Analysis of the Effects of Cell Temperature on the Predictability of the Solar Photovoltaic Power Production

Sameer Al-Dahidi[1*], Salah Al-Nazer[1], Osama Ayadi[2], Shuruq Shawish[1], Nahed Omran[3]

[1]Department of Mechanical and Maintenance Engineering, School of Applied Technical Sciences, German Jordanian University, Amman, Jordan, [2]Department of Mechanical Engineering, Faculty of Engineering, The University of Jordan, Amman, Jordan, [3]Renewable Energy Center, Applied Science Private University, Amman, Jordan. *Email: sameer.aldahidi@gju.edu.jo

ABSTRACT

The use of intermittent power supplies, such as solar energy, has posed a complex conundrum when it comes to the prediction of the next days' supply. There have been several approaches developed to predict the power production using Machine Learning methods, such as Artificial Neural Networks (ANNs). In this work, we propose the use of weather variables, such as ambient temperature, solar irradiation, and wind speed, collected from a weather station of a Photovoltaic (PV) system located in Amman, Jordan. The objective is to substitute the aforementioned ambient temperature with the more realistic PV cell temperature with a desire of achieving better prediction results. To this aim, ten physics-based models have been investigated to determine the cell temperature, and those models have been validated using measured PV cell temperatures by computing the Root Mean Square Error (*RMSE*). Then, the model with the lowest *RMSE* has been adopted in training a data-driven prediction model. The proposed prediction model is to use an ANN compared to the well-known benchmark model from the literature, i.e., Multiple Linear Regression (MLR). The results obtained, using standard performance metrics, have displayed the importance of considering the cell temperature when predicting the PV power output.

Keywords: Renewable Energy, Photovoltaic, Prediction, Cell Temperature, Multiple Linear Regression, Artificial Neural Networks
JEL Classifications: C53, Q47

1. INTRODUCTION

Jordan is a nation lying in the heart of the Middle East, surrounded by Palestine, Iraq, Syria, Saudi Arabia, and shares a water border with Egypt. Unlike the most of the neighboring nations, Jordan does not have enough crude oil to sustain itself. In fact, Jordan relies heavily on the import of the crude oil to satisfy the consumption. This fact meant that Jordan has to import oil at a huge cost which amounts to more than 10% of the total GDP (Department of Statistics 2017; Jaber et al., 2004; Ministry of Energy and Mineral Resources (MEMR) 2017; Ministry of Planning and International Cooperation 2015; National Electric Power Company (NEPCO) 2018).

In order for Jordan to meet its growing energy demand, alternative means of generating energy have been investigated. Jordan's energy strategy has decided to promote Renewable sources of Energy (RE), especially solar and wind; because Jordan lies in the solar belt and has access to strong winds in some parts of the country. According to national vision and strategy (Ministry of Planning and International Cooperation 2015), it was planned to achieve a contribution of 10% related to the total energy mix in 2020. As a result of the implementation of this strategy, the generation capacity of RE projects carried out on the transmission and distribution grids has been increasing from 1.4 MW in 2014 to 980 MW by late 2018, representing about (18.7%) of the total generation capacity (Figure 1) (Ministry of Planning and International Cooperation, 2015).

Among the various RE sources, the Photovoltaic (PV) systems are considered the most popular and strongly attractive source
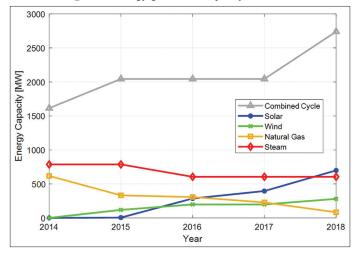
Figure 1: Energy generation capacity since 2014

of energy (Brunet et al., 2018). However, the PV cell relies on sunrays to produce electricity, which poses a problem for the energy distribution companies since the amount of rays hitting the solar panel constantly varies during the day. The power output changes depending on multiple factors affecting the PV, from weather conditions (e.g., wind speed) to the angle of incidence of the solar rays.

With that being said, the energy distribution companies cannot effectively predict the performance of RE and hence do not have the ability to accurately analyze the amount of energy that will be produced by the RE sources and plan accordingly to fulfill the user demand. Therefore, balancing the alternating variable input of the RE becomes a major challenge for the energy supplier, and finding a method or tool to predict (forecast) is necessary in order to aid the implementation of variable RE inputs (Al-Dahidi et al., 2018).

An effective and reliable tool that could be used in power production prediction is the utilization of Artificial Intelligence (AI). AI is a tool that simulates the cognitive behavior of a human brain in machines or computers. The computer or machine initially learns (i.e., Machine Learning [ML]) or in this context data is inputted into the system. The system then utilizes algorithms to attempt to reach a certain target or output. After learning, the system then starts reasoning which algorithm is best to reach the desired output. Finally, the system undergoes a self-correction process, which tries to continually improve the algorithm used to reach the desired output more accurately.

In order to predict the power generated from PV panels, there have been two main types of ML algorithms utilized to determine the power output as accurately as possible. These algorithms can be generally categorized into physics-based and data-driven (Al-Dahidi et al., 2018; Das et al., 2018; Ernst et al., 2009; Moslehi et al., 2018). Physics-based extract a mathematical equation from the collected weather variables (e.g., ambient temperature, irradiation, wind speed, etc.) to find the PV power output. On the other hand, data-driven are appropriated by ML algorithms without the need for any physics-based model. In fact, they exploit pre-existing historical data collected from sensors or a weather station to find a relation between the weather variables and the power output.

In this work, only data driven methods will be analyzed. Most of the previous research works have used the two parameters of irradiation and ambient temperature in their analysis. For example, Fernandez-Jimenez et al., 2012, presented a short term forecasting method that consists of three modules, two of which were Numerical Weather Prediction (NWP) models and the third was an Artificial Neural Network (ANN)-based model. The first two were used to forecast weather variables to be used by the third module. The final value is the hourly power output of the PV plant with a 1-39 h forecast horizon; Liu et al., 2017, proposed the use of BP NN to predict power output up to 24 h-ahead; Zhong et al., 2018, employed the use of both General Regression and BP, and the results were then compared showing more favorable results with BP; Liu et al., 2019, established a Weight Varying Ensemble forecasting model that improved short term power prediction. In (Mellit, 2009), a Recurrent NN (RNN) was used for forecasting the generation of a PV power system; Ding et al., adopted an ANN-based approach. An improved BP learning algorithm is used to overcome the shortcomings of the standard BP learning algorithm; Chow et al., employed ANN to mimic the nonlinear correlation between meteorological factors and power output, and then display that short-term power prediction performance is commensurate to the real-time power prediction performance when ahead solar angles are taken into account; Oudjana et al., adopted NN for one week-ahead prediction using weather variables; Shi et al., proposed a forecasting PV power output approach based on weather classification and Support Vector Machines (SVMs); Hussein et al. (Kazem and Yousif 2017), used neural mathematical models such as Generalized Feedforward Networks (GFF), MultiLayer Perceptron (MLP), Self-Organizing Feature Maps (SOFM) and SVM to predict power produced and compared the results. Al-Dahidi et al., proposed the exploitation of ELM for faster computational speed and better generalization capability and compared the performance of the model with the traditional BP-ANN of literature.

Some other common weather variables used for prediction purposes were the relative humidity and wind speed with the aforementioned variables. For example, Lin et al.,; proposed a unique hybrid prediction model combining improved *K*-means clustering, Grey Relational Analysis (GRA) and Elman NN (Hybrid improved *K* means-GRA-Elman, HKGE) for forecasting the PV power output. The proposed model was implemented using multiple meteorological conditions and history files of PV output.

The main weather variables have been irradiation and ambient temperature. The following research works substituted the ambient temperature with the cell temperature. For example, Ba et al., implemented a statistical approach using Weibull probability distribution function and obtained an accurate relationship for power output between irradiation and the cells' back temperature. The calculated power output was compared to the measured and they obtained a high correlation coefficient. Bouzerdom et al., combined two models: the Seasonal Auto-Regressive Integrated Moving Average method (SARIMA) and the SVM. The hybrid model showed better prediction results. In (Paulescu et al., 2017), two advanced models for predicting the

power output of PV cells were analyzed: a black-box Takagi-Sugeno fuzzy model and a physically inspired, semiparametric statistical model (Generalized Additive Model, GAM) based on smoothing splines. In (Baharin et al., 2016), a Support Vector Regression (SVR) method was used as well as ANN (nonlinear autoregressive), and these methods were compared to a benchmark model using persistence method. In (Yu and Chang 2011), a NN method was implemented using BP algorithms. Al-Bashir et al., employed a Multivariate Linear Regression (MLR) to forecast power output. Moslehi et al., examined various data collection and modelling scenarios for the prediction of the PV power production. In particular, the effect of exploiting measured (or calculated) cell temperatures on the predictability of the PV power production was studied.

So far, the temperature of the module has been underutilized, and few efforts have been made to implement it into the data-driven prediction model. In this work, the cell temperature is derived from ten physics-based models and each result is correlated with the power output, so that the best models will be determined. Afterwards, a validation of the results is carried out and the Root Mean Square Error (*RMSE*) will be compared to choose the best model. Finally, this model will be implemented in developing the Multiple Linear Regression (MLR) and ANN models for the PV power production prediction and evaluating their performances.

The performance of the prediction models is verified with respect to two standard metrics, namely *RMSE* and Coefficient of Determination (R^2).

The remaining of this paper is organized as follows. Section 2 states the PV power production prediction problem. Section 3 presents the ASU solar PV system case study. Section 4 describes the methodology proposed for investigating the effect of incorporating the cell temperature instead of the ambient temperature. Section 5 discusses the obtained results. Finally, some conclusions are drawn in Section 6.

2. PROBLEM STATEMENT

Let us consider the availability of the weather data (W) and the corresponding power productions ($\vec{P}$) of a solar PV system for *Y* years. The former is assumed to combine the hourly values of three main variables: the global solar radiations ($\vec{I}_{rr}$), the ambient temperature ($\vec{T}_{amb}$), and the wind speed ($\vec{v}$). The time stamp in terms of the corresponding hour ($\vec{hr}$) and day ($\vec{d}$) number from the beginning of each year data is also considered. Thus, we can establish an overall matrix $\mathbf{X} = [\vec{hr}\,\vec{d}\,\vec{I}_{rr}\,\vec{T}_{amb}\,\vec{v}\,\vec{P}]$ that will be used to build/develop models for the prediction of the power output of the solar PV system.

In this work, the objective is to substitute the aforementioned ambient temperature ($\vec{T}_{amb}$) with the more realistic PV cell temperature ($\vec{T}_{cell}$), whose values are not measured and, thus, not available during the study period *Y*, and to investigate its importance when predicting the PV power output. To this aim, existing physics-based models have been adopted to determine the cell temperature, and their results have been compared to some cell temperature values measured for a short period of time for validation purposes: the model with the lowest *RMSE* has been adopted and the $\vec{T}_{amb}$ values are replaced with the realistic best obtained cell temperature values ($\vec{T}_{cell}^{best}$). The updated matrix (X′) will be, then, used to build/develop prediction models and the built-models are in need to be evaluated to verify the effectiveness of such a substitution.

The proposed prediction model is to use the Artificial NN (ANN) whose prediction capability is to be compared with the well-known benchmark MLR from the literature.

3. CASE STUDY

The solar PV power grid-connected system of the Applied Science Private University (ASU) of a capacity 264 kWp is presented in this Section. A brief introduction on the site is in order. ASU is a private university located in Amman (Capital of Jordan) at the coordinates 32°2'24.0324" N and 35°54'1.4328" E, latitude and longitude, respectively. The location of the PV cells are found atop the Faculty of Engineering building (Figure 2). The PV array was at an angle of 36° pointing in the direction of southeast, and have a tilt angle of 11°. The inverters connected to the PV panels are of the SMA SUNNY TRIPOWER type and consist of 13 17000W inverters and one 10000 W inverter. The solar panels are of the Yingli Solar: YL 245P-29b-PC type, and those consist of polycrystalline structure (Applied Science Private University, 2019).

The existing weather station in the ASU campus located around 171 m from the Faculty of Engineering helped by tabulating and recording the weather conditions experienced by the PV system, and classifying them into 45 different variables (e.g., solar radiation, ambient temperatures, wind speeds), and gave values for these weather variables every hour for the past *Y*~3.5 years (i.e., May 16, 2015 to December 31, 2018), whereas the inverters connected to the PV panels recorded the corresponding power output delivered by the system (Applied Science Private University (ASU) 2019).

Among the available weather variables, some of them have been excluded from the analysis due to the facts that either their behaviour is constant during the *Y*~3.5 years study period, such as precipitation amounts, or they are irrelevant to the delivered PV power, such as soil surface and subsoil (−10 cm) temperatures, whereas the global solar radiation, ambient temperature at 1m level, and the wind speed at 10 m level have been recommended and utilized for building the prediction models as they have the large influence on the solar PV power productions (Al-Dahidi et al., 2019). In addition to the before-mentioned considered weather variables, the time effect in day hour and number in a year has been also considered while building the prediction models because they represent the diurnal cyclic and the seasonal effects, respectively (Al-Dahidi et al., 2019).

All of the considered hourly weather variables together with the time stamp and the corresponding power productions are

Figure 2: ASU PV panels

stored in the dataset matrix X that is used later on in Section 5 for the purpose of calculating the cell temperature, validating the calculated cell temperatures, building/developing the MLR and ANN prediction models, and comparing their performances.

The whole inputs (weather variables and the time stamp)-outputs (power productions) patterns are divided into (i) training dataset (X_{train}) (it contains N_{train} = 15115 patterns (i.e., 50%) randomly selected from the 30229 inputs-outputs patterns available in the overall dataset matrix), (ii) validation dataset (X_{valid}) (it contains N_{valid} = 7557 patterns (i.e., 25%) randomly selected from the remaining patterns available in the overall dataset matrix), and (iii) test dataset (X_{test}) (it contains N_{test} = 7557 patterns [i.e., the remaining 25%]).

The three datasets will be used to build/develop the prediction models, optimize the models' architectures, and test/evaluate the effectiveness of the predictability of the two prediction models and compare their predictability when the ambient temperature is being replaced with the best obtained cell temperature, respectively.
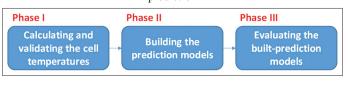
4. METHODOLOGY

In this Section, we describe the methodology proposed for predicting the solar power productions of the ASU PV system. The proposed methodology is structured in three phases and is sketched in Figure 3. The proposed methodology amounts to calculate the ASU cell temperatures by using different physics-based models and validate the calculated values (Phase I – Section 4.1), build/develop two different prediction models (Phase II - Section 4.2), and evaluate the built-prediction models (Phase III - Section 4.3).

4.1. Phase I: Calculating and Validating the Cell Temperatures
4.1.1. Calculating the cell temperatures
Ten different physics-based models (HOMER Pro. 2019; Schwingshackl et al., 2013) are investigated to estimate the PV

Figure 3: The proposed methodology for solar PV power production prediction

cell temperatures (T_{cell}), hereafter denoted as $T_{cell}^1, T_{cell}^2, ..., T_{cell}^{10}$. The models characterize the inherent relationship between the cell temperature, relevant weather variables, such as global solar radiation, wind speed, wind direction, and ambient temperature, and some other characteristics which depend on the PV cell technology under study (i.e., in our case study the polycrystalline silicon (p-Si)).

The different physics-based models adopted in this work are hereafter summarized. For more details on the PV cell temperature physics-based models, the interested reader may refer to (HOMER Pro. 2019; Schwingshackl et al., 2013).

- Standard PV cell temperature model
This is the simplest physics-based model developed for estimating the PV cell temperature (Markvart, 2000). It calculates the cell temperature $\left(T_{cell}^1\right)$ as a function of the ambient temperature (T_{amb}), solar radiation (I_{rr}), and other PV technology dependent characteristics (Eq. (1)).

$$T_{cell}^1 = T_{amb} + \frac{I_{rr}}{I_{NOCT}}.(T_{cell,NOCT} - T_{amb,NOCT}) \qquad (1)$$

where $T_{cell,NOCT}$ is the Nominal Operating Cell Temperature that depends on the PV technology under study whose value is taken at the solar radiation I_{NOCT} = 800 W/m², the ambient temperature $T_{amb,NOCT}$ = 20°C, and wind speed v = 1 m/s. This model is denoted as *Model 1*.

- Skoplaki PV cell temperature model
This model estimates the cell temperatures $\left(T_{cell}^2, T_{cell}^3, T_{cell}^4, \text{and } T_{cell}^5\right)$ by integrating the wind speed and other specific solar cell properties into the standard PV cell temperature model (T_{cell}^1 obtained by Eq. (1)) (Schwingshackl et al., 2013; Skoplaki et al., 2008):

$$T_{cell}^{2,3,4,5} = T_{cell}^1 . \frac{h_{w,NOCT}^{2,3,4,5}}{h_w^{2,3,4,5}(v)} . \left[13\frac{\eta_{STC}}{\tau.\alpha}\left(1 - \beta_{STC}.T_{amb,STC}\right)\right] \quad (2)$$

where η_{STC} and β_{STC} are efficiency and temperature coefficient of maximal power under Standard Test Conditions (STC), respectively, i.e., solar radiation of 1000 W/m², ambient temperature $T_{amb,STC}$ = 25°C, and air mass of 1.5. τ and α are the transmittance of the cover system and absorption coefficient of the PV cells [%], respectively. $h_{w,NOCT}$ is the wind convection heat transfer coefficient for wind speed measured at NOCT conditions, i.e. v = 1 m/s. $h_w^{2,3,4,5}(v)$ are the wind convection heat transfer coefficients which are typically linear functions of the wind velocity (v) as defined in (Skoplaki et al., 2008):

$$h_w^2(v) = 8.91 + 2v_f \tag{3}$$

$$h_w^3(v) = 5.7 + 2.8v_w \tag{4}$$

where v_f is the wind speed whose values are measured at 10m above the ground, whereas v_w is the wind speed whose values are measured close to the PV module. The v_w can be obtained from the v_f through $v_w = 0.68\ vf - 0.5$ (Loveday and Taki 1996; Schwingshackl et al., 2013). The cell temperatures $\left(T_{cell}^2 \text{ and } T_{cell}^3\right)$ obtained using the former equations (Eq. (3) and Eq. (4)) for the wind convection heat transfer coefficient are hereafter denoted as *Model 2* and *Model 3*, respectively.

Other formulations of the $h_w(v)$ have been defined in (Sharples and Charlesworth, 1998) for the wind direction perpendicular and parallel to the PV module's surface as follows, respectively:

$$h_w^4(v) = 8.3 + 2.2v_w \tag{5}$$

$$h_w^5(v) = 6.5 + 3.3v_w \tag{6}$$

The cell temperatures $\left(T_{cell}^{4,5}\right)$ obtained using the former equations (Eq. (5) and Eq. (6)) for the wind convection heat transfer coefficients are hereafter denoted as *Model 4* and *Model 5*.

• Kurtz PV cell temperature model
(Kurtz et al., 2009) estimated the cell temperature $\left(T_{cell}^6\right)$ as follows without distinguishing between the different PV cell technologies:

$$T_{cell}^6 = T_{amb} + I_{rr}.e^{-3.473 - 0.0594v_w} \tag{7}$$

This model is denoted as *Model 6*.

• Koehl PV cell temperature model
This model calculates the cell temperature (T_{cell}^7 – hereafter denoted as *Model 7*) as a function of I_{rr}, T_{amb}, local wind speed (v_w), and other PV cell technology dependent constants (i.e., U_0, U_1) (Koehl et al., 2011):

$$T_{cell}^7 = T_{amb} + \frac{I_{rr}}{U_0 + U_1.v_w} \tag{8}$$

• Mattei PV cell temperature model
(Mattei et al., 2006) estimated the cell temperature as follows:

$$T_{cell}^{8,9} = \frac{U_{PV}^{8,9}(v).T_{amb} + I_{rr}.[\tau.\alpha - \eta_{STC}.\left(1 - \beta_{STC}T_{amb,STC}\right)]}{U_{PV}^{8,9}(v) + \beta_{STC}.\eta_{STC}.I_{rr}} \tag{9}$$

where $U_{PV}^{8,9}(v)$ are the heat exchange coefficients for the total surface of the PV module. Two different formulations for the *UPV* have been defined in (Mattei et al., 2006) for the $U_{PV}^{8,9}(v)$ and adopted in this work, they are:

$$U_{PV}^8(v_w) = 26.6 + 2.3v_w \tag{10}$$

$$U_{PV}^9(v_w) = 24.1 + 2.9v_w \tag{11}$$

The obtained cell temperatures $\left(T_{cell}^8 \text{ and } T_{cell}^9\right)$ using the former two equations for the heat exchange coefficient U_{PV} are hereafter denoted as *Model 8* and *Model 9*, respectively.

• Homer PV cell temperature model
Apart from the above-mentioned equations, another equation was used to determine the cell temperature $\left(T_{cell}^{10}\right)$ taken from (Duffie and Beckman, 1991; HOMER Pro. 2019) and is hereafter denoted as *Model 10*:

$$T_{cell}^{10} = \frac{T_{amb} + \left(T_{cell,NOCT} - T_{amb,NOCT}\right)\left(\dfrac{G_T}{G_{T,NOCT}}\right)\left[1 - \dfrac{\eta_{mp,STC}\left(1 - \alpha_p.T_{cell,STC}\right)}{\tau\alpha}\right]}{1 + (T_{cell,NOCT} - T_{amb,NOCT})(\dfrac{G_T}{G_{T,NOCT}})(\dfrac{\alpha_p.\eta_{mp,STC}}{\tau\alpha})} \tag{12}$$

where $T_{cell,NOCT}$ and $T_{cell,STC}$ are the cell temperature under NOCT and STC, respectively, $T_{amb,NOCT}$ is the ambient temperature at NOCT, G_T and $G_{T,NOCT}$ are the solar radiations striking the PV array and that value at NOCT [kW/m²], η_{mp} and $\eta_{mp,STC}$ are the efficiency of the PV array at its maximum power point in percentage and that value under STC [%], α_p the temperature coefficient of power [%/°C].

Thereafter, the different investigated models are used to estimate the cell temperature of the ASU PV system and, then, correlated with the PV output power to determine the most promising model to be used later in the analysis (Section 4.2). The numerical values and the application results are fully reported in Section 5.

4.1.2. Validating the obtained cell temperatures
A field trip to the ASU was carried out to validate the former findings and find the best model that represents the real values of PV cell temperature. A K-type infrared sensor was initially calibrated and then the readings of the cell temperature were taken at a five-minute interval for two hours. Due to the large number of PV cells available, the cells selected were random and the temperature of the module was measured at the top and bottom to get the average. For each interval, two modules were selected and the average was taken. The 24 results obtained were then compared to the theoretical value based on the physics-based models by calculating the *RMSE*) (Eq. (13)):

$$RMSE = \sqrt{\frac{\sum_{k=1}^{24}\left(\hat{T}_{cell}^k - T_{cell}^k\right)^2}{24}} \tag{13}$$

The lowest *RMSE* value indicates the goodness of the estimated cell temperature (hereafter denoted as T_{cell}^{best}) obtained by most realistic physics-based model among the ten investigated models.

4.2. Phase II: Building the Prediction Models

Two different prediction models are here developed and later evaluated in terms of their prediction performances of the ASU PV power

production (i.e., MLR (Abuella and Chowdhury, 2015) and ANNs (Hornik et al., 1989; Rumelhart et al., 1986) to study the influence of the cell temperature on the solar PV power production prediction.

The ambient temperature (T_{amb}) is replaced with the best obtained cell temperature $\left(T_{cell}^{best}\right)$ and the overall dataset X′ is established that will be used to build/develop the prediction models.

A problem arises when the data is directly used due to the presence of missing values, therefore the data are pre-proceeded as follows (Al-Dahidi et al., 2018):

1. There have been a few errors where the irradiation was measured with negative values during the late evening (6 p.m.-11 p.m.) and early morning (12 a.m.-6 a.m.). These errors were due to an offset in the measurement sensors that measure the irradiation values, and/or inverter failures. These values and their respective power values have been made zeros;
2. Missing data were also found in the data for the T_{amb}, I_{rr}, and power productions due to malfunctioning measurement sensors at the weather station, as well as failure in the inverters. These values have been excluded from the analysis;
3. The final step before being able to properly utilize the data is to normalize the values of time stamp, irradiation, temperature (whether ambient or cell), wind speed and power. These datasets are made to be in the range of [0-1]. The normalization formula is in the form of (Eq. (14)):

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (14)$$

where X, X_{max}, X_{min} are the actual, maximum, and minimum values of the considered variables to be normalized.

It is worth mentioning that the data patterns of the early morning and late evening of each day (i.e., power values available in these periods are zeros) have been used to train/develop the prediction models but, they have been excluded from the evaluation analysis of the prediction models' effectiveness (Section 4.3). This is because the PV system owner is not interested in predicting the power output of PV cells during the early morning or night with no solar irradiance. The two prediction models adopted in this work are hereafter presented (Sections 4.2.1 and 4.2.2).

4.2.1. MLR
The MLR employs a mechanism with which it attempts to model a relationship between the inputs (independent variables), i.e., time stamp and weather variables, with the output (dependent variable), i.e., PV power, by fitting a linear model as per Eq. (15). Each value from the independent variables is assigned to a value of the output. In the least-squares method, the best-fit line is calculated by reducing the sum of the squares of the vertical deviations from each data point to the line.

$$P = a_0 + a_1.hr + a_2.d + a_3.I_{rr} + a_4.T_{cell}(or\ T_{amb}) + \epsilon \qquad (15)$$

where P is the hourly PV power production, hr and d are the hour and day number time stamp parameters from the beginning of each

year data, I_{rr} and T_{cell} (or T_{amb}) are the hourly solar global radiation and cell (or ambient) temperature, a_0, a_1, a_2, a_3, a_4 are the regression coefficients, and $\in$ is the mismatch between the actual (true) and the predicted hourly PV power production of the PV system.
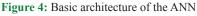
The Minitab (Minitab LLC. 2013) is used to define the optimal relation between the inputs and the output by estimating the regression model intercept and coefficients associated with each variable (Eq. 15). Afterwards, the best regression model function is used to predict the hourly PV power production values of the test dataset (X_{test}) based on the hourly inputs' values. The obtained results will be compared to the predictions obtained by the ANN prediction model.
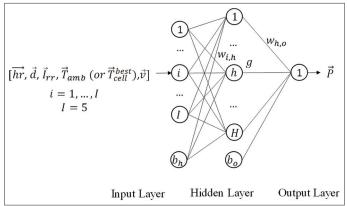
4.2.2. ANNs
A brief explanation will be given for the inner workings of the ANN to aid in understanding the how it works. ANN is a method used for computers to mimic the real world behavior and make it learn by itself. Even though a computer on its own is fast and reliably solves our tasks, but it does not have the capability of solving if the user does not know the problem, or if the data used is incomplete or random. The ANN aids the computer in this regard. ANN was first proposed in 1958 by a psychologist and was meant to see how a human recognized objects and interpreted visual stimuli (Hornik et al., 1989; Rumelhart et al., 1986).

Just as the human brains are connected by the means of neurons where the dendrites take information from other neurons whereas the axon shares the information, so does the ANN function (Hornik et al., 1989; Rumelhart et al., 1986). The ANN is split into three main categories: input layer, hidden layer, and output layer (Muhammad Ehsan et al., 2017). Figure 4 shows a very basic architecture of the ANN.

The schematic above serves to explain the mathematics behind the ANN. The input layer are the $I = 5$ inputs available in the training dataset (X_{train}) used to predict the output, these inputs could be just one or many depending on the application (i.e. in this work for the PV power production prediction, time stamp $([\vec{hr}\ \vec{d}])$, hourly global solar irradiation $\left(\vec{I}_{rr}\right)$, and hourly ambient temperature $\left(\vec{T}_{amb}\right)$ or hourly cell temperature $\left(\vec{T}_{cell}^{best}\right)$, and the

Figure 4: Basic architecture of the ANN

Input Layer   Hidden Layer   Output Layer

hourly wind speed ($\vec{v}$) are used as inputs, whereas the hourly power productions ($\vec{P}$) are used as outputs). Each i-th input is then connected to each h-th hidden neuron in the hidden layer ($h=1,…,H$) with a different weight ($w_{i,h,}$ $i=1,…,I,h=1,…H$). Initially the weights assigned to the connections are random and are changed with each iteration. A multiplication operation is performed such that the input value is multiplied to the weight given to that connection and added to an additional weight (hidden bias [b_h]) of the connection between the bias neuron and the corresponding hidden neuron, and then an addition operation gets carried out to add all the modified inputs that come to the neuron after they are multiplied with the weighted value. The hidden neurons are given an activation function g, which works by transforming the signal or the value coming from the input layer into another to be taken to the outer layer. Each activation function is more or less a graph where the value coming from the input layer is the x-value, and the value leaving the neuron is the respective y-value on the graph. Finally these values are sent to the output layer, multiplied with the weights of connections between the hidden neurons and the output neuron ($w_{h,o,}$ $h=1,…,H,o=1$), added to an additional weight (output bias [b_o]) of the connection between the bias neuron and the output neuron, ultimately all added together to give the final value typically via a linear activation function. This value is then checked with the actual power output and an error value is measured. From this value, the weights that were initially randomly assigned are readjusted and the process is repeated to get a more accurate result (i.e., the so called error Back-Propagation (BP) optimization algorithm) (Rumelhart et al., 1986).

In this work, different candidate numbers of the hidden neurons $h_{candidate}$ and different candidate hidden neuron activation functions $g_{candidate}$ are explored to establish an optimum version of the ANN architecture.

4.3. Phase III: Evaluating the Built-prediction Models
Once the prediction models are built using the training dataset (X_{train}), the prediction models are, then, evaluated on the test dataset (X_{test}), in terms of their prediction performances using two well-known standard performance metrics from the literature, they are (Al-Dahidi et al., 2018; Al-Dahidi et al., 2019):

- *RMSE* [kW] (Eq. (16)) that computes the deviation between the actual (true) and the predicted power productions obtained by the two prediction models. The model with the smallest *RMSE* value means that it is effectively capable of capturing the hidden (unknown) mathematical relationship between the inputs and the output and, thus, of predicting the PV power productions accurately, and vice versa.

$$RMSE = \sqrt{\frac{\sum_{j=1}^{N_{test}}\left(P_j - \hat{P}_j\right)^2}{N_{test}}} \qquad (16)$$

- Coefficient of Determination (R^2) [%] (Eq. (17)) that describes the variability in the outputs of the two prediction models caused by the considered inputs. A value of $R^2 = 100\%$ indicates that the variability in the prediction models' outputs

have been fully justified by the considered inputs used to build/develop the corresponding prediction models, and vice versa: lower R^2 values indicate that, in addition to the considered inputs, other variables need to be taken into account during the development of the prediction models to fully justify their prediction outcomes.

$$R^2 = \left[1 - \frac{\sum_{j=1}^{N_{test}}\left(P_j - \hat{P}_j\right)^2}{\sum_{j=1}^{N_{test}}\left(P_j - \bar{P}\right)^2}\right] \times 100\% \qquad (17)$$

where P_j and $\hat{P}$ are the j-th actual (true) and the predicted PV power production obtained by the two prediction models, $j = 1,…,N_{test}$, N_{test} is the overall test data patterns available in the test dataset (X_{test}), and $\bar{P}$ is the mean value of the obtained power production predictions.

The two considered metrics are calculated on the N_{test} test data patterns for the two prediction models and the obtained values are, then, compared to each other. Furthermore, the performance gain (PG_{Metric}) (Al-Dahidi et al., 2018; Al-Dahidi et al., 2019; Al-Dahidi et al., 2019) of each prediction model for the two cases, i.e., when the T_{amb} and the T_{cell}^{best} are being used, is calculated for the two metrics, using Eq. (18). It highlights the improvements achieved by the prediction models when the T_{cell}^{best} is being used instead of the T_{amb}, as well as it compares the predictability of the prediction models to each other.

$$PG_{Metric} = \left|\frac{Metric_{T_{amb}} - Metric_{T_{cell}^{best}}}{Metric_{T_{amb}}}\right| \times 100\% \qquad (18)$$

where $Metric_{Tamb}$ and $Metric_{T_{cell}^{best}}$ are the two considered performance metrics calculated for each prediction model when the T_{amb} and the T_{cell}^{best} are used in developing, optimizing, and evaluating the prediction models, respectively. Positive/negative values of the PG_{RMSE}/PG_R^2 indicate the benefits of exploiting the cell temperature instead of the ambient temperature, and vice versa.

5. RESULTS

In this Section, the application results of the proposed methodology of Section 4 (Figure 3) on the ASU case study of Section 3 are here presented step-by-step.

5.1. Phase I: Calculating and Validating the Cell Temperatures
5.1.1. Calculating the cell temperatures
The ten physics-based models investigated in this work are used to calculate the cell temperatures of the ASU solar PV system for the $Y\sim3.5$ years (i.e., 16 May 2015 to 31 December 2018) study period.

For p-Si modules of the ASU PV system under study, Table 1 reports the models' parameters values used to calculate the different cell temperatures (Duffie and Beckman, 1991; HOMER

Pro. 2019; Mattei et al., 2006; Schwingshackl et al., 2013; Skoplaki et al., 2008). The cell temperatures obtained by the ten models are denoted as T_{cell}^1 to T_{cell}^{10}.

Once the cell temperatures values are obtained, the correlations of these values with the PV power productions are calculated for each season of each year and for each year of the study period as shown in Figure 5 (top and bottom, respectively).

Looking at Figure 5, one can easily recognize that:
• The correlations vary with season showing highest and lowest values in summer and autumn seasons, respectively (Figure 5 [top]);
• The correlation values obtained by the ten different models can be grouped as follows (Figure 5 [bottom]):

• Correlation values > 0.85 obtained by *Model 10* (i.e., 0.868) and *Model 1* (i.e., 0.862);
• 0.85 > correlation values > 0.80 obtained by *Model 6* (i.e., 0.839) and *Model 7* (i.e., 0.814);
• Correlation values < 0.80 obtained by the remaining six models.
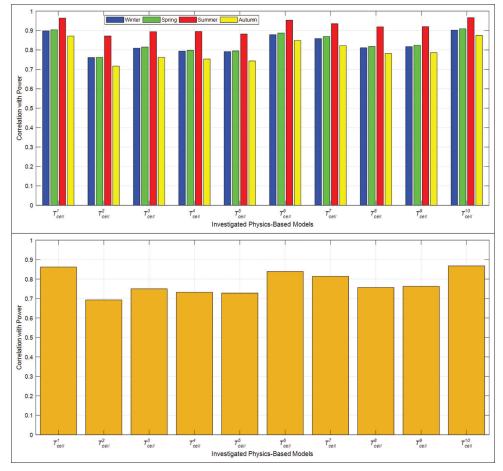
This variation can be justified by whether the wind speed (v_w) is considered in the physics-based models to calculate the cell temperatures or not (Section 4.1.1). Specifically:
• *Model 10* and *Model 1* do not incorporate the wind speed to calculate the cell temperatures;
• *Model 6* and *Model 7* directly incorporate the wind speed to calculate the cell temperatures;

Table 1: The models' parameters values for p-Si PV modules used in this work.

	Parameters values
Standard PV cell temperature *(Model 1)*	$T_{cell,NOCT}$= 46
Skoplaki PV cell temperature *(Models 2-5)*	η_{STC}=14.1%,β_{STC}= −0.45%/K,$\tau.\alpha$ = 0.9
Koehl PV cell temperature *(Model 7)*	U_0 = 30.02, U_1 = 6.28
Mattei PV cell temperature *(Models 8-9)*	η_{STC}=14.1%,β_{STC}= −0.45%/K,$\tau.\alpha$ = 0.81
Homer PV cell temperature *(Model 10)*	$T_{cell,STC} = 25^oC, \tau.\alpha = 0.9, G_{T,NOCT} = 0.8\dfrac{kW}{m^2}, T_{amb,NOCT} = 20^oC$

Figure 5: Correlation between the ten cell temperatures with the power productions for each season (top) and for each year (bottom)

- The remaining six models consider different formulations for the wind convection heat transfer coefficients (h_w) and the heat exchange coefficients for the total surface of the PV module (U_{PV}) to incorporate the wind speed in the calculations of the cell temperatures.

Considering the fact that the weather station is 171 m away from the ASU PV system under study, the available wind speed values might not be fully representative at the PV panels' locations and, thus, the inclusion of the wind speed in calculating the cell temperatures might lead to non-accurate cell temperatures (as we shall see in Section 5.1.2).

For clarification purposes for the importance of calculating the correlation values, Figure 6 shows the hourly global solar radiations (I_{rr}) (top left), ambient temperature (T_{amb}) (top middle), cell temperature obtained by the model that provides the highest correlation values with the power production $\left(T_{cell}^{10}\right)$ (i.e., 0.868 by Model 10) (top right), wind speed (v_w) (bottom left), and the corresponding power productions (P) (bottom right) for the four seasons in one arbitrary day.

Looking at Figure 6, one can notice that even though the irradiation was higher in Summer than in Spring, the power output in Summer was lower than that in Spring due to the higher ambient temperature in Summer with respect to that in Spring, and hence higher cell temperature. In addition, one can also recognize that the cell temperature $\left(T_{cell}^{10}\right)$ has a higher correlation to the power output than the ambient temperature $\left(T_{amb}\right)$.

5.1.2. Validating the obtained cell temperatures
For the 24 measured cell temperatures of the ASU PV system, the corresponding weather variables are recorded from the weather station at the ASU for the estimation of the PV cell temperatures by using the investigated ten models discussed earlier. These

variables were the solar irradiation, ambient temperature at 1 m, and wind speed at 10 m.

Finally, the *RMSE* value is computed for each method to display which model has more accurate results. From Figure 7 it can be inferred that T_{cell}^1 had the lowest *RMSE* (i.e., 2.834), and hence the best representation of the actual PV temperature $\left(T_{cell}^{best}\right)$. This temperature will be used to substitute the ambient temperature and establish the updated dataset X′.

5.2. Phase II: Building the Prediction Models
Once the updated dataset (X′) is established, it is used to build/develop the MLR and ANN prediction models.

5.2.1. Building the MLR Model
With respect to the MLR, the MLR model is built using the training dataset to provide the solar PV power production predictions. The obtained linear regression models using either the T_{amb} or the T_{cell} are given by Eq. (19) and Eq. (20), respectively. It is worth mentioning that the inclusion of the time stamp (i.e., the chronological order of the hour and day number) in the MLR would not be representative in this case. In fact, if one would manipulate the time stamp to be used in the MLR, it would be correlated (and thus, excluded) with the solar irradiation variable (i.e., I_{rr}). However, in this case, the results obtained show that the predictability of the solar power production does not significantly change, which indicates that the MLR cannot capture the hidden "apparently non-linear relationship" between the inputs and the power output.

$$P = -2.3564 + 0.1813.I_{rr} - 0.0078.T_{amb} + 0.731126.v \quad (19)$$

$$P = -2.3095 + 0.1849.I_{rr} - 0.0118.T_{cell} + 0.6347.v \quad (20)$$

In fact looking at Eq. (19) and Eq. (20) one can notice that:
- As the I_{rr} increases, the power production increases due to the increase of energy incident on PV system. This has been
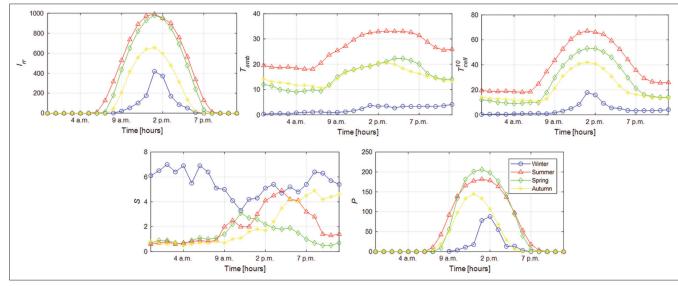
Figure 6: Irradiation (top left), ambient temperature (top middle), cell temperature (top right), wind speed (bottom left), and the corresponding power productions (bottom right) for the four seasons in one arbitrary day

effectively represented by the positive regression coefficient associated with the I_{rr} variable;

- As the T_{cell} (or T_{amb}) increases, the power production decreases due to the significant decrease in output voltage compared to marginal increase in output current (Al-Bashir et al., 2020; Ba et al., 2018). This has been effectively represented by the negative regression coefficient associated with the T_{cell} (or T_{amb}) variable;
- as the v increases, the power production increases due to the cooling of the PV panels, and hence, decreasing the cell temperature. This has been effectively represented by the positive regression coefficient associated with the v variable;

With respect to the ANN prediction model, the model is built (using the training dataset) and optimized (using the validation dataset) in the Matlab NN Toolbox™ (Demuth et al., 2009) in terms of number of hidden neurons, H and hidden neuron activation functions (g), to provide accurate solar PV power production predictions. Specifically, we follow an exhaustive search procedure by considering:

1. Twenty different numbers of hidden neurons that span the interval [2-40] with a step size of 2 for the ANN model development;
2. Twelve different activation functions, g = "*Log-Sigmoid*", "*Tan-Sigmoid*", "*Linear*", "*Triangular Basis*", "*Radial Basis*", "*Elliot Symmetric Sigmoid*", "*Symmetric hard-limit*", "*hard-limit*", "*Positive Linear*", "*Normalized Radial Basis*", "*Saturating linear*", and "*Symmetric Saturating Linear*" functions available in the Matlab NN Toolbox™ (Demuth et al., 2009);

The effectiveness of each ANN architecture established by a combination of the above-mentioned corresponding choices, is examined by quantifying the predictions accuracy of the validation dataset (X_{valid}), using the *RMSE* (Eq. 16) and R^2 (Eq. 17) performance metrics. Specifically, a 5-fold cross validation procedure is used to robustly evaluate the ANN prediction performance in terms of the *RMSE* and R^2: the training and validation patterns are sampled randomly from the inputs-output patterns available in the updated dataset (X') with fractions of 50% (i.e., N_{train} = 15115 patterns) and 25% (i.e., N_{valid} = 7557 patterns),

respectively. The cross validation procedure is then, repeated 5 times, using different patterns for training and validation datasets. The final metrics values are then, computed by averaging the 5 metrics' values of the 5 different trials.

Table 2 reports the modelling parameters of the optimum ANN architecture found at the smallest *RMSE* value, i.e., *RMSE* = 10.9784 kW (using the T_{cell}^{best}) and 11.0150 kW (using the T_{amb}), and largest R^2 value, i.e., R^2 = 96.8593 % (using the T_{cell}^{best}) and 96.8112 % (using the T_{amb}) on the validation dataset. For completeness, the obtained metrics found at H = 25 when the T_{amb} is used are *RMSE* = 11.2532 kW and R^2 = 96.7079 %. This assures the improvement obtained in the prediction accuracy when the T_{cell}^{best} is being used instead of the T_{amb}.

5.3. Phase III: Evaluating the Built-prediction Models

To demonstrate the effectiveness of replacing the T_{amb} with the best obtained cell temperature T_{cell}^{best} (i.e., the use of the updated dataset X′ which contains the T_{cell}^{best} instead of the original dataset X which contains the T_{amb} in developing the prediction models), Table 3 reports the average performance metrics obtained by the 5-fold cross validation using the prediction models for the case of using the T_{cell}^{best} instead of the T_{amb}, on the test dataset, together with the computed performance gains.
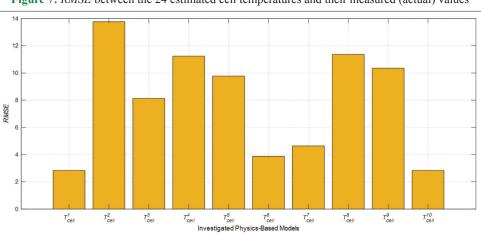
Looking at Table 3 one can easily recognize:
- A small improvement in the prediction accuracy is gained by the ANN prediction model when the T_{cell}^{best} is used instead of the T_{amb}. Specifically, an enhancement reaches up to ~1.93% and 0.11% on the *RMSE* and R^2 performance metrics, respectively. Despite the fact that these improvements in the

Table 2: The modelling parameters of the optimum ANN architecture obtained on the validation dataset

	H	g	RMSE [kW]	R² [%]
T_{amb}	20	"*Radial Basis*"	11.0150	96.8112
T_{cell}^{best}	25	"*Radial Basis*"	10.9784	96.8593

Figure 7: *RMSE* between the 24 estimated cell temperatures and their measured (actual) values

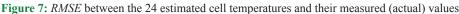**Table 3: Average performance metrics and gains obtained by the prediction models on the test dataset using either the T_{amb} or T_{cell}^{best}**

	MLR		ANN	
	T_{amb}	T_{cell}^{best}	T_{amb}	T_{cell}^{best}
RMSE [kW]	19.81	19.81	11.5867	11.3633
PG_{RMSE} [%]	0	0	0	1.9281
R^2 [%]	90.03	90.00	96.5227	96.6334
$PG_R{}^2$ [%]	0	0	0	−0.1147

prediction accuracy might seem to be relatively small, but they can be considered relevant to the PV system owner;
- The MLR model cannot exploit the benefits of using the T_{cell}^{best} instead of the T_{amb} and yield the same predictability results;
- The worse predictability obtained by the MLR compared to that obtained by the ANN is an indication of the existence of non-linearity relationship between the PV power production and the corresponding weather and time stamp variables that cannot be captured by the simple MLR, and nonlinear prediction models are required.

For completeness, possible sources of errors can be avoided in the future when using the PV cell temperatures to further enhance the obtained improvements in the prediction accuracy, they are:
- During the validation phase of the PV cell temperature models, the PV panels were measured randomly and due to the fact that the PV cells have different conditions (i.e., some are shaded, some have more access to wind, etc.), this could have led to taking a wrong model;
- The weather station is 171 m away from the faculty of engineering and, thus, the weather variables used in calculating the cell temperatures, such as the wind speed, could have contributed to the error.

6. CONCLUSIONS

In this work, the effect of the PV cell temperature for predicting power output has been investigated and the results have been compared to those when ambient temperature has been used instead. The case study was taken at the Applied Science Private University (ASU) where 264 kWp PV panels are installed on the roof of the faculty of engineering, Amman, Jordan. Two main prediction models have been utilized, namely MLR and ANNs. The best performance metrics (the RMSE and Coefficient of Determination) have been exhibited by the ANN and this is mainly due to its capability in capturing the hidden (unknown) nonlinear relationship between the inputs-output patterns of the ASU solar PV system. The results displayed the importance of adopting the PV cell temperature in the prediction models, however more research is needed to find a better way of measuring it.

Finally, some possible improvements for the future. Placing K-type infrared sensors on the back of the PV panels to constantly measure the temperature of the module. The

use of ANN may be limiting and, therefore, using another Machine Learning (ML) approaches such as Long Short-Term Memory (LSTM) may aid in power prediction and improve the performance metrics.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Abuella, M., Chowdhury, B. (2015), Solar power probabilistic forecasting by using multiple linear regression analysis. In: Conference Proceedings-IEEE SoutheastCon, IEEE. p1-5.

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

Al-Dahidi, S., Ayadi, O., Adeeb, J., Alrbai, M., Qawasmeh, R.B. (2018), Extreme learning machines for solar photovoltaic power predictions. Energies, 11(10), 1-18.

Al-Dahidi, S., Ayadi, O., Adeeb, J., Louzazni, M. (2019), Assessment of artificial neural networks learning algorithms and training datasets for solar photovoltaic power production prediction. Frontiers in Energy Research, 7, 1-18.

Al-Dahidi, S., Ayadi, O., Alrbai, M., Adeeb, J. (2019), Ensemble approach of optimized artificial neural networks for solar photovoltaic power prediction. IEEE Access, 7, 81741-81758.

Applied Science Private University (ASU). (2019), PV System ASU09: Faculty of Engineering. Available from: http://www.energy.asu.edu.jo. [Last accessed on 2019 Feb 17].

Ba, M., Ramenah, H., Tanougast, C. (2018), Forseeing energy photovoltaic output determination by a statistical model using real module temperature in the North East of France. Renewable Energy, 119, 935-948.

Baharin, K.A., Rahman, H.A., Hassan, M.Y., Gan, C.K. (2016), Short-term forecasting of solar photovoltaic output power for tropical climate using ground-based measurement data. Journal of Renewable and Sustainable Energy, 8(5), 053701.

Bouzerdoum, M., Mellit, A., Pavan, M.A. (2013), A hybrid model (SARIMA-SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant. Solar Energy, 98(PC), 226-235.

Brunet, C., Savadogo, O., Baptiste, P., Bouchard, M.A. (2018), Shedding some light on photovoltaic solar energy in Africa-a literature review. Renewable and Sustainable Energy Reviews, 96, 325-342.

Chow, S.K.H., Eric, W.M.L., Danny, H.W.L. (2012), Short-term prediction of photovoltaic energy generation by intelligent approach. Energy and Buildings, 55, 660-667.

Das, U.K., Tey, K.S., Seyedmahmoudian, M., Mekhilef, S., Idris, M.Y.I., Van Deventer, W., Horan, B., Stojcevski, A. (2018), Forecasting of photovoltaic power generation and model optimization: A review. Renewable and Sustainable Energy Reviews, 81, 912-928.

Demuth, H., Beale, M., Hagan, M. (2009), Neural Network Toolbox™ 6 User's Guide. Natick, Massachusetts, United States: The MathWorks, Inc.

Department of Statistics. (2017), Jordan in Figures 2017. Amman:

Department of Statistics.

Ding, M., Wang, L., Bi, R. (2011), An ANN-based approach for forecasting the power output of photovoltaic system. Procedia Environmental Sciences, 11, 1308-1315.

Duffie, J.A., Beckman, W.A. (1991), Solar Engineering of Thermal Processes. 2nd ed. Hoboken: John Wiley & Sons.

Ernst, B., Reyer, F., Vanzetta, J. (2009), Wind power and photovoltaic prediction tools for balancing and grid operation. In: CIGRE/IEEE PES Joint Symposium Integration of Wide-Scale Renewable Resources Into the Power Delivery System. Calgary, AB: IEEE. p1-9.

Fernandez-Jimenez, L.A., Muñoz-Jimenez, A., Falces, A., Mendoza-Villena, M., Garcia-Garrido, E., Lara-Santillan, P.M., Zorzano-Alba, E., Zorzano-Santamaria, P.J. (2012), Short-term power forecasting system for photovoltaic plants. Renewable Energy, 44, 311-317.

HOMER Pro. (2019), How HOMER Calculates the PV Cell Temperature. USA: HOMER Pro.

Hornik, K., Stinchcombe, M., White, H. (1989), Multilayer feedforward networks are universal approximators. Neural Networks, 2(5), 359-366.

Jaber, J.O., Badran, O.O., Abu-Shikhah, N. (2004), sustainable energy and environmental impact: Role of renewables as clean and secure source of energy for the 21st century in Jordan. Clean Technologies and Environmental Policy, 6(3), 174-186.

Kazem, H.A., Yousif, J.H. (2017), Comparison of prediction methods of photovoltaic power system production using a measured dataset. Energy Conversion and Management, 148, 1070-1781.

Koehl, M., Heck, M., Wiesmeier, S., Wirth, J. (2011), Modeling of the nominal operating cell temperature based on outdoor weathering. Solar Energy Materials and Solar Cells, 95(7), 1638-1646.

Kurtz, S., Whitfield, K., Miller, D., Joyce, J., Wohlgemuth, J., Kempe, M., Dhere, N., Bosco, N., Zgonena, T. (2009), Evaluation of high-temperature exposure of rack-mounted photovoltaic modules. In: 34th IEEE Photovoltaic Specialists Conference (PVSC). p2399-2404.

Lin, P., Peng, Z., Lai, Y., Cheng, S., Chen, Z., Wu, L. (2018), Short-term power prediction for photovoltaic power plants using a hybrid improved kmeans-GRA-elman model based on multivariate meteorological factors and historical power datasets. Energy Conversion and Management, 177, 704-717.

Liu, L., Liu, D., Sun, Q., Li, H., Wennersten, R. (2017), Forecasting power output of photovoltaic system using A BP network method. Energy Procedia, 142, 780-786.

Liu, L., Zhao, Y., Wang, Y., Sun, Q., Wennersten, R. (2019), A weight-varying ensemble method for short-term forecasting PV power output. Energy Procedia, 158, 661-668.

Loveday, D.L., Taki, A.H. (1996), Convective heat transfer coefficients at a plane surface on a full-scale building facade. International Journal of Heat and Mass Transfer, 39(8), 1729-1742.

Markvart, T., editor. (2000), Solar Electricity. 2nd ed. Hoboken: Wiley.

Mattei, M., Notton, G., Cristofari, C., Muselli, M., Poggi, P. (2006), Calculation of the polycrystalline pv module temperature using a simple method of energy balance. Renewable Energy, 31(4), 553-567.

Mellit, A., Shaari, S. (2009), Recurrent neural network-based forecasting of the daily electricity generation of a photovoltaic power system. In: Proceedings of the Ecological Vehicle and Renewable Energy (EVER'09). Monaco: Monte-Carlo. p265-270.

Ministry of Energy and Mineral Resources (MEMR). (2017), Annual Report 2017. Amman: Ministry of Energy and Mineral Resources.

Ministry of Planning and International Cooperation. (2015), Jordan 2025 a National Vision and Strategy. Amman: Ministry of Planning and International Cooperation.

Minitab LLC. (2013), Regression Analysis Tutorial and Examples. United States: Minitab LLC.

Moslehi, S., Reddy, T.A., Katipamula, S. (2018), Evaluation of data-driven models for predicting solar photovoltaics power output. Energy, 142, 1057-1065.

Muhammad Ehsan, R., Simon, S.P., Venkateswaran, P.R. (2017), Day-ahead forecasting of solar photovoltaic output power using multilayer perceptron. Neural Computing and Applications, 28, 3981-3992.

National Electric Power Company (NEPCO). (2018), Annual Report 2018. Amman: National Electric Power Company.

Oudjana, S.H., Hellal, A., Mahamed, I.H. (2012), Short term photovoltaic power generation forecasting using neural network. In: 2012 11th International Conference on Environment and Electrical Engineering. p706-711.

Paulescu, M., Brabec, M., Boata, R., Badescu, V. (2017), Structured, physically inspired (gray box) models versus black box modeling for forecasting the output power of photovoltaic plants. Energy, 121, 792-802.

Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986), Learning representations by back-propagating errors. Nature, 323(6088), 533-536.

Schwingshackl, C., Petitta, M., Wagner, J.E., Belluardo, G., Moser, D., Castelli, M., Zebisch, M., Tetzlaff, A. (2013), Wind effect on PV module temperature: Analysis of different techniques for an accurate estimation. Energy Procedia, 40, 77-86.

Sharples, S., Charlesworth, P.S. (1998), Full-scale measurements of wind-induced convective heat transfer from a roof-mounted flat plate solar collector. Solar Energy, 62(2), 69-77.

Shi, J., Lee, W., Liu, Y., Yang, Y., Wang, P. (2011), Forecasting power output of photovoltaic system based on weather classification and support vector machine. In: 2011 IEEE Industry Applications Society Annual Meeting. p1-6.

Skoplaki, E., Boudouvis, A.G., Palyvos, J.A. (2008), A simple correlation for the operating temperature of photovoltaic modules of arbitrary mounting. Solar Energy Materials and Solar Cells, 92(11), 1393-1402.

Yu, T., Chang, H. (2011), The forecast of the electrical energy generated by photovoltaic systems using neural network method. In: 2011 International Conference on Electric Information and Control Engineering. p2758-2761.

Zhong, J., Liu, L., Sun, Q., Wang, X. (2018), Prediction of photovoltaic power generation based on general regression and back propagation neural network. Energy Procedia, 152, 1224-1229.