International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 14, No.1 (Oct-23)

http://dx.doi.org/10.12785/ijcds/1401105

Prediction of the Performance of a Sun Tracking Photovoltaic System using different Artificial Intelligence Techniques: Case Study in Zarqa, Jordan

Jaafar Alghazo[1], Ghazanfar Latif[2,3], Bashar Hammad[4], Mohammad Al-Abed[5], Mohammad Al-Dweri[5], Fadi N. Sibai[6] and Wael Al-Kouz[4]

[1]*Artificial Intelligence Research Initiative, College of Engineering and Mines, University of North Dakota, Grand Forks, ND, USA*
[2]*Department of Computer Science, Prince Mohammad bin Fahd University, Saudi Arabia*
[3]*Université du Québec à Chicoutimi, 555 boulevard de l'Université, Quebec, Canada*
[4]*Mechanical and Maintenance Engineering Department, German Jordanian University, Amman, Jordan*
[5]*Center of Renewable Energy, Water, and Environment and Sustainability, The Hashemite University, Zarqa, Jordan*
[6]*College of Engineering and Architecture, Gulf University for Science and Technology, Kuwait*

Received 22 Apr. 2022, Revised 24 Sep. 2023, Accepted 27 Sep. 2023, Published 1 Oct. 2023

Abstract: The article proposes the use of Artificial Intelligence (AI) models to predict the performance of a sun tracking Photovoltaic (PV) system built-in Zarqa City, Jordan. The system is off-grid with various Azimuth angles and tilt angles. The study involved taking various measurements over a 6-month period. The prediction models employed Artificial Neural Networks (ANN) with five different prediction classifiers, namely, random forest, forest tree, multilayer perceptron (MLP), BPF regression, and linear regression, to predict the performance of the sun-tracking PV system using experimental data. Different metrics are used to demonstrate and validate the accuracy of the proposed models. It is found that all proposed prediction models are of great accuracy. The best prediction classifier is found to be a forest tree classifier with an R2 value of 99.79% and a minimum absolute relative error of 2.36%. Moreover, the least accurate prediction classifier is found to be the linear regression with an R2 of 95.27% and an absolute relative error of 25.71%.

Keywords: Photovoltaic Systems, Artificial Neural Network, Ambient Temperature, Extreme Learning Machine, Artificial Intelligence

1. INTRODUCTION

Due to the decreased cost of solar photovoltaic (PV) technologies, it is considered one of the most viable alternative energy techniques. Around the globe, solar PV plants have been installed from very small rooftop systems to mega-plants. However, the efficiency of individual PV solar panels is controlled by many environmental factors such as temperature and dust, among others. It is extremely important to investigate such factors as dust and ambient temperature either when choosing a site or even after the installation to choose an optimal cleaning frequency that will optimize the efficiency of the PV panels. Measurement complexity for the environmental variables is sometimes neglected when choosing a site for large PV power plants. However, the interest in alternative energies, especially solar, due to their abundance, is increasing because it offers a viable solution to the tradeoff between the mitigation of greenhouse emissions and high energy demands. The output of PV power plants is very unpredictable due to the affecting variables such as clouds, humidity, wind speed, temperature, etc. This unpredictability affects the power grid in many different ways, including but not limited to power planning, ancillary generation, reliability, etc. Thus, accurate forecast techniques are needed for both the electric industry and policymakers.

The motivation behind this work is to study the efficiency of PV panels within a PV power plant located in Zarqa, Jordan, to produce measurable results that will assist both investors and policymakers in planning the mobilization and installation of PV power plants. Jordan is under a lot of pressure for electricity demands and with few natural resources that could be used as a source of energy to generate electricity. Alternative Energy sources are a viable economical solution to Jordan's increased demand for power, and with the abundance of solar power, a lot of research is being done on PV power plant efficiency and PV power plant feasibility. The developments in weather prediction now make it easier for researchers to study the quantifiable effects of electricity demand using environmental weather variables. The current research fills a void in the literature related

E-mail address: Jaafar.alghazo@und.edu

to identifying the proneness of the weather trends on electricity consumption.

The main contribution of the current research is the use of neural networks on quantifying the effects of environmental weather conditions such as dusty environment and ambient temperature on sun-tracking PV power panels that could be used as a base for predicting the performance of these systems. The Middle East area including Jordan and the other Middle Eastern countries such as Saudi Arabia, Syria, Iraq, Oman, United Arab Emirates, Bahrain, Qatar, Palestine, etc. offer vast areas of empty land suitable for Solar Power Plant. However, most of these areas are in dessert land that suffer from continuous dust and other conditions that could result in reducing the efficiency of the solar cells and the actual output power. Thus such studies are extremely important especially in a region that could potentially benefit from alternative power plants and could be potentially an area where many of the alternative power research could be financed and done. Five neural networks technique are applied, namely, Multilayer Perceptron (MLP), Random Tree (RT), Random Forest (RF) , Radial Basis Function Regression (RBF Regression), and Linear Regression (LR). The data consisted of real-time experimental data obtained from a PV power plant at Hashemite University, Zarqa, Jordan, over a period of 6 months. Many variables were included in the dataset to include angle, dust accumulation, ambient temperature, etc., taken on a different hourly basis for a period of 6 months. The use of Neural Networks as a technique for forecasting proven accuracy is of great importance to match the electric power supply with consumer demands. The Hashemite Kingdom of Jordan is a country located in the Middle East that has similar characteristics as the neighboring countries but suffers from a lack of natural resources in addition to an influx of refugees from neighboring countries due to the unstable political situation in some of the neighboring countries. The sudden spike in population due to this fact burdens the country with the increased demand for electric power. Thus, this study and similar studies provide a useful guide for energy economics, energy planning, energy policy, and sustainability for Jordan. Countries with similar socio-economic energy consumption and a similar climate to that of Jordan can also benefit greatly from the findings of this study to plan their sustainable energy consumption agenda. Jordan and other nations in the Middle East such as Syria that suffer from poor economy and scarcity in natural resources also the means for offering continuous services to its citizens or offers these services at very high prices. In Syria for example, power is cut off from the citizens for part of the day, in Jordan, it is continuous but higher prices. In Jordan, citizens receive water for example on specific days because the country lacks water resources and lack the economy for offering this service continuously to its citizens. Therefore, it is extremely important to continue the research on alternative power sources so that when the technology of alternative power plants is reduced enough for these economies to benefit from them, then they will be ready with studies such as this that will assist in choosing the location, environment, type, technology and other variables for large power plants. It can be noted here that in Jordan for example the solar technology for individual houses is widely used for heating water.

This study is not without limitation as it does not combine all the different environmental conditions due to the computational complexity of some variables. A more comprehensive study would be required in the future to measure all variables for a long period of time to study their effects and produce a more accurate thermal mapping, including wind direction, humidity, rainfall, and pollution.

The novelty of the presented work and as far as the author's knowledge is that this study proposes five different prediction classifiers for the output power prediction of a PV sun tracking system installed in Zarqa City, Jordan. This is the first time to utilize ANN techniques to predict the output power for PV sun-tracking systems in Jordan with very high accuracy. The measured data are taken through a period of 6 months with variable ambient temperature, wind speed, different azimuth and tilt angles, and the incident irradiance.

2. LITERATURE REVIEW

A lot of research has been done on various aspects of the performance of the sun-tracking Photovoltaic (PV) system, whether through analysis, prediction, or parameter setting for optimal performance. In [1], the authors propose a model in order to investigate the effect of wind speed, cell temperature, and solar irradiance on the performance of PV systems with a case study at the Hashemite University, Jordan. Analysis of variance was the methodology of choice for demonstrating and validating the model. A linear relationship was found to exist between these parameters and the generated power. No correlation between the parameters was reported. The complete model predicts the generated power with an R2 value of 96.5%, and the study noted that solar irradiance was the most effective of the three parameters.

In [2], the authors studied PV system performance in a desert environment. Dust accumulation, in particular, was used to determine the output power loss of the PV system and in order to find the optimal tilt angle. The study was done in Upper Egypt in a desert environment. The power was recorded for the PV panels throughout the test period for various tilt angles for both the dusty and cleaned PV panels. After 10 months of dust accumulation, the power reduction ranged from 43%-25% for tilt angles of 15-45 degrees, respectively. In [3], the authors explore the difference of thermal characteristics between PV systems by varying the ambient temperature from 25 °C to a maximum of 50 °C with and without fins. In [4], the authors propose a thermodynamic model of PV energy conversion system basing the model on both the 1st and 2nd Laws of thermodynamics to include entropy generation, thermal, optical, fill factor losses, and spectral. They classify their proposed model as exoreversible, endoreversible, reversible, and irreversible systems. The authors also present the energetic and exergetic efficiencies as well as the irreversibilities and thermodynamic losses. In [5], the authors utilize GIS special multi-criteria evaluation and fuzzy logic for the integration of the effects of both dust and temperature in siting large

PV power plants. A case study of Oman is used for land suitability analysis to implement large PV plants. After considering both dust and temperature, they found that land suitability has decreased by 81% and they also concluded through comparison that concentrated PV (CPV) technology is better suited for implementing large PV plants. In [6], the authors present a broad review of the different PV monitoring systems. Their review included a general view of all the major PV monitoring evaluation methods with their relative performance. The review provides information that is deemed essential for developing viable, low-cost, and effective PV monitoring systems. In [7], the authors study the impact of humidity levels, air velocity, and dust accumulation separately on PV performance. They additionally study the impact of each of the factors on the other. They conclude that there is a significant correlation between the factors on each other and that all three should be accounted for in parallel whenever designing a PV plant.

In [8], the authors provide a comprehensive review of the aspects that influence the operation and efficiency of the PV electric generation systems, which include but are not limited to selecting the right equipment, ambient conditions, and PV cell technology. They also include proposals for the design of efficient PV systems. In [9], the author studied the factors that have influenced the decline in the cost of producing solar panels. Their findings indicate four key factors that have influenced the decline in cost, and these are increased investment, increased manufacturing capability in China, technological innovations, and finally, the reduction in the cost of the main raw materials. In [10], the authors assess and design a simple automatic self-cleaning PV system to improve the efficiency of the solar panels. There was a noteworthy increase in the solar panel efficiency for a case study done in Alkhobar, Saudi Arabia, which also led to an increase in the average power output as well. A correlation was also found between efficiency and ambient temperature of the solar panel, with panel efficiency increasing when panel temperature decreases. In [11], the authors develop a model for the prediction of power loss at different levels of irradiance and then develop a neural network model using the obtained experimental data. They concluded that regression is a valid method to quantify and analyze the dust particle size and its influence on the losses in PV systems. They also concluded that neural networks were good predictors for the power output of dusty PV Panels. In [12], the authors propose a hybrid clustering method to pre-process data used for training neural networks. They also propose a data division method for predicting the power output of dusty solar panels. They explored different algorithms, but Neural Network Random (NNR) with random division outperformed the other for both known and unknown soil samples. In [13], the authors investigate the effect of change in weather conditions and dust accumulation of the PV panel efficiency using an experimental study done in Surabaya, Indonesia. They base the study on a proposed rule-based model to identify the conditions affecting the PV panel efficiency. Using a Cortex-M4 ARM microcontroller with attached sensors to measure current, voltage, humidity, and temperature to measure the PV power output. Two weeks' accumulation of dust resulted in a reduction of 10.8% in average humidity of 52.24%. Rainy conditions with an average humidity of 76.32% resulted in over 40% reduction, while cloudy conditions with an average humidity of 60.45% resulted in over 45% reduction.

In [14], the authors presented a systematic and complete review of the available PV systems output power forecast. Their review covered the PV output power profile, various factors that affect the PV power forecast, and the performance matrices used in evaluating the forecasting models. AI-based models are included in the review as well. The authors conclude with a discussion of the benefits of using hybrid methods for PV power forecasts. In [15], the authors used statistical analysis and processing for modeling the hourly electricity demands in Jordan between 2007-2016. The degree-day approach was used for actual weather data. The authors were able to determine the hourly, daily, and monthly seasonal variation indices. A thermal map of the consumption in electricity was identified using the elasticity of the polynomial functions. Their model was successful in relation description of the mean ambient temperature and daily demand. The average comfort zone was reported at a width of 4 °C, while the average mean temperature base was reported to be 17.9 °C. Sensitivity was reported to have increased by 11% and 16.4% for hot and cold weather conditions, respectively. For cooling, the electricity demand saturated at 32.9% and 4.7% for heating. In [16], the electrical performance of a PV system is discussed using the operating temperature of a sun commercial-grade silicon solar cell. In [17], the authors present a review paper that shows the long-term dynamic of soft and hard costs that are linked with the deployment of PV systems since the beginning of the 1990s in Germany. Hardware has decreased by 70-87% since the 1990s, while soft factors such as installation and planning have decreased by 65-85%. In [18], the authors show prediction models for other alternative power sources such as wind energy. They propose a hybrid model, the dubbed ESMD-PSO-ELM method. They use an experimental case study in Yunnan, China, to show the validity of their proposed model. In this present paper, Artificial Neural Networks (ANNs) methods are used for predicting the performance of the PV system. Various ANNs methods were studied from previous literature. In [19], the authors propose a modified unified method for handwritten Persian and Arabic numerals recognition using the improved structural feature. A 96.15% average accuracy recognition was reported. In [20], the authors proposed a feature-based method for multi-language handwritten numeral recognition. They presented various classifiers, among which Random Forest (RF) and others were used, and RF achieved the best accuracy of 96.73%. In [21], the authors presented a comparison study of the various combination and arrangements of photoelectric cells in Jordan. The concluded that the landscape arrangement produced better results. In [22], the authors presented a study for the comparison of fixed and tracking solar power systems. They concluded that the tracking solar power systems are far superior and produce annually 31.29% higher than the fixed power systems. More comprehensive studies such as that presented in [23] explore the location as compared to the equator and also explore fixed or tracking power systems. In addition, even in the tracking systems they further explore the comparison between single axis or two axis systems.

As Jordan is under a lot of pressure for electricity demands and with few natural resources that could be used as a source of energy to generate electricity, these findings may be useful to guide the energy planning and sustainability in Jordan.

3. EXPERIMENTAL SETUP AND DATA COLLECTION

A 5-module off-grid connected photovoltaic system is mounted in the Hashemite University located in the city of Zarqa, Jordan. The nominal power of each module is 285 Wp (Suntech, STP285–24/Vd). The modules are connected in series, as pictured in Figure 1. The system is equipped with two actuators operated by two motors used to change periodically and manually the system's azimuth and tilt angles. A rheostat, buried in the ground, is utilized to absorb the complete output power. A data acquisition system monitored the system performance from 04/09/2016 to 23/01/2017. The system parameters and meteorological data are collected each minute and manipulated on an average hourly basis for further analysis, using the following sensors:

- Pyranometer: It is used to measure the total incident irradiance on the PV system, and it is fitted to the panels and moves with them. Further calculations taking into account the azimuth and tilt angles are not needed.
- k-type thermocouple: It is installed at the surface bottom of the module and measures cell temperatures.
- Anemometer: It is mounted close to the PV system, obstacle-free, and measures wind speed.
- DC voltage transducer: measures the output DC voltage of the PV string.
- DC current transducer: measures the DC output current at the output of the PV string.

Weather condition are one variable in the output power for solar power systems but there are many other variables as was discussed in the literature review pertaining to the photovoltaic cells, the cleanliness of the cells, the angle, etc. All these conditions have been accounted for in this study.

4. METHODOLOGY

The collected data is divided into two sets for prediction purposes; 80% for training and 20% for testing. Several neural network algorithms are used in this paper for the prediction of the performance of a sun-tracking PV system. The neural network algorithms are detailed in the references; however, we briefly describe each algorithm below.

Figure 1. Serial One String Installation of PV modules

A. Random Forest (RF)

This method is an example of ensemble learning, i.e., it uses multiple algorithms to construct a number of decision trees at training time [24]. The output of a random forest classifier consists of the mode of the classes. They are superior in correcting the overfitting of decision trees to the training set of the images. RF is governed by rules to control the tree combinations, tree growing, post-processing, and self-testing. A stable method, even in the presence of outliers and high dimensional parameter spaces compared to other machine learning algorithms, is also robust in overfitting [25]. RF works by ranking the features while building the classification method and thus identifies the most effective features.

B. Multilayer Perceptron (MLP)

MLP is a feedforward artificial neural network that uses backpropagation for training [26], [27]. It consists of at least three layers; an input, hidden, and output layer. Except for the input node, what differentiates MLP from other classifiers is that each of its neurons utilizes a nonlinear activation function. Therefore, it is capable of distinguishing data even those not linearly separable. An MLP layer involves several layers of nodes in a directed cyclic graph where every layer is completely connected to the subsequent layer. The MLP used in this study consisted of the following parameters; Batch size = 100, hidden layers= number of feature attributes (6), Epochs=500, learning rate=0.3, and momentum=0.2.

C. Random Trees (RT)

In [28], the authors introduce and detail the Random trees used for the classification and regression problems. In [29], the Random trees are described as a collection (ensemble) of tree predictors that could be referred to as a Forest. Within the RT, all the individual trees are trained using the same parameters but applied to different datasets. A bootstrap procedure is used to generate the sets, which implies the random selection of the same number of vectors for each set as that used in the original set. The vector replacement will also be chosen, which basically means the occurrence of some vectors more than once and the absence of some. At each tree node, a random subset of the variable will be classified, meaning that a new subset will be generated for each new node. The training parameter is set to (number of variables) *0.5 by default.

D. Linear Regression (LR)

Linear regression is a technique commonly used in machine learning problems [30], [31]. Linear regression takes an input feature vector $x = [x_0, x_1, \dots\dots\dots, x_n]$ and attempts to predict a variable y , based on a linear equation in x, such that

$$\dot{y} = X^T.a + b = a_0x_0 + a_1x_1 + ... + a_nx_n + b \quad (1)$$

where $\dot{y}$ is the predictor of y, X^T is the transpose of x, and $a = [a_0\, a_1 \dots a_n]$ and scalar b are the parameters to estimate. Normally, a and b are estimated by minimizing the sum of squared errors given by equation 2.

$$\sum(\dot{y} - y)^2 \quad (2)$$

which is summed over the different data points $(y, x_0, x_1, \dots, x_n)$. One way to avoid model overfitting is to employ regularization such as L1, aka LASSO (Least Absolute Shrinkage and Selection Operator) regularization, when the model complexity increases as a result of having to deal with a large number of features, n. This regularization adds a penalty "$\lambda \sum |a_i|$" to the error shown in equation 3.

$$\sum(\dot{y} - y)^2 + \lambda \sum |a_i| = \sum(\dot{y} - y)^2 + \lambda(|a_0| + |a_1| + ... + |a_n|) \quad (3)$$

where (lambda) is the regularization parameter that penalizes all parameters except the intercept b. This helps the model generalize and become sparser by reducing the importance of the higher terms leading to a less complex equation and avoiding overfitting the data.

E. Radial Basis Function Networks Regression

To implement regression for predictions, Radial Basis Function (RBF) networks have been extensively employed [32]. RBF networks are networks that take n inputs in the input layer and feed them forward or connect them to m neurons in the hidden layer, each representing a basis function. A weight is assigned to the m basis functions output in the hidden layer and then connected to the output neuron in the output layer. The output function is given by equation 4.

$$f(x) = \sum w_i\, \phi_i(x, c_i) + b \quad (4)$$

where w are the weights, x the inputs, b the bias, and ϕ the radial functions.

Parameters of a radial function include the distance scale, the center, and the exact shape of the radial function. One popular radial function is the Gaussian radial function $\phi(x, c, r) = \exp(-\frac{x-c}{r^2})$ where c is the center and r is the radius. By using Gaussian functions, the predicted value (network output) is, therefore, the sum of the various Gaussian bell-shaped curves. Radial functions are not only used in RBF networks but as popular kernels in support vector machines.

5. PERFORMANCE METRICS

The proposed models in predicting the performance of the sun-tracking PV system require various performance metrics:

A. Mean Absolute Error (MAE)

MAE is a metric to calculate the average of all absolute errors shown in equation 5.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |x_i - x| \quad (5)$$

where n is the number of errors, $|x_i - x|$ is the absolute errors.

B. Root Mean Square Error (RMSE)

$RMSE$ is a metric calculated by the standard deviation from the residuals shown in equation 6.

$$RMSE = \sqrt{\overline{(f - o)^2}} \quad (6)$$

where f is forecasted or expected values and o is observed values

C. Relative Absolute Error (RAE)

RAE is a method for measuring the performance of a predictive model shown in equation 7. It compare actual earning (A) with predicted earnings (P).

$$U_1 = \frac{\left[\sum_{i=1}^{n}(P_i - A_i)^2\right]^{\frac{1}{2}}}{\left[\sum_{i=1}^{n} A_i^2\right]^{\frac{1}{2}}} \quad (7)$$

D. Root Relative Square $E_i(RRSE)$

$RRSE$ is shown in equation 8.

$$E_i = \sqrt{\frac{\sum_{i=1}^{n}(P_{ij} - T_j)^2}{\sum_{j=1}^{n}(T_j - \bar{T})^2}} \quad (8)$$

where P_{ij} is the predicted value by the individual program i for same case j (out of n sample cases); T_j is the target value for sample case j; and $\bar{T}$ is given by equation 9.

$$\bar{T} = \frac{1}{n}\sum_{j=1}^{n} T_j \quad (9)$$

E. Correlation Coefficient Equation

Calculated through the division of the covariance by the product of the two variables standard deviation shown in equation 10.

$$\rho_{xy} = \frac{cov(x, y)}{\sigma_x \sigma_y} \quad (10)$$

where

ρ_{xy} = Pearson product-moment correlation coefficient

$cov(x, y)$ = covariance of variable x and y

σ_x= standard deviation of x

σ_y= standard deviation of y

F. Spearman's Rank-Order Correlation

Two methods depending on whether data has tied ranks or does not have tied ranks, shown in equation 11.

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \quad (11)$$

when data does not have tied ranks where d_i = difference in paired ranks and n= number of cases.

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \tag{12}$$

when data has tied ranks where i = paired score, this is shown in equation 12.

G. Kendall's tau

$$\tau a = Nc - \frac{NdN(N-1)}{2} \tag{13}$$

where the Nc and Nd do not add 0.5 to the tied values. Kendall's tau-b is defined as

$$\tau b = Nc - Nd(Nc + Nd + Tx)(Nc + Nd + Ty) \tag{14}$$

where Tx is the number of pairs tied for the first response variable, only Ty is the number of pairs tied for the second variable. Kendall's tau-c is defined as

$$\tau c = 2(Nc - Nd)n2(m-1)/m \tag{15}$$

where m is the minimum of Xd and Yd where Xd is the number of distinct values of X and Yd is the distinct values for Y.

In this paper, the dataset is initially subjected to a preprocessing stage in which the data is cleaned up. The preprocessing step consisted of various algorithms to standardize the dataset which include imputation, normalization, and others in order for the dataset to be ready for the neural network algorithms.

The neural networks were then trained using the training portion of the dataset and tested using the testing portion of the dataset. The results are reported in the results section.

6. RESULTS

Using the seven performance metrics, the five classifiers are evaluated in order to achieve the best possible prediction rate for the performance of the PV system. The performance metrics were carefully chosen for this kind of dataset and this kind of prediction. Table 1 shows the average performance metrics acquired from the five classifiers, with 80% data for training and validation and 20% for testing.

Table 1 shows that the best prediction based on the accuracy criteria discussed before (in the relevant equations) is achieved using the random forest classifier with a correlation coefficient of 99.79% and a relative minimum absolute error among all the proposed classifiers with 2.36%. Also, the table shows that the worst prediction classifier is the linear regression with a relative absolute error of 25.71% and a correlation coefficient of 95.27%. Moreover, the table illustrates the other accuracy metrics that are applied for comparison purposes. Finally, as a conclusion, the table shows the best prediction classifiers in ascending order. In table 1, the random tree is the second classifier in terms of performance. The problem of over-fitting and under-fitting is accounted for in random forest and random trees. We notice that MLP, RBF regression, and linear regression are not well suited for this kind of prediction problem as the performance metrics reported are showing that they are very poor predictors.

For validation purposes, cross-validation with 10 folds results is presented in Table 2. Results assure that the best prediction classifier is the random forest, and the worse one is the linear regression. The Random Tree is still showing to be the second highest performing classifier and predictor. Again, in the cross validation, the MLP, RBF regression, and linear regression are showing to be poor performers for this type of problem and this type of dataset.

Shown in Figure 2 is a comparison between the actual measured power and the predicted output power distributed over the span of the experiment. The graph elucidates that the best graph with a minimum error to the actual measured output power is achieved by the forest tree classifier. And the maximum error is achieved by the linear regression model. However, all classifiers capture the trend of the actual measured output power with a min R2 of 95.38%, which is excellent by all measures. Figure 2 shows the predicted values for the tested portion of the dataset and it is apparent that Random forest in (Orange) is the closest to the actual values from the dataset (Red). The other algorithms are capable to capture the trend of the predicted values but not close enough to the actual values which was also shown in Table 1 and Table 2. The rising and falling edge of the figure have no significant meaning other than the fact that this is the trend of the actual values produced by the actual measurement over a period of 6 months. This is practical comparing that in the summer months we will achieve the highest output power compares to month prior to and after the summer period.

7. CONCLUSION

The current paper aims to present a prediction of the output power of a sun-tracking PV system utilizing various prediction ANN classifiers installed in Zarqa city, Jordan. The measured output power of the system was taken over a period of six months with different variables recorded. These include ambient temperature, wind speed, azimuth and tilt angles of the PV module, and incident irradiance. The classifiers used in the prediction are random forest, forest tree, multilayer perceptron (MLP), BPF regression, and linear regression. Results of the study show that all classifiers achieve excellent accuracy. Moreover, the best classifier with the highest R2 and the minimum relative absolute error is found to be a random forest. Finally, the least accurate prediction classifier is found to be linear regression.

The study of prediction of output power for alternative power sources is extremely important and it is imperative that such studies be promoted because the classical power sources such as oil, coal, and natural gas will eventually be depleted and the human resource reliance on alternative power sources will increase. In addition, the benefits of alternative power source for the environment and the cleanliness of these sources will contribute positively to the release of carbon dioxide and its negative effects on earth atmosphere layers.

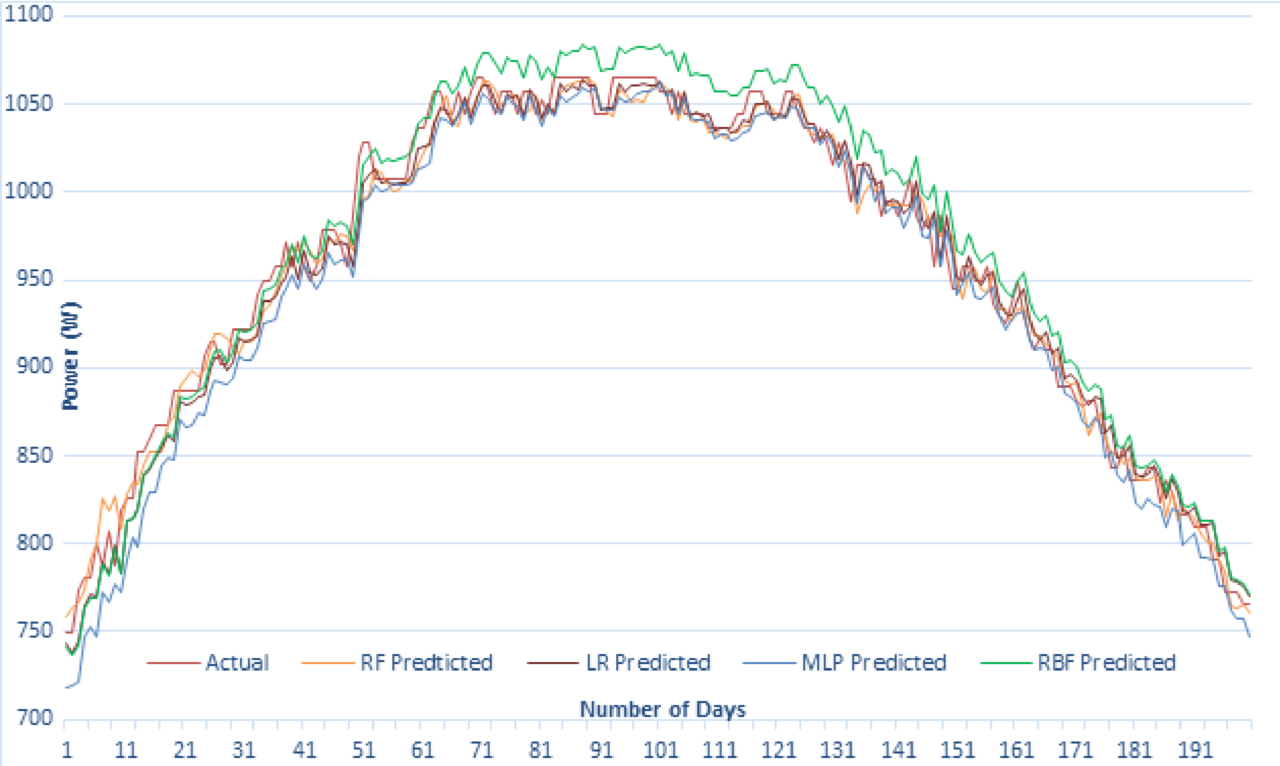


Figure 2. Power output using 5 different prediction classifiers compared with the actual measured one vs. number of days

TABLE I. Comparison results of the prediction classifiers and the Performance Metrics

Regression % Split: 80% Training, 20% Testing							
Classifier	MAE	RMSE	RAE	RRSE	Correlation Coefficient	Spearman RHO	Kendall TAU
Random Forest	8.0118	29.055	2.36%	6.50%	0.9979	0.8324	0.6985
Random Tree	9.3389	40.705	2.75%	9.10%	0.9959	0.7619	0.6646
Multilayer Perceptron	30.778	82.736	9.06%	18.50%	0.9827	0.6325	0.4849
RBF Regression	46.177	92.527	13.59%	20.68%	0.9784	0.6765	0.5069
Linear Regression	87.374	136.020	25.71%	30.41%	0.9527	0.6414	0.4863

TABLE II. Regression based on Cross-Validation with 10 folds

Regression-based on Cross-Validation with 10 folds							
Classifier	MAE	RMSE	RAE	RRSE	Correlation Coefficient	Spearman RHO	Kendall TAU
Random Forest	7.4478	27.385	2.19%	6.11%	0.9981	0.8401	3.3305
Random Tree	9.4724	39.599	2.79%	8.84%	0.9961	0.7325	2.0809
Multilayer Perceptron	29.2823	79.236	8.61%	17.68%	0.9842	0.6917	1.0362
RBF Regression	47.6946	93.388	14.03%	20.84%	0.9780	0.7108	1.2248
Linear Regression	87.3302	134.563	25.69%	30.03%	0.9538	0.6391	0.2847

Therefore, the results obtained in the paper is a step forward in the added body of knowledge for the continuous growth of alternative power sources and further studies in solar power, wind power, etc. will continue until optimal solutions for power predictions is achieved to assist in the design and development of power plants for alternative power sources.

References

[1] A. Al-Bashir, M. Al-Dweri, A. Al-Ghandoor, B. Hammad, and W. Al-Kouz, "Analysis of effects of solar irradiance, cell temperature and wind speed on photovoltaic systems performance," *International Journal of Energy Economics and Policy*, vol. 10(1), pp. 353–359, 2020.

[2] E. Abdeen, M. Orabi, and E.-S. Hasaneen, "Optimum tilt angle for photovoltaic system in desert environment," *Solar energy*, vol. 155, pp. 267–280, 2017.

[3] J. P. Kim, H. Lim, J. H. Song, Y. J. Chang, and C. H. Jeon, "Numerical analysis on the thermal characteristics of photovoltaic module with ambient temperature variation," *Solar Energy Materials and Solar Cells*, vol. 95, no. 1, pp. 404–407, 2011.

[4] S. Kaushik, R. Rawat, and S. Manikandan, "An innovative

thermodynamic model for performance evaluation of photovoltaic systems: Effect of wind speed and cell temperature," *Energy conversion and management*, vol. 136, pp. 152–160, 2017.

[5] Y. Charabi and A. Gastli, "Integration of temperature and dust effects in siting large pv power plant in hot arid area," *Renewable Energy*, vol. 57, pp. 635–644, 2013.

[6] S. R. Madeti and S. Singh, "Monitoring system for photovoltaic plants: A review," *Renewable and Sustainable Energy Reviews*, vol. 67, pp. 1180–1207, 2017.

[7] S. Mekhilef, R. Saidur, and M. Kamalisarvestani, "Effect of dust, humidity and air velocity on efficiency of photovoltaic cells," *Renewable and sustainable energy reviews*, vol. 16, no. 5, pp. 2920–2925, 2012.

[8] M. E. Meral and F. Dincer, "A review of the factors affecting operation and efficiency of photovoltaic based electricity generation systems," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 5, pp. 2176–2184, 2011.

[9] U. Pillai, "Drivers of cost reduction in solar photovoltaics," *Energy economics*, vol. 50, pp. 286–293, 2015.

[10] N. Nader, W. Al-Kouz, and S. Al-Dahidi, "Assessment of existing photovoltaic system with cooling and cleaning system: Case study at al-khobar city," *Processes*, vol. 8, no. 1, p. 9, 2019.

[11] S. Pulipaka and R. Kumar, "Power prediction of soiled pv module with neural networks using hybrid data clustering and division techniques," *Solar Energy*, vol. 133, pp. 485–500, 2016.

[12] S. Pulipaka, F. Mani, and R. Kumar, "Modeling of soiled pv module with neural networks and regression using particle size composition," *Solar Energy*, vol. 123, pp. 116–126, 2016.

[13] M. A. Ramli, E. Prasetyono, R. W. Wicaksana, N. A. Windarko, K. Sedraoui, and Y. A. Al-Turki, "On the investigation of photovoltaic output power reduction due to dust accumulation and weather conditions," *Renewable Energy*, vol. 99, pp. 836–844, 2016.

[14] M. Q. Raza, M. Nadarajah, and C. Ekanayake, "On recent advances in pv output power forecast," *Solar Energy*, vol. 136, pp. 125–144, 2016.

[15] A. Almuhtady, A. Alshwawra, M. Alfaouri, W. Al-Kouz, and I. Al-Hinti, "Investigation of the trends of electricity demands in jordan and its susceptibility to the ambient air temperature towards sustainable electricity generation," *Energy, Sustainability and Society*, vol. 9, no. 1, pp. 1–18, 2019.

[16] E. Skoplaki and J. A. Palyvos, "On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations," *Solar energy*, vol. 83, no. 5, pp. 614–624, 2009.

[17] L. Strupeit and L. Neij, "Cost dynamics in the deployment of photovoltaics: Insights from the german market for building-sited systems," *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 948–960, 2017.

[18] J. Zhou, X. Yu, and B. Jin, "Short-term wind power forecasting: A new hybrid model combined extreme-point symmetric mode decomposition, extreme learning machine and particle swarm optimization," *Sustainability*, vol. 10, no. 9, p. 3202, 2018.

[19] J. M. Alghazo, G. Latif, A. Elhassan, L. Alzubaidi, A. Al-Hmouz, and R. Al-Hmouz, "An online numeral recognition system using improved structural features–a unified method for handwritten arabic and persian numerals," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 9, no. 2-10, pp. 33–40, 2017.

[20] J. M. Alghazo, G. Latif, L. Alzubaidi, and A. Elhassan, "Multi-language handwritten digits recognition based on novel structural features," *Journal of Imaging Science and Technology*, vol. 63, no. 2, pp. 20 502–1, 2019.

[21] S. S. Alrwashdeh, "Investigation of the energy output from pv panels based on using different orientation systems in amman-jordan," *Case Studies in Thermal Engineering*, vol. 28, p. 101580, 2021.

[22] B. Hammad, A. Al-Sardeah, M. Al-Abed, S. Nijmeh, and A. Al-Ghandoor, "Performance and economic comparison of fixed and tracking photovoltaic systems in jordan," *Renewable and Sustainable energy reviews*, vol. 80, pp. 827–839, 2017.

[23] S. A. Alkaff, N. H. Shamdasania, G. Y. Ii, and V. K. Venkiteswaran, "A study on implementation of pv tracking for sites proximate and away from the equator," *Process Integration and Optimization for Sustainability*, vol. 3, pp. 375–382, 2019.

[24] A. P. Bhatkar and G. Kharat, "Detection of diabetic retinopathy in retinal images using mlp classifier," in *2015 IEEE international symposium on nanoelectronic and information systems*. IEEE, 2015, pp. 331–335.

[25] A. A. Mahmoud, I. N. A. Alawadh, G. Latif, and J. Alghazo, "Smart nursery for smart cities: Infant sound classification based on novel features and support vector classifier," in *2020 7th international conference on electrical and electronics engineering (ICEEE)*. IEEE, 2020, pp. 47–52.

[26] G. Latif, M. M. Butt, A. H. Khan, M. O. Butt, and J. F. Al-Asad, "Automatic multimodal brain image classification using mlp and 3d glioma tumor reconstruction," in *2017 9th IEEE-GCC Conference and Exhibition (GCCCE)*. IEEE, 2017, pp. 1–9.

[27] J. D. Labhade, L. Chouthmol, and S. Deshmukh, "Diabetic retinopathy detection using soft computing techniques," in *2016 international conference on automatic control and dynamic optimization techniques (ICACDOT)*. IEEE, 2016, pp. 175–178.

[28] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[29] ——, "Random forests leo breiman and adele cutler," *Random Forests-Classification Description*, vol. 106, 2015.

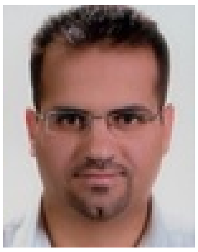
[30] D. Nguyen, N. A. Smith, and C. Rose, "Author age prediction from text using linear regression," in *Proceedings of the 5th ACL-HLT workshop on language technology for cultural heritage, social sciences, and humanities*, 2011, pp. 115–123.

[31] R. L. Ott and M. T. Longnecker, *An introduction to statistical methods and data analysis*. Cengage Learning, 2015.

[32] Y. Wu, H. Wang, B. Zhang, and K.-L. Du, "Using radial basis function networks for function approximation and classification," *International Scholarly Research Notices*, vol. 2012, 2012.

Jaafar Alghazo obtained his Ph.D. in Engineering Sciences from Southern Illinois University-Carbondale (USA). He is an Associate Professor in the Electrical and Computer Engineering Department at the Virginia Military Institute. He previously served as an Associate Professor in the College of Computer Science and Engineering at Prince Mohammad Bin Fahd University (PMU). His research interests are Machine Learning, Artificial Intelligence, and Engineering Education.

Ghazanfar Latif is post-doctoral fellow researcher at at University of Quebec Chicoutimi, Canada and also serving as Research Coordinator at Prince Mohammad bin Fahd University, Khobar, Saudi Arabia. He is a computer scientist whose research interests include medical image processing, deep learning, artificial intelligence, machine learning, data science and computer vision.



Bashar Hammad holds PhD degree in Mechincal Engineering which is earend from Virginia Tech (VT) in 2008. His current research interests are in photovoltaic systems, energy harvesters, and e-learning education. He served as an assistant dean of Engineering for student affairs and training and the director of Academic Quality Assurance Center at HU.

Mohammad Al-Abed is experienced Associate Professor with a demonstrated history of working in the biomedical engineering and engineering management industry. Experience in Project Management, Renewable Energy, and Sustainability. Skilled in MATLAB, Technical Writing, and Engineering Management. Strong professional skills with PMP certification and a high-level education with a Doctor of Philosophy (Ph.D.) focused on Bioengineering and Biomedical Engineering from The University of Texas at Arlington.

Mohammad Al-Dweri is working as part for Center of Renewable Energy, Water, and Environment and Sustainability, The Hashemite University, Zarqa, Jordan.

Fadi N. Sibai received His PhD and MSc degree in Electrical and Computer Engineering from Texas AM University, and the BS degree from the University of Texas at Austin. He holds PMP, CISSP, CQRM, and CCNA certifications.

Wael Al-Kouz is Professor of Mechanical Engineering at American University of the Middle East, Kuwait. His research interest includes Rarefied flows, Thermal Engineering and Renewable Energy.