Digital Twin for Solar Photovoltaic Power Estimations based on an Ensemble of Recurrent Neural Networks

Michael Walters, *IEEE Student Member*
Real-Time Power and Intelligent Systems Laboratory
Holcombe Department of Electrical and Computer Engineering
Clemson University, Clemson, SC 29634, U.S.A
mawalters@ieee.org

Ganesh K. Venayagamoorthy, *IEEE Fellow*
Real-Time Power and Intelligent Systems Laboratory
Holcombe Department of Electrical and Computer Engineering
Clemson University, Clemson, SC 2934, U.S.A
gkumar@ieee.org

***Abstract*—Planning, managing, and maintaining solar photovoltaic (PV) plants is becoming increasingly challenging as a result of their increasing implementation world-wide. Solar PV power generation estimations provide a source of knowledge and certainty, assisting system operators in day-to-day responsibilities. Digital twins (DTs) replicate physical entities within a virtual setting, providing a real-time platform to perform solar PV power generation estimations, further enhancing situational awareness and operational efficiency. In this paper, a DT is developed and implemented for Clemson University's 1 MW solar PV plant located in South Carolina, USA to perform solar PV power generation estimations. An ensemble of Elman recurrent neural networks (ERNNs) is utilized in the DT for solar PV power generation estimations, replicating PV plant behaviors and characteristics. The ERNN ensemble utilizes data collected at the PV plant site, i.e. generated power, solar irradiance and ambient temperature. The DT's performance is evaluated based on different weather conditions and ERNN ensemble's output selection methods. Typical results are presented to show the effectiveness of the neural network ensemble based DT for solar PV power generation estimations.**

Keywords—Computational intelligence, digital twin, ensemble, recurrent neural networks, solar photovoltaic power

I. Introduction

The electric power generation industry is transforming due to the ever-rising demand for solar photovoltaic (PV) power generation. However, with the progressive integration of solar PV plants, the balance and security of the electrical power grid is at risk. Solar PV power generation introduces a source of dynamic power generation, due to a high dependance on volatile weather conditions. Additionally, consistent maintenance and accurate performance monitoring over short and long time periods can further negatively impact efficient widespread PV plant integration Therefore, system operators face greater challenges in the seamless introduction of new solar PV sites and in energy dispatch during daily operation. Solar PV power generation estimations offer a valuable source of information to combat these operational challenges, enhancing situational awareness in distribution control, thus assisting day-to-day operations. Examples include comparisons of measured and estimated data for inferring maintenance or cleaning requirements, evaluating system degradation at regular intervals, providing realistic power generation expectations aiding in new site planning and quality control assurance.

Digital twins (DTs) offer a reliable platform to perform solar PV power estimations. By providing a real-time virtual reproduction of solar PV plants, system characteristics and dynamic responses to variable weather conditions may be captured. DTs developed with computational intelligence techniques may learn solar PV plant behaviors strictly using data gathered from the PV plant and its surrounding environment. With empirical studies, the comparison of multiple digital twin architectures performing solar PV power estimations was made in [1]. This paper concluded that, while DTs offer a reliable reproduction of solar PV plant power generation behaviors, the single multi-layer perceptron and Elman recurrent neural networks (ERNNs) used to create DTs fell short of capturing the vast dynamics and critical responses governed by volatility of a solar PV plant due to limiting factors such as slow convergence speeds, local minima, and low flexibility for adaptable problem solving, especially as data sets and systems increase in size [2].

To improve digital twin architecture for solar PV power generation estimations, the use of a neural network ensemble (NNE) is proposed in this paper. NNEs originate in [3] and consist of a finite number of neural networks that are trained on the same dataset to perform the same task. NNEs offer improvement on computational capabilities when compared to that of a singular NN. Through training, each NN is allowed to specialize on certain patterns in the dataset. As these generalizations develop, their respective errors grow. However, it is the collective decision and resulting error of the ensemble as an entity that has been shown as far less fallible than any individual network [3]. NNEs have been used in both classification and regression applications, including machine learning, pattern recognition, image analysis, medical diagnosis, and weather forecasting [4, 5]. By implementing a NNE as a digital twin, the virtual replication can better capture the dynamic characteristics and attributes of a physical reality. In the context of solar PV plants, NNE DTs can more accurately and reliably learn the dynamic relationship between environmental conditions and generated power, thus providing a more robust application.

This paper presents the development and implementation of a digital twin utilizing an ensemble of Elman recurrent neural networks (ERNNs) for solar photovoltaic power generation estimations of Clemson University's 1 MW solar PV plant located at the R-06 parking site. Two ensemble output calculation methods are presented and compared using different performance metrics.

The remaining sections of this paper are outlined as follows: Section II discusses digital twins for solar PV power estimations, Section III describes the implementation of ensemble neural networks for DTs and Section IV presents results, discussions, and applications for estimation DTs.

This research is funded by US NSF grant CNS 2131070, ECCS 2234032, CNS 2318612 and the Duke Energy Distinguished Professor Endowment Fund. Any opinions, findings and conclusions or recommendations expressed in this material are those of author(s) and do not necessarily reflect the views of NSF and Duke Energy.

2024 IEEE 19th Conference on Industrial Electronics and Applications (ICIEA) | 979-8-3503-6086-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICIEA61579.2024.10665014

Finally, the conclusion and directions for future work are summarized in Section V.

II. Digital Twins for Solar PV Power Estimation

Digital twins have seen increasing deployment in academic, research and industry contexts alike due to their highly versatile and adaptable capabilities seen in a variety of applications. A generalized definition of DTs given in [6] characterizes them as having three primary components: a physical reality, virtual representation, and interconnection between the two. Both physical reality and virtual representation are further decomposed into three subcomponents for each, including relevant systems, processes, and environmental conditions. Each of these elements represents physical and virtual aspects crucial to the framework of DTs. Combining each of these entities allows for the fusion of real-time information. Fig. 1 summarizes the explained DT architecture.

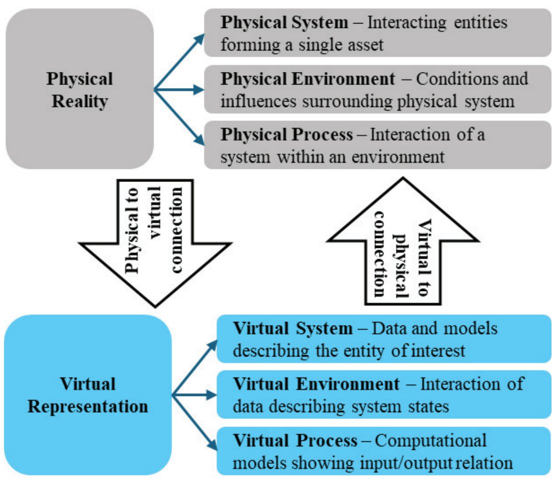

Fig. 1. General digital twin architecture consisting of a physical reality, virtual representation and a bi-directional interconnection.

With increasing levels of solar PV power generation in today's energy market, day-to-day operation and management of these plants is challenged. Solar PV plants are dynamic systems with variable power generation characteristics highly dependent on the present weather conditions. Furthermore, differences in design parameters affecting efficiency and life-expectancy pose challenges to developing a generic, one-size-fits-all model. In short, solar PV plants might have similar characteristics to others, but no two sites are the same. However, digital twins implemented at new or existing sites can provide a greater understanding of their non-linear, volatile power generation characteristics. DTs offer a scalable and adaptable platform to replicate solar PV plants on the basis of historical and real-time data. Therefore, many of the properties, parameters and relationships associated with an individual plant may be captured based on the level of abstraction used. For instance, a DT may be deployed to model the entire energy conversion process that takes place within a solar PV plant. Other DT implementations may involve the individual characterization of components, processes, and machinery within a solar PV site. Regardless of the depth of abstraction within a DT model, an accurate depiction of the physical reality may be obtained based on the nature of data collected [7]. IoT devices and sensors within power systems plants and machinery enable both flexibility for DT development and dependable data sources. Furthermore, artificial intelligence (AI)-based methods aid in power plant modeling in the power systems industry, as they provide a superior computational platform, as compared to physics-based or statistic-based methods [8]. By combining the advantages of big-data collected through IoTs and the robust nature of AI-based DTs, a reliable and accurate realization performing solar PV power generation estimations may be created. Thus, insights to solar PV power generation provide a consistent source of information, improving daily operation, management, and maintenance of solar PV sites.

Fig. 2 summarizes DT performing solar PV power generation estimations developed for this study, and the solar PV plant and physical facilities at the R-06 parking lot provided by Clemson University. The upper portion contains the physical reality, i.e., 1 Mega Watt solar PV plant, weather station containing IoT devices, and micro-PMU. The collected environmental and power generation data are transmitted to the Real-Time Power Systems (RTPIS) Laboratory, located on Clemson University's campus through a wireless data communication network. The lower portion contains the virtual representation, as shown by the data flow and computational models. DT power estimations can then be used in control center applications, such as PV panel cleaning requirements, maintenance notifications, and damage alerts.

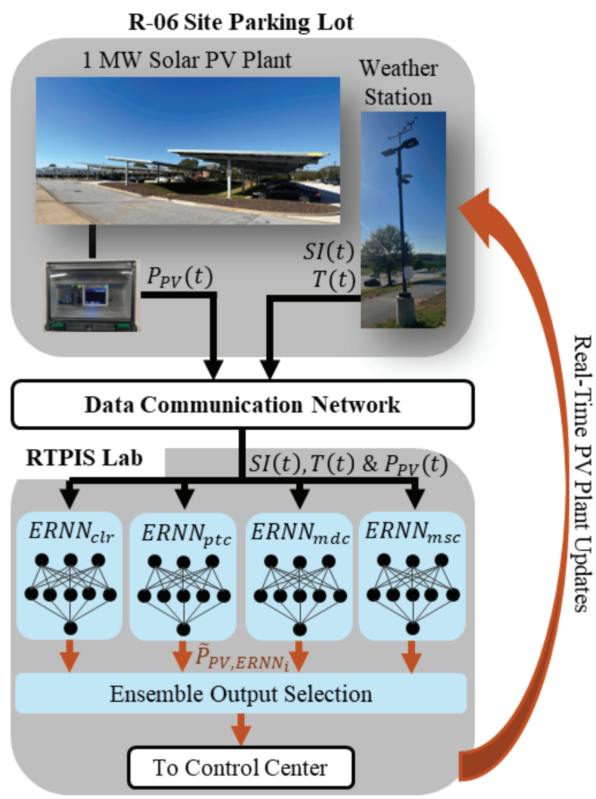

Fig. 2. Overview of solar PV power generation estimation digital twin and Clemson University's 1 MW solar PV plant at the R-06 Site.

III. Neural Network Ensemble Digital Twin Implementation

For solar PV sites, a physics-based approach to estimate PV power at a given instant is given in (1),

$$\tilde{P}_{PV}(t) = \frac{SI(t)}{SI_{ref}} P_{ref,mp}\left[1 + \gamma\left(T(t) - T_{ref}\right)\right] \tag{1}$$

where $\tilde{P}_{PV}$ is the estimated PV power, *SI* is solar irradiance, SI_{ref} is reference solar irradiance, $P_{ref,mp}$ is the maximum PV power reference, γ is the solar array coefficient, *T* is temperature and T_{ref} is reference temperature. Accordingly, $P_{ref,mp}$, SI_{ref} and T_{ref} are all parameters that are subject to change, as they relate to specific PV sites. Often times, these parameters are subject to frequent variations due to an individual PV plant's generation characteristics, including age and efficiency, impacting accurate parametrization and difficulty of modeling with this approach.

On the other hand, NNs may learn PV plant generation characteristics based solely on input environmental conditions and generated output power at a given instant, as seen in (2),

$$\tilde{P}_{PV}(t) = f_{Est}(SI(t),\ T(t),\ W, V) \tag{2}$$

where *W* and *V* represent input and output synaptic weights, respectively derived from training. DTs utilizing NNs provide the capability replicate PV plants as a whole entity. Therefore, variables relevant to a specific PV plant, such as partial shading, cloud coverage, and system degradation are further captured by the DT, as compared to the physics-based approach.

A. Elman Recurrent Neural Network

In previous studies, as well as in recent literature, it has been found that neural networks provide superior performance in modeling and estimating solar PV systems, when compared to other computational intelligence paradigms [9]. For this reason, ERNNs are implemented in this study.

The architecture of an ERNN consists of input, hidden, output and context layers, as seen in Fig. 3. Within each layer, neurons are represented as circles, each containing a transfer function. Linear for input, output, and context layer neurons, and logarithmic sigmoid for hidden layer neurons. Edges signify synaptic weights that connect neurons: input weights *W*, output weights *V*, and intermediate weights *D*. The input layer includes solar irradiance and temperature at time *t*, measured PV power at the previous time instant, $P_{PV}(t-1)$, and time delayed intermediate weight matrix, $D(t-1)$. Including $P_{PV}(t-1)$ adds an element of historical time dependence in the input layer, offering a foundation for power

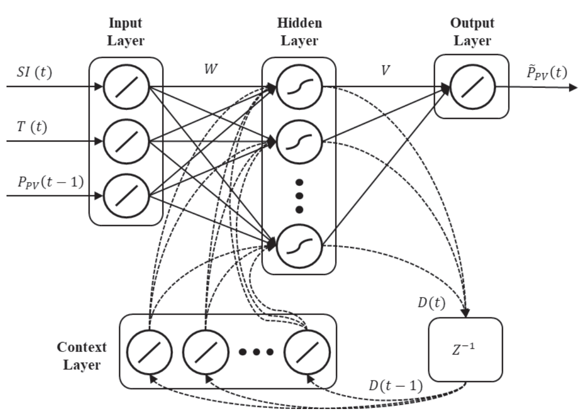

Fig. 3. Elman neural network for solar PV power generation estimation DT implementation.

$$\tilde{P}_{PV,ERNNi}(t) = f_{Est}\begin{pmatrix} SI(t),\ T(t), P_{PV}(t-1), \\ D(t-1), W, V \end{pmatrix} \tag{3}$$

estimations occurring in the present time, *t*. Following (2), the function to estimate PV power of individual NNs within the ensemble is updated in (3). During development, it was determined that a hidden layer of 60 neurons was capable of estimating PV power, while balancing computational performance. Therefore, the resulting ERNN size is $(3 + 60) \times 60 \times 1$, with 3,840 synaptic weights.

B. Training Procedure

Meteorological data and solar PV power generation polled every minute at Clemson University's 1 MW solar PV plant, located in Clemson, South Carolina, USA is used to form the DT training dataset. These specifically include solar irradiance, ambient temperature, and plant-generated power. After a data pre-processing procedure, the resulting dataset includes 86 days ranging from March 2023 to June 2023 (approximately 371,520 datapoints). Due to the range of data collected, the DT primarily captures springtime weather characteristics, whereas dynamics of other seasons, i.e., Summer, Autumn, and Winter may be absent.

The training dataset is categorized daily by cloud-coverage conditions as indicated by PV plant power generation. Characteristics such as peak power generation, frequency of volatility and severity of volatility are prioritized factors while sorting. Of the 86-day dataset, the identified categories, and the number of days within each sub-dataset are as follows: 21 clear, 25 partially cloudy, 27 moderately cloudy, and 13 mostly cloudy days. It is important to note that the mostly cloudy category contains significantly fewer days, impacting the training sub-dataset size. However, sorting prioritized accurate grouping over having an equal distribution of days, ensuring that days within categories have higher correlation.

The NNE implemented in this study utilizes four ERNNs, each training independently on individual sub datasets. Accordingly, the four ERNNs will be referred to and labeled as $ERNN_{clr}$, $ERNN_{ptc}$, $ERNN_{mdc}$ and $ERNN_{msc}$, for clear, partially cloudy, moderately cloudy, and mostly cloudy sub-datasets. By creating sub-datasets classified by weather conditions and solar PV power generation characteristics, and individually training ERNNs within their respective category, specialization within the ensemble occurs. Thus, trends in power generation based on differing meteorological conditions across weather categories may be further exposed. ERNNs are trained using the batch backpropagation algorithm discussed in [10].

C. Ensemble Output Selection

To incorporate the diversity and specialization of individual ERNNs during testing, two ensemble output selection methods are implemented and compared, namely an unweighted averaging and previous best approaches. This further exposes the trends that cause differing relationships between power generation and meteorological conditions for certain weather categories. The overall ensemble output, $\tilde{P}_{PV,Ensemble}$ is determined as a function of the four NN solar PV power estimations, as seen in (4).

$$\tilde{P}_{PV,Ensemble} = \begin{Bmatrix} \tilde{P}_{PV,ERNNclr}, \tilde{P}_{PV,ERNNptc}, \\ \tilde{P}_{PV,ERNNmdc}, \tilde{P}_{PV,ERNNmsc} \end{Bmatrix} \tag{4}$$

Where $\tilde{P}_{PV,ERNNclr}$, $\tilde{P}_{PV,ERNNptc}$, $\tilde{P}_{PV,ERNNmdc}$, and $\tilde{P}_{PV,ERNNmsc}$ are solar PV power estimations corresponding to outputs from $ERNN_{clr}$, $ERNN_{ptc}$, $ERNN_{mdc}$ and $ERNN_{msc}$, respectively.

Ensemble averaging constitutes a computationally inexpensive approach for ensemble output determination, where power estimations of individual ERNNs are averaged at every time instant, *t*. Unweighted averaging produces an equal representation of the four ERNNs. Therefore, strengths and weaknesses of ERNNs are present, but at a diluted level. Ensemble average estimated PV power ensemble output, $\tilde{P}_{PV,Ensemble}$, is determined utilizing (5), where $\tilde{P}_{PV,ERNNi}$ is the estimated PV power of the *i*th ERNN.

$$\tilde{P}_{PV,Ensemble}(t) = \frac{1}{M}\sum_{i=1}^{M}\tilde{P}_{PV,ERNNi}(t) \quad (5)$$

The previous best ensemble output selection method introduces an element of time-dependency, as historical power estimations are utilized for evaluating ensemble output. First, squared errors (SEs) of individual power estimations of ERNNs at the previous time instance, $(t-1)$, are computed and compared. The ERNN with minimum SE is chosen as the ensemble winner for the current time instance (*t*), (6). The estimated PV power of the winning ERNN ($\tilde{P}_{PV,ERNNj}$) is then selected as the ensemble output (7).

$$min\begin{pmatrix} SE\left(\tilde{P}_{PV,ERNNclr}(t-1)\right), \dots, \\ SE\left(\tilde{P}_{PV,ERNNmsc}(t-1)\right) \end{pmatrix} \quad (6)$$

$$\tilde{P}_{PV,Ensemble}(t) = \tilde{P}_{PV,ERNNj}(t) \quad (7)$$

A depiction of the ensemble selection algorithm is provided in Fig. 4. As seen, both selection methods are utilized synchronously while testing.

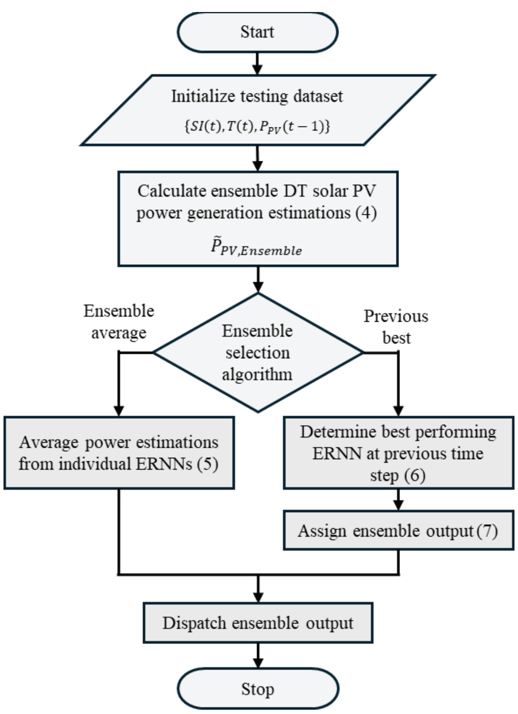

Fig. 4. A flowchart showing two ensemble output selection methods utilized by solar PV power generation estimation DT.

IV. Results & Discussion

Ensemble training is conducted until a 30,000 epoch limit has been achieved. Fig. 5 shows averaged mean square error (MSE) progression of individual NNs within the ensemble through training epochs over twenty training trials. Additionally shown are the error values determined at the conclusion of training.

Certain characteristics of training performance are present for each ERNN, based on assigned weather category. For instance, $ERNN_{clr}$ and $ERNN_{mdc}$ averaged the least error through the initial portion of training, whereas $ERNN_{msc}$ experienced a significant drop in MSE through later portions of training. The final training MSE for $ERNN_{clr}$ and $ERNN_{msc}$ were nearly the same, at 4.5×10^{-4} and 4.0×10^{-4}, respectively. On the other hand, $ERNN_{ptc}$ and $ERNN_{mdc}$, final training MSE were slightly greater, at 3.2×10^{-3} and 1.9×10^{-3}, respectively. These differences can be contributed to the correlation strength between meteorological inputs solar irradiance and temperature and the measured PV power, as a greater number of variations is observed to be a more difficult relationship to learn.

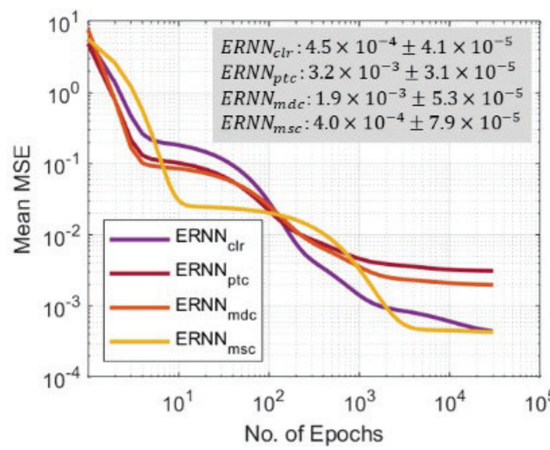

Fig. 5. Averaged mean square error progression for each ERNN ensemble member. Mean square errors are shown after training for 30,000 epochs.

After training is completed, the DT is tested over a sample of 20 testing days gathered in May 2023. The testing dataset contains an even distribution of information for each weather category, i.e., five days for each. Samples of each weather category for both ensemble output selection techniques are shown in Fig. 5. Weather categories are sorted in rows, subplots (*a*), (*b*), (*c*), and (*d*), for clear, partially cloudy, moderately cloudy, and mostly cloudy, respectively. Ensemble outputs are sorted by columns, (*1*) and (*2*), for ensemble average and previous winner methods, respectively.

As seen in Fig. 6, insights to the cloud coverage categorization process are revealed. Clear days, shown in (*1.a*) and (*2.a*) ideally contain a smooth profile and maximum generation during the peak time interval. Partially cloudy days shown in (*1.b*) and (*2.b*), consist of generation volatility through large portions of the sampled day, while still reaching peak generation. Moderately cloudy and mostly cloudy categories, (*1.c*), (*2.c*) and (*1.d*) and (*2.d*), respectively contain

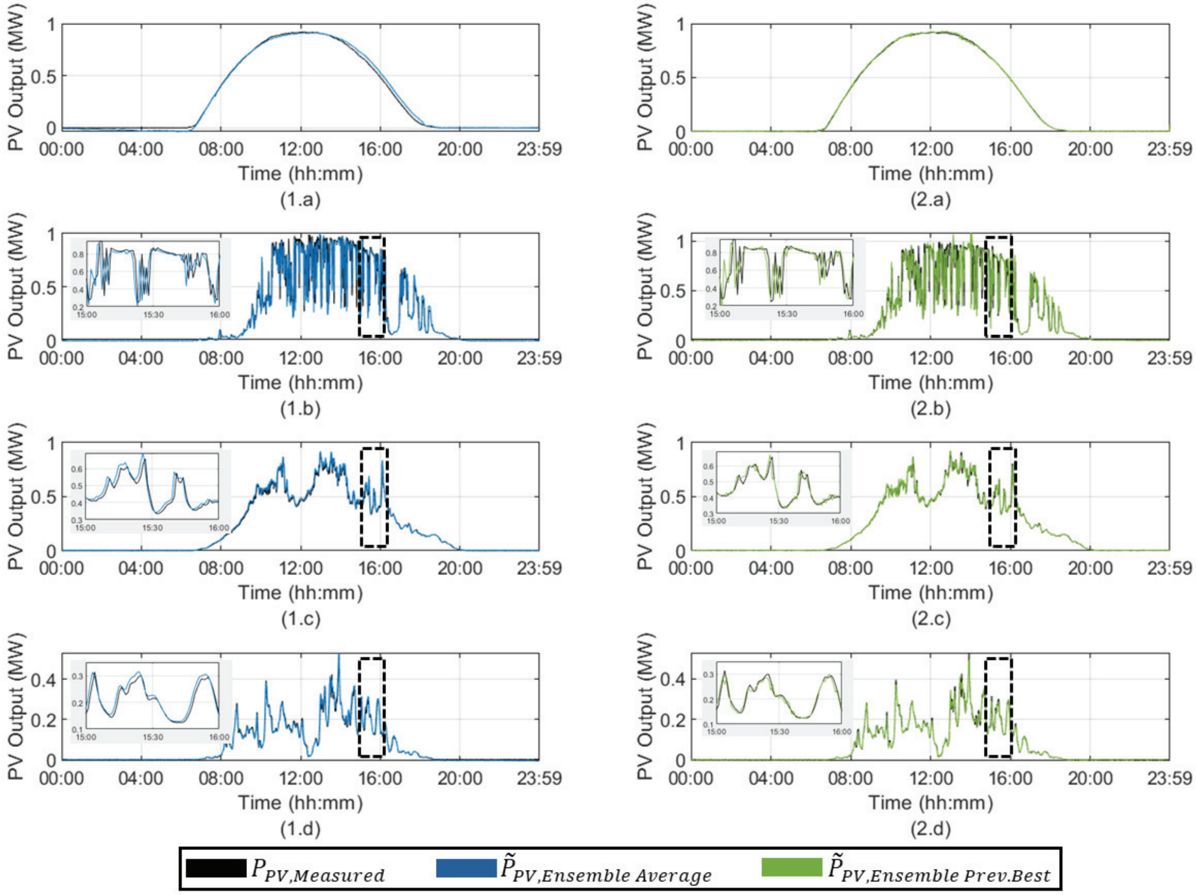

Fig. 6. Sampled DT power estimations comparing measured PV power (black) with ensemble average (blue) and previous best (green) output selection techniques.

TABLE II. DAYTIME PERFORMANCE METRIC COMPARISON OF DTS

	Ensemble Average		Previous Best		Single ERNN		ARIMA	
Weather Profile	*MAPE (%)*	*MSE*	*MAPE (%)*	*MSE*	*MAPE (%)*	*MSE*	*MAPE (%)*	*MSE*
Clear	5.11	1.22×10^{-3}	**1.11**	6.33×10^{-5}	5.27	1.20×10^{-3}	4.42	1.63×10^{-4}
Partly Cloudy	12.89	1.76×10^{-2}	**12.62**	1.78×10^{-2}	12.77	1.79×10^{-2}	18.05	2.20×10^{-2}
Moderately Cloudy	3.23	7.30×10^{-4}	**3.15**	7.34×10^{-4}	3.39	6.71×10^{-3}	6.71	1.18×10^{-3}
Mostly Cloudy	**4.10**	1.62×10^{-4}	4.35	1.67×10^{-4}	4.69	1.33×10^{-3}	11.99	2.86×10^{-4}

similar features in terms of PV power generation volatility but differ significantly in peak power.

Table I summarizes mean absolute percent error (MAPE) and MSE for daytime periods. For further comparison, a single ERNN and an auto-regressive integrated moving average (ARIMA) model are additionally included. Both of these models were trained in the same datasets, with the single ERNN consisting of similar parameters to those in the DT NNE. The ARIMA model utilizes a statistical approach and is frequently implemented for time series forecasting applications [11].

When comparing MAPE for the ensembles, it is observed that the previous best method is significantly more accurate on clear days. Whereas, with partially, moderately and mostly cloudy conditions, both ensemble selection methods feature very similar performance. Previous best tested with slightly lower error on partially and moderately cloudy conditions, and ensemble averaging testing with slightly lower error on mostly cloudy days. When comparing the single ERNN to the ensembles, it is seen that MAPEs are very similar for all conditions except for previous best on clear days. Additionally, both NNEs and single ERNN tested more accurately than the ARIMA model, with the exception of previous best on clear days.

The largest trend observed with testing results is the significantly lower error of the previous best method on clear weather conditions. This can be attributed to the low volatility of power generation on clear days, allowing for the persistence of the NN with the lowest MSE at the previous timestep to dominate in the previous best method. On the other hand, the ensemble average does not consider historical instantaneous error, as reflected in an overall greater error accumulation. Partially, moderately, and mostly cloudy categories involve power generation volatility to some degree, creating a weaker correlation between time-dependent parameters, as shown by the marginal differences in MAPEs exhibited by NNEs during these conditions. In these cases, power generation has a lesser dependence on time-varying persistence, as previously discussed with the clear condition case. Thus, the previous best method is constrained by the limits of the introduced time-dependency for output selection, whereas the ensemble average approach enables an equally weighted, diluted representation of specialized qualities between ERNNs within the ensemble.

Considering the three cloud coverage conditions, the DTs containing NNEs display similar performances to the single ERNN. Specifically with these conditions, the dynamic nature of PV power generation and estimations is uncovered. With these cases, it is important to note that individual NNs within the NNEs still developed specialization qualities as a result of their respective training sub-datasets. The aggregation of NNE PV power estimations with both output selection algorithms enables a generalized representation of these properties, resulting in slight performance gains over the single ERNN. This is particularly evident with the ensemble averaging method, where these specialization properties are equally weighted, resulting in the dilution of strengths developed during individual NN training. Considering the clear cloud coverage condition, it is evident that the specializations developed in individual NNs enables superior performance when utilizing the previous best selection method. It is specifically with this case that specialization of individual NNs dominated PV power estimations. In future exploration of NNEs, the development and exploitation of specialized qualities of all NNs across all weather classifications will lead to greater performance benefits.

Fig. 8 displays the coefficient of determination, R^2, for both ensemble average and previous best selection techniques. Calculated using (8), the coefficient of determination shows the correlation between measured and estimated values.

$$R^2 = 1 - \frac{\sum\left(P_{PV}(t) - \tilde{P}_{PV,Ensemble(t)}\right)^2}{\sum\left(P_{PV}(t) - \bar{P}_{PV}\right)^2} \tag{8}$$

where P_{PV} is measured power, and $\bar{P}_{PV}$ is the mean measured power for the respective ensemble. Additionally, an ideal 1:1 line is plotted to show ideal correlation. In both cases, R^2 is determined to be approximately 98%, indicating strong correlation.

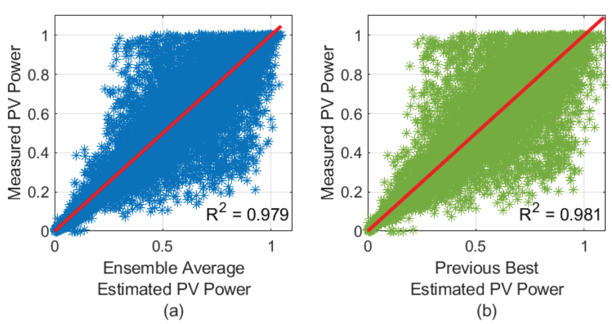

Fig. 8. Coefficient of determination for ensemble average (a) and previous best (b) output selection techniques.

V. Conclusion

Digital twins performing solar PV power generation estimations can provide unique insights for the highly dynamic operating characteristics of a solar PV plant, assisting power system operations. By utilizing neural network ensembles, individual neural networks may specialize on specific trends evident in datasets, further improving accuracy of solar PV power generation estimations. Thus, a trustworthy source of information is introduced for enhanced situational awareness in planning, monitoring, and maintenance applications.

In this study, a digital twin based on an ensemble of recurrent neural networks for solar PV power estimations was developed and implemented for Clemson University's 1 Megawatt solar PV plant. The DT utilized ensembles of Elman recurrent neural networks, each trained on individual sub-datasets according to weather conditions. Different ensemble output aggregation algorithms were compared. The DT featuring NNEs proved capable of accurately replicating solar PV plant dynamics for a variety of weather conditions.

In future work, accuracy and robustness of the DT may be improved by incorporating state-of-the-art computational intelligence paradigms. With this advancement, more opportunities for specialized ensemble output aggregation algorithms to further exploit NNE specialization traits and applications in PV power forecasting can be explored.

Acknowledgment

The authors acknowledge the facilities provided by the Real-Time Power and Intelligent Systems (RTPIS) Laboratory and physical facilities at Clemson University, SC, USA for the research carried out in this paper.

References

[1] M. Walters, J. Yonce and G. K. Venayagamoorthy, "Data-Driven Digital Twins for Power Estimations of a Solar Photovoltaic Plant," in Proc. *2023 IEEE Symposium Series on Computational Intelligence (SSCI)*, Dec. 2023, pp. 246-251

[2] H. Li, X. Wang and S. Ding, "Research and development of neural network ensembles: a survey,"*Artificial Intelligence Review*, vol. 49 pp. 455-479, 2018.

[3] L. K. Hansen and P. Salamon, "Neural Network Ensembles," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 993-1001, Nov. 1990.

[4] I. Maqsood, M. R. Khan and A. Abraham, "An ensemble of neural networks for weather forecasting," *Neural Computing Applications*, vol. 13, pp. 112-122, May 2004.

[5] Z. Zhou, J. Wu and W. Tang, "Ensembling neural networks: Many could be better than all," *Artifical Intelligence*, vol. 137, pp. 239-263, May 2002.

[6] E. VanDerHorn, and S. Mahadevan, "Digital Twin: Generalization, characterization and implementation," *Decision Support Systems*, vol. 145, Jun. 2021.

[7] A. Sleiti, J. Kapat, and L. Vesely, "Digital twin in energy industry: Proposed robust digital twin for power plant and other complex capital-intensive large engineering systems," *Energy Reports*, vol. 8, pp. 3704-3726, Nov. 2022.

[8] X. Zhang, Y. Li, T. Li, Y. Gui, Q. Sun and D. W. Gao, "Digital Twin Empowered PV Power Prediction," *Journal of Modern Power Systems and Clean Energy*, Nov. 2023.

[9] K. Garud, S. Jayaraj, and M. Lee, "A review on modeling of solar photovoltaic systems using artificial neural networks, fuzzy logic, genetic algorithm and hybrid models," International Journal of Energy Research, vol.45, pp. 6-35, Jun. 2021.

[10] V. G. Gudise, and G. K. Venayagamoorthy, "Comparison of particle swarm optimization and backpropagation as training algorithms for neural networks," in Proc. *2003 IEEE Swarm Intelligence Symposium Conf.*, pp. 110-117.

[11] H. Sharadga, S. Hajimirza, and R. S. Balog. " Time series forecasting of solar power generation for large-scale phtotvoltaic platns," *Renewable Energy*, vol. 150, pp. 797-807, May 2020.