A Local Training Strategy-Based Artificial Neural Network for Predicting the Power Production of Solar Photovoltaic Systems

SAMEER AL-DAHIDI, MOHAMED LOUZAZNI, AND NAHED OMRAN

1Department of Mechanical and Maintenance Engineering, School of Applied Technical Sciences, German Jordanian University, Amman 11180, Jordan
2National School of Applied Sciences, Abdelmalek Essaâdi University, Tetouan 93000, Morocco
3Renewable Energy Center, Applied Science Private University, Amman 11931, Jordan

Corresponding author: Sameer Al-Dahidi (sameer.aldahidi@gju.edu.jo)

ABSTRACT Power production prediction from Renewable Energy (RE) sources has been widely studied in the last decade. This is extremely important for utilities to counterpart electricity supply with consumer demands across centralized grid networks. In this context, we propose a local training strategy-based Artificial Neural Network (ANN) for predicting the power productions of solar Photovoltaic (PV) systems. Specifically, the timestamp, weather variables, and corresponding power productions collected locally at each hour interval $h = [1, 24]$ (i.e., an interval of $\Delta h = 1$ hour), are exploited to build, optimize, and evaluate $H = 24$ different ANNs for the 24 hourly solar PV production predictions. The proposed local training strategy-based ANN is expected to provide more accurate predictions with short computational times than those obtained by a single (i.e., $H = 1$) ANN model (hereafter called benchmark) built, optimized, and evaluated globally on the entire available dataset. The proposed strategy is applied to a case study regarding a 264kWp solar PV system located in Amman, Jordan, and its effectiveness compared to the benchmark is verified by resorting to different performance metrics from the literature. Further, its effectiveness is verified and compared when Extreme Learning Machines (ELMs) are adopted instead of the ANNs, and when the Persistence model is used. The prediction performance of the two training strategies-based ANN is also investigated and compared in terms of i) different weather conditions (i.e., seasons) experienced by the solar PV system under study and ii) different hour intervals (i.e., $\Delta h = 2, 3, \text{and } 4$ hours) used for partitioning the overall dataset and, thus, establishing the different ANNs (i.e., $H = 12, 8, \text{and } 6$ models, respectively).

INDEX TERMS Solar photovoltaic system, power prediction, local training strategy, artificial neural networks, real case study.

I. INTRODUCTION

A. BACKGROUND

Sustainable energy sources have shown a keen interest around the world for several important reasons, such as the downturn of fossil fuels, the rise in prices, industrial pollution, the energy crisis, and the surge of ecological concerns [1]–[3]. Renewable Energy (RE) generation has been strongly encouraged and supported by government policies and technology advancements [3]. In 2018, the share of RE (181 gigawatts) in cumulative production capacity worldwide had increased rapidly, contributing to more than 50% of the annual average power production potential added in the same year [4]. Besides, the energy produced by solar irradiation is considered the most promoting and safe energy supplied from the Photovoltaic (PV) systems [5]. PV energy sources are among the most extensively accessible and highly attractive RE sources according to their significant potential for energy production [6]–[9].

The main difficulty in the PV system is the complexity, parasitic capacitance, harmonic distortion, and sophistication of the equation of current-voltage and power-voltage characteristics [10]. The relationship among PV current and voltage is both implicit and complex depending on certain variables, among them are the ambient temperature, solar irradiation, wind speed, and dust accumulation [11], [12]. On hot days,
the cell module temperature can quickly be attained 70°C, where power energy output can drop significantly below nominal values [13]. The production of PV systems mainly depends on the amount of global solar irradiation received by the modules; any change in the power implies that solar irradiation changes during the day or infected by shading. In addition, wind speed can be a significant factor in the occurring of dust, dirt accumulation, and soiling in the PV system [14]. Such a phenomenon prevents the absorption of successful solar irradiance by PV cells and significantly reduces the overall PV power generation. This reduction in power can reach 50% in the arid and semiarid regions, where the solar irradiation is usually high [15]. Thus, the production from such energy sources depends on intermittent (stochastic) weather variables that calls for prediction (forecasting) models and tools capable of accurately estimating the PV power productions by accommodating such inherent stochasticity in the weather variables [16]–[18].

In this context, the prediction of PV power generation would make a significant contribution to the management and maintenance of modern energy systems, such as the connection to microgrids [19]. Prediction plays a critical role in managing the efficiency of the power system [7], [8], [20].

### B. LITERATURE REVIEW AND MOTIVATION

Numerous methods for predicting PV production have been published in the literature. Nevertheless, an effective method is still needed for enhancing the performance of PV prediction to decrease the adverse effects of system instability. However, the prediction methods are generally classified into model-based and data-driven [16], [21]–[23].

Model-based methods are based on analytical equations that focus on the PV power production concept. The equations typically use weather conditions to predict power output [23]. Usually, such methods do not require historical data, but they strongly depend on comprehensive station location details and reliable meteorological data. They could be easily defined and focused on solar irradiation, or they could be more complicated if additional weather variables, like ambient temperature, wind speed, and dust, are used. Thus, the effectiveness of their forecasts is heavily dependent on the precision of the Numerical Weather Predictions (NWP) details. Although such methods increase prediction accuracy, the uncertainty resulting from the approximations and/or assumptions in the adopted models could constitute a limitation on their realistic implementation [18].

Contrarily, data-driven methods (developed using various Machine Learning (ML) techniques) depend uniquely on the availability of the historical pairs of weather variables and the associated solar power productions. They aim to build models (called black-box models) to capture the hidden mathematical relationship between the weather variables and the associated PV power productions [4], [16]–[18].

For example, Ding et al., [24] proposed an improved version of the Back-Propagation (BP) learning algorithm-based Artificial Neural Network (ANN) to predict the power output of a PV system during different environmental conditions. The improved BP algorithm was shown superior compared to the traditional BP algorithm in enhancing the accuracy of the power output prediction.

Zeng and Qiao [25] designed the Radial Basis Function-based Neural Network (RBF-NN) for short-term solar PV power prediction using past values of meteorological data (e.g., sky cover, transmissivity). Results showed that the RBF-NN outperforms the linear autoregressive (AR) and the Local Linear Regression (LLR) models. The authors ended up that the use of transmissivity and other extra meteorological data, particularly the sky cover, could mainly improve the efficiency of the power prediction.

Li et al. [26] predicted the PV output power using the Auto Regressive Moving Average with exogenous inputs (ARMAX) and Auto-Regressive Integrated Moving Average (ARIMA). The two models used as exogenous inputs the ambient temperature, insolation duration, precipitation amount, and relative humidity to predict the power output of a 2.1 kW grid-connected PV system. Results revealed that the ARMAX model significantly enhances the predictability of the power output over the ARIMA.

De Leone et al. [27] used Support Vector Regression (SVR) to predict the energy production of a PV plant located in Italy. The method used the past meteorological data (e.g., solar radiation, ambient temperature) and power outputs to predict future power outputs. The obtained results revealed that the quality of the expected power output depends heavily on the accuracy of the meteorological data.

Yang et al. [28] predicted the PV power in the short-term using Auto-Regressive with exogenous input based Spatio-Temporal (ARX-ST). The evaluation of the results was compared to the conventional Persistence model. The authors addressed that the existing ARX-ST can be expanded with more meteorological data to help boosting the prediction precision.

Khademi et al. [29] proposed a Multi-Layer Perceptron equipped with an Artificial Bee Colony (MLP-ABC) algorithm to predict the power output of a 3.2kW PV plant. The collected data were separated into sunny and cloudy days and used to develop the MLP-ABC prediction model. The findings were compared to the MLP-ABC model when both sunny and cloudy days were used to establish the prediction model. It was concluded that the separation of different weather conditions enhanced the accuracy of the PV power output predictions.

Li et al. [30] used the Multivariate Adaptive Regression Splines (MARS) model for daily power output prediction of a grid-connected 2.1 kW PV system. This model maintains the flexibility of the traditional Multi-Linear Regression (MLR) paradigm; thus, having the ability to handle non-linearity. The obtained results using the MARS model were compared with linear models, such as MLR, ARIMA, and ARMAX, as well as some non-linear models, such as SVR, K-Nearest Neighbors (K-NN), and Classification and Regression Trees (CART). Results showed that non-linear models tend to
The authors concluded that no model could do consistently provide higher performance than linear models, on average. Muhammad Ehsan et al. [31] implemented an MLP-based ANN model for a 1-day ahead power output prediction of a 20 kWp grid-connected solar plant situated in India. Authors examined different combinations of hidden layers, hidden neuron activation functions, and learning algorithms for reliable 1-day ahead power predictions. The authors concluded that the ANN characterized by a single hidden layer, Linear Sigmoid Axon (neuron activation function), and Conjugate Gradient (learning algorithm) was able to deliver reliable power output predictions.

Theocharides et al. [32] examined the performance of three different ML methods, namely ANNs, SVR, and Regression Trees (RTs), with different hyper-parameters and sets of features, in predicting the power production of PV systems. Their success was related to the Persistence model throughout the computation of Mean Absolute Percentage Error (MAPE) and normalized Root Mean Square Error (nRMSE). The obtained enhancements were then evaluated using the Skill Score (SS). It was found that the ANNs outperform other prediction models from the literature.

Alomari et al. [33] proposed an ANN model for PV power production prediction. The proposed model investigated the strengths of two different learning algorithms (i.e., Levenberg-Marquardt (LM) and Bayesian Regularization (BR)), by utilizing different variations of ANN model’s inputs. The conclusions drawn revealed that an ANN-based BR provides more accurate predictions than those obtained by ANNs based LM (i.e., $RMSE = 0.0706$ and 0.0753, respectively).

Al-Dahidi et al. [16] investigated the capability of the Extreme Learning Machine (ELM) in predicting the PV power output. The obtained results revealed that the ELM provides better generalization capability with negligible computational times compared to the traditional BP-ANN.

Later, Al-Dahidi et al. [18] suggested a comprehensive ANN-based set-up solution for enhancing the 24h-ahead solar PV power output predictions. The authors also used the bootstrap technique to quantify the sources of ambiguity that influence the model structure predictions in the form of Prediction Intervals (PIs). The efficacy of the recommended ensemble solution was illustrated by a real case study of a solar PV system (264kWp capacity) located in Amman, Jordan. The suggested method has been shown to be advantageous to various standards in providing more accurate power predictions and accurately quantifying multiple sources of ambiguity.

Behera et al. [34] proposed a prediction technique based on a combination of ELM, Incremental Conductance (IC), and Maximum Power Point Tracking (MPPT) techniques. The obtained results revealed that the ELM provides better performance compared to the standard BP-ANN and that performance can be further enhanced using the PSO technique.

Huang and Kuo [35] proposed a high-precision PVPNet model-based Deep-Learning Neural Networks (DLNNs) for 1-day ahead power output prediction. The prediction results obtained by the proposed PVPNet model were evaluated (in terms of RMSE and MAE) and compared to other ML techniques of literature. Authors concluded that the proposed PVPNet model has an excellent generalization capability and can boost the prediction performance, while reducing monitoring expenses, initial costs of hardware components, and long-term maintenance costs of future PV plants.

Catalina et al. [36] proposed two linear ML models (i.e., Least Absolute Shrinkage and Selection Operator (LASSO) and linear SVR), and two non-linear ML models (i.e., MLPs and Gaussian SVRs) with satellite-measured radiances and clear sky irradiance as inputs to nowcast the PV energy outputs over peninsular Spain. Results revealed that the two non-linear ML models were better than the two linear ML models.

From the above research works, it is apparent that the efforts were mainly dedicated to enhancing the employed data-driven prediction model or investigating other advanced models from the literature. Differently, this work aims to propose a local training strategy applicable to any data-driven prediction model for ultimately boosting the prediction accuracy of the solar PV power outputs, while reducing the computational times. Specifically, the hour-by-hour variability (i.e., the 24-hour seasonality patterns of each day) arise in the solar data, of both the weather variables and the corresponding power productions, has never been explored while developing the prediction models. The consideration of such seasonality while developing the prediction models is expected to be beneficial in enhancing the prediction accuracy while reducing the computational times.

C. CONTRIBUTIONS

The proposed training strategy requires splitting the available inputs-output patterns collected from the actual operation of a PV system based on an hour interval of $\Delta h = 1$, into $H = 24$ datasets, each dataset represents the data collected at each hour interval $h$, $h \in [1, 24]$. The established datasets are then used to build $H = 24$ feedforward ANNs models. The selection of the ANNs is driven by the fact that they are simple, easy to understand and to implement, and capable of solving non-linear interpolation problems [37], [38].

Each built-ANN is initially optimized on a validation dataset in terms of the number of hidden neurons to enhance the prediction accuracy further and, then, utilized online to estimate the corresponding hourly production of a day on a test “unseen” dataset.

The effectiveness of the proposed training strategy-based ANN is examined on a grid-connected solar PV system (264kWp capacity) located in the Applied Science Private University (ASU), Amman, Jordan [4], [16], [18], [39]. Specifically, the accuracy of the power production predictions and the computational times required to develop and evaluate the built-ANNs are verified by resorting to three performance
metrics from the literature [4], i.e., the RMSE, the MAE, and the Weighted MAE (WMAE), and to the computational time in minutes, respectively.

For comparison and validation, a single prediction model (i.e., for a fair comparison, an ANN model is considered) developed and optimized, in terms of the number of hidden neurons, globally on the entire dataset is used as a benchmark to verify the effectiveness of the proposed strategy on the ASU solar PV system. Moreover, the ELMs are used instead of the ANNs, and the Persistence prediction model is adopted further to verify the superiority of the proposed training strategy-based ANN.

Therefore, the significant contributions of the present work are two-fold:

- The development of a local training strategy-based ANN for an accurate estimation of the solar PV power productions with short computational times;
- The comparison of the obtained results to the global and local training strategies-based ANN and ELM, respectively, as well as to the Persistence prediction model of literature, to further explore the effectiveness of the proposed local training strategy.

The remaining of this article is organized as follows. In Section II, the work objectives are illustrated, and the problem of predicting the solar PV power output is stated. In Section III, the ASU solar PV system case study is described, and the proposed local training strategy-based ANN is illustrated, also providing an essential background of ANN. In Section IV, the application of the proposed training strategy-based ANN to the ASU case study is shown, and the obtained results are discussed and compared with those obtained by the global and local training strategies-based ANN and ELM, respectively, as well as to the Persistence model of literature. Section V investigates the influence of using different hour intervals on the prediction performance. Lastly, conclusions are drawn, and future works are recommended in Section VI.

### II. WORK OBJECTIVES

This work aims to develop a data-driven model for an accurate estimation of the power productions of a solar Photovoltaic (PV) system with convenient computational times. We consider the availability of historical weather data (W) and corresponding power production data (P) of the PV system collected during Y years (or D days) (see Fig. 1). The weather data (W) consist of four weather variables collected at hour h, h ∈ [1, 24] of each day during the period Y; they are:

- wind speed (S<sub>h</sub>);
- relative humidity (RH<sub>h</sub>);
- ambient temperature (T<sub>amb</sub>);
- global solar irradiation (I<sub>rr</sub>).

Together with the timestamp and the corresponding power data (P<sub>h</sub>), one can establish an overall inputs-output dataset from the previous data vectors:

\[ X = \left[ \overrightarrow{hr} \overrightarrow{d} \overrightarrow{S RH T_{amb} T_{rr}} \right] \]  

(1)

The timestamp is here represented by the chronological hour (hr) and day number (d) from the beginning of each year data during the period Y, hr = 1, . . . , 8760, d = 1, . . . , 365.

We aim to effectively exploit the pre-processed dataset during the development/training stage of a data-driven prediction model for providing accurate predictions of the PV power production, with convenient computational times. To this aim, the available dataset is divided into H = 24 datasets. Each dataset is composed of the timestamp and the corresponding weather variables and power productions collected locally at each h-i-th hour of a day, h ∈ [1, 24]. The datasets are used to develop and optimize H = 24 data-driven prediction models. Feedforward Artificial Neural Networks (ANNs) are employed as prediction models due to their simplicity and convenient computational efforts required [40]. Still, the proposed training strategy is general and could be applied to any data-driven ML techniques from the literature (e.g., ELMs, SVMs, etc.). Finally, the built-prediction models are individually used to predict the h-th power production of the solar PV system. Thus, our contribution entails proposing an intuitive way of handling the available dataset to accurately build/develop a data-driven prediction model, such as the ANN, in this work.

The proposed local training strategy is expected to provide more accurate power production predictions with short computational efforts, compared to the traditional global training strategy, in which the available dataset is used, entirely,
to develop and optimize a single \( (H = 1) \) ANN prediction model. Additionally, further comparisons and analyses are carried out to explore the effectiveness of the proposed training strategy-based ANN.

### III. MATERIAL AND METHODOLOGY

In this Section, the real data of a solar PV system used in this work, together with the proposed (local) and benchmark (global) training strategies-based ANN, are presented in Section III.A and Section III.B, respectively.

#### A. MATERIAL

This Section shows the real case study of a solar PV system of a capacity of 264kWp mounted on the rooftop of the Faculty of Engineering (Al-Khawarizmi Building) at Applied Science Private University (ASU) in Shafa Badran, Amman, Jordan (Latitude = 32.042044 and Longitude = 35.900232) (Fig. 2).

![ASU PV system map](image1)

The dataset utilized in this work comprises real weather data, \( W \), (i.e., inputs) (measured by a weather station located far away by around 172 m from Al-Khawarizmi Building (Fig. 2) and the corresponding PV power productions (i.e., output) (measured by the inverters of the PV system), \( P \) (in kW) [39]. This dataset has been collected for \( Y = 3.625 \) years (from 16th May, 2015 to 31st December, 2018) with a time step \( \Delta t = 1 \) hour, from 12 a.m. to 11 p.m. daily, i.e., \( D = 1326 \) days with \( N = 31824 \) inputs-output patterns.

1) ASU SOLAR PV SYSTEM

The ASU PV system comprises 14 SMA sunny tri-power inverters (13 inverters with a power of 17kW and 1 inverter with a power of 10kW) attached with Yingli Solar panels (of a type YL 245P-29b-PC) tilted by 11° and oriented 36° (from S to E). This orientation is chosen to accumulate as much optimum solar radiation as possible during the day (as depicted in Fig. 3 [39]).

The design characteristics of the ASU PV system are reported in Table 1.

2) ASU WEATHER STATION

The ASU weather station (depicted in Fig. 4) is 36m high equipped with latest instruments used to measure 45 different weather variables, such as global solar irradiations, relative humidity, precipitation amounts, wind speeds and directions, barometric pressure, and ambient temperatures collected at various levels from the ground.

Among the 45 weather variables, engineering and professional opinion suggested to use the highly correlated weather variables to the PV power productions as inputs to the prediction model. Those variables, together with the instruments installed for their measurements and their detailed characteristics are [42]:

- the wind speed at 10m (\( S \)) (m/s). The wind speed transmitter is used for measuring the horizontal component of the wind speed with high accuracy. The transmitter is equipped with an electronically regulated heating for providing a smooth running of ball bearings during winter operations, and for avoiding the shaft and slot from icing-up. The technical specifications of this instrument are reported in Table 2;
- the relative humidity at 1m (\( RH \)) (%). The hygro-thermo transmitter with capacities sensing element is used for measuring the relative humidity with high accuracy. The transmitter is equipped with weather and thermal radiation shield to protect the humidity sensor against radiation, precipitation, and mechanical damages.

![FIGURE 3. ASU 264 kWp PV panels installed at the rooftop of the Faculty of Engineering (left) and the connected inverters (right).](image2)

**FIGURE 2. ASU PV system map (retrieved and adapted from Google Maps [41]).**

**FIGURE 3.** ASU 264 kWp PV panels installed at the rooftop of the Faculty of Engineering (left) and the connected inverters (right).

**TABLE 1.** The design characteristics of the ASU PV system.

| Parameter                              | Numerical values
|----------------------------------------|-------------------|
| Power output, \( P_{max} \) (Watt)      | 245               |
| Maximum voltage, \( V_{max} \) (V)      | 30.2              |
| Maximum current, \( I_{max} \) (A)      | 8.11              |
| Open circuit voltage, \( V_{OC} \) (V)  | 37.8              |
| Short circuit current, \( I_{SC} \) (A) | 8.63              |

| Electrical parameter at NOCT | Numerical values |
|-----------------------------|------------------|
| Power output, \( P_{max} \) (Watt) | 177.9 |
| Maximum voltage, \( V_{max} \) (V) | 27.2 |
| Maximum current, \( I_{max} \) (A) | 6.54 |
| Open circuit voltage, \( V_{OC} \) (V) | 34.5 |
| Short circuit current, \( I_{SC} \) (A) | 6.99 |

**Thermal characteristics**

| Temperature coefficient of \( P_{max} \) (\%\(^{\circ}\)C) | -0.455\(^{\circ}\)C |
| Temperature coefficient of \( V_{max} \) (\%\(^{\circ}\)C) | -0.33\(^{\circ}\)C |
| Temperature coefficient of \( I_{max} \) (\%\(^{\circ}\)C) | 0.065\(^{\circ}\)C |

The technical specifications of this instrument are reported in Table 2:

- the ambient temperature at 1m ($T_{amb}$) (in °C). The hygro-thermo transmitter with RTD is used for measuring the temperature of ambient air with high accuracy. The transmitter is equipped with weather and thermal radiation shield to protect the temperature sensor against radiation, precipitation and mechanical damages. The technical specifications of this instrument are reported in Table 2; and
- the global solar irradiations ($I_{rr}$) (in W/m²). The pyranometers are used for measuring the global (total) irradiation on a plane surface with high accuracy. The technical specifications of this instrument are reported in Table 2.

Additionally, the corresponding timestamps (i.e., number of hours and day from the beginning of each year data) are also considered in inputs [4], [16], [18], [33], [43]. The remaining variables have been excluded from the analysis.

3) DATA PRE-PROCESSING
For proper utilization of the dataset, it has been pre-processed following the guidelines reported in [4], [16], [18], [33], [43] using the same case study. For example:

- missing values have been excluded from the analysis;
- negative $I_{rr}$ values and corresponding missing $P$ values (recognized in the early morning (i.e., 12 a.m. – 6 a.m.) and late evening (i.e., 6 p.m. – 11 p.m.)) have been set to zero; and
- the overall data have been normalized between 0 and 1.

Fig. 5 shows the pre-processed hourly $I_{rr}$ values (left-top), $T_{amb}$ values at 1m (right-top), $RH$ values (left-middle), $S$ values at 10m (right-middle), and the corresponding solar $P$ values (bottom) of few days selected throughout the 2017-year data. From Fig. 5, one can notice that:

a) the different variability of the reported parameters’ values collected at each $h$-th hour interval of 1 hour, $h \in [1, 24]$, of the days. This is highly necessary for a notification here to justify the motivation of using the collected $h$-th parameters’ values solely in developing $H = 24$ different ANN prediction models. Each model is dedicated to estimating the corresponding $h$-th power production value;

b) the extensive and random variability in the reported parameters’ values collected at the entire hours of the days. It is essential to mention here that, typically, an ANN prediction model built on the whole dataset is expected to estimate the PV power productions at each $h$-th hour interval less accurately, $h \in [1, 24]$ compare to a).

The pre-processed $Y = 3.625$ years’ inputs-output patterns are appended in an overall matrix $X$. This matrix will be used later on for the development of the ANN and ELM prediction models using both the proposed (local) and the benchmark (global) training strategies, as well as for the implementation of the Persistence prediction model.

B. METHODOLOGY
In this Section, the proposed local and the benchmark global training strategies-based ANN performed with a MATLAB code that has been developed in-house are presented in Section B.1 and Section B.2, respectively.

1) THE PROPOSED LOCAL TRAINING STRATEGY-BASED ANN
The proposed training strategy-based ANN is sketched in Fig. 6, and it goes along the following four steps:
Step 1 (Establishing $H = 24$ Different Datasets): This step entails partitioning the overall available pre-processed dataset ($X$) into $H = 24$ different datasets ($X_h$) of equal size. Each dataset represents the timestamp ($h r_h$, $d_h$), weather variables ($S_h$, $RH_h$, $T_{amb} h$, $I_{rr} h$), and the corresponding power productions ($P h$) collected at each hour interval $h$ (i.e., $h \in [1, 24/\Delta h]$), using an hour interval $\Delta h = 1$ hour) during the period $Y = 3.625$ years.
Thus, the $H = 24$ different datasets can be written as follows:

$$X_h = \left[ h \left[ \bar{h}_h \bar{d}_h \bar{S}_h \bar{R}_{\text{RH},h} \bar{T}_{\text{amb},h} \bar{T}_{\text{rr},h} \right] \right], \quad h \in [1, 24]. \quad (2)$$

This partitioning is motivated by the fact that the input-output patterns collected at different hours and different days are mostly variable. Thus, utilizing solely the data collected at the $h$-th hour, for the estimation of the corresponding $h$-th power production, in the development of the ANNs prediction models, is expected to produce more accurate power production predictions. Additionally, this partitioning will, indeed, reduce the computational efforts required by the ANNs prediction models in capturing the hidden “unknown” relationship between the inputs and the output power.

**Step 2 (Train the $H = 24$ ANNs Prediction Models With Different Configurations):** Once the $H = 24$ different datasets are established, training ($X_{\text{train}}^h$), validation ($X_{\text{valid}}^h$), and test ($X_{\text{test}}^h$) datasets are extracted randomly from the established datasets $X_h$, $h \in [1, 24]$, with arbitrary fractions of $\gamma = 50\%$, $\beta = 20\%$, and $\alpha = 30\%$, respectively. Such datasets are used to train, validate (Step 3), and test (Step 4) the $H = 24$ different feedforward ANNs, respectively, whose detailed characteristics are hereafter described. Specifically:

- the training datasets ($X_{\text{train}}^h$, $h \in [1, 24]$), each is formed by $N_{\text{train}} = 663$ patterns (i.e., 663 days), are used for building/training the $H = 24$ ANNs prediction models;
- the test datasets ($X_{\text{test}}^h$, $h \in [1, 24]$), each is formed by $N_{\text{test}} = 398$ patterns (i.e., 398 days), are used to evaluate the performance of the optimum ANNs prediction models. It is, indeed, important to mention that $N_{\text{test}}$ test patterns have never been used during the building/training and the optimization of the ANNs.

In other words, the extracted datasets $X_h$, $h \in [1, 24]$, obtained in Step 1 are divided into three disjoint datasets, namely, training ($X_{\text{train}}^h$), validation ($X_{\text{valid}}^h$), and test ($X_{\text{test}}^h$), by sampling their patterns randomly from the patterns of the corresponding datasets $X_h$, $h \in [1, 24]$, with arbitrary fractions of $\gamma = 50\%$, $\beta = 20\%$, and $\alpha = 30\%$, respectively. The motivation of using such fractions is to assure that the annual seasonality appears in the datasets has been sufficiently captured while developing the ANNs prediction models.

To further assure that while sampling randomly, a Cross-Validation (CV) procedure is being employed, as we shall see in Step 3. Thus, other arbitrary fractions can be considered, and the obtained conclusions would be, indeed, the same.

Each $h$-th ANNs prediction model ($ANN_h$, $h \in [1, 24]$) comprises three main layers, as shown in Fig. 7:

- Input layer. It receives the $j$-th pattern, $x_j^h$, that comprises the $h$-th six inputs (i.e., $[h \bar{d}_h \bar{S}_h \bar{R}_{\text{RH},h} \bar{T}_{\text{amb},h} \bar{T}_{\text{rr},h}]$), $j = 1, \ldots, N_{\text{train}}$;
- Hidden layer. It comprises $N_h$ hidden neurons that used to process the received inputs via a hidden neuron activation function, $f_i()$, and send the processed information to the output layer. In practice, the hidden neuron activation

![FIGURE 7. The basic architecture of the ANN model.](image-url)

$X_j^h$, $j = 1, \ldots, N_{\text{train}}^h$, $h \in [1, 24]$
function is a continuous non-polynomial function (e.g., "Log-Sigmoid", "Linear", "Radial Basis", etc.) established to capture the mathematical hidden non-linear “unknown” relationship between the inputs and the outputs;

- Output layer. It provides an estimation of the corresponding h-th power production \( \hat{P}_h \) via an output neuron activation function, \( f_2() \), which is typically a linear transfer function ("Purlin") [4], [44]. The estimated h-th power production \( \hat{P}_h \) of the j-th input pattern, \( x_{h}^j \), can be written as follows:

\[
\hat{P}_h = f_2 \left( \sum_{i=1}^{N_{\text{train}}} \beta_{j}^{i} (\bar{w}_{j}^{i} x_{h}^i + b_{j}) + b_{\text{v}} \right),
\]

\( h \in [1, 24], \quad j = 1, \ldots, N_{\text{train}} \) (3)

where \( n \) and \( j \) are the indexes of the number of hidden neurons (\( n_{\text{candidate}} \)) and the number of available training patterns (\( j = 1, \ldots, N_{\text{train}} \)), respectively. \( b_{j} \) and \( b_{\text{v}} \) are the weights (biases) of the connections established between the bias neurons and each \( n \)-th hidden neuron and output neuron, respectively. \( \bar{w}_{j} \) and \( \beta_{j}^{i} \) are the weights vectors of the connections established between the input and output neurons with each \( n \)-th hidden neuron, respectively.

To adequately define the ANNs configurations, different candidate numbers of hidden neurons (\( n_{\text{candidate}} \)) are explored. The \( n_{\text{candidate}} \) is considered to span the interval [5-30] with a step size of 5. For each possible configuration, the h-th power production (whose actual value is \( P_h \)) is estimated \( \hat{P}_h \) and the Levenberg-Marquardt (LM) error BP learning algorithm is adopted to minimize the mismatch (typically by calculating the Mean Square Error (MSE)) between actual and estimated power productions by exploring different random initializations of the ANN internal parameters (i.e., \( \beta_{j}^{i}, \bar{w}_{j}^{i}, b_{j}, b_{\text{v}} \)). The built-ANNs (\( \text{ANN}_{\text{opt}}^{\text{h}}, h \in [1, 24] \)) are those whose internal parameters are optimally selected to minimize the MSE on the \( N_{\text{train}} \) training inputs-output patterns.

Step 3 (Validate the Built-ANNs With Different Configurations): Once the \( H = 24 \) different ANNs are built with the different possible configurations, the best models for each h-th hour (\( \text{ANN}_{\text{opt}}^{\text{h}}, h \in [1, 24] \)) are selected by evaluating their prediction performances on the validation dataset (\( \text{X}_{\text{valid}}^{h} \)). To this aim, three standard performance metrics from the literature are computed. The considered performance metrics are, then, calculated by computing the average error between the actual (true) and estimated power productions produced by the built-ANNs models. Small \( \text{MAE} \) values indicate that the predictions are accurate, and vice versa;

\[
\text{MAE}_{h} = \frac{1}{N_{\text{valid}}} \sum_{j=1}^{N_{\text{valid}}} \left| P_h^j - \hat{P}_h^j \right| (5)
\]

- **Weighted Mean Absolute Error (WMAE)** (Eq. (6)). This metric computes the average relative error between the actual (true) and estimated power productions produced by the built-ANNs models. Similar to the previous two metrics, small \( \text{WMAE} \) values indicate that the predictions are accurate, and vice versa. In practice, this metric is of interest to compare the prediction accuracy when the production capacities are changing;

\[
\text{WMAE}_{h} = \frac{1}{\sum_{j=1}^{N_{\text{valid}}} P_h^j} \sum_{j=1}^{N_{\text{valid}}} \left| P_h^j - \hat{P}_h^j \right| (6)
\]

where \( \text{RMSE}_{h}, \text{MAE}_{h}, \) and \( \text{WMAE}_{h} \) are the performance metrics computed for each h-th built-ANN prediction model, and \( \hat{P}_h \) is the average actual production collected at the h-th hour.

A 100-fold CV procedure is employed to evaluate the different ANNs with the various possible configurations robustly. It entails establishing the datasets 100 different times, each with the same fractions of \( \gamma = 50\%, \beta = 20\%, \) and \( \alpha = 30\% \) for the training, validation, and test datasets, respectively. The simulations are then repeated 100 times, and the performance metrics are evaluated 100 times. The ultimate performance metrics are, then, calculated by computing the average and standard deviation values and the optimum number of hidden neurons (\( N_{\text{opt}}^{h} \)) are reported for each h-th ANN prediction model. For each h-th hour, the best ANNs are selected to be those that minimizing the multiplied metric of the \( \text{RMSE}_{h}, \text{MAE}_{h}, \text{WMAE}_{h} \), i.e., \( \text{RMSE}_{h} \ast \text{MAE}_{h} \ast \text{WMAE}_{h} \).

**Step 4 (Test the Optimum ANNs):** The optimum ANNs whose configurations are the best selected among the whole possible different combinations as reported in Step 3 and obtained on the validation dataset are to be evaluated on the unseen test datasets (\( \text{X}_{\text{test}}^{h} \)). The predictability of the optimum ANNs is evaluated using the above-mentioned performance metrics, and their average and standard deviation results calculated over the 100 CV trials are reported.

2) THE BENCHMARK GLOBAL TRAINING STRATEGY-BASED ANN

The benchmark global training strategy-based ANN entails building/training (using a training dataset), optimizing (using a validation dataset), and evaluating (using a test dataset) an \( H = 1 \) single ANN prediction model. To this aim, the overall available pre-processed dataset \( \text{X} \) is divided into the following datasets:

- the training dataset (\( \text{X}_{\text{train}} \)). It is formed by \( N_{\text{train}} = 15912 \) patterns (collected from 663 days, each comprises
24 inputs-output patterns, thus establishing $24 \times 366 = 15912$ patterns. The $X_{train}$ is used for building/training the single $H = 1$ ANN model;

- the validation dataset ($X_{valid}$). It is formed by $N_{valid} = 6360$ patterns (collected from 265 days, each comprises 24 inputs-output patterns, thus establishing $24 \times 265 = 6360$ patterns). The $X_{valid}$ is used for optimizing the configuration of the single $H = 1$ ANN model in terms of number of hidden neurons;

- the test dataset ($X_{test}$). It is formed by $N_{test} = 9552$ remaining patterns (collected from 398 days, each comprises 24 inputs-output patterns, thus establishing $24 \times 398 = 9552$ patterns). The $X_{test}$ is used to evaluate the performance of the optimum single $H = 1$ ANN model.

In other words, the benchmark global training strategy aims at exploiting the complete inputs-output patterns collected at the different hour intervals of different days for the development of the single ANN prediction model (thus called “global” training strategy).

For a fair comparison with the proposed approach, the 663, 265, and 398 days considered here in each simulation (CV) trial are the same as those that have been obtained (using the arbitrary fractions $\gamma = 50\%$, $\beta = 50\%$, and $\alpha = 50\%$) and used in the proposed training strategy for establishing the training, validation, and test datasets, respectively. The simulations are then repeated 100 times, and the ultimate average performance metrics and their standard deviations are reported and compared with those obtained by the proposed training strategy (Section IV).

IV. APPLICATION RESULTS

The application results of the proposed local training strategy-based ANN (Section III.B.1) on the ASU real case study (Section III.A) are here described and compared with the results obtained by the benchmark (Section III.B.2). Further, its effectiveness is verified and compared when Extreme Learning Machines (ELMs) are adopted instead of the ANNs, and when the Persistence model is used.

Because the early morning and late evening hours are with no need to estimate the power productions since there are no solar radiations, the reported results are shown solely for $h \in [7, 21]$.

A. APPLICATION RESULTS OF THE PROPOSED LOCAL TRAINING STRATEGY-BASED ANN

The optimum configurations obtained for each $h$-th optimum ANN ($ANN_{h}^{opt}$; $h \in [7, 21]$) on the validation datasets ($X_{h}^{valid}$) in terms of the number of hidden neurons ($N_{h}$) are reported in Table 3, together with the best obtained average performance metrics. The results are obtained by using the “Radial basis” and “Purlin” hidden/output neuron activation functions ($f_{1}(\cdot)$ and $f_{2}(\cdot)$, respectively) and the 100-fold CV procedure. The “Radial basis” takes in input values between $(-\infty, \infty)$ and produces in output values between (0,1) [44], [45], whereas “Purlin” transfers the inputs to the outputs without any change [44], [45].

It is worth mentioning that one could follow an exhaustive search procedure to select the best-hidden/output neuron activation functions (i.e., $f_{1}(\cdot)$ and $f_{2}(\cdot)$, respectively) among the other available activation functions in the literature [44]. However, to reduce the complexity of the optimization step, the “Radial basis” and “Purlin” functions are selected in this work for the different ANN models, following the recommendations and guidelines reported in [16] on the same dataset. The number of hidden neurons for each $h$-th optimum ANN ($N_{h}^{opt}$) are selected at which the overall multiplied performance metrics is minimized (i.e., $\min(RMSE_{h} \ast MAE_{h} \ast WMAE_{h})$).

Notice that:

- The optimum number of hidden neurons obtained for each $h$-th ANN model is, in general, shown to be proportional to the level of variability exhibited by the data collected at the corresponding hour interval, $h \in [7, 21]$. For example, at $h = 7$ and $h = 12$ (small and large data variability, respectively, as depicted in Fig. 5), small and large numbers of hidden neurons ($N_{h}^{opt} = 5$ and $N_{h}^{opt} = 20$, respectively) are selected instead of the other candidate numbers, $n_{candidate}$, to accurately represent the inputs-output relationship;

- the prediction performance obtained for each $h$-th ANN model is, in general, shown to be proportional to the level of variability exhibited by the data collected at the corresponding hour interval, $h \in [7, 21]$. For example, at $h = 7$ and $h = 12$ (small and large data variability, respectively), small and large performance metrics are obtained, respectively.

| $h$ | $N_{h}^{opt}$ | $RMSE_{h}$ [$\text{[kW]}$] | $MAE_{h}$ [$\text{[kW]}$] | $WMAE_{h}$ |
|-----|-------------|----------------|----------------|--------|
| $h = 7$ | $5$ | $0.064$ | $0.026$ | $0.471$ |
| $h = 8$ | $10$ | $1.507$ | $0.956$ | $0.274$ |
| $h = 9$ | $10$ | $5.685$ | $3.908$ | $0.192$ |
| $h = 10$ | $10$ | $12.248$ | $8.507$ | $0.150$ |
| $h = 11$ | $10$ | $17.482$ | $12.147$ | $0.122$ |
| $h = 12$ | $20$ | $18.213$ | $12.122$ | $0.092$ |
| $h = 13$ | $15$ | $17.660$ | $11.636$ | $0.077$ |
| $h = 14$ | $5$ | $16.477$ | $10.429$ | $0.068$ |
| $h = 15$ | $5$ | $17.673$ | $11.151$ | $0.078$ |
| $h = 16$ | $10$ | $15.057$ | $10.315$ | $0.085$ |
| $h = 17$ | $10$ | $13.701$ | $9.408$ | $0.107$ |
| $h = 18$ | $20$ | $9.443$ | $6.456$ | $0.123$ |
| $h = 19$ | $15$ | $5.002$ | $3.302$ | $0.127$ |
| $h = 20$ | $10$ | $1.445$ | $0.793$ | $0.110$ |
| $h = 21$ | $10$ | $0.257$ | $0.110$ | $0.100$ |
B. APPLICATION RESULTS OF THE BENCHMARK GLOBAL TRAINING STRATEGY-BASED ANN

Similarly, the optimum configuration of the \((H = 1)\) single ANN obtained on the validation dataset \((X_{\text{valid}})\) in terms of the number of hidden neurons \((N_h)\) is found at \(N_{h,\text{opt}} = 5\) with optimum average performance metrics values equal to 14.404 kW (RMSE), 10.203 kW (MAE), and 1.802 (WMAE). For a fair comparison, the “Radial basis” and “Purlin” hidden/output neuron activation functions \(f_1()\) and \(f_2()\), respectively, and the 100-fold CV procedure are used.

For clarification purposes, Fig. 8 shows the evolution of the overall multiplied performance metrics \((RMSE \ast MAE \ast WMAE)\) obtained by the benchmark over the 100-fold CV procedure concerning the candidate numbers of hidden neurons, i.e., \(n_{\text{candidate}}\), that span the interval \([5-30]\) with a step size of 5. One can recognize that the optimum number of hidden neurons is obtained at \(N_{h,\text{opt}} = 5\) (star), and as long as the number of hidden neurons increases, the overall multiplied performance metrics increases (i.e., less prediction performance).

![Figure 8](image-url)

**FIGURE 8.** The evolution of the overall multiplied performance metrics concerning the candidate numbers of hidden neurons.

C. COMPARISONS AND DISCUSSIONS

Table 4 reports the average performance metrics obtained by the optimum ANN configurations of the proposed model (as reported in Table 3) accompanied with those produced by the optimum ANN configuration of the benchmark model (as shown in Figure 8) on the entire day hours \((h \in [7, 21])\) of the validation datasets \((X_{\text{valid}}^h)\) and \((X_{\text{valid}})\), respectively) and training datasets \((X_{\text{train}}^h)\) and \((X_{\text{train}})\), respectively) over the 100-fold CV. Results show that the proposed training strategy-based ANN (Section IV.A) outperforms the performance of the benchmark training strategy-based ANN (Section IV.B).

The performances of the two models are also verified using the test “unseen” datasets. In this regard, Table 5 reports the average performance metrics obtained by the proposed model accompanied with those produced by the benchmark model on the entire day hours \((h \in [7, 21])\) of the test datasets \((X_{\text{test}}^h)\) and \((X_{\text{test}})\), respectively) over the 100-fold CV.

![Table 4](image-url)

**TABLE 4.** The overall performance metrics obtained by using the proposed and the benchmark training strategies-based ANN on the training and validation dataset.

| Benchmark | Proposed |
|-----------|----------|
| Training  | Validation | Training | Validation |
| RMSE [kW] | 12.51 | 14.404 | 9.095 | 10.128 |
| MAE [kW]  | 6.623 | 10.203 | 6.068 | 6.751 |
| WMAE      | 0.147 | 1.802 | 0.124 | 0.145 |

To effectively compare the performance metrics from the two prediction models for the test datasets, the Performance Gain \((PG_{\text{Metric}})\) of each performance metric \((\text{Metric})\) is calculated as per Eq. (7) [4], [17]:

\[
PG_{\text{Metric}} = \frac{\text{Metric}_{\text{Benchmark}} - \text{Metric}_{\text{Proposed}}}{\text{Metric}_{\text{Benchmark}}} \times 100\% \quad (7)
\]

This gain describes the performance gain obtained by the proposed model to the performance of the benchmark for each of the computed performance metrics [4], [17]. Typically, positive values of the \(PG\) calculated for the \(RMSE, MAE, WMAE\) indicate that the proposed approach outperforms the benchmark model.

Also, the computational efforts required by the prediction models for their developments, validations, and evaluations are reported as well in Table 5.

![Table 5](image-url)

**TABLE 5.** The overall performance metrics, computational time, and the performance gains obtained by using the proposed and the benchmark training strategies-based ANN on the test dataset.

| Benchmark | Proposed |
|-----------|----------|
| Training  | Validation | Training | Validation |
| RMSE [kW] | 13.357 | 10.011 | 25.051 |
| \(PG_{\text{RMSE}}\) [%] | 0 | 10.302 | 7.169 |
| \(PG_{\text{MAE}}\) [%] | 0 | 30.412 |
| \(PG_{\text{WMAE}}\) [%] | 0 | 0.122 |
| \(PG_{\text{Time}}\) [Minutes] | 77.060 | 46.116 |
| \(PG_{\text{Time}}\) [%] | 0 | 40.156 |

Looking at Table 5, one can easily recognize that the proposed prediction model outperforms the benchmark significantly. Specifically:

- The proposed approach enhances the prediction performance by \(\sim 25\%\) (for the \(RMSE\)), \(\sim 30\%\) (for the \(MAE\)), and \(\sim 22\%\) (for the \(WMAE\));
- Additionally, the computational efforts required by the proposed approach are significantly reduced by \(\sim 40\%\), as expected;
- Thus, the proposed training strategy boosts the prediction performance than that of the benchmark with short computational efforts.
Specifically, Fig. 9 shows the evolutions of the average performance metrics (Fig. 9 (left)) and the corresponding standard deviations (Fig. 9 (right)) at each hour $h$, $h \in [7, 21]$ using the proposed (circles) and the benchmark (squares) models computed over the 100-fold CV on the test datasets.

Looking at Fig. 9 (left), one can recognize that:

- The predictions provided by the two models are comparable. Specifically, the benefit in accuracy of the proposed model to the benchmark (in particular at the early and late days’ hours) for the three performance metrics is justified by the use of solely the hourly data in building/developing the ANNs for the prediction of the corresponding hourly power productions. In contrast, the entire dataset of the whole days’ hours is used to build/develop the single ANN model. For example, at $h = 12$, the performance metrics obtained by the proposed training strategy (i.e., $RMSE = 19.05\text{kW}$, $MAE = 12.52\text{kW}$, and $WMAE = 0.095$) are smaller and, thus, superior to those obtained by the benchmark (i.e., $RMSE = 20.92\text{kW}$, $MAE = 14.63\text{kW}$, and $WMAE = 0.111$).

Whereas looking at Fig. 9 (right), one can recognize that:

- the variability (standard deviation) in the three performance metrics obtained by the proposed model is smaller than that obtained by the benchmark. Similarly, this is because the proposed model exploits solely the data collected at each hour $h$ to predict the corresponding power production predictions, whereas the benchmark utilizes the entire data collected at the whole hours to predict the power productions at each hour $h$;

- the variability is reduced between the two models at the middle days’ hours, as expected, due to the large variability in the collected data at these particular hours utilized in the proposed model compared to the early and late days’ hours;

Further insights on the superiority of the proposed model can be seen by looking at Fig. 10 that shows the computed average performance metrics over the 100-fold CV at each season (i.e., different weather conditions) on the test datasets (Fig. 10 (left)), together with the obtained performance gains (Fig. 10 (right)). It can be seen that:

- the proposed model (a dark shade of color) provides more satisfactory performance in terms of prediction accuracy of the power productions, i.e., lower metrics, for all seasons, compared to the benchmark (a light shade of color) (Fig. 10 (left));

- the highest performance gains achieved by the proposed model to the benchmark using the three metrics are obtained at the Summer season, whereas the lowest performance gains are obtained at the Winter season, with almost equal intermediate performance gains obtained at both Autumn and Spring seasons (Fig. 10 (right)). This indicates the superiority of the proposed model in achieving more accurate predictions for significant power production season (Summer season) compared to the small power production season (Winter season).

For clarification purposes, Fig. 11 shows four examples of the best (Fig. 11 (top)) and worst (Fig. 11 (bottom)) power production predictions obtained by the proposed model (circles) of four different days (one day for each season) compared with the corresponding predictions obtained by the benchmark model (squares) together with the actual productions (solid lines), respectively. The predictions provided by the two prediction models are comparable: the benefit in the
prediction accuracy of the proposed model to the benchmark is justified by the use of the solely hourly data for training the proposed model to predict the corresponding hourly power productions, whereas the complete hourly data are used to train the benchmark model.

For completeness, Table 6 reports the average performance metrics and the corresponding performance gains obtained by the proposed model to the benchmark of these particular days in the four seasons for one CV trial, i.e., CV = 6. One can recognize the superiority of the proposed model compared to the benchmark for all of the selected days across the four seasons. For example, the most significant enhancement obtained by using the proposed training strategy reaches up to \( \sim 58\% \) (RMSE), \( \sim 60\% \) (MAE), and 60% (WMAE) for Day 4 (6th June, 2015 - Summer), whereas the lowest enhancement obtained by the proposed training strategy reaches up to
to $\sim 2\%$ (for the three performance metrics) for Day 1 (25th February, 2017 - Winter). This indicates the capability of the proposed training strategy in enhancing the prediction performance across the four seasons, even for the worst predictions obtained by the proposed training strategy.

### D. COMPARISONS WITH OTHER PREDICTION TECHNIQUES

In this Section, the effectiveness of the proposed local training strategy with respect to the global training strategy when other ML techniques are adopted is investigated (refer to Fig. 6). Mainly, the Extreme Learning Machines (ELMs) are employed as prediction models instead of the ANNs. In addition, the consideration of using the ANNs is justified by comparing the prediction performances obtained by using the ANNs to those obtained by using the ELMs. Further, the prediction performances obtained by using the ANNs and the ELMs are compared to the well-known Persistence prediction model of literature, for completeness.

The ELM, initially developed by [46], is a new learning algorithm for single-hidden layer neural networks. Similar to ANN model architecture, the ELM comprises an input layer, a hidden layer that consists of $N_h$ hidden neurons, and an output layer. The idea underpinning the development of the ELM is two-fold: i) it randomly chooses the inputs’ parameters (i.e., weights and biases) of the hidden neurons, instead of using the iterative (traditional) Back-Propagation learning algorithm and, then ii) it determines the output weights, analytically. The application of ELM in different industrial fields show that it has better generalization capability and requires shallow computational efforts [46].

The Persistence model [47] for PV power production prediction is an intuitive and straightforward approach commonly used as a benchmark for evaluating the effectiveness of any proposed prediction techniques. Basically, it assumes that the PV power production at time $h, h \in [1, 24]$, of the next day will be the same as the present PV power production collected at the same time $h$.

Both the proposed (local) and the benchmark (global) training strategies are employed for developing $H = 24$ and $H = 1$ ELMs, respectively, following the steps reported in Section III.B. For the two training strategies, the ELMs are built using the training datasets ($X_{\text{train}}$ and $X_{\text{train}}^\text{valid}$, respectively), optimized in terms of the number of hidden neurons using the validation datasets ($X_{\text{valid}}$ and $X_{\text{valid}}^\text{valid}$), and evaluated and compared using the test datasets ($X_{\text{test}}^\text{h}$ and $X_{\text{test}}^\text{h}$), respectively. It is worth mentioning that different numbers of hidden neurons ($n_{\text{candidate}}$) are considered and examined, they are $n_{\text{candidate}} = [25, 50, 100, 500, 900, 1300, 1700, 1900]$. In addition, the “Radial Basis” is used as a hidden neuron activation function [16].

Table 7 reports the average performance metrics obtained by using the proposed (local) training strategy-based ELM together with those obtained by using the benchmark (global) training strategy-based ELM on the entire day hours ($h \in [7, 21]$) of the test datasets ($X_{\text{test}}^\text{h}$ and $X_{\text{test}}^\text{h}$), respectively over the 100-fold CV.

![Table 7](image-url)
Looking at Table 7, one can easily recognize that:
- the utilization of the local training strategy for the development of the ELMs largely enhances the solar PV power production prediction compared to the use of the global training strategy. Accurately, enhancements reach up to \(\sim 34\%\), \(\sim 36\%\), and \(\sim 30\%\) for the RMSE, MAE, and WMAE, respectively.
- the predictability obtained by using the ELMs is slightly lower than that obtained by the adoption of the ANNs. Future works can be devoted to enhancing the adopted prediction model embedded with the proposed local training strategy of this work.

In addition, Fig. 12 shows the evolutions of the average performance metrics (Fig. 12 (left)) and the corresponding standard deviations (Fig. 12 (right)) at each hour \(h\), \(h \in [7, 21]\) using the proposed (local) training strategy-based ANN (circles), the proposed (local) training strategy-based ELM (squares), and the Persistence model (diamonds) computed over the 100-fold CV on the test datasets. Looking at Fig. 12 (left), one can recognize that:
- The predictions provided by the three models are comparable. Specifically, the Persistence model seems to slightly/largely outperform the proposed (local) training strategy-based ANN/ELM, respectively, at the early morning (i.e., \(h = 7, 8, 9, 10\)) and late evening hours (i.e., \(h = 18, 19, 20, 21\)). This can be justified by the fact that at those time hours, the variability of the weather conditions is small that makes the intuitive justification of the Persistence model valid (i.e., the PV power production at time \(h\), \(h \in [1, 24]\), of the next day will be the same as the present PV power production collected at the same time \(h\)). However, the performance of the Persistence model starts to notably decrease (the RMSE, MAE, and WMAE start to increase) at the middle days’ hours (i.e., \(h \in [11, 17]\)) with respect to the performances obtained by using both the ANN and the ELM. This is due to the large variability of the weather conditions experienced by the ASU PV plant at those time hours;
- The proposed (local) training strategy-based ANN allows obtaining more accurate power predictions throughout the whole time hours than the proposed (local) training strategy-based ELM;
- Whereas looking at Fig. 12 (right), one can recognize that:
  - the variability (standard deviation) in the three performance metrics obtained by the Persistence model is the smallest among the three models. Similarly, this is due to the intuitive operation of the Persistence model that considers solely the data collected at each hour \(h\) to predict the corresponding power production predictions;
  - the variability obtained by the utilization of the ANN is lower than that obtained by the ELM, in particular at the middle days’ hours. This assures the effectiveness of the proposed (local) training strategy-based ANN in providing accurate solar PV power production predictions with small variability (i.e., confidence bounds).

V. INFLUENCE OF USING DIFFERENT HOUR INTERVALS FOR DATASET PARTITIONING ON THE PREDICTION PERFORMANCE
In this Section, the influence of using different hour intervals, \(\Delta h\), for partitioning the overall available pre-processed...
dataset (X) into H different datasets, on the predictability of the
ASU power production, is investigated.

Specifically, three different hour intervals are considered in this work; they are:

- $\Delta h = 2$ hours. This interval entails partitioning the
dataset (X) into $H = \frac{24}{\Delta h} = 12$ different datasets,
and thus, $H = 12$ ANNs will be built, optimized,
and evaluated. Each dataset represents the timestamps,
weather variables, and the corresponding power pro-
ductions collected at each hour interval $h$ during the
$Y = 3.625\text{ years, } h \in [1, 12]$;

- $\Delta h = 3$ hours. This interval entails partitioning the
dataset (X) into $H = \frac{24}{\Delta h} = 8$ different datasets,
and thus, $H = 8$ ANNs will be built, optimized,
and evaluated. Each dataset represents the timestamps,
weather variables, and the corresponding power pro-
ductions collected at each hour interval $h$ during the
$Y = 3.625\text{ years, } h \in [1, 8]$;

- $\Delta h = 4$ hours. This interval entails partitioning the
dataset (X) into $H = \frac{24}{\Delta h} = 6$ different datasets,
and thus, $H = 6$ ANNs will be built, optimized,
and evaluated. Each dataset represents the timestamps,
weather variables, and the corresponding power pro-
ductions collected at each hour interval $h$ during the
$Y = 3.625\text{ years, } h \in [1, 6]$.

Once the overall dataset (X) is partitioned using the three
considered hour intervals, i.e., $\Delta h = 2, 3$, and 4 hours,
into $H = 12, 8$, and 6 different datasets ($X_h$), respec-
tively, the proposed training strategy is applied following
the steps illustrated in Section III.B (Fig. 6). Specifically,
training ($X_h^{\text{train}}$), validation ($X_h^{\text{valid}}$), and test ($X_h^{\text{test}}$) datasets
are extracted randomly with arbitrary fractions of $\gamma = 50\%$,
$\beta = 20\%$, and $\alpha = 30\%$, and are used to train, validate/optimize,
and evaluate/test the $H = 12, 8$, and 6 different
ANNs, respectively.

Fig. 13 shows the evolutions of the average performance
metrics at each hour $h, h \in [7, 21]$ using the proposed (circles) and
the benchmark (squares) prediction models (as depicted in Fig. 9), and the proposed training strategy-based
ANN using the three different hour intervals, i.e., $\Delta h = 2, 3,$
and 4 hours, (diamonds, triangles, and stars, respectively)
computed over the 100-fold CV on the test datasets.

For more clarifications, Fig. 14 shows the average perfor-
mance metrics obtained by using the proposed training
strategy-based ANN with different hour intervals (light
dark shade of color) accompanied with those produced by using
the benchmark training strategy-based ANN (dark shade of
color) on the entire day hours ($h \in [7, 21]$) of the test datasets
($X_h^{\text{test}}$ and $X^{\text{test}}$, respectively) over the 100-fold CV.

The analysis of Fig. 13 and Fig. 14 conclude that:

- The predictions provided by the proposed (local) and
the benchmark (global) training strategies-based ANN
are comparable. Specifically, the benefit in accuracy
obtained by the proposed training strategy using the
different hour intervals to the benchmark (in particular at
the early and late days’ hours) for the three performance
metrics is still justified by the use of the local hourly data
for building/developing the ANNs for the prediction of
the corresponding local hourly power productions;

- The prediction performance seems to be shifted towards
the global training strategy (i.e., the benchmark with
$H = 1$) when large hour intervals are being used (e.g.,
$\Delta h = 4$ hours) and, thus, less number of ANN models
are being used.

For completeness, Fig. 15 shows the computational efforts
in minutes required by using the proposed and the benchmark
strategies during the training, optimization, and evaluation
phases. One can notice that the computational efforts needed
by the proposed approach using the three different hour inter-
vals (i.e., $\Delta h = 2, 3$, and 4 hours) and using the suggested
hour interval (i.e., $\Delta h = 1$ hour) are mostly less than those required by the benchmark model. In fact, the variable computational efforts needed when using the local training strategy is due to the differences in the number of ANN models ($H$) established and the amount of data used to train and optimize the $H$ models.

To conclude, the consideration of the hour-by-hour variation (i.e., the 24-hour seasonality patterns of each day) while building/training a data-driven prediction model is shown beneficial in enhancing the predictability of the solar PV power productions while reducing the computational efforts required by the adopted model compare to the traditional global training strategy. In practice, these enhancements are valuable in balancing power supplies and demands across centralized grid networks through economic dispatch decisions between the energy sources.

VI. CONCLUSION AND FUTURE WORKS

In this work, a local training strategy-based Artificial Neural Network (ANN) is proposed to enhance the solar PV power productions with short computational efforts. Specifically, the proposed training strategy is local in a sense that solely the timestamp, weather variables and the corresponding power productions collected at each $h$ hour interval of size $\Delta h = 1$ hour, $h \in [1, 24]$, are used to build, optimize, and evaluate $H = 24$ ANN prediction models each to be used for the estimation of the $h$-th hour power production. The proposed strategy is validated on a solar PV system of Applied Science Private University (ASU) located in Amman, Jordan. The effectiveness of the proposed strategy is evaluated to a state-of-the-art ANN model built and optimized using the entire available dataset. Three performance metrics are used for the comparisons, namely the Root Mean Square Error ($\text{RMSE}$), the Mean Absolute Error ($\text{MAE}$), and the Weighted MAE ($\text{WMAE}$), in addition to the computational times ($\text{Time}$) measured in minutes. Results show that the proposed training strategy-based ANN outperforms the benchmark training strategy-based ANN with performance gains reach up to 25% ($\text{RMSE}$), 30% ($\text{MAE}$), 22% ($\text{WMAE}$), and 40% (computational training and test times). Further, the effectiveness of the proposed strategy is evaluated when Extreme Learning Machines (ELMs) are adopted instead of the ANNs and when the Persistence prediction model is being used. Lastly, the proposed training
strategy-based ANN is evaluated in terms of i) different weather conditions (i.e., seasons) to also confirm its superiority to the benchmark training strategy-based ANN and ii) different hour intervals (i.e., $\Delta h = 2, 3, \text{and} 4$ hours) used for partitioning the overall dataset and, thus, establishing the different ANNs (i.e., $H = 12, 8, \text{and} 6$ models).

Future works can be devoted to the application of deep learning techniques, e.g., Long Short Term Memory (LSTM) and/or Echo State Networks (ESNs), instead of the employed ANN to further enhance the prediction performance.

**NOMENCLATURE**

**A. ABBREVIATIONS**

| Acronym | Description |
|---------|-------------|
| RE      | Renewable Energy |
| PV      | Photovoltaic |
| ASU     | Applied Science Private University |
| NWP     | Numerical Weather Prediction |
| ML      | Machine Learning |
| ANNs    | Artificial Neural Networks |
| BP      | Back-Propagation |
| RBF-NN  | Radial Basis Function based Neural Network |
| AR      | Auto-Regressive |
| LLR     | Local Linear Regression |
| ARMAX   | AR Moving Average with exogenous inputs |
| ARIMA   | AR Integrated Moving Average |
| SVR     | Support Vector Regression |
| ARX-ST  | AR with exogenous input based Spatio-Temporal |
| MLP-ABC | Multi-Layer Perceptron-Artificial Bee Colony |
| MARS    | Multivariate Adaptive Regression Splines |
| MLR     | Multi-Linear Regression |
| CART    | Classification and Regression Trees |
| RTs     | Regression Trees |
| ELMs    | Extreme Learning Machines |
| PSO     | Particle Swarm Optimization |
| IC      | Incremental Conductance |
| MPPT    | Maximum Power Point Tracking |
| DLNNs   | Deep-Learning Neural Networks |
| LASSO   | Least Absolute Shrinkage and Selection Operator |
| MLPs    | Multilayer Perceptrons |
| LM      | Levenberg-Marquardt |
| BR      | Bayesian Regularizations |
| PIs     | Prediction Intervals |
| RMSE    | Root Mean Square Error |
| nRMSE   | Normalized RMSE |
| MAPE    | Mean Absolute Percentage Error |
| MAE     | Mean Absolute Error |
| WMAE    | Weighted MAE |
| MSE     | Mean Square Error |
| SS      | Skill Score |
| CV      | Cross-Validation |
| LSTM    | Long Short Term Memory |
| ESN     | Echo State Network |

**B. NOTATIONS**

| Symbol | Description |
|--------|-------------|
| $W$    | Weather data |
| $Y$    | Number of available years data |
| $D$    | Number of available days data |
| $j$    | $j$-th inputs-output patterns |
| $N$    | Number of available inputs-output patterns |
| $I_{rr}$ | Global solar radiation |
| $S$    | Wind speed at 10m |
| $RH$   | Relative humidity at 1m |
| $T_{amb}$ | Ambient temperature at 1m |
| $h$    | Hour number from the beginning of each day, $h = 1, \ldots, 8760$ |
| $d$    | Day number from the beginning of each year data, $d = 1, \ldots, 365$ |
| $H$    | Number of hours in a day, $H = 24$ |
| $h$    | Hour instant, $h \in [1, 24]$ |
| $I_{rr, h}$ | $I_{rr}$ value collected at hour $h$ of each day during the period $Y$ |
| $T_{rr, h}$ | $T_{rr}$ data vector collected at hour $h$ during the period $Y$ |
| $S_{h}$ | $S$ value collected at hour $h$ of each day during the period $Y$ |
| $\bar{T}_{rr, h}$ | $T_{rr}$ data vector collected at all $H = 24$ hours |
| $\bar{S}_{h}$ | $S$ data vector collected at all $H = 24$ hours |
| $RH_{h}$ | RH value collected at hour $h$ of each day during the period $Y$ |
| $\bar{RH}_{h}$ | RH data vector collected at hour $h$ during the period $Y$ |
| $\bar{T}_{amb, h}$ | $T_{amb}$ value collected at hour $h$ of each day during the period $Y$ |
| $\bar{P}_{h}$ | $P_{h}$ data vector collected at hour $h$ during the period $Y$ |
| $\bar{P}_{h}$ | $P_{h}$ data vector collected at hour $h$ during the period $Y$ |
| $\bar{X}_{h}$ | Overall inputs-output dataset available at $h$-th hour |
| $X_{train}$ | Inputs-output training dataset |
| $X_{train, h}$ | Training dataset available at $h$-th hour |
| $X_{valid}$ | Inputs-output validation dataset |
| $X_{h}$ | Validation dataset available at $h$-th hour |
| $X_{test}$ | Inputs-output test dataset |
| $X_{test, h}$ | Test dataset available at $h$-th hour |
| $N_{train}$ | Number of training inputs-output patterns |
| $N_{valid}$ | Number of validation inputs-output patterns |
| $N_{test}$ | Number of test inputs-output patterns |
| $\gamma$, $\beta$, $\alpha$ | Arbitrary fractions used for extracting training, validation, and test datasets, respectively |
| $x_{j, h}$ | Generic $j$-th input pattern collected at hour $h$ |
| $\hat{p}_{h}$ | The $j$-th power prediction obtained at hour $h$ |
| $p_{h}$ | The $j$-th actual power collected at hour $h$ |
\( f_1() \) Hidden neurons activation function

\( f_2() \) Output neuron activation function

\( N_h \) Number of hidden neurons

\( n \) Index of hidden neuron, \( n = 1, \ldots, N_h \)

\( n_{\text{candidate}} \) Possible (candidate) number of hidden neurons

\( b_n, b_o \) Hidden and output bias neurons, respectively

\( \overrightarrow{w}_n, \overrightarrow{\beta}_n \) Hidden and output connection weights, respectively

\( \text{ANN}_{opt}^h \) Optimum ANN model obtained at hour \( h \)

\( \Delta h \) Hour interval

\( \text{Metric}_{h}^\text{PGMetric} \) Prediction performance metric calculated over the CV trials at hour \( h \)

\( \text{Metric}_{h}^\text{Benchmark} \) Prediction performance metric obtained by the Benchmark global training strategy

\( \text{Metric}_{h}^\text{Proposed} \) Prediction performance metric obtained by the Proposed local training strategy

\( \text{PGMetric} \) Prediction performance gain obtained for the performance metric Metric

ACKNOWLEDGMENT The authors would like to thank the Renewable Energy Center at the Applied Science Private University for sharing with them the Solar PV data. The authors would like to thank all the reviewers for their valuable comments to improve the quality of this article.

REFERENCES

[1] F. R. Arroyo M. and L. J. Miguel, “The trends of the energy intensity and CO2 emissions related to final energy consumption in ecuador: Scenarios of national and worldwide strategies,” Sustainability, vol. 12, no. 1, p. 20, Dec. 2019.

[2] Y. El Mghouchi, E. Chham, E. M. Zennmouri, and A. El Bouardi, “Assessment of different combinations of meteorological parameters for predicting daily global solar radiation using artificial neural networks,” Building Environ., vol. 149, pp. 607–622, Feb. 2019.

[3] H. Gohar Ali, R. Vilanova Arbos, J. Herrera, A. Tóbón, and J. Peláez-Restrepo, “Non-linear sliding mode controller for photovoltaic panels with maximum power point tracking,” Processes, vol. 8, no. 1, p. 108, Jan. 2020.

[4] S. Al-Dahidi, O. Ayadi, J. Adeeb, and M. Louzazni, “Assessment of artificial neural networks learning algorithms and training data sets for solar photovoltaic power production prediction,” Frontiers Energy Res., vol. 7, pp. 1–18, Nov. 2019.

[5] A. Sow, M. M. Relief, D. R. Rousse, and D. Haillot, “Economic analysis of residential solar photovoltaic electricity production in canada,” Sustain. Energy Technol. Assess., vol. 33, pp. 83–94, Jun. 2019.

[6] C. Brunet, O. Savadogo, P. Baptiste, and M. A. Bouchard, “Shedding some light on photovoltaic solar energy in africa – a literature review,” Renew. Sustain. Energy Rev., vol. 96, pp. 325–342, Nov. 2018.

[7] H. Wang, H. Yi, J. Peng, G. Wang, Y. Liu, H. Jiang, and W. Liu, “Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network,” Energy Convers. Manage., vol. 153, pp. 409–422, Dec. 2017.

[8] M. Rana, I. Koprinska, and V. G. Ageidis, “Univariate and multivariate methods for very short-term solar photovoltaic power forecasting,” Energy Convers. Manage., vol. 121, pp. 380–390, Aug. 2016.

[9] A. T. Eseye, M. Lehtonen, T. Tukia, S. Uimonen, and R. J. Millar, “Adaptive predictor subset selection strategy for enhanced forecasting of distributed PV power generation,” IEEE Access, vol. 7, pp. 90652–90665, 2019.

[10] N. Vázquez and J. Vázquez, 24—Photovoltaic System Conversion, E. R. U. Oxford, U.K.: Butterworth-Heinemann, 2018, pp. 767–781.

[11] E. Roumpakia and T. Stamateles, “Surface dust and aerosol effects on the performance of grid-connected photovoltaic systems,” Sustainability, vol. 12, no. 2, p. 569, Jan. 2020.

[12] X. Huang, J. Shi, B. Gao, Y. Tai, Z. Chen, and J. Zhang, “Forecasting hourly solar irradiance using hybrid wavelet transformation and elman model in smart grid,” IEEE Access, vol. 7, pp. 139909–139923, 2019.

[13] A. Bai, J. Popp, P. Balogh, Z. Gabnai, B. Pálai, I. Farkas, G. Pintér, and H. Zsiroborács, “Technical and economic effects of cooling of monocrystalline photovoltaic modules under hungarian conditions,” Renew. Sustain. Energy Rev., vol. 60, pp. 1086–1099, Jul. 2016.

[14] R. Conceição, H. G. Silva, L. Falhao, F. M. Lopes, and M. Collares-Pereira, “PV system design with the effect of soiling on the optimum tilt angle,” Renew. Energy, vol. 133, pp. 787–796, Apr. 2019.

[15] P. G. McCormick and H. Suehrcke, “The effect of intermittent solar radiation on the performance of PV systems,” Sol. Energy, vol. 171, pp. 667–674, Sep. 2018.

[16] S. Al-Dahidi, O. Ayadi, J. Adeeb, M. Alrbai, and B. R. Qawasmeh, “Extreme learning machines for solar photovoltaic power predictions,” Energies, vol. 11, no. 10, p. 2725, Oct. 2018.

[17] S. Al-Dahidi, P. Baraldi, E. Zio, E. Zio, and E. Legnani, “A dynamic weighting ensemble approach for wind energy production prediction,” in Proc. 2nd Int. Conf. Syst. Rel. (IFCSRS), Dec. 2017, pp. 296–302.

[18] S. Al-Dahidi, O. Ayadi, M. Alrbai, and J. Adeeb, “Ensemble approach of optimized artificial neural networks for solar photovoltaic power prediction,” IEEE Access, vol. 7, pp. 81741–81758, 2019.

[19] G. Pillai and H. Y. Naser, “Assessing the technical impact of integrating largescale photovoltaics to the electrical power network of babriain,” Sustain. Energy Technol. Assessments, vol. 20, pp. 78–87, Apr. 2017.

[20] A. A. Babatunde and S. Abbasoglu, “Predictive analysis of photovoltaic plants specific yield with the implementation of multiple linear regression tool,” Environ. Prog. Sustain. Energy, vol. 38, no. 4, p. 13098, Jul. 2019.

[21] U. K. Das, K. S. Tey, M. Seyedehmansaudian, S. Mekhilef, M. Y. Idris, W. Rahim, D. Deventer, R. Horan, and A. Stojcevski, “Forecasting of photovoltaic power generation and model optimization: A review,” Renew. Sustain. Energy Rev., vol. 81, pp. 912–928, Jan. 2018.

[22] B. Espinar, J. L. Aznarte, R. Girard, A. Mbaïaridj Moussa, and G. Karinotakis, “Photovoltaic Forecasting: A state of the art,” in Proc. 5th Eur. PV-Hybrid Mini-Grid Conf., 2010, pp. 1–7.

[23] C. Monteiro, L. A. Fernandez-Jimenez, I. J. Ramirez-Rosado, A. Muñoz-Jimenez, and P. M. Lara-Santillan, “Short-term forecasting models for photovoltaic plants: Analytical versus soft-computing techniques,” Math. Problems Eng., vol. 2013, Nov. 2013, Art. no. 767284.

[24] M. Ding, L. Wang, and R. Bi, “An ANN-based approach for forecasting the power output of photovoltaic system,” Procedia Environ. Sci., vol. 11, pp. 1308–1315, 2011.

[25] J. Zeng and W. Qiao, “Short-term solar power prediction using an RBF neural network,” in Proc. IEEE Power Energy Soc. General Meeting, Jul. 2011, pp. 1–8.

[26] Y. Li, Y. Su, and L. Shu, “An ARMAX model for forecasting the power output of a grid connected photovoltaic system,” Renew. Energy, vol. 66, pp. 78–89, Jun. 2014.

[27] R. De Leone, M. Pietrini, and A. Giovannelli, “Photovoltaic energy production forecast using support vector regression,” Neural Comput. Appl., vol. 26, no. 8, pp. 1955–1962, Nov. 2015.

[28] C. Yang, A. A. Thatte, and L. Xie, “Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation,” IEEE Trans. Sustain. Energy, vol. 6, no. 1, pp. 104–112, Jan. 2015.

[29] M. Khadem, M. Moadel, and A. Khosravi, “Power prediction and technoeconomic analysis of a solar PV power plant by MLP-ABC and COMFAR III, considering cloudy weather conditions,” Int. J. Chem. Eng., vol. 2016, Jun. 2016, Art. no. 1031943.
[30] Y. Li, Y. He, Y. Su, and L. Shu, “Forecasting the daily power output of a grid-connected photovoltaic system based on multivariate adaptive regression splines,” Appl. Energy, vol. 180, pp. 392–401, Oct. 2016.

[31] R. Muhammad Ehsan, S. P. Simon, and P. R. Venkateswaran, “Day-ahead forecasting of solar photovoltaic output power using multilayer perceptron,” Neural Comput. Appl., vol. 28, no. 12, pp. 3981–3992, Dec. 2017.

[32] S. Theocharides, G. Makrides, G. E. Georgiou, and A. Kyprianou, “Machine learning algorithms for photovoltaic system power output prediction,” in Proc. IEEE Int. Energy Conf. (ENERGYCON), Jun. 2018, pp. 1–6.

[33] M. H., O. Younis, and S. M., “A predictive model for solar photovoltaic power using the Levenberg-Marquardt and Bayesian regularization algorithms and real-time weather data,” Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 1, p. 2, 2018.

[34] M. K. Behera, I. Majumder, and N. Nayak, “Solar photovoltaic power forecasting using optimized modified extreme learning machine technique,” Eng. Sci. Technol., Int. J., vol. 21, no. 3, pp. 428–438, Jun. 2018.

[35] C.-J. Huang and P.-H. Kuo, “Multiple-input deep convolutional neural network model for short-term photovoltaic power forecasting,” IEEE Access, vol. 7, pp. 74822–74834, 2019.

[36] A. Catalina, A. Torres-Barrán, C. M. Alaiz, and J. R. Dorronsoro, “Machine learning nowcast of PV energy using satellite data,” Neural Process. Lett., vol. 52, no. 1, pp. 97–115, Aug. 2020.

[37] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” Neural Netw., vol. 2, no. 5, pp. 359–366, Jan. 1989.

[38] A. Saberian, H. Hizam, M. A. M. Radzi, M. Z. A. Ab Kadir, and M. Mizraei, “Modelling and prediction of photovoltaic power output using artificial neural networks,” Int. J. Photoenergy, vol. 2014, pp. 1–10, Oct. 2014.

[39] PV System ASU09: Faculty of Engineering. Accessed: May 5, 2020. [Online]. Available: http://energy.asu.edu.jo/

[40] G. de Freitas Viscondi and S. N. Alves-Souza, “A systematic literature review on big data for solar photovoltaic electricity generation forecasting,” Sustain. Energy Technol. Assessments, vol. 31, pp. 54–63, Feb. 2019.

[41] G. Maps. Applied Science Private University, Al Arab St, Amman. Accessed: May 5, 2020. [Online]. Available: https://goo.gl/qA4h3w

[42] ASU Weather Station. Accessed: Jul. 17, 2020. [Online]. Available: http://energy.asu.edu.jo/index.php/projects/asu-weather-station

[43] M. H. Alomari, J. Adeeb, and O. Younis, “Solar photovoltaic power forecasting in jordan using artificial neural networks,” Int. J. Electr. Comput. Eng., vol. 8, no. 1, p. 497, Feb. 2018.

[44] H. Demuth, M. Beale, and M. Hagan, Neural Network ToolboxTM 6 User’s Guide. Natick, MA, USA: The MathWorks, 2009.

[45] N. Akkouche, K. Loubar, F. Neboua, M. E. A. Kadi, and M. Tazerout, “Micro-combined heat and power using dual fuel engine and biogas from discontinuous anaerobic digestion,” Energy Convers. Manage., vol. 205, Feb. 2020, Art. no. 112407.

[46] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: Theory and applications,” Neurocomputing, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.

[47] E. Ogliari, A. Niccolai, S. Leva, and R. Zich, “Computational intelligence techniques applied to the day ahead PV output power forecast: PHANN, SNO and mixed,” Energies, vol. 11, no. 6, p. 1487, Jun. 2018.

SAMEER AL-DAHIDI was born in Kuwait, in 1986. He received the B.Sc. degree (Hons.) in electrical engineering from The Hashemite University, Zarqa, Jordan, in 2008, the M.Sc. degree in nuclear energy (operations specialty) from Ecole Centrale Paris and Universite Paris-Sud 11, Paris, France, in 2012, and the Ph.D. degree (Hons.) in energetic and nuclear science and technology from the Politecnico di Milano, Milan, Italy, in 2016. He has been an Assistant Professor with the Department of Mechanical and Maintenance Engineering, School of Applied Technical Sciences, German Jordanian University, Amman, Jordan, since February 2018. He has published several articles in high-quality journals, such as Applied Soft Computing, Expert Systems with Applications, Advances in Mechanical Engineering, the International Journal of Prognostics and Health Management, Reliability Engineering and System Safety, Optik, Energies, Applied Sciences, and Entropy. His current research interests include the development of analytics and models for prognostics and health management (PHM), operation, maintenance and reliability, availability, maintainability and safety (RAMS) analysis of engineering systems, and the development of artificial intelligence (AI)-based methods for renewable energy production prediction. His research interests also include renewable energy systems and mechanical engineering fields, such as thermal sciences, heating, ventilation and air-conditioning (HVAC), and others.

MOHAMED LOUZAZNI received the B.Sc. degree in electronics from the Faculty of Science, Ibn Tofail University, the M.Sc. degree in electronics and telecommunication from the Faculty of Science, Abdelmalek Essaâdi University, and the Ph.D. degree in electrical and renewable energy engineering from the National School of Applied Science, Tangier. He was a visiting Ph.D. student with the Faculty of Electrical Engineering, University Polytechnic of Bucharest, in 2015. In 2017, he started his research at the Solar Tech Laboratory, Energy Department, University Polytechnic of Milan, Italy. In 2018, he was a Ph.D. visitor at the Higher Polytechnic School of Algeciras, University of Cadiz, Spain, for five months. He is currently a Researcher with Abdelmalek Essaâdi University. His research interests include mathematical modeling, optimization, meta-heuristic algorithm, computational intelligence, photovoltaic and power energy, forecasting, fuel cell, GPR, radio frequency, electromagnetic, and electronic.

NAHED OMARAN received the B.Sc. degree in mechanical engineering from the Faculty of Engineering Technology, Al-Balqa Applied University, Amman, Jordan, in 2015. She is currently a full-time Teacher Assistant with Applied Science University, Jordan. She is also working as a Renewable Energy Engineer with the Renewable Energy Center, ASU. Her research interests include PV panels’ technologies, machine learning applications in renewable energy, and thermal science.