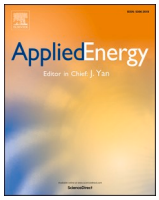
Applied Energy 402 (2025) 126855

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy



Situational awareness indices of solar PV power generation under temporal weather conditions for near real-time planning and operation

Michael Walters [a,*], Ganesh K. Venayagamoorthy [a,b]

[a] Real-Time Power and Intelligent Systems Laboratory, Holcombe Department of Computer and Electrical Engineering at Clemson University, Clemson 29634, SC, USA
[b] Department of Electrical, Electronic and Computer Engineering at University of Pretoria, South Africa

HIGHLIGHTS

- Three situational awareness indices (SAIs) for solar PV plants are proposed.
- SAIs characterize solar PV plant performance based on environmental conditions.
- Historical PV power generation data provides high spatial-temporal SAI resolutions.
- Near real-time planning and operational contexts of solar PV plants are enhanced.
- Practical application of SAIs with energy dispatch PV-BESS controller study.

ARTICLE INFO

Keywords:
Fuzzy inference system
Performance state
Situational awareness
Solar PV plant operations & planning

ABSTRACT

Solar photovoltaic (PV) plant development and utilization is increasing worldwide but remains intrinsically challenged by its large dependence on highly variable weather conditions and operating states. This paper presents a framework to leverage three new situational awareness indices (SAIs), namely: weather condition index (WCI) to gauge operational performance based on environmental states, operational complexity index (OCI) to indicate the severity of power generation reductions, and photovoltaic generation index (PVGI) to provide a final determination of the impact on power generation and to bolster situational awareness in planning and operational contexts for solar PV plants. This is accomplished by exploiting the effects of weather conditions, operating states, and solar PV power generation performance in high spatial-temporal resolution contexts residing in solar PV power generation data with independent fuzzy inference systems (FISs) for each index. SAIs provide additional operational insights to evaluate solar PV plant performance over both short-term (minute(s)) and long-term (24 h) time intervals in a variety of areas, including weather condition classification studies, energy dispatch controllers, and power system voltage and frequency stability assurance. The proposed SAI framework is developed, demonstrated, and evaluated for a 1MWp solar plant located in Clemson, South Carolina, USA.

1. Introduction

Electric power generation markets worldwide are experiencing ever-increasing energy demands. In fact, projections from the U.S. Energy Information Administration (EIA) account for at least a 55 % increase in global electric power generating capacity by 2050 [1]. Supporting this additional power generation capacity amidst the industry-wide transition from fossil fuel-based generation to renewable energy sources (RESs), is a crucial target. A primary driver of this transformation is the urgent need to mitigate harmful effects of climate change by reducing greenhouse gas emissions, particularly with coal and natural gas power generation. Other influences on the transition to renewable energy generation, highlighted by the United Nations include carbon-neutral governmental regulations, sustainability initiatives, and growing requirements for energy security [2].

Among widely available RES technologies, notably, solar and wind energy are at the forefront. Benefiting from supporting policies and Industry 4.0 advancements, the cost of solar power generation has dropped dramatically, with the International Energy Agency (IEA) reporting solar generating capacity tripling from 2018 to 2024 and projecting the technology to account for 80 % of renewable generation growth by 2030

* Corresponding author.
E-mail address: mawalters@ieee.org (M. Walters).

https://doi.org/10.1016/j.apenergy.2025.126855
Received 11 April 2025; Received in revised form 27 August 2025; Accepted 1 October 2025
Available online 18 October 2025

Nomenclature

Abbreviations

AI	Artificial intelligence
BESS	Battery energy storage system
CPF	Cumulative power fluctuation (MWh)
CPG	Cumulative power generation (MWh)
DP	No. polled daytime samples
EIA	U.S. Energy Information Administration
FIS	Fuzzy inference system
GHI	Global horizontal irradiance (W/m^2)
H	High
IEA	International energy agency
IoT	Internet-of-things
L	Low
M	Medium
MH	Medium-high
ML	Medium-low
MLP	Multi-layer perceptron
MPPT	Maximum power point tracking
NWP	Numerical Weather prediction
OC	Operating condition
OCI	Operational complexity index
PFD	Power fluctuation density (MWh)
PFF	Power fluctuation frequency
PMU	Phasor measurement unit
PV	Photovoltaic
PVGI	PV generation index
PVGI-DT	PVGI-based decision tree
PVPS	PV power generation signature (MW)
RES	Renewable energy source
SA	Situational awareness
SAI	Situational awareness index
SCADA	Supervisory control and data acquisition
SoC	State of charge
TF	Timeframe
T-DT	Traditional decision tree
VH	Very-high
VL	Very-low
WC	Weather condition
WCI	Weather condition index

Variables

d	Distance
N	Total days in historical dataset
n	Arbitrary day
P	24 h Time-series power generation (MW)
p	Power generation sample (MW)
T	Total samples in 24 h period
t	Arbitrary timestamp
w	FIS rule output weight
X	Arbitrary OC
z	FIS rule output level
α	Feature scaling value
ϵ	Distance proportion

[3]. In particular, solar photovoltaic (PV) technologies offer flexible, scalable, and adaptive designs to fit robust applications in varying climates worldwide, boosting their appeal. These include grid-connected utility-scale PV plants on and off-shore, distributed PV systems in microgrids, and standalone residential contexts [4–6]. Additionally, technological developments have further enhanced solar PV implementations by enabling the synergy of state-of-the-art technologies with solar PV systems, namely battery storage [7,8], and information gathering and processing with internet-of-things (IoT) devices, big data centers, and artificial intelligence (AI) [9–11].

Variability and intermittence of solar power generation pose a significant challenge to widespread utilization of PV systems in the power grid. Weather condition (WC) volatility, especially cloud-layer coverage, presents a direct influence on global horizontal irradiance (GHI) capture, thus impacting the power generating capacity of a solar PV system. Consequently, there exists a large difference between solar PV power generation during clear days with minimal cloud coverage obstructions, and days with overcast, rain and/or foggy conditions. This resource variability has placed greater challenges in system operation and energy dispatch control, potentially threatening power grid reliability and stability, especially in regions with greater levels inverter-based resources featuring low-inertia power generation [12,13]. As a result, significant research spanning beyond the past decade has been devoted toward developing methods to obtain information sources to improve intelligence in solar PV system operation. Namely, solar PV power estimation and forecasting has developed into a substantial field, evolving with a multitude of varying methodologies. These include physics-based, statistical, sky-based imaging, and AI based learning models [14–18]. Regardless of the approach, a large portion of this literature focuses on predictive modeling driven by historical data.

It is noted that GHI exposure, and thus solar PV plant power generation, has a high correlation to the concurrent weather conditions, especially cloud coverage. Therefore, characterizing the relationship between GHI exposure and corresponding solar PV output with highly volatile cloud coverage is increasingly difficult. In surrounding literature, this relationship is widely observed. Consequently, some studies have targeted this relationship within their predictive models by further introducing weather classification studies. Generally, these methods utilize a preprocessing stage where historical data is analyzed, and a WC label is assigned to a time-series sample. Commonly, these studies utilize AI-based methods [19–25], or a lookup table [26–28] to make inferences on weather conditions for historical solar PV plant power generation. Other forms of WC based studies utilize numerical weather prediction (NWP) data provided by local or regional weather stations [29–32], or sky image-based analysis [33]. In all cases, a set of WCs are identified, often two, three, or four classifications, that are utilized to create sub datasets. Predictive models are trained independently on their corresponding sub dataset. Including a WC classification stage in preprocessing typically yielded higher accuracy and precision than that of a single, uniform model.

While WC classifications have been primarily utilized for improving solar PV plant power forecasting accuracy, other applications have emerged. For instance, in [34], an hourly weather status pattern recognition system is developed for measurement anomaly detection, and error recognition and mitigation in solar PV systems. However, in this surrounding literature, a few shortcomings with WC-based studies have been identified, as they can provide the additional value of massive data source in solar PV plant planning and operation. The first relates to the type of data utilized to determine WC classifications. Primary methods of data acquisition include historical PV plant power generation measurements with supervisory control and data acquisition (SCADA) or Phasor Measurement Units (PMUs), weather data collected at the site, and that provided by NWP. However, it is noted that the latter two data sources often contain deficiencies, either with WC identification definition, spatial-temporal resolution, and investment/operating costs [30]. On the other hand, utilizing historical solar PV plant power generation measurements allows a more targeted approach in terms of geographical scope, temporal characteristics, and high resolution

analysis.

Secondly, the nature of which classifications and labels are derived and assigned also pose challenges to the WC ideology itself. Significant efforts in literature have been devoted toward complex analysis for generalizing various weather conditions, besides the two most essential weather types, i.e. sunny and cloudy, based on a variety of environmental factors [35]. These include studies on cloud condition identification based on motion, formation and density, wind fields, temperature, precipitation, atmospheric pressure and event duration. In turn, a multitude of classification labels exist, such as rainy, snowy, overcast, foggy, partly/mostly cloudy, with some additionally including temperature (hot/cold) and duration (short-term/continuous rainfall) effects. Classification labels become more complex when considering seasonal impacts, like temperature, daylight duration, solar zenith angle, precipitation frequency, and sky conditions, and may greatly vary based on geographical location throughout the year.

On the other hand, generalizing WCs over 24 h intervals also restricts the level of situational awareness (SA) when considering real-time planning and operation of solar PV plants. WC labels are typically restricted to a 24 h period; very few studies have taken a more in-depth approach to higher resolution WC classification, i.e. intra-day contexts [35]. This can be attributed to factors such as slow data sampling rates causing data scarcity in intra-day contexts, and the type of application for the WC classification labels. However, focusing on the time interval of WC classification on historical solar PV power generation data can constitute massive source of knowledge considering near real-time planning and operating procedures. For instance, standardizing WC classifications based on an intra-day, intra-hour, or sampling rate basis can add great insights to PV plant power generation performance in ultra short-term contexts.

Considering the volatile nature of WCs, and the difficulty of assigning a descriptive label to accurately reflect the impact of WCs on PV power generation, the use of fuzzy inference systems (FISs) are explored in this study. FISs utilize fuzzy logic in an expert-driven modeling approach to assign classifications and particularly excel when distinct relationships between input and output variables are not present. There are six categories that feature FIS-type architecture applications with solar PV plants: solar irradiance forecasting, PV output power estimation, parameter identification for PV system sizing, maximum power point tracking (MPPT), inverter control and fault diagnosis [36]. FISs also provide a unique opportunity to expand on labels with an index-based system, rating outputs on a scale, [0,1], instead of subjective linguistic labels. Thus, the utilization of FISs for WC-type identification in a data-driven scene constitutes an intuitive approach for WC label assignment.

In this study, three new situational awareness indices (SAIs) are proposed to not only improve WC identification and classification, but to provide further insights on solar PV plant performance based on concurrent operational and environmental states. The methodology features a strictly data-driven approach, utilizing solar PV power generation data collected at the site for individual FISs. The SAIs specifically target an independent aspect relevant to near real-time planning and operation of solar PV plants as follows:

- Weather Condition Index (WCI): Reverse-engineers a rating for the concurrent WCs, specifically targeting sky-clearness, with respect to its impact on solar PV plant power generation. Thus, WCs may be inferred by the WCI at the highest spatial-temporal resolution, bypassing the inaccuracies of other methods. Use cases may include WC delineation in power generation datasets and WC-based performance evaluation algorithms.
- Operating Complexity Index (OCI): Takes into account the impact of both magnitude and frequency of power generation deviations to quantify the severity of non-ideal power generating circumstances on PV plant performance. The OCI provides a timely metric indicative of reductions or deficiencies in power generation. Use cases may include predictive assessments for power system voltage and frequency stability assurance.
- PV Generation Index (PVGI) provides a holistic comparison of power generation output with respect to solar PV plant performance and the power grid. The PVGI leverages real-time temporal impacts of solar PV power generation with the characteristics identified in the OCI to provide grid coordinators with a ‘usability’ determination of solar PV power. Use cases may include predictive assessments for energy dispatch and electric vehicle charge scheduling.

To summarize, with improved means of WC identification and overall solar PV performance evaluation, provided by the SAIs, information sources in operational control centers are enhanced with greater levels of situational awareness. With greater levels of SA, operators will better perceive, interpret and react to events, bolstering the stability, reliability and safety of solar PV plants in the power grid.

The primary contributions of this paper are as follows:

1. Three new situational awareness indices; WCI, OCI, and PVGI, are developed to characterize performance of solar PV plants. Each SAI utilizes PV plant power generation measurements and individual FISs to define a relationship between concurrent operational and environmental states with power generation measurements on a scale [0,1] via two representations, namely: i) as an averaged value over a 24 h interval and ii) as a real-time value expressed at the power generation sampling rate.
2. Solar PV power generation signatures (PVPSs) are derived for each of the SAIs with respect to an operating condition, indicative of typical PV plant performance based on the utilized SAI. A PVPS is a synthetic representation of expected power generation characteristics and performance for a given operating condition for a long-term timeframe, i.e. monthly, seasonally, yearly.
3. The proposed framework is implemented for a 1MWp solar PV plant located in Clemson, South Carolina, USA to demonstrate SAI capabilities. Utilizing the SAIs, applications to enhance solar PV system planning operation are presented and discussed.

The remaining sections of the paper are organized as follows: Section 2 describes the development and determination of SAIs, Section 3 describes how SAIs are utilized to perform classifications on solar PV power generation data for both 24 h and sample-wise contexts, Section 4 details implementation of the proposed framework utilizing Clemson University’s 1MWp solar PV plant as a case study, Section 5 presents results, discussions, limitations and applications, and finally, concluding remarks are provided in Section 6.

2. Situational awareness indices

In this section of the paper, a framework to determine the SAIs strictly utilizing solar PV plant generation data is proposed. The 24 h SAI representation, SAI_{24h}, provides a singular, generalized value to summarize the performance of a solar PV plant for a given day, with respect to the mentioned index. SAI_{24h} provides a foundation for further analysis in classifications, determining PVPSs, and sample-based SAI values, $SAI(t)$, as discussed in later sections of this paper. In Section 2.1, features describing PV power generation are derived, Section 2.2 introduces the Takagi-Sugeno fuzzy inference system, and Section 2.3 details the development of individual FISs for each SAI.

Fig. 1 depicts the general, comprehensive algorithm flow for determining SAI values and PVPSs. In the first stage, sufficient solar PV power generation data is collected from a solar PV site that is then preprocessed and stored in a historical database. Preprocessed data is utilized to calculate feature vectors that target specific characteristics and behaviors of solar PV plant generation over a 24 h interval. Feature vectors are then leveraged as inputs in an FIS tree to determine the corresponding SAI_{24h} values. The PVPS algorithm uses PV plant power generation

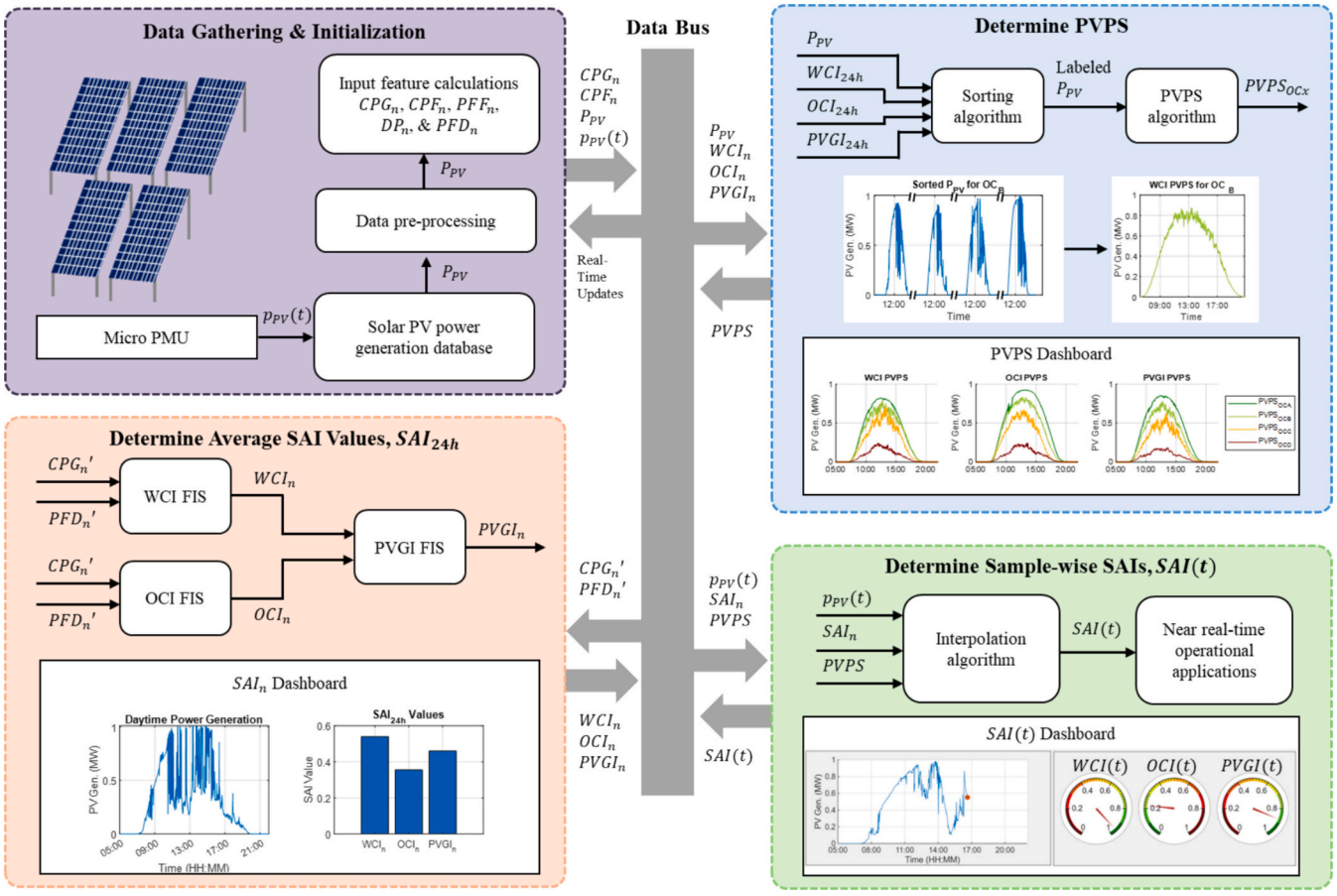


Fig. 1. Situational awareness index (WCI, OCI and PVGI) determination representation. Utilizing features calculated from historical PV power generation, SAI_{24h} are calculated, each with an individual FIS. PVPSs are generated using SAI_{24h} for offline analysis and are further utilized to determine sample-wise SAI values.

measurements with the SAI_{24h} values as inputs to both sort 24 h time-series power generation measurements and create the respective PVPS signatures based on the represented SAI and operating condition. Finally, sample-based SAI values are formulated as a function of power generation measurements, SAI_{24h}, and the previously calculated PVPSs. $SAI(t)$ presents a near-real time evaluation criteria that may be utilized throughout planning and operational applications for solar PV plants.

2.1. Feature calculation

The selected features are derived from a database of historical solar PV power generation. This allows for a highly descriptive, data-driven approach, as real-time influences from environmental factors are encompassed with the data. As opposed to utilizing site-collected environmental data or NWP, the historical power generation database provides the greatest definition and spatial-temporal resolution, while simultaneously capturing the effects of weather based features like solar irradiance variance and temperature. Two features, namely cumulated power generation (CPG), and power fluctuation density (PFD) are chosen to provide insights into the quality and quantity of solar power generation over a 24 h time period. After the features are calculated, both CPG and PFD are normalized according to a theoretical maximum value for each. As determined in empirical studies, these features provide sufficient power generation performance data, while maintaining minimal computational resource requirements.

To start, solar PV plant power generation for a given day, P_n, $n \in [1:N]$ for a total of N days, is characterized (1). Here, p_t represents measured power at the time, $t \in [1:T]$, over a 24 h interval.

$$P_n = \{p_1 \;\; p_2 \;\; \dots \;\; p_t \;\; \dots \;\; p_T\} \tag{1}$$

CPG is calculated as the arithmetic sum of measured data points (2). It provides insight into the quantity of power over the respective 24 h time interval, capturing the effects of cloud coverage, partial shading and other obstructions. Once normalized, the CPG operates similar to other metrics, such as a performance ratio or utilization factor [37]. Similar to CPG, both of these metrics express measured power over time, or energy, with a theoretical reference quantity. Days with the greatest CPG value are characterized by long intervals between sunrise and sunset, sustained peak generation at solar noon, and minimal power fluctuations throughout the day.

$$CPG_n = \sum_{t=1}^{T} p_t \tag{2}$$

The PFD feature characterizes power generation volatility for a given day. More specifically, PFD ensures that weather effects of irradiance variances and partial/sustained cloud coverage on factors like power ramp rates are captured. PFD is calculated as the proportion of two power fluctuation indicators to measure the magnitude and frequency of power generation volatility, namely cumulated power fluctuation (CPF) and power fluctuation frequency (PFF), divided by the number of samples collected during daylight-hours (DP) for that day (3). Including the DP factor ensures the PFD is not biased by different sampling rates and allows for accurate comparisons of PFD especially in instances where fluctuation behavior remains constant but for different intervals of time.

$$PFD_n = CPF_n * PFF_n / DP_n \tag{3}$$

The magnitude of power generation fluctuations is measured in CPF, calculated as the cumulation of the absolute value of the difference between two consecutive measurements (4). The frequency of power generation fluctuations is measured with power fluctuation frequency (PFF), indicating the number of dips (5). A dip in solar power generation is defined as when a power generation measurement is less than the preceding measurement before solar noon, or greater than the preceding power generation measurement after solar noon.

$$CPF_n = \sum_{t=1}^{T} |p_t - p_{t-1}| \tag{4}$$

$$PFF_n = \sum_{t=1}^{T_s} \mathbb{1}(p_t < p_{t-1}) + \sum_{t=T_s+1}^{T} \mathbb{1}(p_t > p_{t-1}) \tag{5}$$

where $\mathbb{1}(\cdot)$ is an indicator function, returning 1 if the condition is true, otherwise zero, and T_s is the timestamp corresponding to solar noon for a given day. Visualizations of CPG and PFF calculations for an arbitrary day of power generation measurements are provided in Fig. 2(a) and Fig. 2(b), respectively.

After calculating CPG and PFD features for all N days, a given feature, X (i.e. CPG or PFD), is normalized within the bounds $[0, 1]$ based on a feature scaling value, α_X (6).

$$X_n' = X_n / \alpha_X \tag{6}$$

Rather than simply utilizing the maximum feature value calculated, α_X is artificially generated, representing the maximum possible value for each feature. This allows for the addition of new data in the future, without necessitating constant renormalization of the entire dataset, in the event a new maximum feature value be calculated. α_{CPG} represents the maximum possible CPG for a given solar plant, calculated by determining the CPG of an artificially generated 24 h interval of ideal solar power generation on a day with maximum daylight-generating hours, i.e. $PFD = 0$ on the summer equinox. α_{PFD} represents a realistic maximum PFD value for a given solar plant, determined by assessing maximum possible values for both CPF and PFF.

The two resulting scaled feature vectors, CPG' and PFD' are characterized in (7.a) and (7.b), respectively.

$$CPG' = \{CPG_1' \;\; CPG_2' \;\; \dots \;\; CPG_n' \;\; \dots \;\; CPG_N'\} \tag{7.a}$$

$$PFD' = \{PFD_1' \;\; PFD_2' \;\; \dots \;\; PFD_n' \;\; \dots \;\; PFD_N'\} \tag{7.b}$$

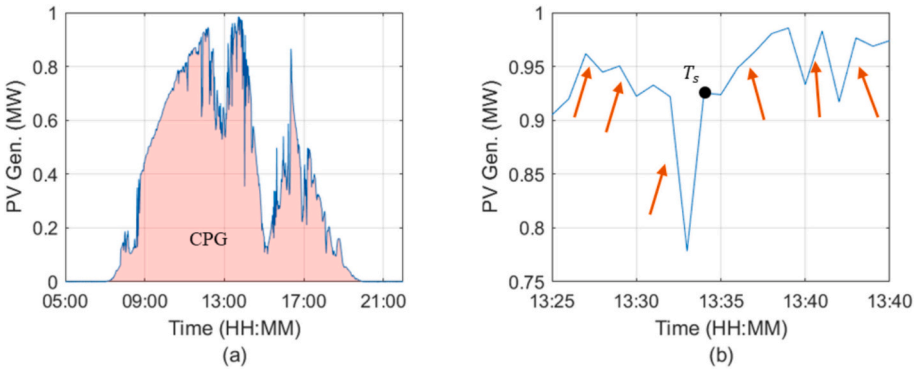


Fig. 2. Visualization of feature calculation for an arbitrary day of solar power generation. In (a), CPG is calculated as the total power generation. In (b), solar noon labeled as T_s, occurs at 13:34; examples of dips in power generation before and after are indicated by arrows.

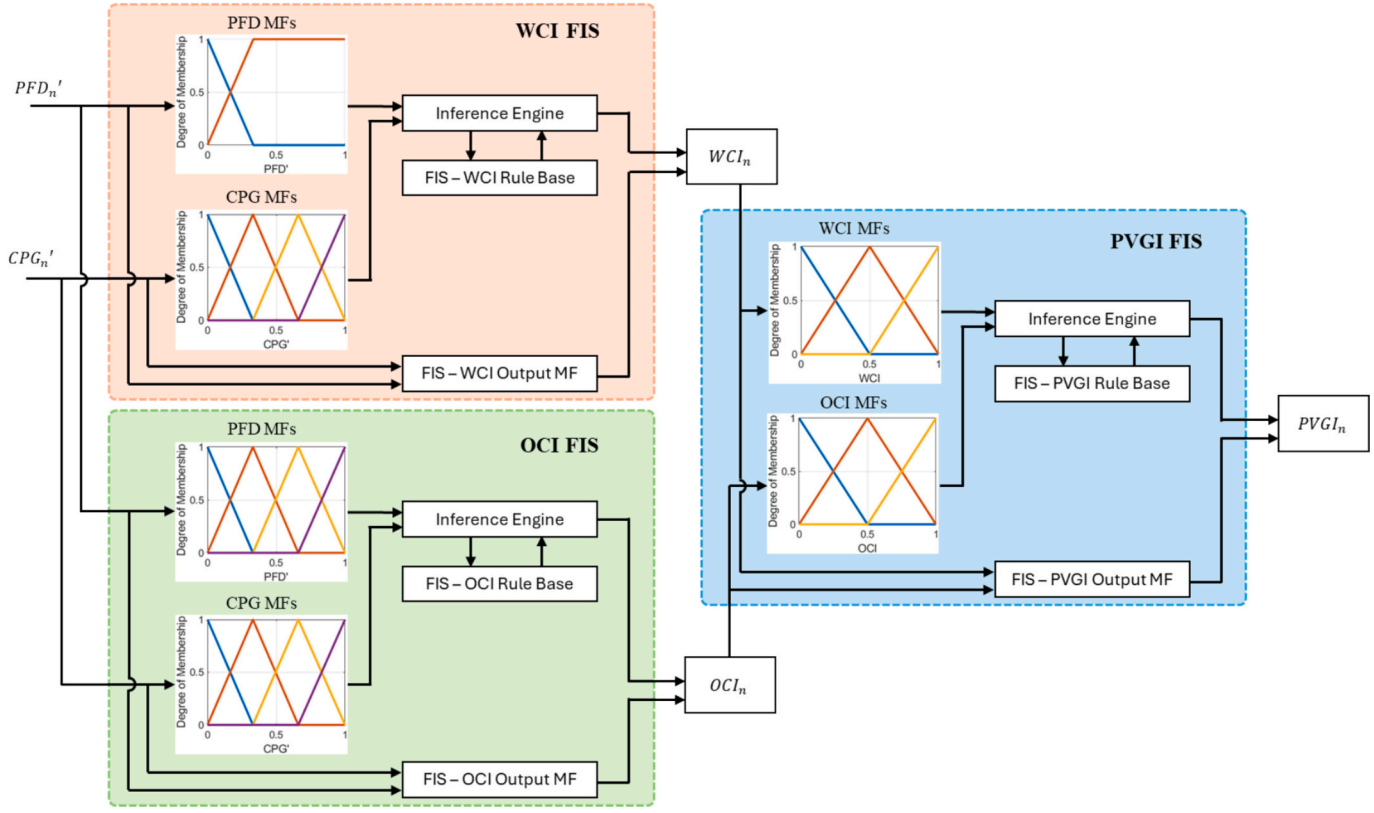


Fig. 3. Overview of three FISs to determine the SAIs, WCI, OCI and PVGI in the left-upper, left-lower and right portions, respectively. Each FIS contains the displayed input membership function sets, inference engine, rule base, and output membership function.

2.2. Takagi-Sugeno fuzzy inference system

In recent years, deployment of fuzzy logic based systems has significantly increased, spanning across a wide variety of industries for a plethora of different applications. These can include controls, scheduling and optimization, and signal analysis tuning and interpretation, among many others [38]. In particular, FIS is an ideal platform for considering classification type problems, where distinctions between different groups are often unclear. As opposed to classical sets, where elements are assigned a Boolean "1" or "0" label, fuzzy sets allow for partializing the degree of membership an element has to its related set. Therefore, fuzzy logic enables the representation of an input variable on a scale, e.g. [0, 1], rather than strictly Boolean values.

At a high level, FISs aim to analyze a set of crisp inputs and based on some set of rules, assign crisp values in an output set. The fuzzifier first converts the crisp input values into fuzzy sets using membership functions. A membership function provides a degree of similarity rating between the input and individual fuzzy subsets, commonly described by linguistic variables. The inference engine takes these fuzzy input sets and maps into fuzzy output sets, by referencing the rule base, to determine the degree of activation of each rule. These rules contain a collection of IF-THEN type statements that combine to define relationships between inputs and outputs. Finally, crisp outputs are obtained through mathematical output functions, as opposed to fuzzy sets, with a Takagi-Sugeno (TS) type FIS architecture.

The TS-FIS architecture provides an excellent option for characterizing the relationship between input variables (scaled feature vectors) and an output set (any of the three SAIs). For example, considering the case of weather condition classifications, it is very seldom that a given weather condition is exclusive to any particular classification. As in the case of a typical "partially cloudy day" varying densities of cloud coverage and cloud formations are a few of many factors that are represented in the level of cloudiness. Thus, for a given "partially cloudy

Table 1
WCI rule base and weights.

Rule No.	CPG'	PFD'	w	Value
1	L	L	VL	0
2	ML	L	H	1
3	MH	L	H	1
4	H	L	H	1
5	L	NOT L	L	0.2
6	ML	NOT L	ML	0.4
7	MH	NOT L	M	0.6
8	H	NOT L	MH	0.8

Table 2
OCI rule base and weights.

Rule No.	CPG'	PFD'	w	Value
1	L	L	VH	1
2	ML	L	MH_1	0.667
3	MH	L	ML	0.333
4	H	L	VL	0
5	L	ML	VH	1
6	ML	ML	MH_2	0.722
7	MH	ML	M_1	0.444
8	H	ML	L	0.167
9	L	MH	VH	1
10	ML	MH	MH_3	0.778
11	MH	MH	M_3	0.556
12	H	MH	ML	0.333
13	L	H	VH	1
14	ML	H	H	0.833
15	MH	H	MH_1	0.667
16	H	H	M_2	0.5

Table 3
PVGI rule base and weights.

Rule No.	WCI	OCI	w	Value
1	L	H	L	0
2	M	H	ML	0.25
3	H	H	M	0.5
4	L	M	ML	0.25
5	M	M	M	0.5
6	H	M	MH	0.75
7	L	L	M	0.5
8	M	L	MH	0.75
9	H	L	H	1

day" there exists a range of cloudiness levels that differentiate it from mostly cloudy and clear conditions. In this respect, the WCI utilizes the scaled features with a TS-FIS architecture to instead interpret a weather condition rating on a scale, rather than with discrete labels, further capturing weather condition dependent relationships within the solar power generation dataset. The same reasoning may be applied to the OCI and PVGI, where a multitude of external factors, present within the scaled feature inputs, impact the operating complexity solar PV plants and total grid impact of power generation, respectively.

2.3. FIS for WCI, OCI & PVGI

To determine an individual SAI, the FIS must be comprised of the following parts: two input membership function sets for fuzzification, a sufficient rule base to analyze all possible combinations of inputs, and an output membership function for defuzzification, resulting in the SAI. This process is generalized as a function of the corresponding FIS and inputs for WCI, OCI and PVGI in (8.a), (8.b) and (8.c), respectively.

$$WCI_n = FIS_{WCI}(CPG_n', PFD_n') \tag{8.a}$$

$$OCI_n = FIS_{OCI}(CPG_n', PFD_n') \tag{8.b}$$

$$PVGI_n = FIS_{PVGI}(WCI_n, OCI_n) \tag{8.c}$$

An overview of FISs with each step for the WCI, OCI and PVGI is shown in Fig. 3. Given that the input feature vectors are calculated over a 24 h period of time, the resulting indices calculated represent an averaged value for the corresponding time period. It is important to note differences in the input membership functions, rule bases, and output membership functions.

Fuzzification first occurs to determine the degree of membership of each input variable to a corresponding membership function. In this case, two, three, or four symmetric, triangular input membership functions are utilized to describe the input feature vectors. Each input membership function has an attached linguistic variable to describe its relative location, such as low (L), medium-low (ML), medium-high (MH), etc. It is noted that input values may be determined to have a non-zero degree of membership with respect to more than one membership function. It is in these cases where fuzzy logic excels, accounting for scenarios where variables may belong to more than one set.

For each input variable, the number of input membership functions (MFs) may vary for different FIS implementations. The main reasoning is mostly decided by the desired level of resolution, while balancing minimal computational requirements. Increasing the number of input MFs provides a more detailed comparison of input variables, as a greater number of possible combinations are available. However, as additional input MFs are added, the rule base exponentially expands, adding computational complexity. Therefore, the smallest rule bases are utilized, while simultaneously maintaining a sufficient resolution for each SAI to accurately fulfill their desired target.

The rule base of the FIS must contain an evaluation of each combination of input variables with their corresponding membership function possibilities. In this case, 2 input variables are present for each index. The largest rule base possible for each FIS is then calculated as the product of the number if input functions for each input variable, $2 \times 4 = 8$ rules for WCI, $4^2 = 16$ rules for OCI, and $3^2 = 9$ rules for PVGI.

To formulate the rule base, all possible combinations of each input membership function must be considered for each scaled input. Tables 1, 2 and 3 summarize the rule bases for WCI, OCI and PVGI, respectively, according to their input vectors, and resulting rule weight, w. To determine the rule base, first, the most optimal and least optimal input combinations for a given index are considered. Then, subsequent combinations of inputs may be interpreted based on the effects present in input combinations.

WCI provides a metric to infer weather and cloud coverage conditions based on power generation measurements. Large values indicate ideal WCs for power generation, and small values indicate sub-optimal conditions. Therefore, the PFD feature, indicative of power generation volatility, is a primary driver in deciding whether or not an arbitrary day maintains optimal solar PV power generation weather. Days exhibiting low PFD, given a sufficient level of CPG is maintained, are characterized as optimal conditions, as reflected in Rules 2–4. On the other hand, if significant levels of PFD are present (i.e. NOT L), then CPG is the deciding factor for the resulting WCI rating, seen in Rules 5–8. From the WCI, it is possible to interpret weather conditions, i.e. high scores indicate clear operating conditions with direct solar irradiance exposure. As cloud coverage increases, the resulting WCI decreases. Once CPG drops below a certain threshold, power generating conditions are at a minimum, reflected in a low WCI.

The OCI provides a metric to gauge the quantity and severity of power generation reductions that occur for a solar PV plant in a given day. Therefore, ideal conditions (high CPG and low PFD) are indicated by a low OCI score, as in Rule 4, and least ideal conditions (low CPG) are indicated by a high OCI score, as in Rules 1, 5, 9, 13. Considering the optimal case in Rule 4 as a starting point, OCI increases as a function of decreasing CPG and/or increasing PFD. Additionally, as power generation reductions due to volatility may occur as a combination of quick dips and/or larger sustained deviations, CPG is assigned a larger significance when determining the resulting weights for subsequent input combinations. This trend may be revealed when comparing the output values of certain rules, such as Rules 6 and 7, with Rules 10 and 11. In any case, the OCI indicates when operational complexity is heightened, due to sustained power generation reductions occurring, or power generation volatility, providing a degree of necessary response from operational control centers.

As the PVGI represents a holistic grid integration impact score for a solar PV plant, the two previously mentioned SAIs, WCI and OCI, are utilized as inputs. A high WCI, indicating clear-sky conditions, and a low OCI, indicating low operational complexity, is taken as the ideal PVGI inputs to result in the highest rating, as in Rule 9. Conversely, the opposite combination, i.e. minimal WCI and maximum OCI, represent the least optimal input combination for the PVGI in Rule 1. Starting from Rule 1, the PVGI improves for subsequent WCI and OCI input values as a function of either increasing WCI or decreasing OCI. Observing Rules 3, 5 and 7, an interesting trend emerges, where several different input combinations result in the same PVGI evaluation. This is attributed to the degree of separation that WCI/OCI have from the optimal case, i.e. H/H, M/M, and L/L all contain equivalent levels of difference.

The FIS calculates its respective 24 h SAI value as described in [39]. Utilizing the corresponding rule base, the inference engine calculates the rule weight, as in (9). Here, $F_1(\cdot)$ and $F_2(\cdot)$, represent a membership function for a given input variable.

$$w_n = AndMethod(F_1(\cdot), F_2(\cdot)) \tag{9}$$

The rule output level, z, is calculated as a linear function of input variables (10). Here, x and y are the values of both input variables, and a, b, and c are constant coefficients.

$$z_n = ax_n + by_n + c \tag{10}$$

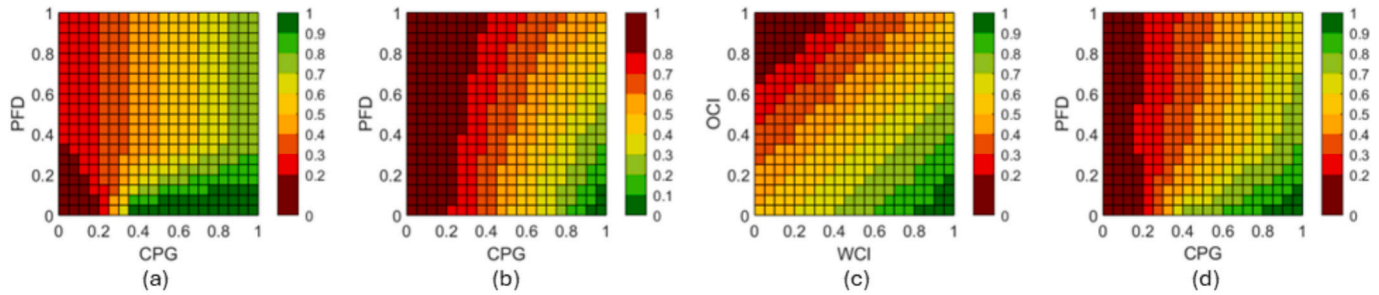


Fig. 4. SAI heatmaps generated from FISs (Fig. 3). (a) WCI heatmap – shows outputs based on CPG and PFD inputs, (b) OCI heatmap – shows outputs based on CPG and PFD inputs, (c) PVGI output heatmap – shows outputs based on WCI and OCI inputs, and (d) PVGI output heatmap – shows outputs based on CPG and PFD.

The resulting 24 h averaged index value for a given input set is taken as the ratio of the sums of all weighted averages and all rule outputs (11).

$$SAI_n = \sum_{n=1}^{N} w_n z_n \Big/ \sum_{i=n}^{N} w_n \tag{11}$$

Once the entire dataset has been considered, the resulting index vector may be characterized as in (12.a), (12.b) and (12.c).

$$WCI_{24h} = \{\, WCI_1 \quad WCI_2 \quad \dots \quad WCI_n \quad \dots \quad WCI_N \,\} \tag{12.a}$$

$$OCI_{24h} = \{\, OCI_1 \quad OCI_2 \quad \dots \quad OCI_n \quad \dots \quad OCI_N \,\} \tag{12.b}$$

$$PVGI_{24h} = \{\, PVGI_1 \quad PVGI_2 \quad \dots \quad PVGI_n \quad \dots \quad PVGI_N \,\} \tag{12.c}$$

For visualization, control surfaces detailing the combinations of the two scaled feature vectors and the resulting SAI are provided in Fig. 4. WCI is shown in (a) to illustrate the outcome of Table 1, OCI is shown in (b) to illustrate the outcome of Table 2, and PVGI is shown in (c) to illustrate the outcome of Table 3. Values of each SAI are interpreted by the respective color identified in the plot legend.

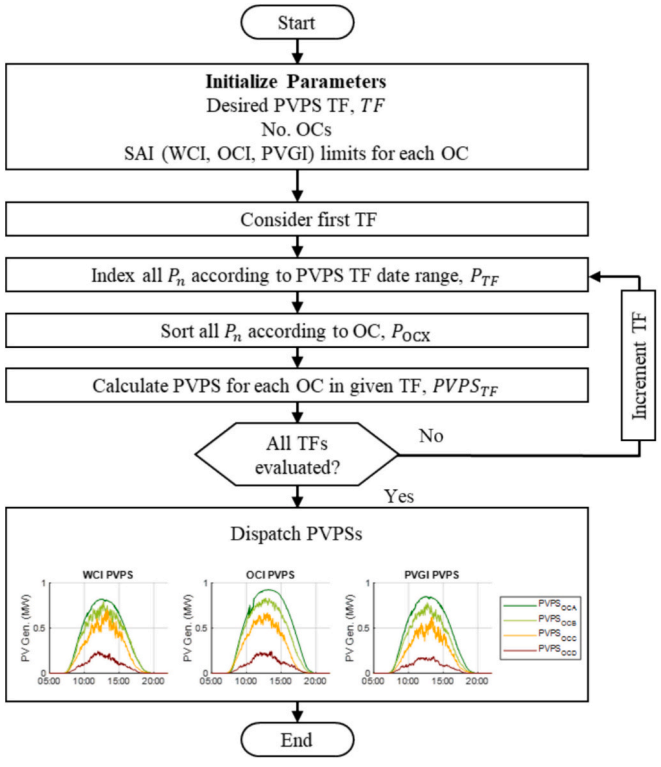


Fig. 5. PVPS determination flowchart representation for a user-selected PVPS TF and four identified OCs. PVPSs are calculated for all TFs and OCs.

Table 4
Lookup table for OC classification based on SAI thresholds.

SAI Limits	Condition
$\alpha_{A,min} < SAI_n < \alpha_{A,max}$	OC_A
$\alpha_{B,min} < SAI_n < \alpha_{B,max}$	OC_B
⋮	⋮
$\alpha_{X,min} < SAI_n < \alpha_{X,max}$	OC_X

3. Solar PV generation signatures & classifications

In many applications, the 24 h SAI representation provides sufficient resolution for WC-based classification and analysis. For instance, in the application of solar PV plant power generation forecasting, it is commonly observed that parallel ML or AI-based architectures are deployed on historical sub datasets with independent WC classifications. Alternatively, numerous other applications may require discrete labels attached to the time-series solar PV power generation representations on a revolving 24 h interval. In any case, for WC classification labels for 24 h time-series, the procedure requires establishing WCI maximum and minimum thresholds in a lookup table for each of the desired WCs, based on the previously discussed FIS rule base. To further include the OCI and PVGI, operating condition notation $(OC_A, OC_B, \dots, OC_X)$ will be utilized instead of WC to provide better generalization.

Alternatively, considering near real-time planning and operational aspects of solar PV plants, defining the solar PV power generation indices over a 24 h interval does not provide the necessary resolution. In this respect, PVPSs are proposed as a means for further solar PV plant performance analysis, and further, as a method for deriving SAIs on a sample-wise basis.

In this section, PVPSs are proposed to generalize solar PV generation according to an arbitrary operating condition utilizing the 24 h SAI representations. In essence, PVPSs are a synthetic representation of typical power generation characteristics, with respect to the SAI being utilized, over specified duration of time. More description on the development of PVPSs are provided in Section 3.1, and Section 3.2 details the process for sample-wise SAI derivation and classification.

3.1. Solar PV power signatures

Solar PV power signatures form the basis for which many operations, comparisons and analysis may be conducted with. PVPSs may be derived from a historical solar PV power generation database, utilizing any of the three SAIs to interpret power generation for the OCs. In simple terms, PVPSs represent a culmination of solar PV power generation profiles for a specified OC over a desired timeframe. When determining PVPSs, two parameters must be established, namely, the desired timeframe (TF) for which PVPSs will be considered, and the number of OCs to classify within the desired TF. Timeframes may be selected at an arbitrary length as to provide a temporal basis for classifications. For example, designations of yearly, quarterly, monthly, etc., TF intervals

may be chosen for this step. Additionally, the number of selected OCs decides the granularity and resolution of which to operate on the SAI. Fig. 5 represents a flowchart for the creation of PVPSs utilizing a SAI.

Given the desired signature TF, a subset of 24 h time-series historical power generation measurements are indexed, as shown in (13), where $n \in [1 : TF]$ total days within the TF range.

$$P_{TF} = \{P_1, \quad P_2, \quad \ldots, \quad P_{TF}\} \tag{13}$$

The procedure then follows a similar process to the 24 h designation for grouping OCs, by creating thresholds as in Table 4. By considering all the previously labeled timeseries data with idx_n, days may be sorted according to identified OCs over the selected TF. The process is repeated until all historical data has been considered. Following this process, resulting datasets are characterized in (14).

$$P_{TF} = \begin{Bmatrix} P_{OCA} = \{P_{A1}, \quad P_{A2}, \quad \ldots, \quad P_{AN}\} \\ P_{OCB} = \{P_{B1}, \quad P_{B2}, \ldots, \quad P_{BN}\} \\ \vdots \end{Bmatrix} \tag{14}$$

P_{TF} represents the matrix containing all sorted historical 24 h time-series solar PV power generation measurements for the given TF, P_{Xn} represents measured power generation P_n, for a day, n, sorted a given OC, X, and N represents the total number of days sorted into the respective OC.

To calculate the PVPS for a given TF and OC, each OC within a TF interval is considered independently. For a given OC, X, the PVPS is calculated as the point-wise arithmetic average of each measurement, across all sorted days (15).

$$PVPS_{OCX} = \frac{1}{N}\sum_{n=1}^{N} P_{Xn} \tag{15}$$

The result from repeating (15) for each OC within a signature TF is characterized in (16).

$$PVPS_{TF} = \begin{Bmatrix} PVPS_{OCA} = \{p_{A1}, \quad p_{A2}, \quad \ldots, \quad p_{AT}\} \\ PVPS_{OCB} = \{p_{B1}, \quad p_{B2}, \quad \ldots, \quad p_{BT}\} \\ \vdots \end{Bmatrix} \tag{16}$$

Here, each $PVPS_{OCX}$ contains a time-series vector for the average power measurements, p_{Xt}, for OC_X, where $t = [1 : T]$, polled measurements.

Accordingly, SAI values corresponding to an OC within the PVPS TF, idx_{TF}, are calculated by averaging all of the index values of corresponding days for the given OC and TF, as characterized in (17).

$$idx_{TF} = \begin{Bmatrix} idx_{OCA} = avg(\, idx_{PA1}, \quad idx_{PA2}, \quad \ldots, \quad idx_{PAN}\,) \\ idx_{OCB} = avg(\, idx_{PB1}, \quad idx_{PB2}, \quad \ldots, \quad idx_{PBN}\,) \\ \vdots \end{Bmatrix} \tag{17}$$

The PVPS timeframe designation allows for spatial-temporal characteristics within solar PV power generation to be further exploited. It is widely known that solar PV power generation is highly dependent on the climate and weather patterns associated with different regions around the world. For instance, significant differences between power generation are present when comparing a solar PV plant located in dry, arid regions seen in the Southwest, USA, and humid continental regions in the Northeast, USA. Incorporating TFs allows for the analysis of year-to-year, season-by-season, and even month-to-month PV power generation characteristics for any solar PV plant based on within its corresponding climate region. Depending on the TF interval chosen, some of these characteristics may be more evident than others. Comparatively smaller TF intervals, e.g. on a monthly or weekly basis, lead to more descriptive analysis at a higher resolution than comparatively longer TF intervals, e.g. seasonally or yearly, providing a more generalized approach. It is additionally important to note the relationship between data scarcity and TF interval. It is entirely possible that for a solar PV plant in a given region, some OCs occur at very low frequencies given the time of year. Expanding the TF interval may further include low frequency OCs for a given time of year, but at the loss of resolution and definition.

Selecting the number of OCs constitutes another important step in PVPS determination. In essence, the number of OCs will directly pertain to the resolution of which the corresponding index is applied. Increasing the number of OCs will decrease the range of Table 4 SAI thresholds. Consequently, with a greater number of specified OCs, higher resolution PVPSs can be obtained, but at the loss of generalization. Therefore, when deciding these parameters, it is crucial to be mindful of the inverse relationship between resolution and generalization.

3.2. Generating sample-wise SAIs from PVPSs

Determining sample-wise SAI values for time-series solar PV power generation measurements is a crucial step in utilizing these indices for near-real time planning and operating procedures. These sample-wise SAI values provide the highest definition analysis of solar PV power generation with respect to the individual index being utilized, as measurements can be examined at the frequency of the measurement equipment. In turn, this allows for solar PV plant power generation performance monitoring at a near real-time rate, further encompassing the highly volatile nature of solar PV power generation.

The proposed algorithm utilizes 24 h average SAI, PVPSs and interpolation to determine sample-by-sample SAI value for a given solar PV power generation measurement. The algorithm is represented in Fig. 6.

The PVPS TF for sample-wise SAI determination is utilized as a moving window, such that historical solar PV generation data is only considered over the TF range, with respect to the day of the current power measurement, n. As a result, (13) is modified, (18).

$$P_{TF} = \{P_n, \quad P_{n-1}, \quad \ldots, \quad P_{n-TF}\} \tag{18}$$

As previously discussed, P_{TF} is then sorted by OC, as in (14). By applying (15) and (16), signatures for the given TF are calculated with respect to d, and corresponding SAI values for this TF are calculated in (17).

Interpolating the sample-wise SAI value utilizes the calculated signatures, and resulting average SAI value for each OC. For a given measurement, p_t, Euclidean distances to the nearest two adjacent PVPSs

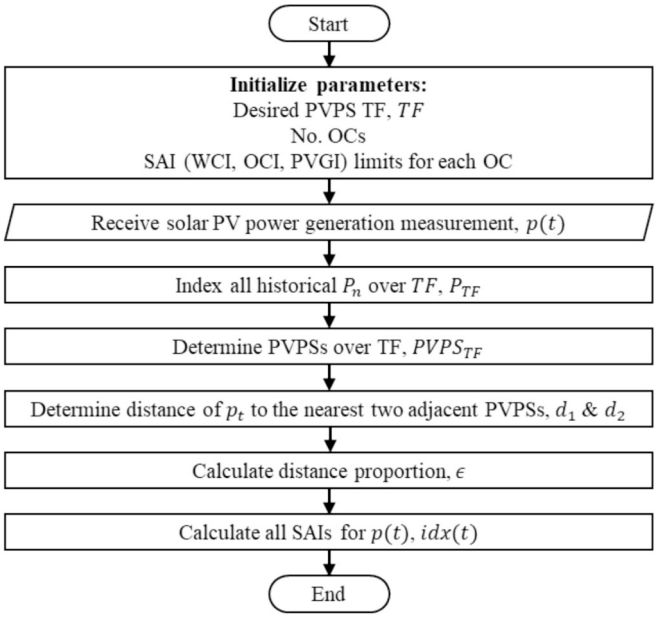


Fig. 6. Sample-wise SAI determination utilizing interpolation of PVPSs for a sample, p_t, where PVPSs are calculated over the TF range with respect to the day of the sampled power generation measurement.

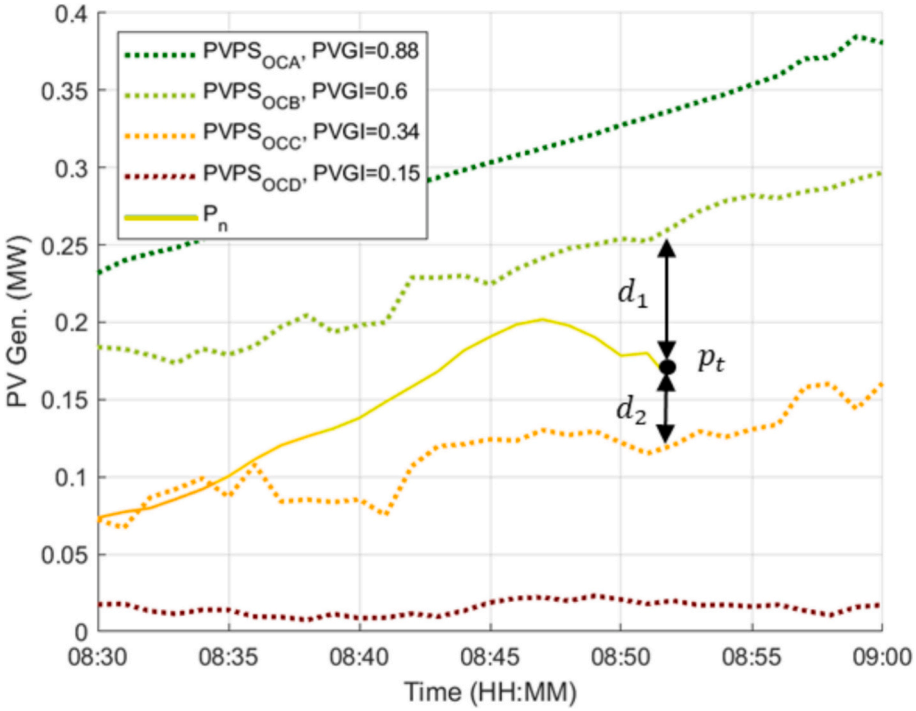


Fig. 7. Illustration of sample-wise interpolation for PVGI. For the power generation time-series, the nearest adjacent PVPSs for sample p_t, $PVPS_{OCB}$ and $PVPS_{OCC}$, are utilized for interpolation of the sample-wise SAI.

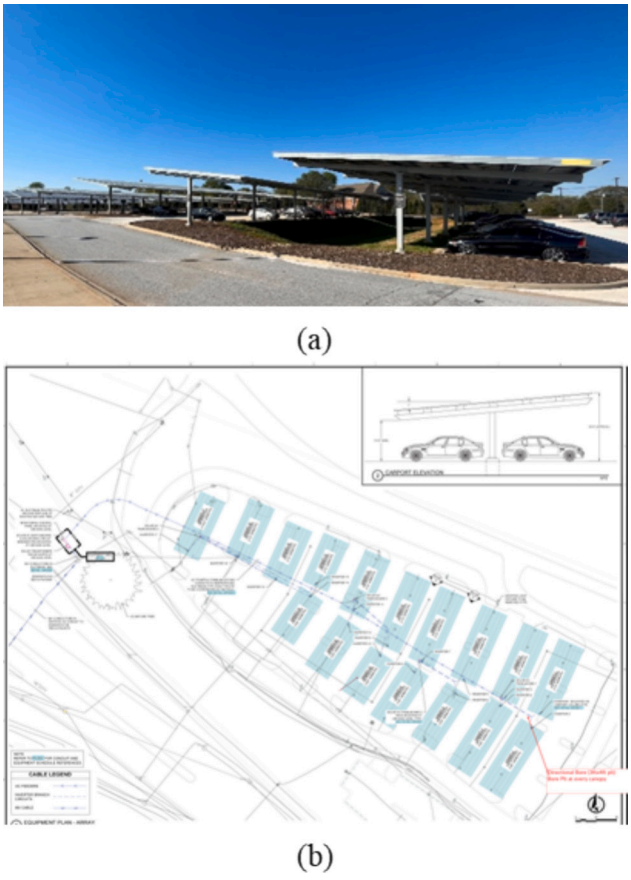


Fig. 8. Picture of the Clemson University 1 MW solar PV plant car park located at the R06 parking lot in (a). Layout of the 17-solar canopies and power flow in (b) [40].

for the same instance in time are calculated (19).

$$d_i(t) = |p(t) - PVPS_{OCi}(t)| \tag{19}$$

Next, the proportion of the distance between p_t and the two adjacent PVPSs is calculated (20).

$$\epsilon(t) = d_1(t)/d_1(t) + d_2(t) \tag{20}$$

The resulting SAI value for p_t, $SAI(t)$, is then interpolated as a function of the distance and the average SAI values of the two adjacent PVPSs (21).

$$SAI(t) = SAI_{OC1} - \epsilon(SAI_{OC1} - SAI_{OC2}) \tag{21}$$

In the event that p_t is either greater than or less than all corresponding values of PVPSs at the same instance, distances may be interpolated between maximum/minimum power generation with injected values at 1 and 0, respectively.

Fig. 7 contains a representation of the interpolation. As pictured, PVPSs are determined for four OCs based on PVGI, and the distances are calculated between $PVPS_{OC1} = PVPS_{OCB}$ and $PVPS_{OC2} = PVPS_{OCC}$.

4. SAI implementation

This section presents the implementation of SAIs for a 1MWp solar PV plant located in Clemson, South Carolina, USA. The solar PV plant contains 17 solar arrays arranged in canopies operating as a multi-functional carpark. Individual canopies contain anywhere from 138 to 204 modules, depending on their size, tilted anywhere in the range of 113° to 123° azimuth, and a power inverter. Electricity from each canopy is transmitted to an on-site junction point, where a micro-PMU measures generated AC power. A picture of the solar PV plant is shown in Fig. 8(a), and schematic diagram provided in Fig. 8(b).

A micro-PMU is a device that measures voltage, current, and power phasors at a high spatial-temporal resolution for distribution grids. Each measurement is synchronized both with GPS and a universal clock, ensuring the highest precision. Collected data is then archived with cloud storage software openHistorian. In this case, measurements are

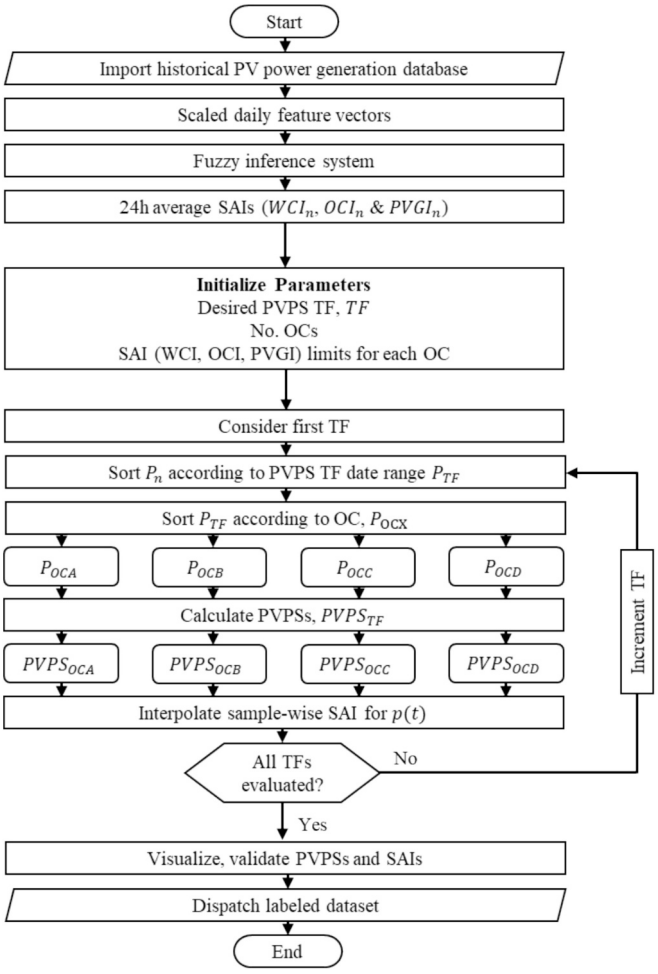


Fig. 9. Full implementation of the proposed framework to determine WCI, OCI and PVGI on both 24 h average and sample-wise contexts based on four OCs.

Table 5
WCI, OCI & PVGI Ranges for OCs.

WCI_n Range	OCI_n Range	$PVGI_n$ Range	Operating Condition
$0.875 \leq WCI \leq 1$	$0 \leq OCI \leq 0.25$	$0.75 < PVGI \leq 1$	OC_A
$0.625 < WCI \leq 0.875$	$0.25 < OCI \leq 0.5$	$0.5 < PVGI \leq 0.75$	OC_B
$0.375 < WCI \leq 0.625$	$0.5 < OCI \leq 0.75$	$0.25 < PVGI \leq 0.5$	OC_C
$0 \leq WCI \leq 0.375$	$0.75 < OCI \leq 1$	$0 \leq PVGI \leq 0.25$	OC_D

polled at the rate of 1 sample per minute, dating from January 1, 2024, to December 31, 2024.

Full implementation of the proposed algorithm is presented in Fig. 9, summarizing the combination of previously discussed steps. For the implementation detailed, four OCs are utilized, with corresponding ranges outlined in Table 5. Operating condition ranges for WCI, OCI and PVGI are determined by analyzing the FIS rule base, determined in empirical studies, to provide the best fit for the given data.

5. Results and discussion

This section provides some visualizations, comparisons, and evaluations of the SAIs for the SAI_{24h}, PVPS, and $SAI(t)$ representations. In Section 5.1, 24 h average SAI values are presented for a diverse selection of power generation characteristics. Section 5.2 provides visualizations and discussions of quarterly PVPS calculated for the entire dataset. In Section 5.3, sample-based SAIs are presented and evaluated. Section 5.4 details some discussions on the applicability and limitations of the SAIs in relevant areas.

5.1. 24 h SAIs

The 24 h WCI, OCI, and PVGI representations provide a generalized metric to quantify solar PV power generation characteristics and behaviors over the duration of a singular day. To visualize these representations, eight days of solar PV power generation are plotted in Fig. 10 (a)–(h). It is important to note the specific power generating qualities of these days independently, such as daylight duration, peak power generation, and magnitude/quantity of power generation fluctuations, as each of these properties influence the SAIs in different manners. Specific WCI_n, OCI_n, and $PVGI_n$ values are provided in Fig. 10(i) for each of the respective subplots.

Some patterns clearly emerge while analyzing Fig. 10. For instance, it is generally supported through visual inspection that WCI, OCI, and PVGI are at optimal values for optimal power generation weather, and accurately reflect non-optimal conditions in their respective manners. However, evaluating each of the SAIs for a non-optimal time-series power generation measurements provides crucial aspects to points of analysis.

As the WCI primarily targets the clearness aspect of a given day, this trend is accurately supported by visual inspection of Fig. 10. For instance, with Fig. 10(a) and (b), two distinctively clear days of power generation are plotted with a corresponding high WCI rating, indicated in (i). Additionally, upon comparison of (c) and (d), it is evident that (d) experiences greater levels of sustained power generation deviations from an arbitrary clear-day reference, that is also accurately expressed in the WCI_n. This trend may be further expanded through the remaining subplots, such that, as cloud coverage increases and resulting sky-clearness dissipates, i.e. (e), (f), (g), and (h), the WCI performs as expected with corresponding decreasing scores.

The OCI is utilized to quantify the magnitude, frequency and duration of power generation deviations, all of which are reflected in the metric visualization of Fig. 10. While (a) and (b) are both rated nearly optimal WCI scores, the OCI reveals a different comparison. It is evident that (a) experiences much shorter daylight duration and features peak reduction much more significant from the 1 MWp than experienced(b), demonstrated by substantially different OCI scores. This phenomena likely occurs due to differences in power generation characteristics between difference seasons. In this case, the OCI reveals that (a) is considerably *more challenging* from an operator's perspective than (b) due to this comparison. In a different observation, OCI_n in (a) and (d) are approximately similar values, despite quite different power generation characteristics. In this case, OCI_n is impacted in (a) due to the overall power generation experienced throughout the day, while OCI_n in (d) is impacted due to the high levels of volatility and cloud coverage in the initial portion of the day. Subsequent evaluations of OCI_n for the remaining examples follow a similar and expected trend.

As the PVGI features a combination of both WCI and OCI evaluations, it is expected to follow similar, yet diluted patterns that are prevalent within both indices, but some important distinguishments should be noted. For instance, $PVGI_n$ in (a) takes into account the near-perfect sky-clearness, while also reflecting the substantial reduction in overall power generation in the resulting score, thus providing a balanced metric, not too heavily weighed by a singular factor. As anticipated, the highest $PVGI_n$ occurs in (b), as shown by an appropriate allowance of power generation peak behavior/interval and marginal fluctuations. As shown in (c)–(h), $PVGI_n$ again follows an expected trend to decrease as a function of decreasing overall power generation and increasing power generation intermittence.

While evaluating SAI_n for specific days helps expose many of the tedious properties that exist with solar PV power generation behavior, it is hard to dismiss many of the similarities that emerge between the WCI, OCI and PVGI. For each of the indices, optimal and least-optimal power

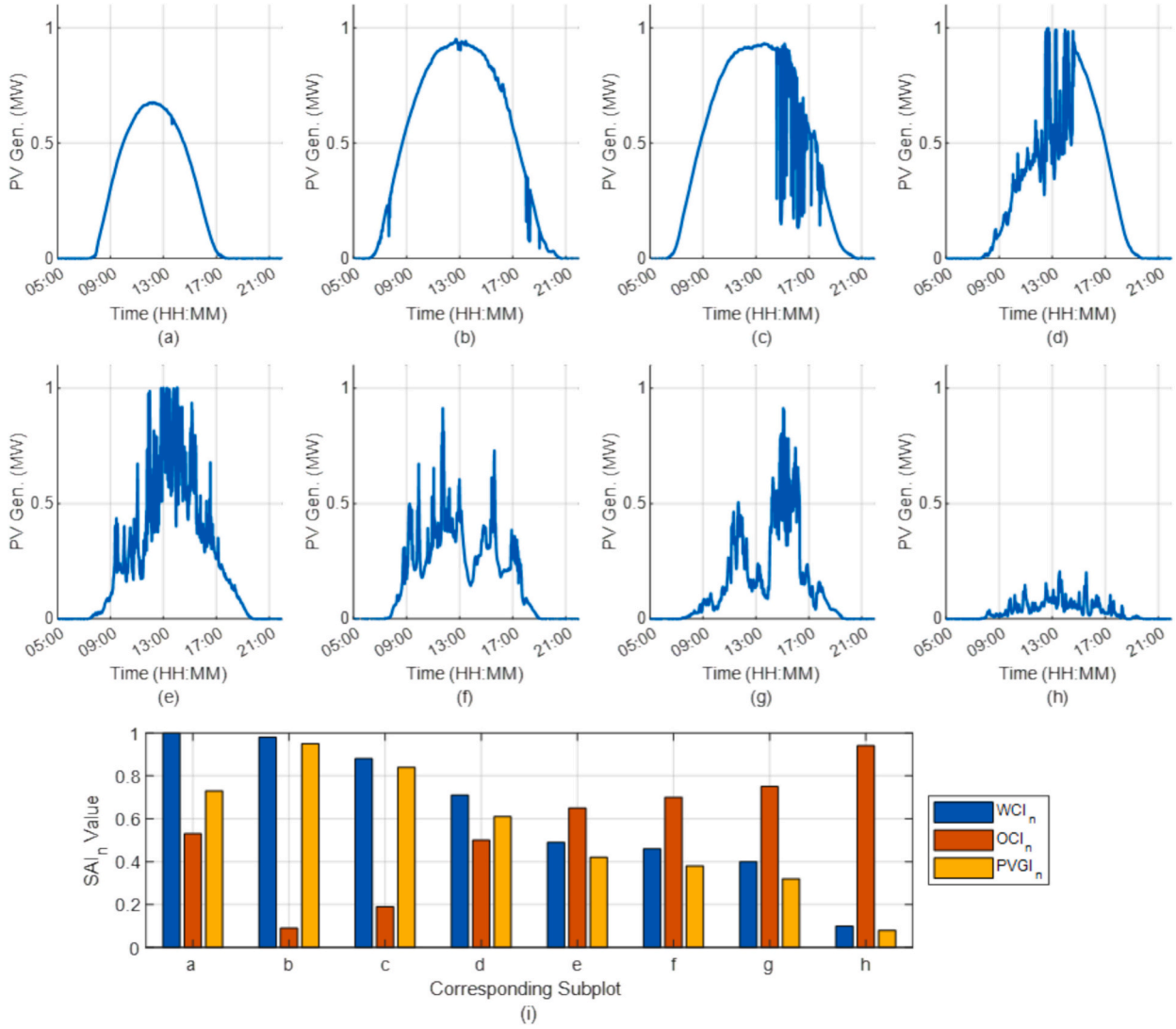


Fig. 10. Eight days of solar PV power generation, each representing different, distinctive power generation characteristics in (a)–(h). Respective SAI_n values for each subplot are represented by the bar-value in (i).

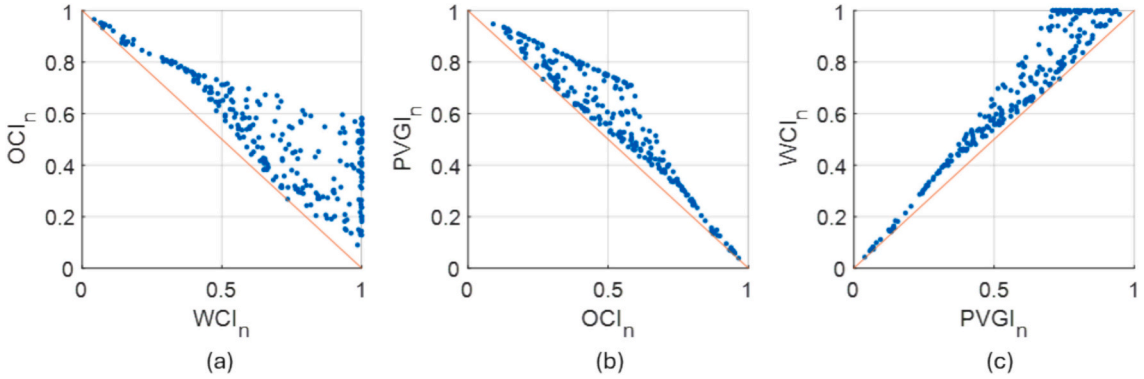


Fig. 11. WCI_n and OCI_n correlation in (a), OCI_n and $PVGI_n$ correlation in (b), and WCI_n and $PVGI_n$ correlation in (c) for the entire historical dataset. Additionally included is the ideal correlation line.

generation occur for similar conditions, which is a property inherent to solar generation itself. However, the SAIs target different power generation characteristics with varying degrees of significance that are particularly revealed during non-optimal power generation conditions, i.e. with some level of power generation intermittence. Fig. 11 displays correlation plots for WCI_n, OCI_n, and $PVGI_n$, where each of the plotted point represent the corresponding SAI_n combination for an arbitrary day.

Fig. 11 reveals that the most significant differences between SAI ratings exist for every case when non-optimal power generation

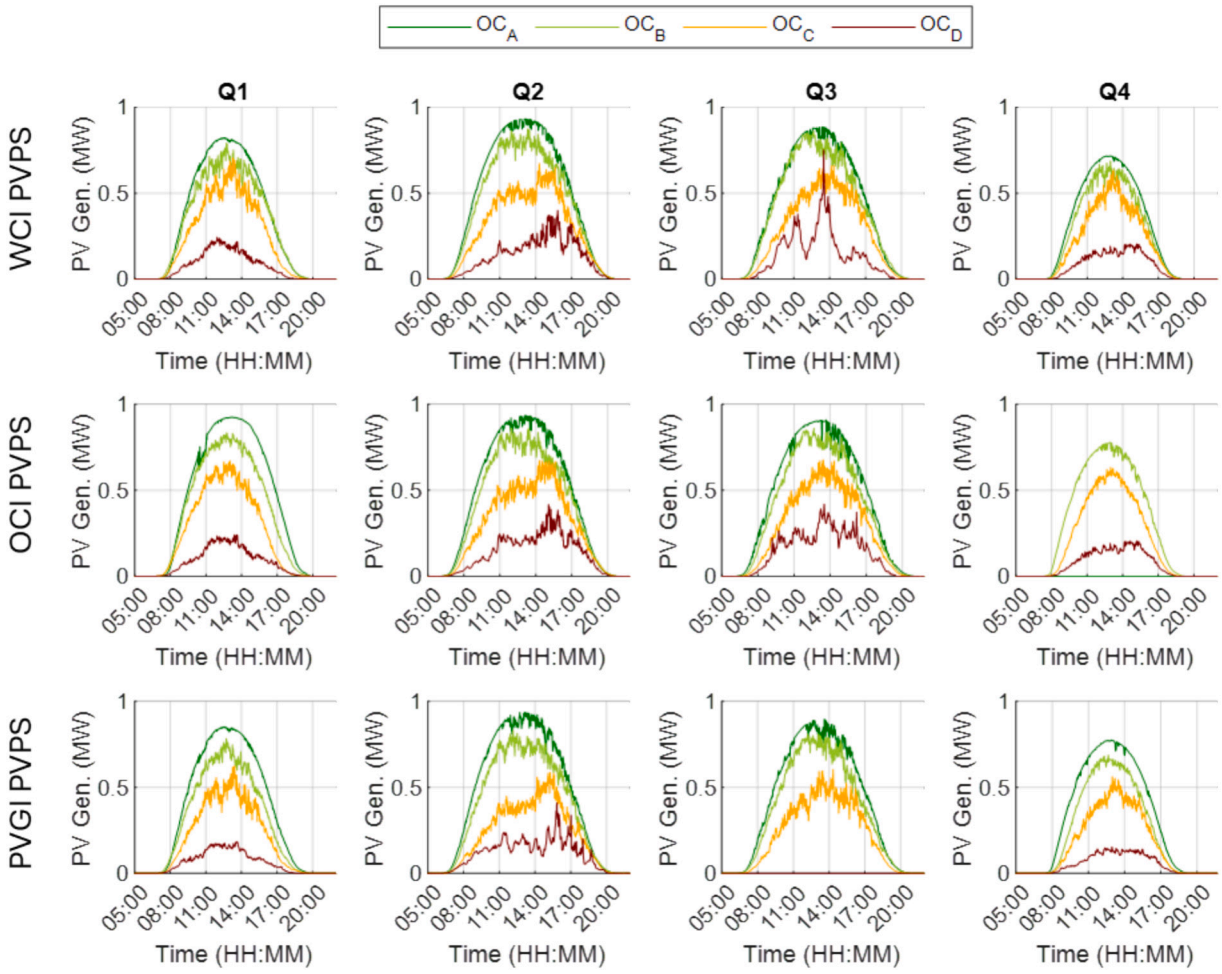


Fig. 12. Quarterly PVPS, SAIs are organized by row, and quarters of 2024 are organized by quarters, Q1, Q2, Q3, Q4, and are plotted in dark green, light green, yellow, and red colors for operating conditions, OC_A, OC_B, OC_C and OC_D, respectively.

Table 6
Quarterly PVPS average SAI.

TF	WCI				OCI				PVGI			
	OC_A	OC_B	OC_C	OC_D	OC_A	OC_B	OC_C	OC_D	OC_A	OC_B	OC_C	OC_D
Q1	0.98	0.74	0.5	0.19	0.17	0.34	0.61	0.84	0.81	0.63	0.39	0.14
Q2	0.96	0.73	0.53	0.32	0.15	0.38	0.58	N/A	0.85	0.63	0.39	N/A
Q3	0.95	0.74	0.53	0.34	0.18	0.36	0.55	N/A	0.80	0.62	0.41	N/A
Q4	0.98	0.76	0.47	0.18	0.21	0.34	0.58	0.85	N/A	0.62	0.40	0.13

Table 7
Quarterly PVPS sorted proportions.

TF	WCI				OCI				PVGI			
	OC_A	OC_B	OC_C	OC_D	OC_A	OC_B	OC_C	OC_D	OC_A	OC_B	OC_C	OC_D
Q1	0.37	0.16	0.24	0.23	0.30	0.40	0.18	0.12	0.18	0.46	0.21	0.16
Q2	0.20	0.38	0.35	0.07	0.38	0.45	0.17	0	0.38	0.42	0.20	0
Q3	0.20	0.44	0.35	0.02	0.40	0.45	0.15	0	0.40	0.44	0.16	0
Q4	0.40	0.19	0.20	0.21	0.19	0.53	0.16	0.13	0	0.57	0.29	0.14

conditions occur. Specifically in Fig. 11(a) and (c), this trend becomes more obvious for WCI/PVGI values greater than 0.5 and OCI values less than 0.5. In Fig. 11(b), differences in OCI and PVGI values exist in a slightly different area on the solution plane, primarily existing around the OCI range [0.3, 0.7] and PVGI range [0.8, 0.4]. In this area, the greatest differences between input features (CPG and PFD) exist, as volatility occurs for a multitude of reasons, including the frequency/magnitude of power generation fluctuations, and its resulting impact on power generation reductions. These relationships are further magnified when considering changes in seasonal weather patterns and frequency of certain operating conditions.

An exception to this observation is with the WCI, where a large portion of the data is populated for near ideal scores at the maximum boundary position, seen in both (a) and (c). This relationship follows a similar approach to the analysis between Fig. 10(a) and (b), where optimal sky conditions were present, but substantial differences in peak power generation values and daylight intervals were recorded. It becomes evident that the WCI is more inclusive on identifying what days exhibit optimal power generation conditions, as opposed to the stricter optimality identification present with the OCI and PVGI.

5.2. Solar PV power signatures

Visualizations for the WCI, OCI and PVGI PVPSs calculated for a quarterly TF of the 2024 year are shown in Fig. 12. Each quarter, i.e. Q1: Jan. 1–Mar. 31, Q2: Apr. 1–Jun. 30, Q3: Jul. 1–Sep. 30, Q4: Oct. 1–Dec.

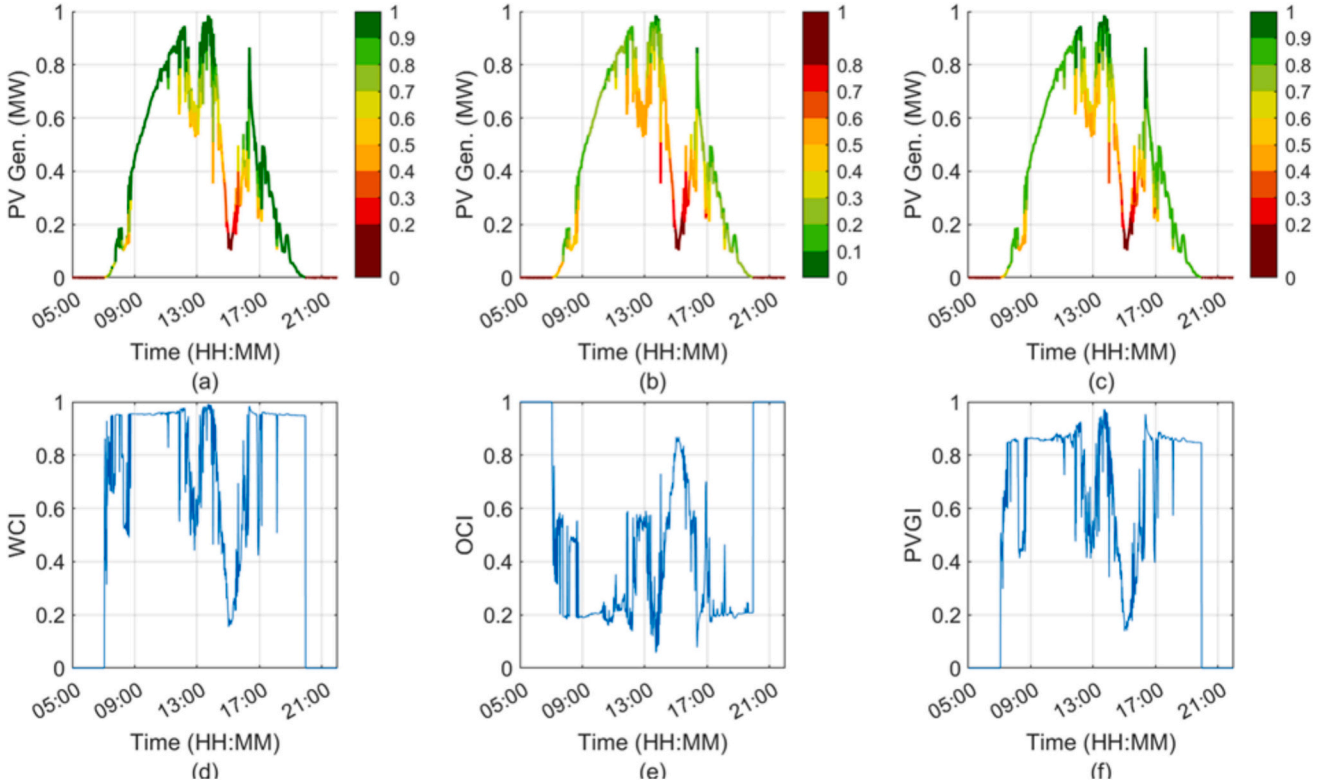


Fig. 13. Measured PV power generation for a single day, with $WCI(t)$, $OCI(t)$, and $PVGI(t)$ values indicated by color in (a), (b), and (c), respectively. $WCI(t)$, $OCI(t)$, and $PVGI(t)$ plotted as a function of time for the same power generation data in (d), (e), and (f), respectively.

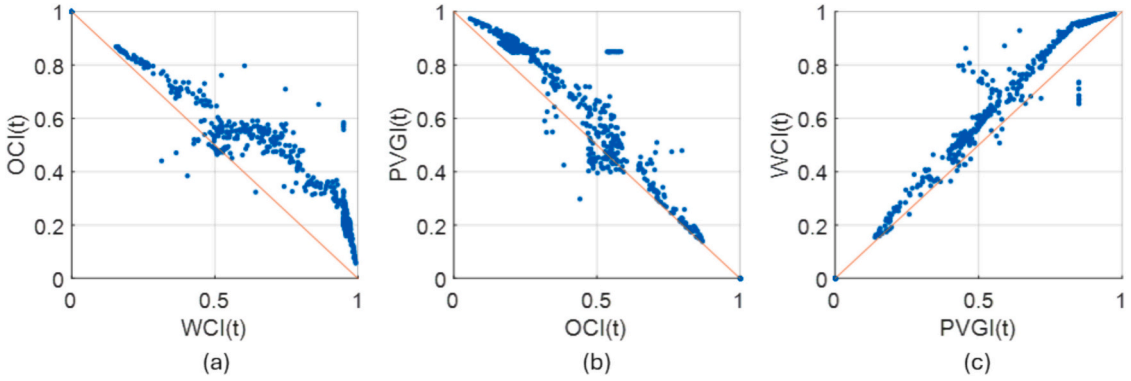


Fig. 14. $WCI(t)$ and $OCI(t)$ correlation in (a), $OCI(t)$ and $PVGI(t)$ correlation in (b), and $WCI(t)$ and $PVGI(t)$ correlation in (c) for the plotted day of power generation of Fig. 13. Additionally included is the ideal correlation line.

Table 8
SAIs applied for near real-time solar PV plant planning and operational applications.

SAI	Interpretation	Application	Ref
WCI	• Precise weather condition inference • Discrete weather condition labelling	• Dataset division for AI-based power estimations/predictions with parallel/ensemble architectures based on weather condition • Parameterizing solar PV performance based on concurrent weather conditions	[28]
OCI	• Integration complexity based on volatility and actual power magnitude	• PV power generation estimation/prediction assessments • Power system bus voltage and frequency regulation • Electro-mechanical oscillation assessment	[28,41,42]
PVGI	• Total impact/value of PV power generation	• Dispatch controllers for demand response and electric vehicle integration requirements • Power system bus voltage and frequency regulation • Automated generation control (AGC) tuning • Anomaly detection in measurement equipment	[41–47]

31, is organized by column, and rows 1–3 show plots for the calculated WCI, OCI and PVGI PVPSs for the described four operating conditions, respectively.

Analyzing the PVPSs quarterly timeframe, trends relating to seasonal behaviors emerge in solar PV power generation. Some key features include daylight generating hours, peak power generation, and the sparsity of PVPSs for some OCs. It is additionally important to remember that PVPSs represent averaged measurements for each sample instance. As a result, the magnitude and frequency of power generation fluctuations, especially in OC_B and OC_C, are diluted. Averaged index values of WCI_{24h}, OCI_{24h}, and $PVGI_{24h}$ for each PVPS are presented in Table 6, and calculated proportions of the number of classified days within each operating condition for each quarter are presented in Table 6.

Observing the visualizations, it is easiest to infer power generating characteristics of OC_A. With the WCI, typical clear weather conditions are expressed by OC_A. Both daylight power generating hours and peak power generation of clear days are significantly less for Q1 and Q4 compared to Q3 and Q4. These two factors are primarily represented in the CPG feature vector, as the number of daylight power generating hours and peak power generation directly influence the total power generated over a given interval. However, as defined in the WCI rule base, OC_A is primarily determined by PFD, given a sufficient level of CPG, hence the seasonal characteristics are interpretable. On the other hand, CPG has a significant influence on the OC_A classification for OCI. Therefore, days classified within OC_A must contain a tighter correlation of similar features, regardless of the time of year. The same trend is observed in the PVGI in all quarters, except for Q4, primarily due to the fact CPG is significantly lower during that time of year. Therefore, during Q4, power generating characteristics of OC_A never occurred.

As for the remaining OCs, most of their characteristics are clearly differentiable for each index across different quarters. Generally, OC_B contains a slightly smaller CPG value, but with greater PFD, reflected by high power generation at peak times with notable fluctuations. As mentioned, since the PVPS visualizations are average measurements across all timestamps for each operating condition, the fluctuation frequency and magnitude are diluted, but overall power generating characteristics are preserved. Similar trends appear for OC_C, but with comparatively smaller CPG values. With OCI and PVGI, it is observed that OC_D, does not have any classified instances during Q2 and Q3 timeframes. This can be attributed to the fact that during these quarters, the power generating characteristics never fall within the OC_D classification, thus becoming an infrequent operating condition.

Analyzing Table 6, the presented average index values for each SAI for each OC accurately reflect the thresholds determined previously in Table 5.

Analyzing Table 7, inferences on the frequency of sorted OC types may be inferred. From this analysis, it is possible to determine what operating conditions occur at a certain frequency for a given time of year, with respect to each index.

Additional PVPS determinations for monthly timeframes with corresponding SAI values and sorted proportions are provided in Appendix B.

5.3. Sample-wise SAIs

Sample-wise SAIs provide the highest resolution available for analyzing solar PV power generation characteristics and provide unique opportunities for in-depth analysis. Provided in Fig. 12 is a visualization of sample-wise SAI quantities. In the upper row, Fig. 12(a)–(c), a single day of solar PV power generation is selected and plotted. Each power

generation measurement is color coordinated with the corresponding $WCI(t)$, $OCI(t)$, and $PVGI(t)$ values indicated in the respective color legend. In the lower row, Fig. 12(d)-(f), $WCI(t)$, $OCI(t)$, and $PVGI(t)$ are plotted as a function of time for the same power generation data.

For the day of power generation measurements of Fig. 13, a wide variety of characteristics are present. This is represented by the relatively smooth, predictable segments in the first portion of the day, followed by abrupt volatility and a very significant drop in power generation around 15:00, while later returning to a preferred generation profile. As expected, the $SAI(t)$ representations follow somewhat expected trends, but with some few key differences.

For the WCI, the first and last portions of the day are identified as being optimally clear in Fig. 13(a), represented by the dark green data points, and further validated by $WCI(t)$ representation in (d). $WCI(t)$ also accurately reflects differences in power generation behavior during the mid-day power generation drop. Accordingly, OCI and PVGI indicate similar results, but at varying values. For instance, $OCI(t)$ and $PVGI(t)$ indicate some degree of non-optimality during the first and last portions of the day with average values of approximately 0.2 and 0.85, respectively. As opposed to the nearly optimal $WCI(t)$ average of 0.97 during the same periods, some PV power generation shortcomings were relevant. In this case, the added value of the SAIs from an operational perspective come from the higher definition analysis of data points.

This observation is further justified by analyzing the $SAI(t)$ correlations of Fig. 14, where a trend similar to that observed in Fig. 11 occurs. Shown in Fig. 14 are the $SAI(t)$ correlation plots for the power generation data of Fig. 13. Considering the most and least optimal combinations of $SAI(t)$ correlations, it is expected to have a degree of correlated results. However, it is in the middle ranges of $SAI(t)$ the lowest levels of correlation occur, indicating that SAIs are independently indicating their targeted metric.

As the sample-based SAI representations provide the highest level of resolution for analysis, it is more appropriate to evaluate $SAI(t)$ to a higher precision. To expand, $SAI(t)$ enables the evaluation of individual power generation measurements, rather than emphasizing the generalization of an entire day of volatile power generation to a strict, discrete label. Thus, for an arbitrary power generation measurement and/or interval of power generation measurements, the SAIs add value in the form of near real-time, time-dependent information. This is particularly evident in two example scenarios: a. gauging rising edge (morning) or falling edge (evening) solar PV plant power generation performance and b. quantifying the impact of any power generation intermittence, further factoring influences of magnitude, duration or frequency at which they occur. For the first case, when power generation is climbing in the morning, or dropping in the evening, it can be difficult to distinguish the operational implications, or impact of environmental conditions during these times of rapid dynamics strictly with a power generation measurement. In the second case, SAIs can well-exploit severity of sudden generation swings and be utilized to formulate responses given any degree of volatility in solar PV power generation.

Furthermore, as sample-wise SAIs are derived from both SAI_{24h} and PVPSs, the impact of longer-term dependencies may be further included at the discretion of the user. For instance, PVPSs may be calculated for TFs ranging potentially from days to years of historical data. Thus, the influences of monthly, seasonally, and yearly weather patterns for a given region, at the highest spatial-temporal resolution may be further encapsulated in the classifications given for $SAI(t)$.

To provide a metric for gauging algorithm runtime, the sample-wise SAI determination is timed on a desktop workstation with the following components: Windows 11 Enterprise with 12th Gen Intel i7-12700KF CPU, NVIDIA RTX 4500 GPU and 64GB RAM capacity. Once the 24 h average SAI and PVPS interpretations are established in the offline algorithm, the sample-wise SAIs take approximately 0.95 ms calculation time per sample.

5.4. Applications & limitations

It is shown that WCI, OCI and PVGI provide massive insights into the operating conditions of a solar PV plant, for both very short (sample-wise) to longer time intervals (24 h). These insights may be applied throughout PV plant planning and operational contexts, some examples are highlighted in Table 8.

As previously discussed in the literature review, WC based studies are utilized throughout solar PV power generation forecasting applications. It is widely known that the largest errors occur in solar PV power generation forecasting during volatile weather conditions, especially when cloud coverage is intermittent. During these instances, GHI is influenced by a variety of environmental factors, thus impacting solar PV output. Implementation of WCI provides improvements in this regard, as PV output characteristics based on specific weather conditions are distinguished. Additionally, the WCI can be utilized as a metric to quantify performance of PV power prediction algorithms, with respect to WC.

Considering the OCI, complexity of PV system operation with respect to power generation volatility and actual output can further bolster PV power generation estimations and predictions as an assessment metric. The OCI can also be implemented for assessments of power system bus voltage and frequency regulation with PV-integrated schemes. For instance, if the OCI increases in value, symbolizing reduced power output as a result of fluctuations, operators can accurately assess the situation and act to maintain system stability.

Since the PVGI represents a comprehensive impact/value of PV power generation, applications throughout power systems are available. For example, PVGI implementation can be explored for flexible dispatch controllers incorporating some combination of battery storage systems, demand response requirements, or electric vehicle smart park integration. Dynamic voltage and frequency regulation can be enhanced with informative power compensation strategies to maintain system stability through significant variations in PV power generation. In a similar manner, automated generation control practices can be improved by proactively tuning AGC set points based on PVGI operating points. PVGI can further be utilized to assess, maintain and validate measurement equipment by providing a metric for anomaly detection.

It is important to note some limitations of the proposed SAIs. As previously discussed, solar PV plant power generation characteristics contain a high correlation to the concurrent weather and environmental conditions experienced. Thus, the SAI determinations can only be as accurate as the power generation data. The same is true for insufficient data sampling rates. On the other hand, the proposed methodology has not been implemented and assessed with solar PV plant power generation data from other regions/climates around the world. While the FIS provides a robust platform for determining the SAIs, it is possible that other locations may have other influences in SAI determination, given differences in weather event frequency and climate differences.

6. Conclusion

The development of three SAIs, weather condition index, operational complexity index and solar PV power generation index for near real-time planning and operation of solar PV plants is presented in this paper.

Features describing the operational behaviors of solar PV plants under dynamic operating conditions are derived from a historical power generation database of a solar PV plant. Three FISs utilize the scaled feature vectors to independently determine their corresponding index. From the index, typical power generation characteristics based on four operating conditions are further determined in the PVPS. Based on the PVPSs, it is possible to interpolate sample-wise index values for an arbitrary power generation measurement of a solar PV plant. Utilization of the WCI, OCI, and PVGI provides significant insights into the stochastic operating characteristics of solar PV plants, offering further opportunities for enhancing planning and operational procedures.

Each of the proposed SAI, WCI, OCI and PVGI, offer significant benefits, based on the underlying principle of spatial-temporal resolution and the nature of site-specific operating parameters of solar PV plants. The proposed framework is adaptive for different locations, sizes, and configurations of solar PV plants, where each index may be re-evaluated, re-optimized and continuously updated with new data.

Future work will include the implementation of the WCI, OCI and PVGI in grid simulation contexts. Following studies will examine the impact each SAI has on a targeted operating scenario to demonstrate their performance enhancing capabilities.

CRediT authorship contribution statement

Michael Walters: Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Ganesh K. Venayagamoorthy:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by US National Science Foundation (NSF) under grants CNS#2318612 and ECCS#22344032, and the Duke Energy Distinguished Professor Endowment Fund.

Appendix A

A.1. SAI-based PV dispatch controller

To demonstrate a practical application of SAIs in near real-time planning and control scenarios for solar PV plants, two dispatch controllers for a solar PV plant with connected battery energy storage system (BESS) are developed in this supplementary section.

Fig. A.1 shows a representation of the proposed system, including Clemson University's 1MWp solar PV plant and BESS with the energy management controller, battery bank, supporting power electronic equipment and new connection with existing facilities. The micro-PMU communicates power generation data to the energy management controller which decides to charge/discharge the battery bank to meet the dispatch objective – power firming. Power firming refers to maintaining a certain level of combined PV and BESS output power for a specified time interval during the day. Measured PV power (P_{PV}), state of charge (SOC), power level commitment (P_{com}), $PVGI$, predicted PVGI ($\widehat{PVGI}$), combined PV-BESS power dispatched to the grid (P_{grid}) and battery bank power (P_{BB}) represent input and output variables for the controller. Note that PVGI related variables are only utilized in the SAI controller.

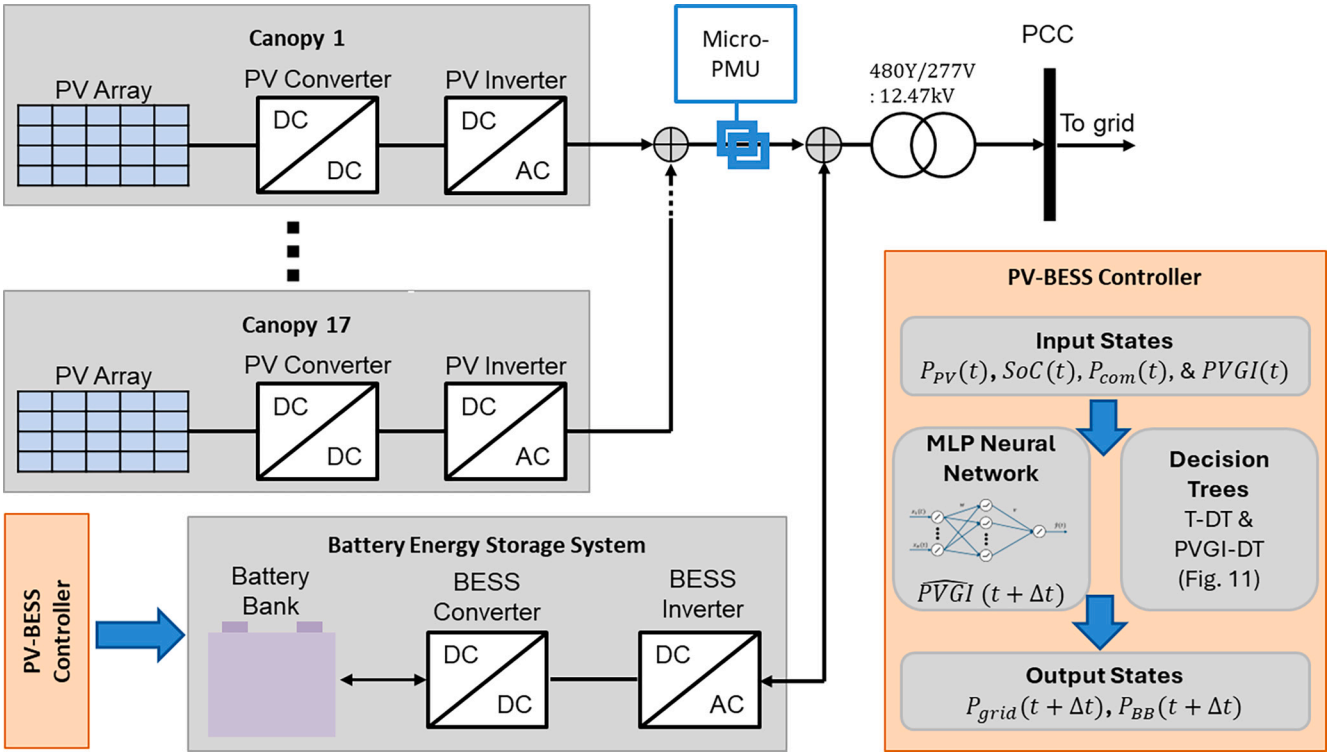


Fig. A.1. Depiction of the solar PV plant, proposed BESS with battery bank and power electronics, and a model flow of the PV-BESS controller with input states, internal systems, and output states illustrated.

To accomplish PV power firming, the battery bank must be sized accordingly to support energy requirements, while maintaining economical reason. As such, the WCI is utilized to evaluate the frequency of OCs to evaluate these requirements. This is an additional insight the SAIs provide, as both over-loading and under-loading a battery bank could be detrimental to battery life. For example, if a solar PV plant is located in a climate that is frequently known for having experienced consistent OC_A weather conditions, then the power commitment requirements and battery bank must be tuned accordingly as to ensure over-power event do not occur.

The PV-BESS controller utilizes and compares two decision tree approaches, namely a traditional decision tree (T-DT) and PVGI decision tree (PVGI-DT). Both decision trees utilize two primary decision factors (i.e. PV power generation and state of charge measurements) to achieve the multiple objectives listed below:

1) Maintain P_{grid} within the power commitment interval during the identified time-span.
2) Maintain a BESS state of charge (SoC) near to the SoC set point, such that sustained BESS charge deviations either higher or lower from the desired value are minimized.

These objectives are realized through controlling the battery bank power flow to realize the PV power firming requirements.

As compared to the T-DT, the PVGI-DT contains an additional decision factor to better achieve the PV power firming requirements. As indicated, the PVGI-DT evaluates current and predicted PVGI values, provided by a single feed-forward, time-series multi-layer perceptron (MLP) neural network at a prediction horizon of 30 min. Depictions of the full algorithm and different subroutines are provided in Fig. A.2.

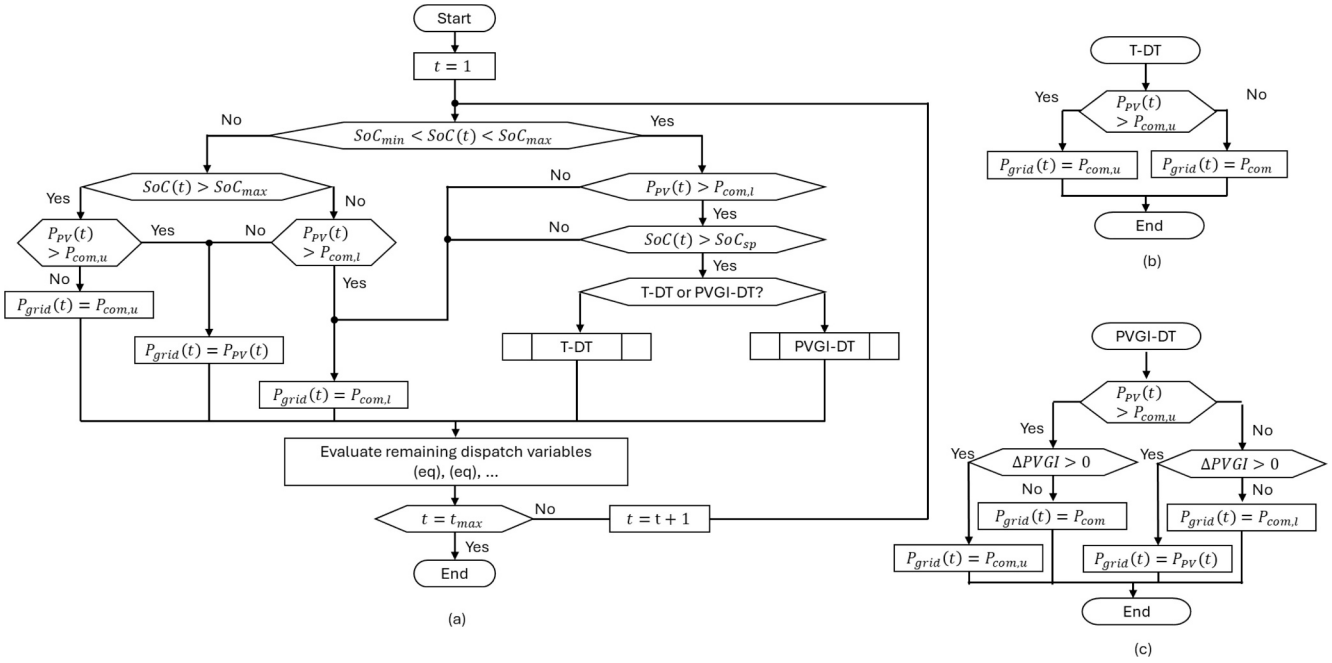


Fig. A.2. (a) PV-BESS full algorithm implementation to realize control requirements based on input states, (b) T-DT basic subroutine and (c) PVGI-DT subroutine with integrated foresight capabilities.

In the PVGI-DT methodology, the change in PVGI ($\Delta PVGI$) is included to quantify the expected change in solar PV plant performance. $\Delta PVGI$ is calculated as the difference between PVGI predictions ($PVGI'(t+\Delta t)$) and the calculated PVGI at time t is given in (A.1)

$$\Delta PVGI(t) = PVGI'(t+\Delta t) - PVGI(t) \tag{A.1}$$

To reasonably compare the two decision tree approaches, performance metrics such as average dispatch power ($P_{grid,avg}$), and SoC set point standard deviation, (SoC_{dev}) are calculated in (A.2) and (A.3), respectively, where $t \in [1:T]$ samples included within the power commitment interval.

$$P_{grid,avg} = \frac{1}{T}\sum_{t=1}^{T} P_{grid}(t) \tag{A.2}$$

$$SoC_{dev} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(SoC(t) - SoC_{sp}\right)^2} \tag{A.3}$$

A.2. Results

First, by evaluating Table A.1, insights to the probability of the next day's operating condition based on the current day's operating condition is shown for Q2 of the previously discussed dataset. Here, an interesting phenomenon is noted by the succession of OCs. Given that a current day is an arbitrary OC, there is at least a 50 % chance the next day is the same day type, showing that OCs commonly repeat for the region where the data is collected. Utilizing this analysis, a battery bank size for the PV-BESS may be determined considering the probability of successive events.

Table A.1
Frequency of operating conditions based on WCI.

		Current Day		
		OC_A	OC_B	OC_C
Next Day	OC_A	53 %	31 %	21 %
	OC_B	42 %	56 %	29 %
	OC_C	5 %	13 %	50 %

With this additional information, for battery bank sizing, additional design parameters may be found in Table A.2. As noted, the BESS size is 4MWh, with a minimum and maximum permitted range allotted by a difference of 10 % of full-charge and no-charge conditions. This battery size is determined to sustain the committed power interval of ±10% of 700 kW between 09:00 and 17:00 for the majority of days within the Q2 dataset.

Table A.2
PV-BESS design parameters.

Parameter	Value
BESS nominal energy	4 MWh
BESS SoC range	[10% – 90 %]
BESS SoC set point	50 %
Power commitment	700 kW
Power commitment range	±10% [630 kW – 770 kW]
Energy commitment	5.6 MWh
Commitment start time	9:00
Commitment end time	17:00

To demonstrate the effectiveness of the PVGI-DT compared to the T-DT baseline approach, output plots are presented in Fig. A.3 for three test cases. For each test case (organized by columns), measured PV power generation is plotted in the first row in green to show when solar PV plant generation exceeds the minimum commitment level, and red otherwise. The second row shows $PVGI$ and $\widehat{PVGI}$ time series, third row shows P_{grid} for both decision trees, and the fourth row shows the SoC time series. The aforementioned performance metrics are provided in Table A.3, with an additional metric, number of battery switching events (BB_{sw}) also included indicative of the number of charge/discharge switching events.

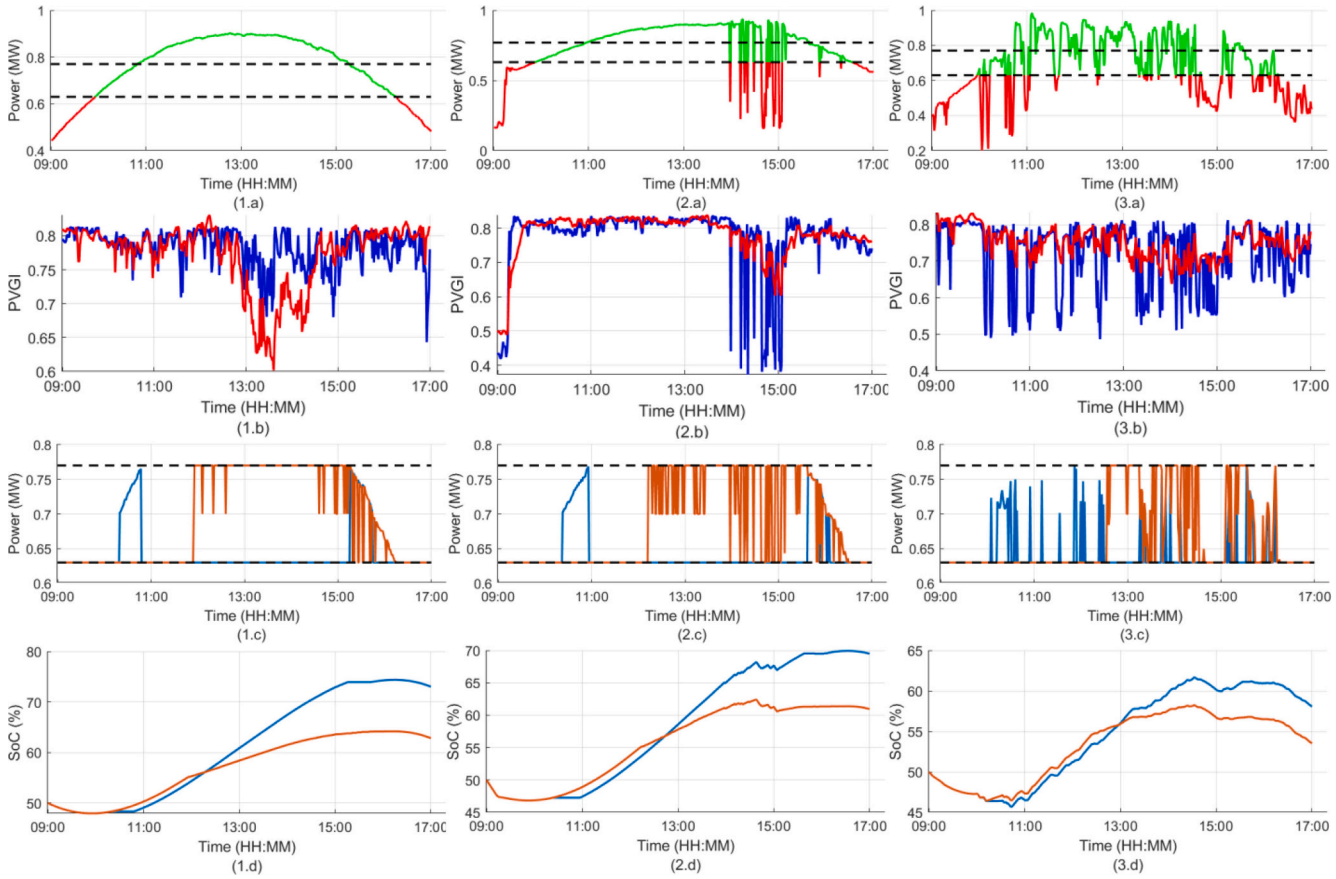


Fig. A.3. PV-BESS controller output comparison for three test cases. Measured PV power for each test case in plots (1.a), (2.a), and (3.a). *PVGI* and $\widehat{PVGI}$ time series in blue and red, respectively, in (1.b), (2.b), and (3.b). Combined power dispatch, P_{grid}, in plots (1.c), (2.c), and (3.c), and SoC in plots (1.d), (2.d), and (3.d) each with T-DT in blue and PVGI-DT in orange.

Table A.3
PV-BESS test case performance metrics.

Performance Index	Test Case 1		Test Case 2		Test Case 3	
	T-DT	*PVGI-DT*	*T-DT*	*PVGI-DT*	*T-DT*	*PVGI-DT*
$P_{grid.avg}$	0.64	0.69	0.64	0.68	0.64	0.66
SoC_{dev}	0.12	0.064	0.97	0.58	0.057	0.044
BB_{sw}	2	1	24	21	43	32

In all the test cases shown, the PVGI-DT outperforms the T-DT in each metric. Supported by Table A.3, $P_{grid.avg}$ is consistently maximized, and both SoC_{dev} and $BB_{sw}$are consistently minimized by the PVGI-DT, compared to the T-DT. Therefore, even in a simple application scenario, such as a decision tree style PV-BESS energy dispatch controller, the SAIs offer excellent sources of information to enhance planning and operational contexts of solar PV plants.

A.3. Summary

PV-BESS technology demonstrate strong potential to address the inherent variability and intermittency of solar PV power generation. Thus, distribution systems stand to benefit from greater efficiency, reliability and resiliency. This case study presented the development of an intelligent PV-BESS controller designed to enhance the power output of a solar PV plant with PV power firming. By integrating SAIs, namely the WCI for battery bank sizing evaluations, and the PVGI for advanced intelligence sources in dispatch control, more informed and planning and operational decisions occurred. The PVGI-DT further demonstrated improved dispatch control by supporting optimal SoC management and reducing switching cycles, ultimately extending battery bank lifetime.

Appendix B

This appendix provides supplementary visualizations and comparisons of SAI values for comparison. Days with corresponding SAI values are plotted in Fig. B.1.

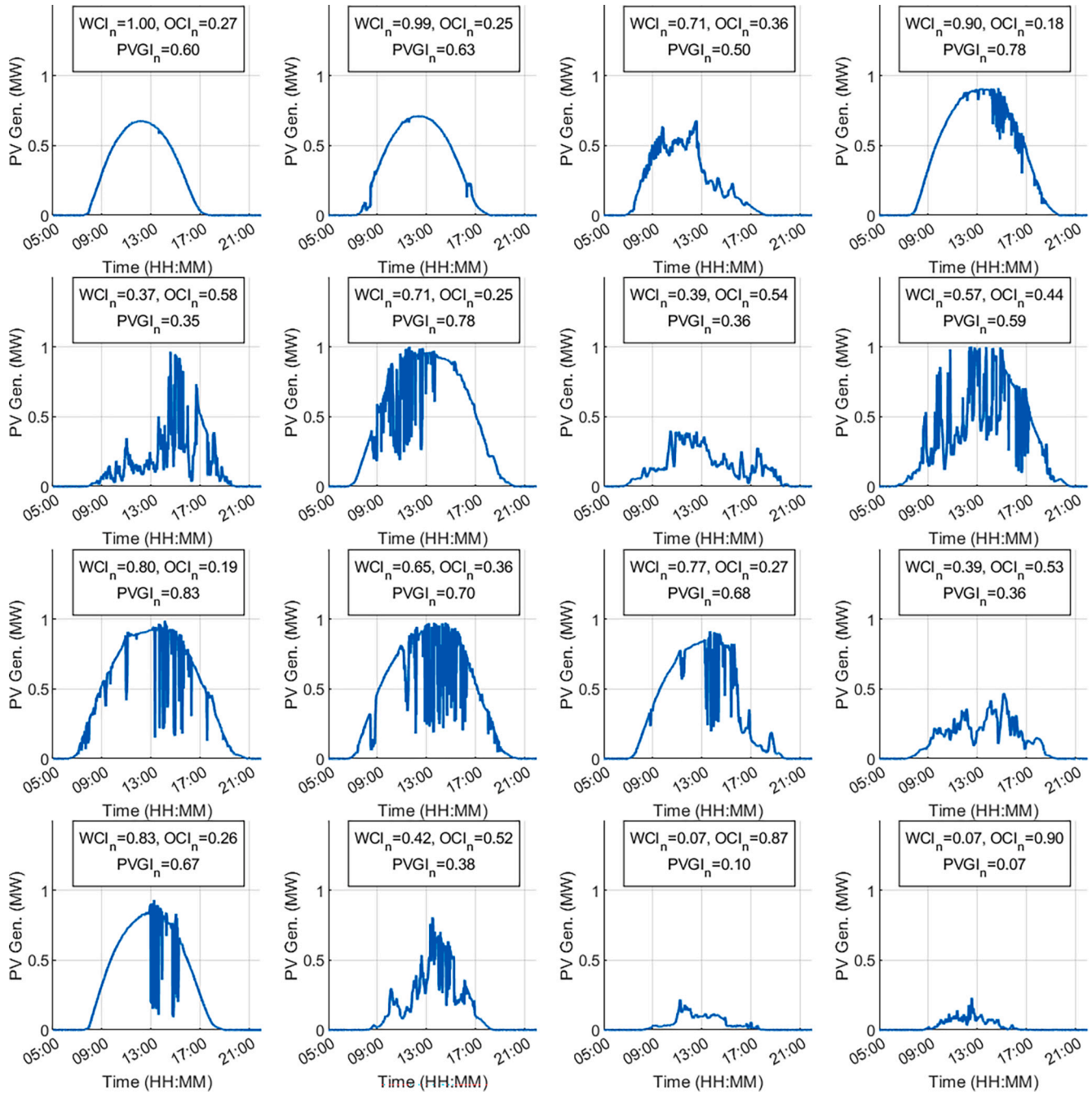


Fig. B.1. Visualization of actual solar PV power generation for Clemson University 1MWp solar PV plant measured at a 1-min sampling rate for randomly selected days during the 2024 year.

Appendix C

This appendix provides supplementary visualizations of the PVPSs. Monthly PVPSs are shown for WCI, OCI, and PVGI in Fig. C.1, Fig. C.2, and Fig. C.3, respectively. Average SAI values for each index are shown in Table C.1, with their sorted frequency proportions for each OC summarized in Table C.2. It is observed that the monthly PVPSs provide higher resolution analysis given the smaller TF, but at the cost of sparsity, as some OCs do not occur within a given month

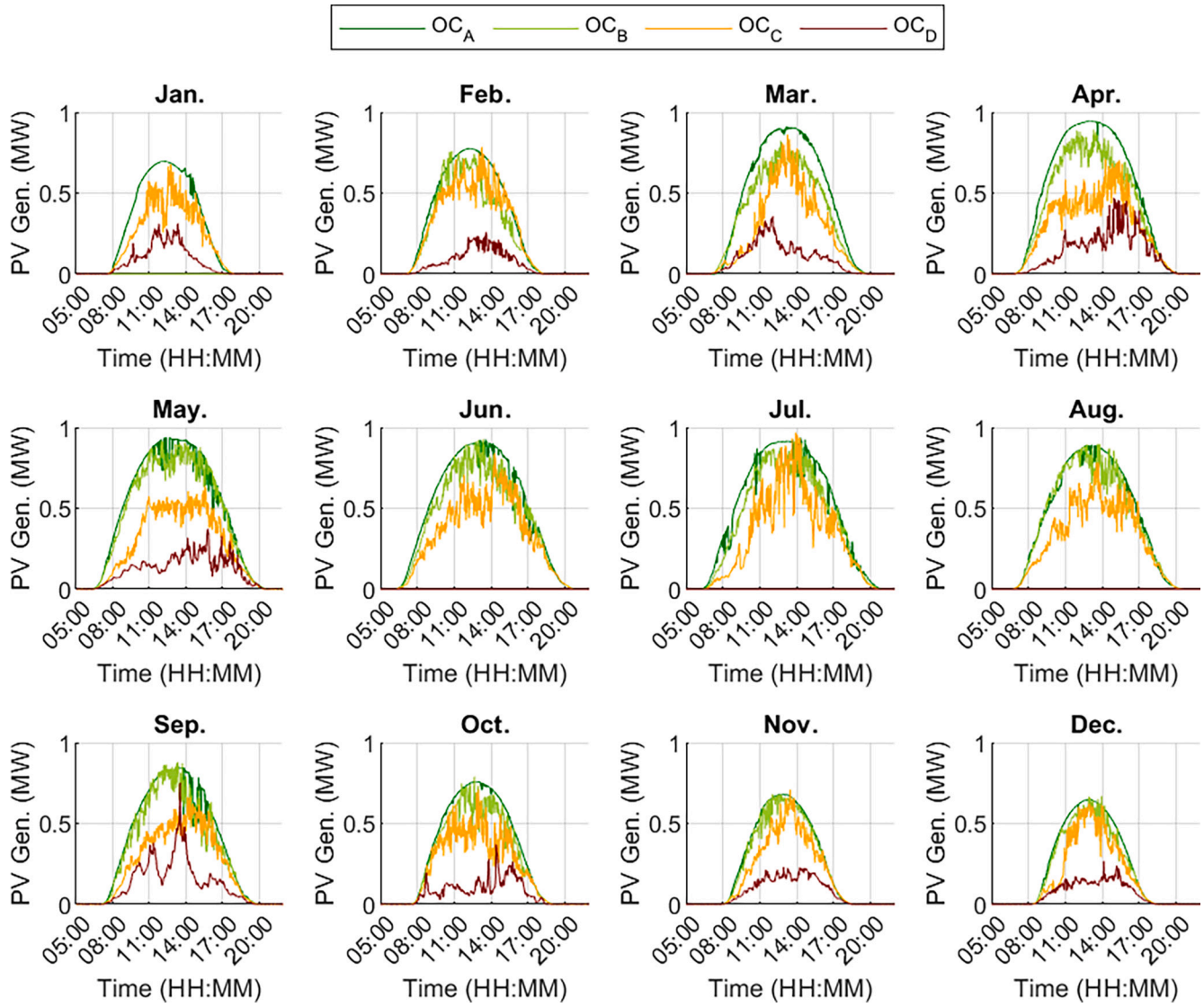


Fig. C.1. Monthly WCI PVPSs. As pictured, the month represented in each plot contains comparative differences with respect to power peak power generation and volatility for each OC.

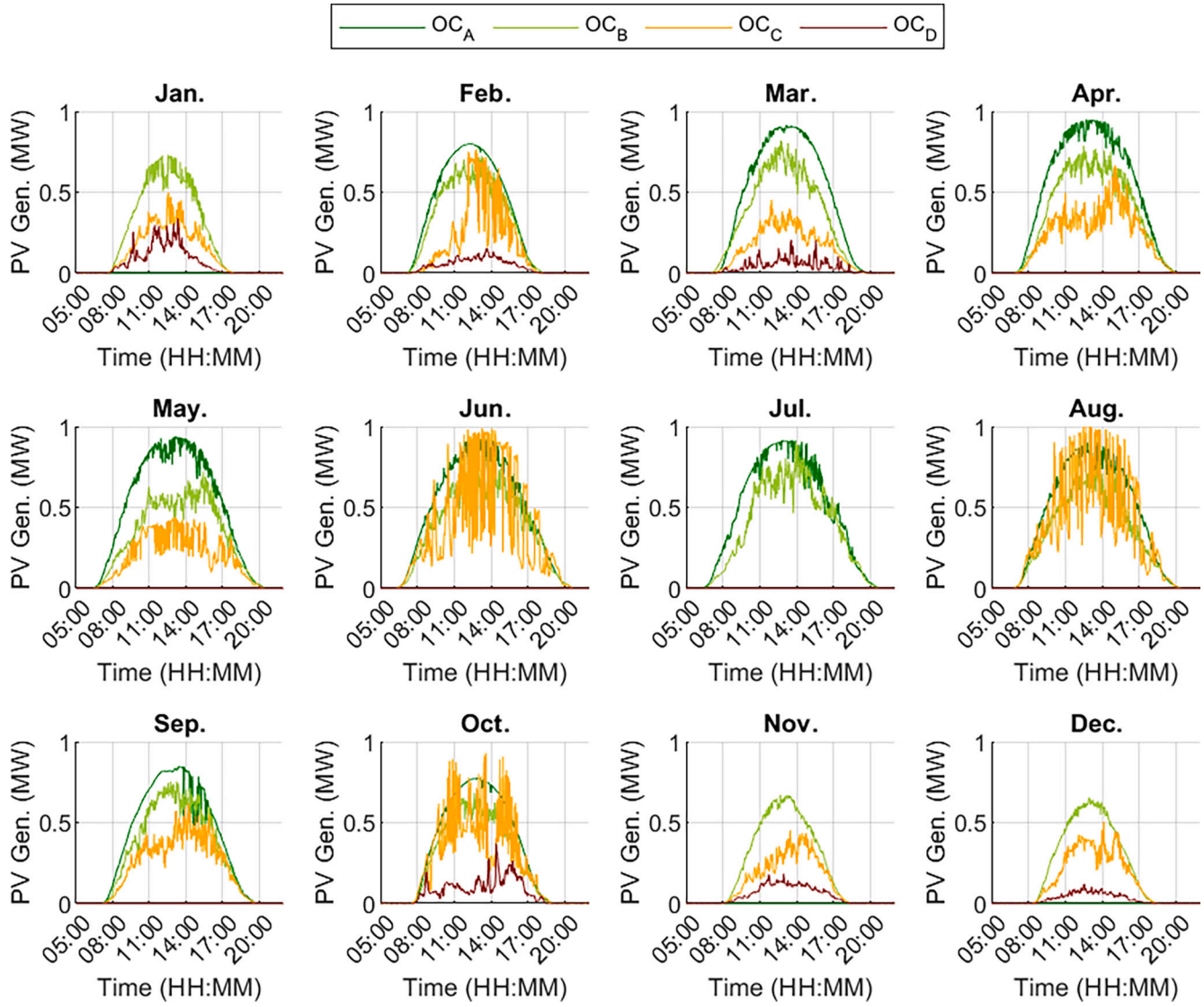


Fig. C.2. Monthly OCI PVPSs. As pictured, the month represented in each plot contains comparative differences with respect to power peak power generation and volatility for each OC.

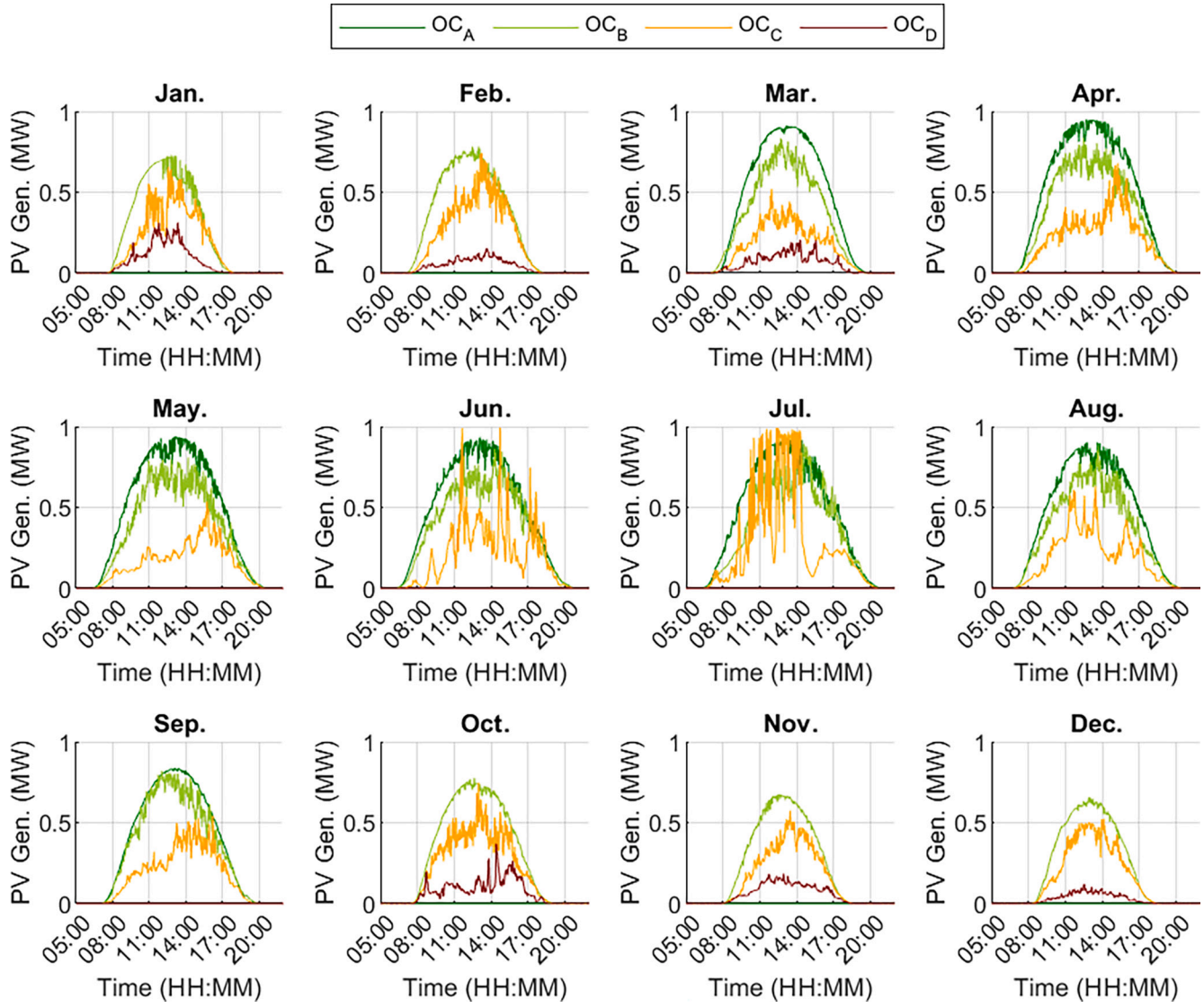


Fig. C.3. Monthly PVGI PVPSs. As pictured, the month represented in each plot contains comparative differences with respect to power peak power generation and volatility for each OC.

Table C.1
Monthly PVPS average SAI.

TF	*WCI*				*OCI*				*PVGI*			
	OC_A	OC_B	OC_C	OC_D	OC_A	OC_B	OC_C	OC_D	OC_A	OC_B	OC_C	OC_D
Jan.	0.97	N/A	0.45	0.16	N/A	0.34	0.59	0.78	N/A	0.58	0.4	0.18
Feb.	0.99	0.77	0.52	0.15	0.21	0.36	0.58	0.86	N/A	0.64	0.44	0.11
Mar.	0.97	0.73	0.53	0.24	0.15	0.32	0.62	0.88	0.81	0.63	0.34	0.15
Apr.	0.98	0.71	0.5	0.33	0.13	0.35	0.58	N/A	0.85	0.64	0.39	N/A
May	0.94	0.76	0.53	0.31	0.16	0.40	0.60	N/A	0.85	0.62	0.38	N/A
Jun.	0.95	0.72	0.55	N/A	0.15	0.39	0.51	N/A	0.83	0.64	0.44	N/A
Jul.	0.92	0.75	0.54	N/A	0.18	0.35	N/A	N/A	0.81	0.65	0.49	N/A
Aug.	0.95	0.76	0.59	N/A	0.18	0.34	0.53	N/A	0.8	0.62	0.46	N/A
Sep.	0.95	0.68	0.49	0.34	0.18	0.39	0.55	N/A	0.76	0.61	0.39	N/A
Oct.	0.98	0.76	0.45	0.16	0.21	0.34	0.56	0.76	N/A	0.67	0.45	0.18
Nov.	0.98	0.78	0.5	0.19	N/A	0.33	0.59	0.83	N/A	0.57	0.38	0.15
Dec.	1	0.73	0.47	0.17	N/A	0.33	0.56	0.91	N/A	0.56	0.4	0.07

Table C.2
Monthly PVPS sorted proportions.

TF	WCI				OCI				PVGI			
	OC_A	OC_B	OC_C	OC_D	OC_A	OC_B	OC_C	OC_D	OC_A	OC_B	OC_C	OC_D
Jan.	0.30	0	0.40	0.30	0	0.50	0.30	0.20	0	0.40	0.30	0.30
Feb.	0.37	0.10	0.29	0.23	0.30	0.48	0.05	0.19	0	0.62	0.19	0.19
Mar.	0.38	0.27	0.16	0.19	0.42	0.31	0.23	0.04	0.38	0.35	0.19	0.08
Apr.	0.23	0.38	0.27	0.12	0.35	0.38	0.27	0	0.35	0.38	0.27	0
May	0.17	0.33	0.42	0.08	0.42	0.42	0.16	0	0.42	0.33	0.25	0
Jun.	0.21	0.42	0.37	0	0.37	0.58	0.05	0	0.37	0.58	0.05	0
Jul.	0.14	0.64	0.20	0	0.50	0.50	0	0	0.57	0.36	0.07	0
Aug.	0.32	0.45	0.23	0	0.59	0.36	0.05	0	0.59	0.36	0.05	0
Sep.	0.11	0.26	0.58	0.05	0.11	0.53	0.36	0	0.05	0.58	0.37	0
Oct.	0.62	0.19	0.15	0.04	0.50	0.42	0.04	0.04	0	0.77	0.19	0.04
Nov.	0.24	0.16	0.24	0.36	0	0.56	0.24	0.2	0	0.44	0.32	0.24
Dec.	0.32	0.21	0.21	0.26	0	0.63	0.21	0.16	0	0.47	0.37	0.16

Data availability

Data will be made available on request.

References

[1] Higginbotham C. EIA projections indicate global energy consumption increases through 2050, outpacing efficiency gains and driving continued emissions growth. eia.gov; 2025. Accessed: Feb. 6. [Online]. Available: https://www.eia.gov/pressroom/releases/press542.php.

[2] United Nations, "The Sustainable Development Goals Report 2024," United Nations Statistics Division, New York, NY, USA. Accessed: Apr. 3, 2025. [Online]. Available: https://unstats.un.org/sdgs/report/2024/The-Sustainable-Development-Goals-Report-2024.pdf.

[3] Bojek P. Solar PV is set to become the largest renewable energy source by 2029. iea.org; 2025. Accessed: Feb 6. [Online]. Available: https://www.iea.org/energy-system/renewables/solar-pv.

[4] Nasir M, Khan HA, Hussain A, Mateen L, Zaffar NA. Solar PV-based scalable DC microgrid for rural electrification in developing regions. IEEE Trans Sustain Energy 2018;9(1):390–9.

[5] Hossain MS, Wadi Al-Fatlawi A, Kumar L, Fang YR, El Haj Assad M. Solar PV high-penetration scenario: an overview of the global PV power status and future growth. Energy Syst 2024;1(57).

[6] Lopez M, Claus R, Soto F, Hernandez-Garrastacho ZA, Cebada-Relea A, Simancas O. Advancing off-shore energy generation: the HelioSea concept. Appl Energy 2024;359.

[7] Panagoda SS, et al. Advancements in photovoltaic (Pv) technology for solar energy generation. J of Res Technol Eng 2023;4(3):30–72.

[8] Khezri R, Mahmoudi A, Aki H. Optimal planning of solar photovoltaic and battery storage systems for grid-connected residential sectors: review, challenges and new perspectives. Renewable and Sustain. Energy Rev. 2022;153.

[9] Mellit A, Kalogirou SA. Artificial intelligence and internet of things to improve efficacy of diagnosis and remote sensing of solar photovoltaic systems: challenges, recommendations and future directions. Renew Sustain Energy Rev 2021;143.

[10] Youssef A, El-Telbany M, Zekry A. The role of artificial intelligence in photovoltaic systems design and control: a review. Renew Sustain Energy Rev 2017;78:72–9.

[11] Mellit A, Kalogirou SA. Artificial intelligence techniques for photovoltaic applications: a review. Prog Energy Combust Sci 2008;34(5):574–632.

[12] Shah R, Mithulananthan N, Bansal RC, Ramachandaramurthy VK. A review of key power system stability challenges for large-scale PV integration. Renew Sustain Energy Rev 2015;41:1423–36.

[13] Shafiullah M, Ahmed SD, Al-Sulaiman FA. Grid integration challenges and solution strategies for solar PV systems: a review. IEEE Access 2022;10:52233–57.

[14] Meenal R, et al. Weather forecasting for renewable energy system: a review. Arch Comput Methods Eng 2022;29:2875–91.

[15] Markovics D, Mayer MJ. Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction. Renew Sustain Energy Rev 2022;161.

[16] Liu J, Zang H, Cheng L, Ding T, Wei Z, Sun G. A transformer-based multimodal-learning framework using sky images for ultra-short-term solar irradiance forecasting. Appl Energy 2023;342.

[17] Nespoli A, Niccolai A, Ogliari E, Perego G, Collino E, Ronzio D. Machine learning techniques for solar irradiation nowcasting: cloud type classification forecast through satellite data and imagery. Appl Energy 2022;305.

[18] Kumari P, Toshniwal D. Long short term memory- convolutional neural network based deep hybrid approach for solar irradiance forecasting. Appl Energy 2021;295.

[19] Chen C, Duan S, Cai T, Liu B. Online 24-h solar power forecasting based on weather type classification using artificial neural network. Solar Energy 2011;85(11):2856–70.

[20] Wang F, Zhen Z, Wang B, Mi Z. Comparative study on KNN and SVM based weather classification models for day ahead short term solar PV forecasting. Appl. Sci. 2017;8(1):28–51.

[21] S. Nitisanin and N. Hoonchareon, "Solar power forecast with weather classification using self-organized map," in *Proc. 2017 IEEE Power & Energy General Meeting*.

[22] Zafar R, Vu BH, Husein M, Chung I. Day-ahead solar irradiance forecasting using hybrid recurrent neural network with weather classification for power system scheduling. Appl Sci 2021;11(15):6738–61.

[23] Wang F, Zhen Z, Mi Z, Sun H, Su S, Yang G. Solar irradiance feature extraction and support vector machines based weather statues pattern recognition for short-term photovoltaic power forecasting. Energ Buildings 2015;86:427–38.

[24] Haljasmaa KI, Bramm AM, Matrenin PV, Eroshenko SA. Weather condition clustering for improvement of photovoltaic power plant generation forecasting accuracy. Algorithms 2024;19(9):419–39.

[25] Wu Y, Phan Q, Zhong Y. Overview of day-ahead solar power forecasts based on weather classifications and a case study in Taiwan. IEEE Trans. on Indust. Appl. 2024;60(1):1409–23.

[26] Walters M, Venayagamoorthy GK. Digital twin for solar photovoltaic power estimations based on an ensemble of recurrent neural networks. In: Proc. 19th Annul. IEEE Conf. Ind. Electron. Appl.; Aug. 2024.

[27] Wang X, Sun Y, Luo D, Peng J. Comparative study of machine learning approaches for predicting short-term photovoltaic power output based on weather classification. Energy 2022;240(1).

[28] Zheng L, Su R, Sun X, Guo S. Historical PV-output characteristic extraction based weather-type classification strategy and its forecasting method for the day-ahead prediction of PV output. Energy 2023;271.

[29] Gao M, Li J, Hong F, Long D. Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM. Energy 2019;187.

[30] Naware D, Mitra A. Weather classification-based load and solar insolation forecasting for residential applications with LSTM neural networks. Elect Eng 2021;104:347–69.

[31] Shi J, Lee W, Liu Y, Yang Y, Wang P. Forecasting power output of photovoltaic systems based on weather classification and support vector machines. IEEE Trans Indust Appl 2012;48(3):1064–9.

[32] Dou W, et al. Day-ahead numerical weather prediction solar irradiance correction using a clustering method based on weather conditions. Appl Energy 2024;365.

[33] Nie Y, et al. Sky image-based solar forecasting using deep learning with heterogeneous multi-location data: dataset fusion versus transfer learning. Appl Energy 2024;369.

[34] Qu J, Qian Z, Pei Y, Wei L, Zareipour H, Sun Q. An unsupervised hourly status pattern recognition and blending fitting model for PV system fault identification. Appl Energy 2021;319.

[35] Kousounadis-Knousen MA, Bazionis IK, Georgilaki AP, Catthoor F, Georgilakis PS. A review of solar power scenario generation methods with a focus on weather classifications, temporal horizons, and deep generative models. Energies 2023;16(15):5600–29.

[36] Guerra MI, Araujo FM, Neto JT, Vieira R. Survey on adaptative neural fuzzy inference systems (ANFIS) architecture applied to photovoltaic systems. Energy Sys 2022;15:505–41.

[37] Khalid AM, Mitra I, Warmuth W, Schacht V. Performance ratio – crucial parameter for grid connected PV plants. Renew Sustain Energy Rev 2016;65:1139–58.

[38] Mendel JM. Fuzzy logic systems for engineering: a tutorial. Proc IEEE 1995;83(3):345–77.

[39] "Mamdami and Sugeno Fuzzy Inference Systems." Mathworkscom, Accessed: Feb 14, 2025. [Online]. Available: https://www.mathworks.com/help/fuzzy/types-of-fuzzy-inference-systems.html.

[40] Gouch J. 17 Parking lot solar canopies will generate power and research opportunities. ClemsonNews; 2025. Accessed: Feb 25. [Online]. Available: https://news.clemson.edu/17-parking-lot-solar-canopies-will-generate-power-and-research-opportunities/.

[41] Arzani A, Venayagamoorthy GK. An empirical approach to frequency droop characterization from utility-scale photovoltaic plants operation in a power system. IET Gener Transmiss Distrib 2021;15:1539–51.
[42] Dharmawardena H, Venayagamoorthy GK. Distributed volt-var curve optimization using a cellular computational network representation of an electric power system. Energies 2022;15(12):4438–56.
[43] Liu Y, Venayagamoorthy GK. An intelligent energy management system for a photovoltaic-battery system. IFAC Proc 2012;45(21):115–20.
[44] Venayagamoorthy GK, Welch R. Energy dispatch controllers for a photovoltaic system. Eng. Appl. A.I. 2010;23(2):249–61.
[45] Welch R, Venayagamoorthy GK. Energy dispatch fuzzy controller for a grid-independent photovoltaic system. Energy Convers Manage 2010;51(5):928–37.
[46] Venayagamoorthy GK, Subasinghe I, Naidoo R. Power dispatch estimation from electric vehicles in a distribution system. In: *Proc.* 7th Intl. Conf. on Elect. Power and Energy Convers. Syst.; Nov. 2024. p. 117–22.
[47] Ratnakumar R, Venayagamoorthy GK. Adaptive automatic generation control for improved stability of power systems with utility-scale photovoltaic plants. In: Proc. IEEE IAS Global Conf. on Renewable Energy and Hydrogen Technol.; Mar. 2023. p. 1–6.