasets. The results show that the developed AMPGS simulation is able to replicate the measured SPG behaviour with good statistical agreement and that Extra Trees is an accurate data-driven model for SPG prediction.

Keywords: Hybrid Energy System, Machine Learning, Extra Trees Algorithm, Statistical Validation, Renewable Energy Management.

Introduction

The growing concern on the need for reliable, environmentally friendly and sustainable energy around the world, especially due to its reduced impact on human health and the atmosphere, has led the global market to step towards the adoption of renewable energy technologies (Schleifer et al., 2023). Solar Photovoltaic (PV) power systems (one of the solar energy technologies) have therefore gained special attention given the widely available and lean the Sun's energy, coupled with decreasing costs of solar photovoltaic

installations (Onawumi et al., 2022). However, with many positive points, the variability and intermittent characteristics of electricity generation from solar systems still pose difficulties in maintaining the reliability and efficiency of power supply, stable grid operation and suitable load matching (Adeyinka et al., 2024). Therefore, the new hybrid energy solutions, which are the combinations between solar PV and energy storage systems, utility grid and back-up generator, emerged as essential alternatives to smart power generation systems to enhance their operational flexibility, robustness and energy dependability (Rajendran et al., 2025).

Automated Multi-source Power Generating system (AMPGS) generally involves various sources of generation and intelligent control schemes which ensure the continuous power supply at comparatively cheaper operating cost and minimum environmental pollution. These systems employ various sources on demand depending on the various resource availability, load demand and the state of charge of the battery (Akinrinade et al., 2026a). The use of renewable sources with conventional sources makes the system resilient, as well as maintaining the sustainability of the world. In a recent study, a hybrid solar/battery/generator system showed better energy utilisation and less dependency on conventional fuel, with less disturbance on supply reliability (Ijeoma et al., 2024). Nevertheless, the modelling and simulation of AMPGS design and usage require proper characterisation and validation such that the simulation results may be truly representative of the operational system in conditions similar to those of the real system (Ambelu et al., 2025). In the simulation of AMPGS, the commonly used tools are MATLAB/Simulink for the assessment of control operation of AMPGS, prediction of operation, and estimation of dynamic response. So, statistical validation is still necessary for checking the accuracy, effectiveness and practicality of the simulation outputs (Golestan et al., 2024). Histogram comparison, correlation analysis, the Kolmogorov-Smirnov (KS) test, and hypothesis testing are generally used in the comparison of the experimental and simulated data sets (Mundu et al., 2024).

Over the last decade, Machine Learning (ML) technology has been widely used for intelligent energy management and forecasting. ML can also describe nonlinear correlations between multiple variables in energy systems to predict output accurately and outperform the analytic method in these aspects (Akinrinade et al., 2026b). This has already been applied to ML in renewable energy management to successfully forecast power generation, energy load, and system output in various ambient situations. The performance of ensemble learning methods includes Extra Trees, Random Forest (RF), Gradient Boosting (GB) and XGBoost because of their robust model structure, resistance to overfitting, capability for complex nonlinear inter-variable relations, and high prediction accuracy (Banik and Biswa, 2024). Although there is abundant literature on prediction or hybrid models, none of them combines the rigorous statistical verification of simulation with a comparative study between multi-machine learning algorithms in a multi-source power generation automated system. Hence, the study combined validation and prediction that greatly benefit the enhancement of the reliability and smartness of the hybrid energy generation system.

This work, therefore, offers an end-to-end statistical validation and a machine learning-based prediction methodology for solar power forecasting in an AMPGS. The experimental and simulated datasets were validated with standard statistical measures, and eight machine learning methods were used and compared to ascertain which machine learning model offers the highest prediction capabilities. The results of this study provide a thorough foundation for hybrid energy system simulations and enhance intelligent forecasting in energy management.

Materials and Methods

Data acquisition and preprocessing

Data set of 1000 observations from the experimental system as shown in Figure 1 were utilised while simulated data were produced from the MATLAB/Simulink models of the developed AMPGS including the solar, grid, generator, battery storage and load management strategies was supplied with experimentally

measured inputs including Solar Panel Current (SPC), Battery Level (BL), Irradiance (Ir), Temperature (Temp), Response Time (RT), and Solar Power Generation (SPG) under the same operational conditions as the physical system. The model outputs, SPC, BL, Temp, Ir, RT, and SPG, were produced for 1000 observations and exported for statistical validation against the corresponding experimental dataset. The corresponding simulated data generated from the validated model were used for the purpose of statistical validation. Statistical validation of the designed system involved distribution comparison, the KS-test, correlation analysis and t-test to check the congruency between the experimental and simulation data. Analyses were performed to investigate whether there was any significant difference between the two data sets. Models based on ensemble methods were used for prediction accuracy. The procedural steps for the statistical validation are demonstrated in Figure 2.



Figure 1: Experimental setup

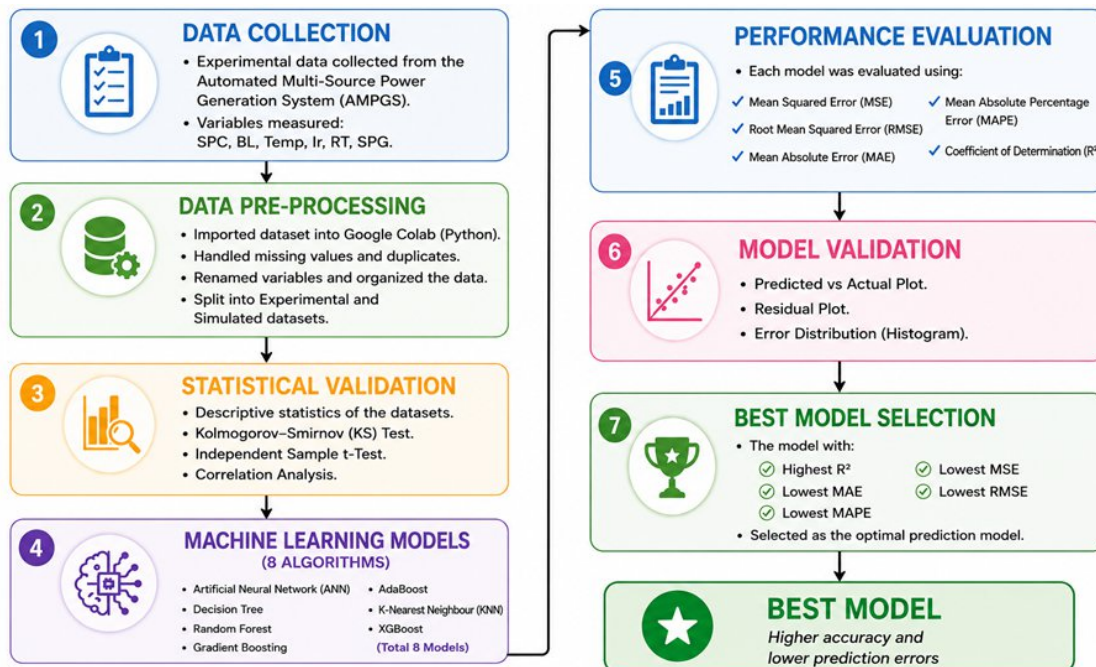


Figure 2: Procedural Steps for the Statistical Validation

Statistical validation of datasets

Statistical validation was performed to assess the similarity between the experimental and simulated datasets using distribution analysis and hypothesis testing techniques. This ensured that the simulated data reliably represent the behaviour of the real hybrid energy.

Distribution similarity analysis

This section deals with the approaches used for quantitatively comparing the experimental and simulated datasets using both visual and statistical methods to ensure consistency, reliability, and accurate comparison of distribution characteristics. The parameters used are histogram comparison, the KS-test and correlation.

Histogram comparison

Histogram plots were used as a graphical comparison tool in order to see the shape of the distribution between experimental and simulated data. The two datasets are plotted on two individual histograms, and the comparison of these datasets within the same graphical space was made based on the overall shape of the distribution, spread of data and location of the data in order to make the first comparison between the two datasets and check if these values match.

KS- test

The test was applied to quantitatively evaluate the similarity between the distributions of the two datasets. The KS statistic D is expressed by Equation 1 (Tanguet and Njomen, 2021).

$$D_{n,m} = \sup_x \left| \frac{1}{n} \sum_{i=1}^n I(X_i \leq x) - \frac{1}{m} \sum_{j=1}^m I(Y_j \leq x) \right| \quad (1)$$

Where $D_{n,m}$ is the two-sample Kolmogorov–Smirnov (KS) statistic, n and m are the sample sizes of the two datasets, X_i and Y_j are the observations, and $I(\cdot)$ is the indicator function, and x is the evaluation point. The KS statistic represents the maximum absolute difference between the two empirical cumulative distribution functions.

Correlation Analysis

The relationship between input and output variables was evaluated using Pearson correlation as presented in Equation 2. This identifies dominant parameters influencing solar power generation (Memić *et al.*, 2024).

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (2)$$

Where x_i , y_i are individual data points in the sample, $\bar{x}$, $\bar{y}$ is the mean (average) of all x values and all y values.

Hypothesis Testing

Statistical hypothesis testing was conducted using the independent t-test to further validate dataset similarity by assessing whether there is a statistically significant difference between the mean values of the experimental and simulated datasets. The hypotheses are defined as:

H_0 is no significant difference between datasets, H_1 is a significant difference in the datasets. An independent t-test was applied using Equation 3 (Brahim *et al.*, 2025)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (3)$$

Where $\bar{x}_1, \bar{x}_2$ are the sample means (averages) for groups 1 and 2, S_1^2, S_2^2 are the sample variances for each group. They measure how spread out the data points are, n_1, n_2 are the sample sizes in each group. The decision rule is that when $P > 0.05$, i.e., accept H_0

Machine learning model development

The machine learning analysis was implemented using Python in Google Colab. Eight regression models were developed and evaluated for predicting SPG. These models include Artificial Neural Network (ANN), RF, Extra Trees, Boosting, AdaBoost, Decision Tree, K-Nearest Neighbour (kNN), and Extreme Gradient Boosting (XGBoost). In accordance with Al-Dahidi et al. (2024), the algorithms listed above were chosen because their capabilities have been shown to excel when it comes to simulating nonlinear relationships, utilising the characteristic properties of engineering datasets and delivering trustworthy predictions for the purpose of power forecasting in renewable energies. The chosen algorithms exemplify several different Machine Learning methodologies. Tree-based ensemble algorithms (RF, Extra Tree, GB, XGBoost) are best utilised for learning highly non-linear relations and reducing high variances by combining estimates. The ANN is chosen for its ability to learn complex relationships, although it generally relies on large dataset sizes. kNN presents an easy instance-based learning method that considers nearby local data and a Decision Tree, for its ease of interpretation and usage as a benchmark model.

Training and testing procedure

Training on the original dataset – 80% training, 20% testing, respectively. The models which were created were tested using 80% of the original datasets for machine learning algorithms (training set), while 20% were held out from usage (test set) for purposes of evaluation. An 80:20 split, a split that should lead to a functional model learning as well as effective performance evaluation on data that the algorithms would likely never see during training, hence no overfitting.

Hyperparameter tuning and model configuration

Before the final model training phase, all selected machine learning algorithms, namely ANN, RF, Extra Trees, GB, AdaBoost, Decision Tree, kNN, and XGBoost, were subjected to a systematic hyperparameter tuning procedure. This process was undertaken to minimise prediction error and improve model generalisation in regression tasks and mitigate the risk of overfitting, particularly given the moderate dataset size of approximately 1000 observations. Hyperparameters are a predefined set of configuration settings that are not learnt by the machine learning model but control the learning process. They need to be chosen carefully since poor values can lead to underfitting or overfitting. Hyperparameter optimisation was performed in this study using the grid search cross-validation method, which tests a grid of hyperparameter value combinations and looks for the best using cross-validation. Given that this study involves a regression task, model evaluation during tuning was conducted using standard regression performance metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and the Coefficient of Determination (R^2). These performance metrics give good insight into how accurately a prediction can be made and how much explanatory power the model possesses. Tuning of the model was performed in accordance with the nature of the algorithm; the hyperparameters are adequate considering the model architecture and dataset size. The chosen optimised hyperparameter settings, with Python libraries (Scikit-learn and XGBoost), are summarised as follows:

ANN was built using a Multilayer Perceptron Regressor. Hyperparameters such as hidden layer sizes (100, 50), ReLU activation, learning rate (0.001) and maximum iterations (1000) were trialled and optimised to maximise the ANN's ability to learn non-linearly, generalise, converge and predict accurately.

RF: The RF was set to the following combination of hyperparameters: `n_estimators=200`, `max_depth=20`, `min_samples_split=2`. The values of these parameters were selected by grid-search cross-validation, comparing different parameters by how well they predicted the validation dataset. The chosen set of parameter values provided a balance between prediction accuracy and generalisation ability, while reducing the model variance and overfitting.

Extra Trees: The Extra Trees Regressor(ET) model with parameters `n_estimators=300`, `max_depth=15`, and `min_samples_split=2`. With `n_estimators=300`, the number of trees is higher, so the randomness involved is larger, which increases the generality of the model. A max depth of 15 limits the complexity of the tree and should be suitable for the size of the data.

GB: The GBRegressor model was used and tuned with `n_estimators = 200`, learning rate = 0.05 and `max_depth = 4` in order to perform gradual learning and reduce overfitting.

AdaBoost: The AdaBoost model was trained using the AdaBoostRegressor algorithm with parameters set to `n_estimators=200` and learning rate=0.1 to perform iterative correction to the predictions of weak learners.

Decision Tree: The model (DecisionTreeRegressor) is set with `max_depth=10` and `min_samples_split=2`. The depth constraint prevents overfitting, limiting the tree's growth while enabling the model to learn non-linear patterns in the data. The model is a simple and interpretable benchmark.

kNN: For the KNN model (K-Neighbours Regressor), the hyperparameter `n_neighbors` was set to 5. This quite small neighbour size allows the model to capture local behaviour in the data, and given the dataset size of 1000 samples, this value provides a reasonable amount of tuning for the model.

XGBoost: For this model (XGBoostRegressor), `n_estimators=300`, learning rate=0.05, and `max_depth=6` were used to optimise the predictions by making GB with regularisation more capable in terms of performance. The low learning rate and medium depth made it possible to avoid any overfitting; the number of estimators makes the model capable of learning more data, and hence reduces the error.

Following the well-defined hyperparameter optimisation, optimal values were obtained for each parameter of the respective models, which are used in the further stage of training the models. The balance of model complexity to dataset size helps to find the model that has less bias and variance, and it will provide accurate predictions in the case of predicting continuous values, which results in robust predictions as shown in Table 1.

Table 1: Optimised hyperparameters

Model	Hyperparameters
ANN	Hidden layers=(100,50), activation='relu', learning_rate_init = 0.001, max_iter=1000
RF	<code>n_estimators=200</code> , <code>max_depth=20</code> , <code>min_samples_split=2</code>
Extra Trees	<code>n_estimators=300</code> , <code>max_depth=15</code> , <code>min_samples_split=2</code>
GB	<code>n_estimators=200</code> , learning_rate=0.05, <code>max_depth=4</code>
AdaBoost	<code>n_estimators=200</code> , learning_rate=0.1
Decision Tree	<code>max_depth=10</code> , <code>min_samples_split=2</code>
kNN	<code>n_neighbors=5</code>
XGBoost	<code>n_estimators=300</code> , learning_rate=0.05, <code>max_depth=6</code>

Comparative model assessment

The prediction abilities of the developed ML models were compared based on experimental and simulation data by evaluating their predictive performance using statistical performance metrics (MSE, RMSE, MAE, MAPE and R^2) to find the best model which has prediction accuracy, lower error, and good generalisation ability. The best model was chosen according to its accuracy, reliability and consistency in system prediction.

Error analysis

To further analyse the predictive capability and dependability of the machine learning models generated, residuals were analysed, and error analysis was carried out. Residuals are calculated as the actual value minus the predicted value, and a plot of residuals was generated to evaluate the distribution of the prediction errors across the observations. A randomly scattered plot of residuals around the zero point would suggest that the model has no bias and that the data are appropriately fitting the trend line. An error distribution histogram was plotted to investigate the distribution and frequency of the errors.

Results and Discussion

Validation of the Experimental and Simulation Results

Figure 2 shows the distributions of SPG for both experimental (blue) and simulated (orange) data. Both distributions are approximately normal (bell-shaped) with a similar shape. The experimental data is concentrated and centred about 150, with the values mostly ranging from 140-165, with low variance. The simulated data is more dispersed toward the higher values up to about 200, implying high variance. Both distributions have strong overlap, showing similar trends, with simulated data being slightly shifted to the right, predicting higher SPG values.

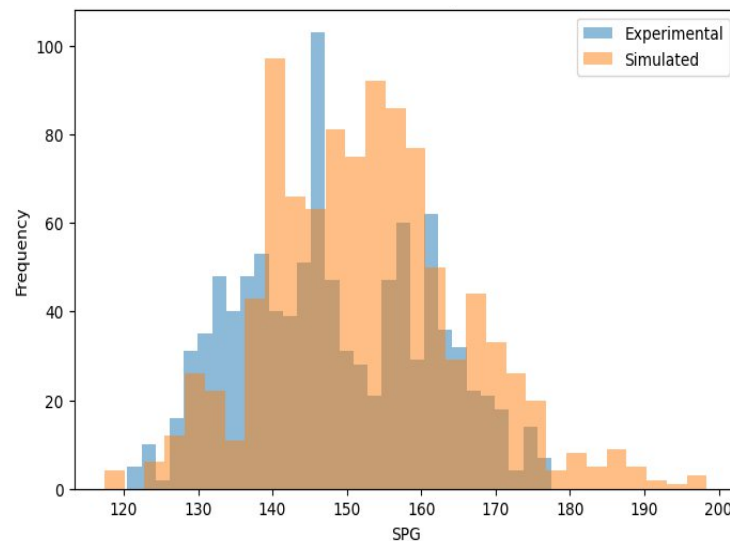


Figure 2: Distribution comparison between the experimental and simulated SPG data

KS-test results

The KS test was employed to assess the similarity between the original dataset and the generated dataset. The test produced a KS statistic of 0.045 and a p-value of 0.1383, as shown in Table 2. The KS statistic, which is the maximum difference between the cumulative distribution functions of the two datasets, is relatively small, which demonstrates the similarity between the two distributional structures, but importantly, the p-value exceeds 0.05, the typical significance value for rejecting a null hypothesis, which in this context would mean that there is no statistically significant difference between the two. This thus verifies that the generated data successfully captures the statistical properties of the original dataset, such as its distribution and volatility and thus demonstrates the confidence of the data reproduction method and

the validity of the synthetic dataset as statistically similar to the original, and as such is capable of use within the current system performance model and predictive model construction for this study.

Table 2: KS-test results

Test	Value
KS-statistics	0.045
p-value	0.1383

Correlation analysis results

Figure 3(a) presents the heatmaps showing the correlation between the major system parameters (SPC, BL, Temp, Ir, RT, SPG) for experimental and simulated datasets. It can be observed that SPC shows a positive correlation to SPG, BL and Ir ($r=0.61$), ($r=0.073$), and ($r=0.055$), emphasising irradiance as the major influence on photovoltaic energy generation. These findings are in accordance with Al-Dahidi *et al.* (2024) and Ndiaye *et al.* (2025), who claim that irradiance is the leading parameter that influences both PV energy yield and the current produced. BL($r=0.76$) and SPC($r=0.61$) were also highly positively correlated with SPG, which means battery level and current generated will also influence the overall system (Bamisile *et al.*, 2025). On the contrary, the relationship between SPG and temperature is weakly negatively correlated ($r = -0.15$) due to the negative influence of temperature on the overall efficiency of PV at high-temperature operating points.

As depicted in Figure 3(b), the simulated dataset reflected correlation profiles that were very similar to the experimental dataset; BL had the highest positive correlation with SPG ($r=0.70$), followed by SPC ($r=0.54$). Temp, Ir and RT had weak inverse correlations with SPG, with correlation coefficients of -0.10 , -0.14 and -0.09 , respectively. Similar correlation profiles for the experimental and simulated data suggest that the MATLAB/Simulink model can recreate the relationships between the major variables of the system and solar power generation, thus validating the developed simulation model.

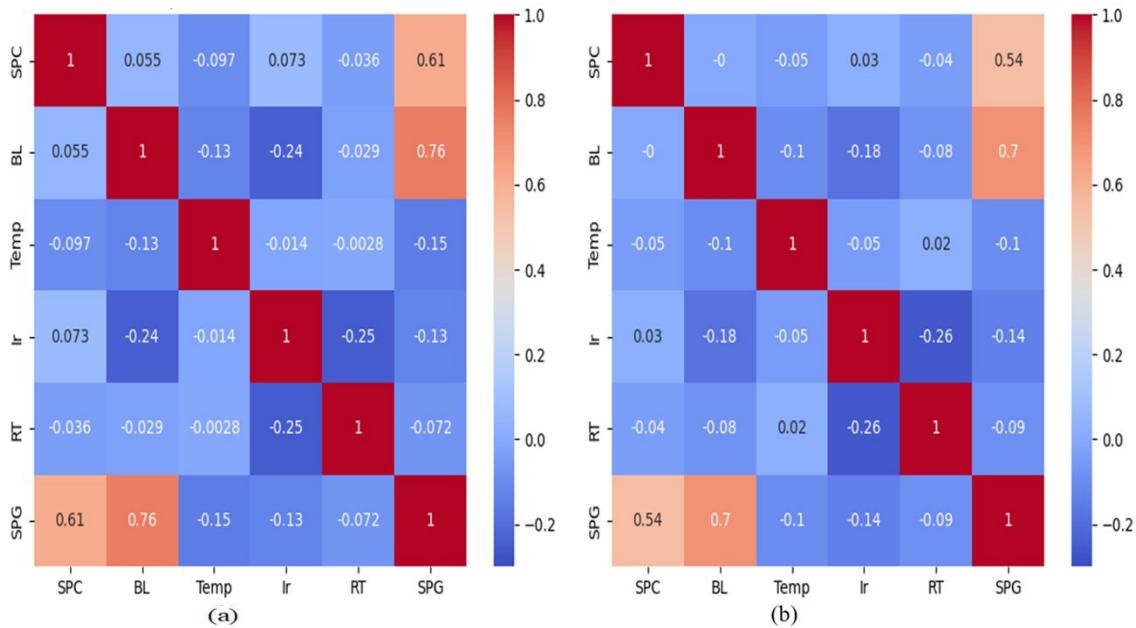


Figure 3: (a) Experimental data correlation (b) Simulated data correlation

Hypothesis test results

A sample t-test was conducted to validate the similarity between the datasets as represented in Table 3. The t-test produced a t-statistic of 0.0001 and a p-value of 0.332934, which is significantly greater than 0.05.

This indicates that there is no statistically significant difference between the mean SPG values of the experimental and simulated datasets. Both the KS test and t-test results consistently support the acceptance of the null hypothesis, confirming that the simulated data is statistically indistinguishable from the experimental data in terms of distribution and mean values.

Table 3: Hypothesis results

T-test	Value
t-statistics	0.0001
p-value:	0.332934

Comparative analysis of experimental and simulated results

Table 4 presents the performance evaluation of eight machine learning models developed using both experimental and simulated datasets. The results indicate that ensemble learning algorithms generally outperform other models in predicting system response. Among all models, the Extra Trees algorithm achieved the highest prediction accuracy, with an R^2 value of 0.993 for both datasets and the lowest error metrics. The experimental MSE, RMSE, MAE, MAPE and R^2 were 1.041, 1.020, 0.476, 0.324% and 0.993, respectively. However, the MSE, RMSE, MAE, MAPE and R^2 of the simulation were calculated as 1.627, 1.276, 0.297, 0.198% and 0.993, respectively. Thus, it can be inferred that the developed model performed extremely well in replicating the nonlinear relation between output and the input parameters in the developed system. The enhanced prediction accuracy of Extra Trees was explained by its ensemble structure with a large degree of randomness introduced into feature selection and split determination while constructing trees, leading to low model variance, no overfitting and better model generalisation performance for the prediction of unseen data. This similar behaviour was also reported in previous studies, with Extra Trees outperforming standard machine learning approaches for regression due to the robustness of the model and its capability to represent engineering systems (Galelli and Castelletti, 2013; Campagner *et al.*, 2023). In addition, XGBoost and Random Forest achieved good prediction accuracy for both cases ($R^2 > 0.98$) with a small number of errors, which implies good reliability of the model. The GB model performed well, especially for simulated data ($R^2 = 0.989$).

In contrast, ANN showed a lower value of prediction accuracy with R^2 values of 0.652 and 0.689 and a higher error of MSE with 55.214 and 67.721 for the experiment and simulation datasets, respectively. The relatively inferior ANN prediction can also be attributed to the limited data used and its sensitivity to hyperparameters. Neural networks require a sufficient amount of data to efficiently learn the relationships between independent and dependent variables in order to make adequate predictions without underfitting or overfitting.

Table 4: Evaluation Metrics on Experimental and Simulated Datasets

Model	Experimental					Simulated				
	MSE	RMSE	MAE	MAPE	R^2	MSE	RMSE	MAE	MAPE	R^2
ANN	55.214	7.431	5.104	3.103	0.652	67.721	8.229	5.952	3.875	0.689
RF	2.185	1.478	1.017	0.685	0.984	3.308	1.819	0.700	0.455	0.985
Extra Trees	1.041	1.020	0.476	0.324	0.993	1.627	1.276	0.297	0.198	0.993
Gradient Boost	5.596	2.366	1.790	1.207	0.960	2.321	1.524	0.793	0.523	0.989
AdaBoost	11.777	3.432	2.917	1.966	0.915	10.371	3.220	2.498	1.667	0.952
Decision Tree	3.727	1.930	0.725	0.496	0.973	6.736	2.595	0.723	0.469	0.969
kNN	12.412	3.523	2.792	1.896	0.911	6.078	2.465	1.646	1.077	0.972
XGBoost	1.356	1.164	0.605	0.413	0.990	2.889	1.700	0.455	0.300	0.987

Performance assessment of the extra trees model

Figure 4 shows the validation plot comparing the predicted SPG values computed using the Extra Trees machine learning technique with the measured SPG values taken from the AMPGS. The actual SPG values are plotted on the horizontal axis, while the predicted SPG values are plotted on the vertical axis. The dotted diagonal line is representative of perfect prediction, where the predicted value is equal to the actual value. It is apparent from the plot that most of the data points are clustered closely to the ideal prediction line, implying high agreement between the model and the system.

The R^2 of 0.993 reflects that about 99.3% of the variability in SPG can be attributed to the Extra Trees model. This extremely high value suggests the very high accuracy, reliability and robustness of the proposed model for predicting the generation power in AMPGS. The high correlation of validation points to the reference line suggests that the prediction errors are very low and the generalisation of the model is very good. The proposed model perfectly captures the complex nonlinear interactions of critical parameters such as solar power available, battery SOC, generator running condition, load required and availability of the grid. Therefore, the proposed prediction framework can support smart energy management, automatic source switching and real-time supervision in AMPGS.

The generated outcome is in line with the result provided by Geurts *et al.* (2006) in their study, in which they indicated that the use of the Extra Trees algorithm enhanced the prediction of accurate outcomes by heavily randomised decision trees to lower variance, avoid over-fitting and maintain overall good generalisation. Similarly, more recently, the reviews of renewable energy forecasting have listed ensemble learning methods as useful instruments in accurate solar power and hybrid power generation prediction (Rahimi *et al.*, 2023). Moreover, from an engineering standpoint, this increased accuracy in the prediction of SPG contributes to the reason behind a decision to equip the built AMPGS with the data logging equipment and Machine Learning technologies, as the SPG predictions allow for early manipulation of energy generated to optimise charging and discharging cycles for the battery and thus reduce the operating time of the generator, contributing to a more reliable system. Thus, the result supported the use of Extra Trees as a suitable predictive methodology for the smart operation of an AMPGS.

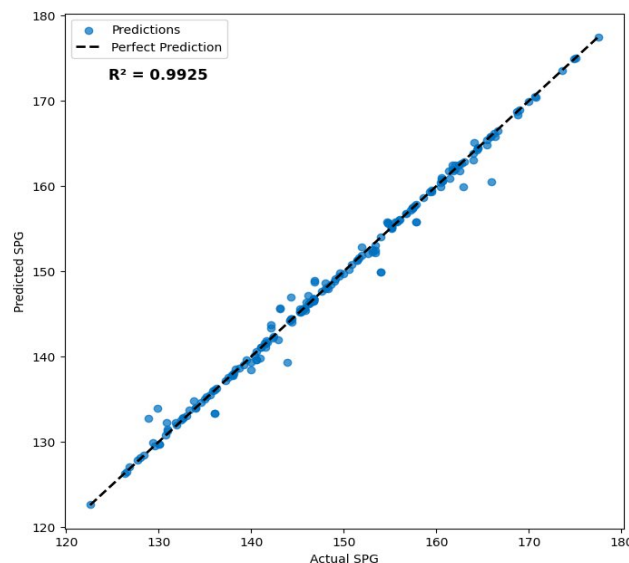


Figure 4: Predicted versus actual SPG responses for the Extra Trees model

Analysis of error plot and error distribution results

Figure 5(a) shows the error plot, which displays the differences between predicted values and actual values and highlights the prediction performance of the model. The observed residuals were randomly dispersed and clustered around zero, a strong indicator that predictions generated by this model were almost always

very close to their actual observed values and therefore had extremely strong predictive power for nearly every observation. If a model is well fit according to statistical learning principles, typical observed residuals are randomly dispersed around zero (James *et al.*, 2013).

In spite of this pattern, some points exhibited relatively large residuals (ranging approximately between -4 and +5). Such points can indicate outlier or influential cases. These may occur as a result of measurement noise, errors in the recorded data or weaknesses inherent in the selected statistical model. According to prior studies, large residuals such as these are sometimes indicative that the model is unable to reveal some intricate relationship (Feng *et al.*, 2023). The absence of any structured patterns within the residual graph is a justification of the assumption of independence and randomness, and therefore, of the model performance. Figure 5(b) displays the histogram of prediction error frequencies. From the histogram, it was observed that the majority of prediction errors are concentrated at or near zero residuals, implying that for most observations the model delivered good predictions. However, the histogram also reveals that the residuals are relatively widely dispersed, with a few errors spanning far into the tail region. These observations are corroborated by the residual plot in Figure 5(a), where sporadic but noticeable positive and negative residuals are readily visible (Gerber and Craig, 2024). With these residuals, it was concluded that the Extra Trees model demonstrates significant predictive power but produces varying errors across observations.

The distribution of the errors shows a broader distribution of residuals, with some observations demonstrating relatively large errors of prediction. As shown in the residual plot, the positive and negative values of residuals support the fact that the model both overpredicted and underpredicted the values of power across various observations. However, given that there is no specific form of systematic pattern in the residual plot, the Extra Trees model has demonstrated the ability to accurately model the dependent variable. Therefore, the results of the residual plots appear satisfactory for the prediction model, which lines up with the R^2 value of 0.993 to indicate excellent prediction.

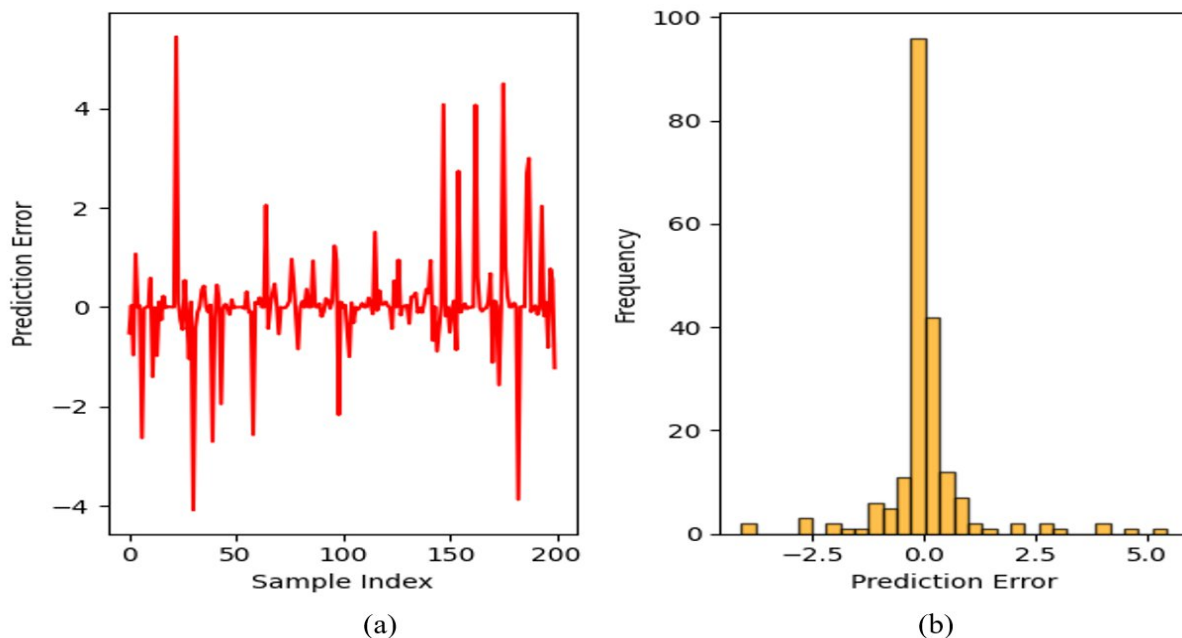


Figure 51: (a) Error plot, (b) Error distribution

Model Validation with Experimental Data

The model was therefore verified over three iterations to determine its prediction accuracy, validity and range of applications. The performance was quantified using various statistical measures such as MSE,

RMSE, MAE, MAPE and R^2 by comparing the predictions from the MATLAB/Simulink model with the available experimental data sets. The Hypothesis test conducted gave a p-value of 0.138, indicating no significant deviation between experimental and simulation datasets, while the KS test ensured the distribution sameness. Statistical validation between the two datasets yielded an R^2 of 0.993, indicating higher accuracy compared to the other models studied. However, the ANN model gave R^2 of 0.652 and 0.689 for the experimental and simulated datasets, respectively, which can be regarded as relatively poor predictive accuracy compared with the other two models studied. As stated by Ogbonna *et al.* (2023), the ANN model needs a larger dataset for its training to attain the desired level of performance. Therefore, the relatively low accuracy observed may be attributed to the limited size of the training data.

These results validate the robustness of the AMPGS model in replicating real-world performance. Also, to strengthen confidence in the model, results were compared with findings from published studies. Garces-Palacios *et al.* (2024) reported that hybrid solar–diesel–battery systems achieved daily solar contributions of 55–65%, which aligns with the AMPGS simulation results. Similarly, Aghenta and Iqbal (2019) highlighted that automated switching times in hybrid systems typically range between 2 and 5 seconds, consistent with the switching response observed in this study. However, validation to the experimental and simulated results in this study confirms that the developed MATLAB/Simulink model is robust, reliable, and consistent with global research trends in multi-source renewable energy systems. The validation approach ensures that the system can be confidently deployed for real-time energy management in practical applications, fulfilling the study’s objectives of accuracy, reliability, and sustainability.

Conclusion

This study demonstrated a framework for statistical verification and a ML-based solar generation forecasting for the AMPGS. The MATLAB/Simulink model simulation data were validated by comparing them with the experimental data of the built AMPGS. Results for the validation of these datasets in terms of histogram matching, the KS test, correlation analysis and t-test were also described. It was observed that the KS test value was 0.045 and the p-value was 0.1383, whereas for the t-test, the p-value was 0.3329, and no statistical differences were found between the two data sets, which also confirmed the simulation results. ML modelling showed the superiority of a data-driven methodology for accurate forecasting in solar power generation. Eight ML algorithms have been built and evaluated against appropriate validation criteria, namely MSE, RMSE, MAE, MAPE, and R^2 values. The Extra Trees algorithm showed the superior predictive ability among all the evaluated algorithms by reaching a maximum R^2 value of 0.993 on both experimental and simulated datasets, while yielding minimum error metrics. This obtained good generalisation ability and was proven by the residual plot and error distribution. Furthermore, correlation analysis identified solar irradiance as the most influential factor affecting solar power generation, followed by battery level and solar power current. Overall, the results confirm that the proposed validation and prediction framework is robust, accurate, and suitable for intelligent energy management applications. The developed approach can support real-time monitoring, predictive control, and optimal source switching in hybrid renewable energy systems, thereby enhancing system reliability, efficiency, and sustainability.

References

- Adeyinka, A. M., Esan, O. C., Ijaola, A. O., and Farayibi, P. K. (2024). Advancements in Hybrid Energy Storage Systems for Enhancing Renewable Energy-to-Grid Integration. *Sustainable Energy Research*, 11(1). <https://doi.org/10.1186/s40807-024-00120-4>
- Aghenta, L. O. and Iqbal, M. T. (2019). Design and Dynamic Modelling of a Hybrid Power System for a House in Nigeria. *International Journal of Photoenergy Research Article*, 13 (1):1-13. doi.org/10.1155/2019/6501785
- Akinrinade, N. A., Onawumi, A. S., Ajayeoba, A. O., Sangotayo, E. O., Olanrewaju, O. A., Bello, K. A., Maladzhi, R. W. and Daniyan, I. A. (2026b). Application of Machine Learning Models for the

- Prediction of Solar Panel Wattage for Rotating and Static Panel Systems. *Advances in Artificial Intelligence and Machine Learning*, 6 (3):1-24. <https://dx.doi.org/10.54364/AAIML.2026.63311>
- Akinrinade, N. A., Onawumi, A. S., Sangotayo, E. O., Omoniyi, E. B., Ajayi, S. T. and Ihekweba, G. C. (2026a). Review on Automated Multi-Source Power Generating Systems: A Mechanical Engineering Perspective. *Nigerian Journal of Applied Science and Innovative Technology (NiJASIT)*, 12 (1):31-49. <https://nijasit.bellsuniversity.edu.ng>
- Al-Dahidi, S., Madhiarasan, M., Al-Ghussain, L., Abubaker, A. M., Ahmad, A. D., Alrbai, M., Aghaei, M., Alahmer, H., Alahmer, A., Baraldi, P., and Zio, E. (2024). Forecasting Solar Photovoltaic Power Production: A Comprehensive Review and Innovative Data-Driven Modelling Framework. *Energies*, 17 (16):1-38. <https://doi.org/10.3390/en17164145>
- Ambelu, M., Woldeyohannes, A. D., and Gebeyehu, S. G. (2025). Simulation Models Verification and Validation: Recent Development and Challenges - A Review. *International Journal of Modelling, Simulation, and Scientific Computing*, 16 (1). <https://doi.org/10.1142/S1793962325300018>.
- Bamisile, O, Acen, C., Cai, D., Huang, Q. and Staffell, I. (2025). The Environmental Factors Affecting Solar Photovoltaic Output. *Renewable and Sustainable Energy Reviews*, 208 (11):1-20. <https://doi.org/10.1016/j.rser.2024.115073>
- Banik, R. and Biswas, A. (2024). Enhanced Renewable Power and Load Forecasting using RF-XGBoost Stacked Ensemble. *Electrical Engineering*. 106, 4947–4967. <https://doi.org/10.1007/s00202-024-02273-3>
- Brahim, R., Ahlam, L., Hamza, A. and Güler, I. (2025). Power Forecasting Using ANN and ELM: A Comparative Evaluation of Machine Learning Approaches. *International Information and Engineering Technology Association*, 12 (1):1-8. doi.org/10.18280/mmep.120101
- Campagner, A., Ciucci, D. and Cabitza, F. (2023). Aggregation Models in Ensemble Learning: A Large-Scale Comparison. *Science Direct*, 90, 241-252. [10.1016/j.inffus.2022.09.015](https://doi.org/10.1016/j.inffus.2022.09.015)
- Feng, C., Li, L. and Sadeghpour, A. (2023). A Comparison of Residual Diagnosis Tools for Diagnosing Regression Models for Count Data. *BMC Medical Research Methodology*, 20 (175):1-22. <https://doi.org/10.1186/s12874-020-01055-2>
- Galelli, S. and Castelletti, A. (2013). Assessing The Predictive Capability of Randomised Tree-Based Ensembles in Streamflow Modelling. *Hydrology and Earth System Sciences*, 17 (7):2669–2684. <https://doi.org/10.5194/hess-17-2669-2013>
- Garces-Palacios, L. E., Rodríguez-Gallegos, C. D., Vaca-Urbano, F., Alvarez-Alvarado, M. S., Gandhi, O. and Rodríguez-Gallegos, C. A. (2024). Optimal Design of a Hybrid Solar–Battery–Diesel System: A Case Study of the Galapagos Islands. *Solar*, 4 (2):232–245. <https://doi.org/10.3390/solar4020011>
- Gerber, E. A. E. and Craig, B. A. (2024). Residuals and Diagnostics for Multinomial Regression Models. *ASA Data Science Journal*, 17, 1-14. doi: 10.1002/sam.11645.
- Geurts, P., Ernst, D., & Wehenkel, L. (2006). *Extremely Randomised Trees*. *Machine Learning*, 63(1):3–42. <https://doi.org/10.1007/s10994-006-6226-1>.
- Golestan, S., Golmohamadi, H., Sinha, R., Iov, F., and Bak-Jensen, B. (2024). Real-time simulation and hardware-in-the-loop testing based on OPAL-RT ePHASORSIM: A review of recent advances and a simple validation in EV charging management systems. *Energies*, 17(19), 4893. <https://doi.org/10.3390/en17194893>
- Ijeoma, M. W., Lewis, C. G., Chen, H., Chukwu, B. N. and Carbajales-Dale, M. (2024). Technical, Economic, and Environmental Feasibility Assessment of Solar-Battery-Generator Hybrid Energy Systems: A Case Study in Nigeria. *Frontier Energy Research*, 12:1397037. doi: 10.3389/fenrg.2024.1397037
- James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer New York Heidelberg Dordrecht London. Doi: 10.1007/978-1-4614-7138-7.
- Memić, A., Dedović, M., Dautbašić, N., Kapo, M. (2024). Application of Neural Networks for Predicting Energy Production from Hybrid Power Systems Considering the Influence of Stochastic Weather Changes. *B&H Electrical Engineering*, 18 (2):11-17. doi: 10.2478/bhee-2024-0006

- Mundu, M. M., Nnamchi, S. N., Sempewo, J. I. and Uti, D. E. (2024). Simulation Modeling for Energy Systems Analysis: A Critical Review. *Energy Informatics*, 7 (75):1-20. <https://doi.org/10.1186/s42162-024-00374-8>
- Ndiaye, E. H. M., Ndiaye, A., Faye, M., Gueye, D., Ba, A., and Traore, M. (2025). Analysis of the Impact of Irradiance and Temperature on Photovoltaic Production: A Statistical and Machine Learning Approach Author Links Open Overlay Panel. *MethodsX*, 15 (11):1-16. <https://doi.org/10.1016/j.mex.2025.103716>
- Ogbonna, C. C., Eze, V. H. U., Ikechuwu, E. S., Okafor, W. O., Anichebe, O. C. and Oparaku, O. U. (2023). A Comprehensive Review of Artificial Neural Network Techniques Used for Smart Meter-Embedded Forecasting Systems. *Journal of Applied Sciences*, 8 (1):13-24. doi: [10.59298-IDOSR/2023/12.1.7895](https://doi.org/10.59298/IDOSR/2023/12.1.7895)
- Onawumi, A. S., Akinrinade, N. A., Olojede, M. A. and Ajayeoba, A. O. (2022). Development of a Solar Power Generating System with Auto-Tracking and Data Logging Devices. *Journal of Engineering Research and Reports*, 23 (12):212–222. doi:10.9734/JERR/2022/v23i12779. <http://journaljerr.com>
- Petersen, L., Iov, F., Tarnowski, G. C., Gevorgian, V., Koralewicz, P. and Stroe, D. I. (2019). Validating Performance Models for Hybrid Power Plant Control Assessment. *Energies*, 12 (22):4289. <https://doi.org/10.3390/en12224330>
- Rahimi, N., Park, S., Choi, W., Oh, B., Kim, S., Cho, Y., Ahn, S., Chong, C., Kim, D., Jin, C. and Lee, D. (2023). A Comprehensive Review on Ensemble Solar Power Forecasting Algorithms. *Journal of Electrical Engineering & Technology*, 18, 719–733. <https://doi.org/10.1007/s42835-023-01378-2>
- Rajendran, G., Raute, R. and Caruana, C. (2025). A Comprehensive Review of Solar PV Integration with Smart-Grids: Challenges, Standards, and Grid Codes. *Energies*, 18(9), 2221. <https://doi.org/10.3390/en18092221>
- Ramasubramanian, D., Pourbeik, P., Farantatos, E. and Gaikwad, A. (2021). Simulation of 100% Inverter-Based Resource Grids with Positive Sequence Modeling. *IEEE Electrification Magazine*, 9 (2):62–71. <https://doi.org/10.1109/MELE.2021.3070938> (doi.org in Bing)
- Resch, J., Schuiki, B., Schöndorfer, S., Brandauer, C., Panholzer, G., Prästl-Andrén, F. and Strasser, T. I. (2020). Engineering and Validation Support Framework for Power System Automation and Control Applications. *Elektrotechnik & Informationstechnik*, 137(8):470–475. <https://doi.org/10.1007/s00502-020-00852-3>
- Salah, O., Shamayleh, A. and Mukhopadhyay, S. (2021). Energy Management of a Multi-Source Power System. *Algorithms*, 14 (7):206. <https://doi.org/10.3390/a14070206>
- Schleifer, A. H., Harrison-Atlas, D., Cole, W. J. and Murphy, C. A. (2023). Hybrid Renewable Energy Systems: The Value of Storage as a Function of PV-Wind Variability. *Frontier Energy Research*, 11:1-20. doi: [10.3389/fenrg.2023.1036183](https://doi.org/10.3389/fenrg.2023.1036183)
- Tanguet, E. and Njomen, D. (2021). Kolmogorov-Smirnov APF Test for Inhomogeneous Poisson Processes with Shift Parameter. *Applied Mathematics*, 12 (4):322-335. doi: [10.4236/am.2021.124023](https://doi.org/10.4236/am.2021.124023)
- Yu, J., Li, X., Yang, L., Li, L., Huang, Z., Shen, K., Yang, X., Yang, X., Xu, Z., Zhang, D., and Du, S. (2024). Deep Learning Models for PV Power Forecasting: Review. *Energies*, 17 (16):1-35. <https://doi.org/10.3390/en17163973>
- Zhu, S., Piper, D., Ramasubramanian, D., Quint, R., Isaacs, A. and Bauer, R. (2018). Modeling Inverter-Based Resources in Stability Studies. *IEEE Power & Energy Society General Meeting*. Pp. 1-5. <https://doi.org/10.1109/PESGM.2018.8586157>.