All Sciences Proceedings
http://as-proceeding.com/

1st International Conference on Recent Academic Studies

May 2-4, 2023 : Konya, Turkey

https://as-proceeding.com/index.php/icras

An efficient energy prediction model for solar energy power system using Artificial Intelligence technique

Muhammad Imran Munawar *, Imtiaz Alam [2] and Muhammad Sharjeel Ali [3]

[1]Departmen of Electrical Engineering, Bahria University, Pakistan
[2]Departmen of Electrical Engineering, Bahria University, Pakistan
[3]Departmen of Electrical Engineering, Bahria University, Pakistan

*(01-244201-015@student.bahria.edu.pk) Email of the corresponding author

Abstract – Prediction of Solar power generation plays an important role to improve the efficiency of economic dispatch function and reduce the dependence on fossil fuels and help in the energy management system. For time series solar energy prediction multiple models were introduced but these model trains are based on yearly historical data. A big data collection containing many missing values makes these model training more complicated that's why In this paper, an efficient energy prediction model is proposed for the prediction of time series solar energy based on short predicted weather training data. Two complimentary models are based on linear regression and a knowledge based neural network is exploited to predict future solar power, with offline training. The LR is structured under the direction of the proposed input method parameter selection and used when training data is enough. KBNN is used for existing advantages predictive models are also very important when training data is not enough. According to test findings using real data sets. An LR model can deal effectively with linear data, but a KBNN model can cope effectively with nonlinear behavior. Additionally, the results demonstrate the effectiveness of LR showing a correlation coefficient (R2) is 98% with a root mean square error of 45 and KBNN shows a correlation coefficient (R2) is 99% with a root mean square error of 44 in providing a reliable version, The results additionally show the functionality of LR and KBNN in imparting a dependable version, especially when the short training dataset is available.

Keywords – Knowledge-Based Neural Network (KBNN), Linear Regression (LR), Solar Energy Prediction, Energy Computing

I. Introduction

In the current era commonly known as the energy era, electrical energy is the most essential requirement for living a comfortable and standard lifestyle for everyone, and with every passing day the demand for it is increasing enormously. Fossil fuels, a most important source of electrical power technology that is diminishing day by day, and their usage reason extreme environmental worries. Switching to renewable power assets like solar, water, wind, geothermal, and biomass electricity assets will assist to keep our environment clean. Solar electricity is the rising and maximum promising renewable energy resource of the energy era across the globe, the crucial future manufacturing approach in the move to easy energy [1][2]. Solar strength has a mixed features of normality and abnormality. On the one side, as the solar rises and solar sets every day the solar radiation takes place as pulses on a day-by-day foundation. On the other side, because of the climate and time change, the acquired sun power on this planet within a day and throughout days can range considerably vary[3].

Numerous factors influence sun electricity technology, along with solar radiation, cloud insurance, temperature, humidity, atmospheric pressure, wind velocity, and so on. These climatic elements may alter drastically at any time because of the unpredictable weather system on Earth, making it difficult for dependable and accurate solar power predictions [4]. The current solar power forecast frameworks are essentially built on the (AI) algorithm. AI is a recent technology that simulates human reasoning on computers to learn more about it, as well as other computer science topics like group dynamics [5]. The most popular AI techniques are heuristic optimization, fuzzy logic, expert systems, and Machine Learning (ML) [6]. AI can be used to achieve high levels of intelligence for solar forecasters, which typically include powerful functions for feature extraction and nonlinear mapping, good compatibility with a variety of PV power prediction situations, and some capacity for logical reasoning. As a result, the estimation of solar radiation and PV power using AI algorithms has proved effective [7]. Barbieri et. al. examined cloud-based models for predicting solar energy and discovered that the temperature of the battery and solar radiation were the key factors [8]. A summary of current deep learning-based techniques for predicting wind and solar power was published by Wang et al [9]. Some scientists use factual techniques to predict Solar Power generation, such as Autoregressive Moving Normal (ARMA) [10], Autoregressive Incorporated Moving Normal (ARIMA) [11], and Autoregressive Moving Normal Model with Exogenous Data Sources (ARMAX) [12]. However, these models are insufficient to increase the nonlinear time series data of SPG's forecasting precision. To support the prediction accuracy of nonlinear time series SPG, experts are driven by artificial intelligence and a combination of factual computerized reasoning processes. Scientists like the brain network (NN) as a taught system for several applications. Analysts favored NN for image processing because of its advantages, including its capacity for self-learning, ability to handle challenges with information loss, versatility [13].However, despite the proliferation of similar studies, the analysis of solar energy prediction from the perspective of AI has not yet been done in recent years. Scientists and engineers can use the review to examine the characteristics of various solar prediction models and identify whether AI can improve their prediction tools to fully use AI's potential for solar energy prediction [14]

II. MATERIALS AND METHOD

A. *Existing Methodology*

In my examination of the literature, I looked at several solar energy forecasting models. All of these models base their predictions on those power plants' data that have at least a year's worth of historical data. There has been very little research on solar plant data that are new or that have minimal generation and sensor weather data. So it's necessary to build a forecasting model that has fewer complications and forecasts results with maximum accuracy on fewer training data.

B. *Proposed methodology*

To forecasting time series solar energy based on upcoming meteorological conditions, two complementing models are proposed in this work. With offline training, a model based on linear regression (LR) and a knowledge-based neural network (KBNN) is used to forecast solar power. Under the guidance of the specified input parameter selection approach, LR is constructed after there is sufficient training data. KBNN is utilized to make use of the existing prediction models when there is a lack of training data. A KBNN model can be a useful technique to increase the predicted accuracy produced by any models used for brief training data, according to test findings using real data sets. An LR model can deal effectively with linear data, but a KBNN model can cope effectively with nonlinear behavior. Figure **1** shows the block diagram of the current two complementary models we have utilized Linear Regression and Knowledge-based neural network for the best solar energy prediction for solar energy power systems

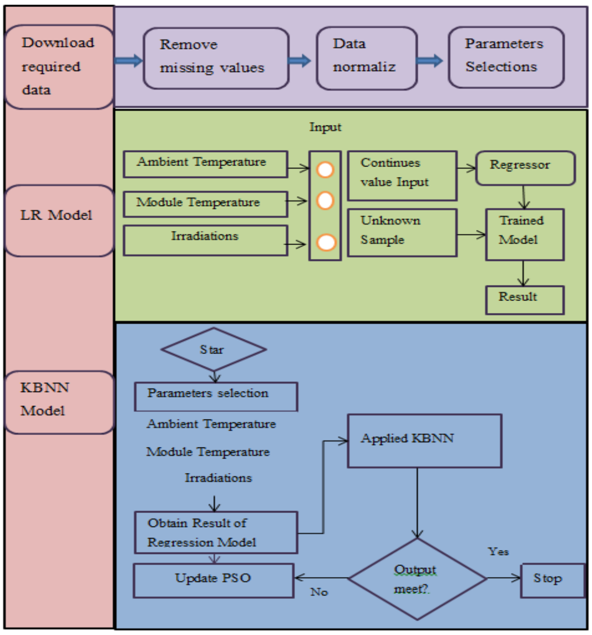

Figure 1: Block diagram of two complementary purposed models

C. *Experiment on dataset*

The data set used in this work is available on https://www.kaggle.com/ and offers two different forms of information. Information gathered from a solar power station about sensor panels and electricity production. The inverters, each of which is connected to a bank of solar panels, are used to collect data on generation. The sensor data is gathered at the plant using a single array of well-placed sensors. Every 15 minutes, the solar power plant's sensor data and power generation were collected. Plant Generation Data and Plant Weather Sensor Data are two CSV files that make up the solar energy projection dataset. The statistics were collected in India between May 15 and June 18, 2020. The dataset was opened in Excel as a CSV file. 22 inverters' power generation and sensor data collected for 34 days are combined based on time and date to ascertain the relationship between the sensor readings and the DC power produced by the solar panel array at a particular moment.

D. *Selection of feature*

The inputs for the planned task were chosen based on the degree of correlation with the AC power generated by the inverters. There is a linear relationship between features such as ambient temperature, module temperature, irradiation, and AC power, according to the scattering matrix of the features in Figure **2**. Time is also regarded as one of the input features for training the model because DC power generation peaks in the afternoon and is non-existent in the late evening and early morning. The inverter's effectiveness in converting the received DC power to AC power will determine how much AC power is generated, hence it was disregarded.

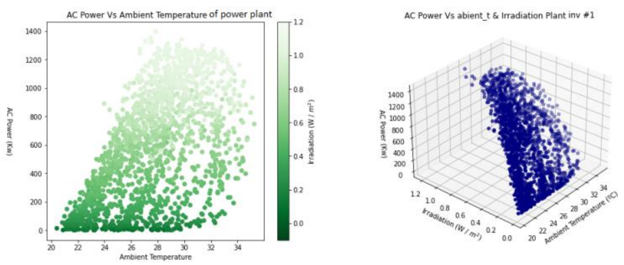

Figure 2: Power Output with Temperature and Irradiation

E. *Physical Model*

The most common prognostic and classification model, particularly in forecasting, uses regression functions. Regardless of where the fresh input vector is placed in the issue space, they are used to generate the output value since they come from data obtained through inductive learning from the full problem domain. This can lead to multiple formulae for regression with the same issue for many problems by using diverse datasets. As a result, these methods only partially account for new data that differ significantly from the ones used for the initial modeling. In Equation 1 open circuit voltage is multiplied by V_{th} (Thèvenin voltage) and the short-circuit current I_{no} can be used to model the solar power output of Pac (taking the inverter into account) (Mayer-Norton current).

$$P_{ac} = V_{th} \cdot I_{no} \tag{1}$$

$$V_{th} = V_0[1 + \beta(T_m - T_0)] \tag{2}$$

$$I_{no} = I_0[1 + \alpha(T_m - T_0)]\left(\frac{G_{ir}}{G_0}\right) \tag{3}$$

The module temperature is directly proportional to the Irradiation and the ambient temperature and can be represented by the following empirical Formula in Equation 4

$$T_m = 30 - 0.0175(G_{ir} - 300) + 1.14(T_a - 25) \tag{4}$$

We obtain a third-degree polynomial by substituting Equation 2, Equation 3, and Equation 4 into Equation 1. This polynomial has the form:

$$P_{ac} = K_1 G_{ir}^3 + K_2 G_{ir}^2 + K_3 G_{ir}^2 T_a + K_4 G_{ir} T_a^2 + K_5 G_{ir} T_a + K_6 G_{ir} \tag{5}$$

The constants are K1, K...., and K6. We now know which new properties we need to include in designing a Regressor that predicts AC Power by using equation 5.

F. *Knowledge-Based Neural Networks*

Connectionist learning techniques are the foundation of the hybrid learning system known as KBANN (Knowledge-Based Artificial Neural Network). We represent "domain theories" for particular problems as propositional logic expressions on neural networks, and then utilize backpropagation to improve this knowledge reconstruction. Radial-based function (RBF) kernels, which are affixed to the nodes of kernel-based ANNs, are changed in terms of their centers and radii through learning from data. They are developed as a collection of interconnected local models that are trained. In Song et al. 2006, they present a technique for integrating kernel functions and regression formulae into a knowledge-based neural network (KBNN) model, which improves local knowledge precision and accuracy. Figure **3** provides a block schematic structure

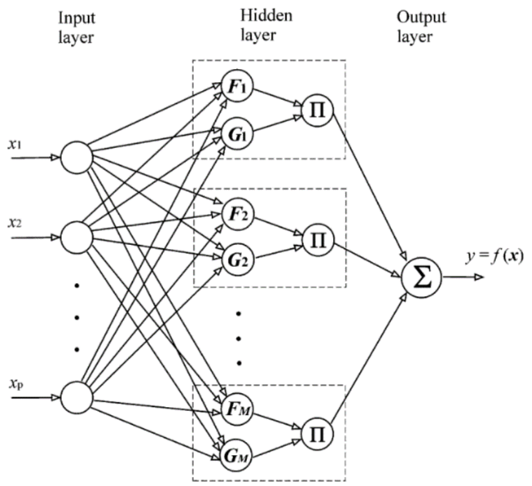

Figure 3: Regression function insert into KBNN for prediction

$$y(x) = G_1(x)F_1(x) + G_2(x)F_2(x) + \cdots + G_M(x)F_M(x) \quad 6$$

A. *Assessment Metrics*

The proposed models' capacity to project the generation of PV energy in the system was assessed using performance measurement techniques including the Root Mean Square Error (RMSE), (R^2 or the Correlation Coefficient).

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \quad 7$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad 8$$

III. Results

A. *Predicted Results of LR and KBNN*

The two complementary models LR and KBNN are executed for time series solar energy prediction for 7 days, this result shows in **Figure 4a**. Comparing the outcomes of the forecast model's 7-day PV power predictions with the actual PV power outcomes for various 7-day periods when the input sequence is 15 minutes. It is discovered that the two models' forecast outcomes generally follow the same pattern as the actual findings. Additionally, the ambient temperature, module temperature, and irradiation figures for the chosen 7 days are displayed and can observe from June 12th, 2020 to June 18th, 2020 in **Figure 4b** and **Figure 4c** respectively. The Result shows that temperature and irradiance, particularly irradiance, which has a trend that is quite comparable to the trend of PV power generation, have an important impact on PV power generation. KBNN takes advantage of LR predictions and enhances the prediction accuracy because it deals with nonlinear behavior in weather conditions if happened.

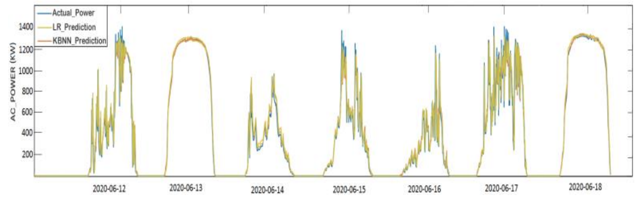

Figure 4a: Prediction and actual AC power of LR and KBNN Model

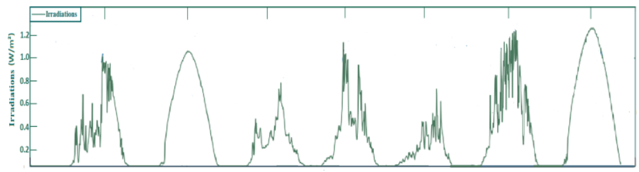

Figure 4b: Irradiation for sensor 1 from June 12th, 2020 to June 18th, 2020

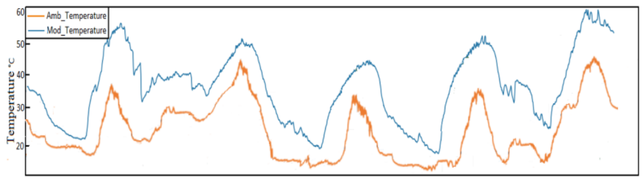

Figure 4c: Ambient temperature and Module Temperature for sensor 1 from June 12th, 2020 to June 18th, 2020

B. *Performance of LR model*

For the model's development and testing the Split Validation Technique is used with ratios of 80% and 20%. For the 34 days, there are virtually the same amounts of readings taken each day. In Table, 1date selection is taken randomly, and 34 days of data were divided into training and testing periods, out of which 27 days of data are used for training and 7 days of data are used for testing. The model predicts time series forecasting using linear regression and measures the model's effectiveness using the root mean squared error and coefficient of determination

Table 1: Performance of LR Energy Prediction Model

	Training	**Testing**
No of days data	27	7
No of samples	47,116	13,072
RMSE(KW)	46	45
Correlation coefficient	99%	98%

The prediction model performance demonstrates how well the training was done. Neither over-fitted nor under-fitted is comparable to the testing performance as depicted in Table 1. RMSE 45KW and Correlation Coefficient R2 98%, of the testing model, and 46KW, 99% of the training model, respectively. The residual factor between the actual value and the projected value for training and testing as shown respectively

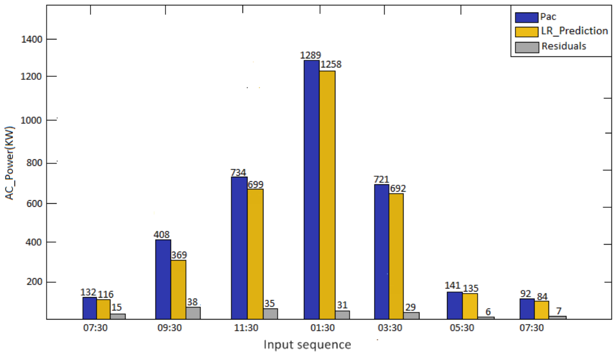

Figure 5: LR Prediction and actual AC power on June 12th, 2020 of inverter 1

Model prediction for inverter 1 of a single day (June 12th, 2020) with an input sequence of 2 hours intervals represents in Figure 5. Residual represents the difference between actual ac power and predicted Ac power. From 07:30 am to 07:30 pm residual fluctuates due to non-linear behaviour in weather conditions. This non-linear behavior affects the LR predictions across the day. Overall, the LR model's forecast is likely to be accurate.

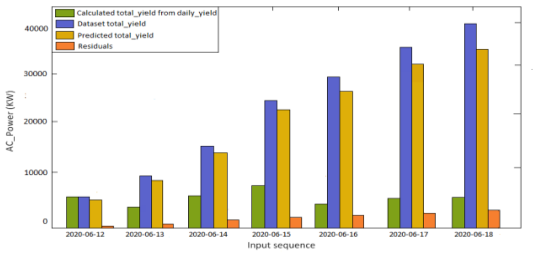

Figure 6: Prediction and actual total yield from June 12th, 2020 to June 18th, 2020 of inverter 1

The prediction and actual total yield from June 12th, 2020 to June 18th, 2020 inverter 1 is represented in Figure 6. Results elaborate the efficiency of the predicted model for calculation of daily yield and total yield for the upcoming 7 days and express the residuals between actual data set ac_power and predicted ac_power. The total yield is calculated from the daily yield of inverter 1. Residuals slightly increase from June 12th to June 18th due to some nonlinear behavior in weather conditions or equipment that is malfunctioning or not working at its best.

C. *Performance of KBNN model*

KBNN model takes the advantages and trained itself on existing predictions. The following 7 days of data were used in the training and testing periods in Table 2. The technique for integrating regression formulae into a knowledge-based neural network model is adopted. The model predicts time series forecasting using KBNN and measures the model's effectiveness using the root mean squared error and coefficient of determination. The prediction model's performance demonstrates how well the training was done. RMSE 44KW and R2 99% measured in the testing model and RMS 45KW and R2 98% in the training model, respectively. RMSE and R2 show KBNN slightly enhances the prediction accuracy of the existing LR model on fewer training data

Table 2: Performance table of KBNN Model

	Training	Testing
No of days data	7	7
No of samples	13,072	13,072
RMSE(KW)	45	44
Correlation coefficient	98%	99%

KBNN algorithm Prediction and actual Ac_power on 2020_06-12 of the inverter 1 taken with an input sequence of 2 hours. If Figure 5 and Figure 7 compare together then we observe that KBNN reduces the residual and enhances the prediction accuracy. KBNN has the ability its increase accuracy day by day because it trained itself on its previous prediction

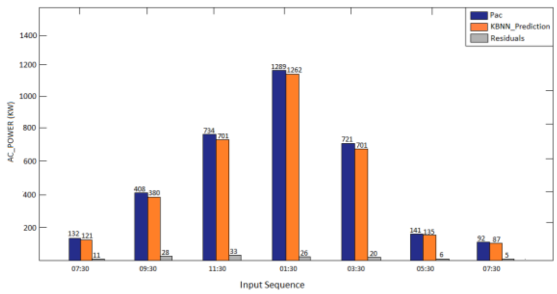

Figure 7: KBNN Prediction and actual ac_power on 2020_06-12 of inverter 1

IV. DISCUSSION

When these two complimentary models are used together the Prediction results are in good accordance with the observation. When that result compares in Table 3 with recent short-term power prediction techniques then we can see the efficiency of two complementary LR and KBNN models. There have been numerous research studies on solar power forecasting, however, certain major problems have yet to be fully solved. It's difficult to predict solar power. It is related to weather and environmental elements in addition to the PV cells' actual operating circumstances. One of the major scientific challenges to be overcome in the future is how to effectively combine the physical model of the battery, weather, and environmental elements. The difficulty of predicting solar energy is currently nearly universally expressed as a "black box" model. It is not entirely clear how the input variables and PV power output relate mathematically. Additionally, it is not clear which input parameter has the greatest influence on how accurately a prediction is made. In conclusion, one of the primary areas of future research will be on how to explain the solar power prediction model.

Table 3**:** Comparison table with recently existing short-term solar power prediction techniques

Article	**Year**	**Methods**	**RMS(KW)**	**R2**
[15]	2021	SVM GPR BP BSAELM IBSAELM	49.3 47.20 93.6 85.8 64.3	98.42 83.48 98.78 98.83 99.35
[16]	2021	KNN LR SVM ANN	92.857 94.583 93.644 86.466	97 96 95 98
My Work	2023	LR KBNN	45 44	98 99

V. CONCLUSION

The experiment demonstrates that compared to existing methods, the proposed LR and KBNN model increases the solar energy prediction accuracy with the appropriate input parameters. Even with a small amount of previous solar data, the KBNN can still provide a reliable prediction. Using ground irradiation and ambient temperature as anticipated weather forecast data received before the power generation, regression models can be developed to predict the AC Power, Daily yield, and Total Yield. Therefore, proving as a proof of concept that this kind of data may be utilized to anticipate precisely with at least 7 days of anticipation if precise enough ambient temperature and irradiation forecasts are provided.

ACKNOWLEDGMENT

We would like to extend our sincere gratitude and appreciation to everyone who helped this research project be completed successfully. Firstly, we would like to thank my Supervisor We also want to express our gratitude to the volunteers who made this study possible by voluntarily sharing their solar energy data. Their excellent data gave us the knowledge we needed to verify and test the suggested LR and KBNN models.

REFERENCES

[1] F. Alassery, A. Alzahrani, A. I. Khan, K. Irshad, and S. R. Kshirsagar, "An artificial intelligence-based solar radiation prophesy model for green energy utilization in energy management system," *Sustain. Energy Technol. Assessments*, vol. 52, p. 102060, Aug. 2022, doi: 10.1016/j.seta.2022.102060.

[2] Z. Liu *et al.*, "Artificial intelligence powered large-scale renewable integrations in multi-energy systems for carbon neutrality transition: Challenges and future perspectives," *Energy AI*, vol. 10, no. August, p. 100195, 2022, doi: 10.1016/j.egyai.2022.100195.

[3] E. Jumin, F. B. Basaruddin, Y. B. M. Yusoff, S. D. Latif, and A. N. Ahmed, "Solar radiation prediction using boosted decision tree regression model: A case study in Malaysia," *Environ. Sci. Pollut. Res.*, vol. 28, no. 21, pp. 26571–26583, 2021, doi: 10.1007/s11356-021-12435-6.

[4] J. Wang, H. Zhong, X. Lai, Q. Xia, Y. Wang, and C. Kang, "Exploring Key Weather Factors From Analytical Modeling Toward Improved Solar Power Forecasting," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1417–1427, Jul. 2019, doi: 10.1109/TSG.2017.2766022.

[5] A. Sharma and A. Kakkar, "Forecasting daily global solar irradiance generation using machine learning," *Renew. Sustain. Energy Rev.*, vol. 82, pp. 2254–2269, Aug. 2018, doi: 10.1016/j.rser.2017.08.066.

[6] E. Yiğit, U. Özkaya, Ş. Öztürk, D. Singh and H. Gritli, "Automatic detection of power quality disturbance using convolutional neural network structure with gated recurrent unit." Mobile Information Systems, 2021, 1-11.

[7] L. B. Yang, "Application of Artificial Intelligence in Electrical Automation Control," *Procedia Comput.*

Sci., vol. 166, pp. 292–295, Aug. 2020, doi: 10.1016/j.procs.2020.02.097.

[8] F. Barbieri, S. Rajakaruna, and A. Ghosh, "Very short-term photovoltaic power forecasting with cloud modeling: A review," *Renew. Sustain. Energy Rev.*, vol. 75, pp. 242–263, Aug. 2017, doi: 10.1016/j.rser.2016.10.068.

[9] H. Wang *et al.*, "Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network," *Energy Convers. Manag.*, vol. 153, no. October, pp. 409–422, 2017, doi: 10.1016/j.enconman.2017.10.008.

[10] W. Ji and K. C. Chee, "Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN," *Sol. Energy*, vol. 85, no. 5, pp. 808–817, Jul. 2011, doi: 10.1016/j.solener.2011.01.013.

[11] E. Pieri, A. Kyprianou, A. Phinikarides, G. Makrides, and G. E. Georghiou, "Forecasting degradation rates of different photovoltaic systems using robust principal component analysis and ARIMA," *IET Renew. Power Gener.*, vol. 11, no. 10, pp. 1245–1252, Jul. 2017, doi: 10.1049/iet-rpg.2017.0090.

[12] Y. Li, Y. Su, and L. Shu, "An ARMAX model for forecasting the power output of a grid connected photovoltaic system," *Renew. Energy*, vol. 66, pp. 78–89, Jul. 2014, doi: 10.1016/j.renene.2013.11.067.

[13] C.-C. Lu and Y. H. Shin, "A neural network based image compression system," *IEEE Trans. Consum. Electron.*, vol. 38, no. 1, pp. 25–29, Jul. 1992, doi: 10.1109/30.125078.

[14] D. W. van der Meer, J. Widén, and J. Munkhammar, "Review on probabilistic forecasting of photovoltaic power production and electricity consumption," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1484–1512, Aug. 2018, doi: 10.1016/j.rser.2017.05.212.

[15] D. Wu, J. Kan, H. C. Lin, and S. Li, "Hybrid Improved Bird Swarm Algorithm with Extreme Learning Machine for Short-Term Power Prediction in Photovoltaic Power Generation System," *Comput. Intell. Neurosci.*, vol. 2021, 2021, doi: 10.1155/2021/6638436.

[16] L. Gutiérrez, J. Patiño, and and E. Duque-Grisales, "A Comparison of the Performance of Supervised Learning Algorithms for Solar Power Prediction," no. 2021, doi: doi.org/10.3390/en14154424.