Renewable Energy 236 (2024) 121493

Contents lists available at ScienceDirect

Renewable Energy

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

Power modeling of degraded PV systems: Case studies using a dynamically updated physical model (PV-Pro)

Baojie Li, Xin Chen, Anubhav Jain *

Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

ARTICLE INFO

Keywords:
Power modeling
Physical model
Power prediction
Health monitoring
Machine learning
Photovoltaic
PV system

ABSTRACT

Power modeling, widely applied for health monitoring and power prediction, is crucial for the efficiency and reliability of Photovoltaic (PV) systems. The most common approach for power modeling uses a physical equivalent circuit model, with the core challenge being the estimation of model parameters. Traditional parameter estimation either relies on datasheet information, which does not reflect the system's current health status, especially for degraded PV systems, or requires additional I-V characterization, which is generally unavailable for large-scale PV systems. Thus, we build upon our previously developed tool, PV-Pro (originally proposed for degradation analysis), to enhance its application for power modeling of degraded PV systems. PV-Pro extracts model parameters from production data without requiring I-V characterization. This dynamic model, periodically updated, can closely capture the actual degradation status, enabling precise power modeling. PV-Pro is compared with popular power modeling techniques, including persistence, nominal physical, and various machine learning models. The results indicate that PV-Pro achieves outstanding power modeling performance, with an average nMAE of 1.4 % across four field-degraded PV systems, reducing error by 17.6 % compared to the best alternative technique. Furthermore, PV-Pro demonstrates robustness across different seasons and severities of degradation. The tool is available as a Python package at https://github.com/DuraMAT/pvpro.

Nomenclature

GHI	Global horizontal irradiance (W/m^2)	ANN	Artificial neural network
G_{POA}	Global plane of Array irradiance (W/m^2)	c-Si	Crystalline silicon
I	Current (A)	DC	Direct current
I_o	Saturation current (A)	DST	Daylight saving time
I_{ph}	Photocurrent (A)	PV	Photovoltaic
I_{DC}	DC current (A)	I-V curve	Current-voltage characteristic
I_{sc}	Short-circuit current (A)	KR	Kernel Ridge
n	Diode factor	LR	Linear regression
nMAE	Normalized mean absolute error	ML	Machine learning
nBE	Normalized bias error	MLP	Multilayer Perceptron
P	Power (W)	MPP	Maximum power point
R_s	Series resistance (Ω)	NWP	Numerical weather prediction
R_{sh}	Shunt resistance (Ω)	PV	Photovoltaic
V	Voltage (V)	RF	Random forest
V_{DC}	DC voltage (V)	SDM	Single diode model
V_{oc}	Open-circuit voltage (V)	STC	Standard test condition
		SVR	Support vector regression

1. Introduction

Photovoltaic (PV) power modeling converts the measured or predicted weather data into the expected output power of PV systems [1]. It has two major application scenarios: health monitoring [2] and power prediction [3,4], as illustrated in Fig. 1. Their main difference lies in the type of weather data used. The weather data generally include irradiance (like global horizontal irradiance or plane-of-array irradiance), temperature (ambient or module), wind speed, humidity [5]. When using measured weather data (real-time or historical), the expected output power of the system can be computed and compared with the actual measured power to assess the operational status of the PV system [6], which is essential for efficient and safe operation [7]. When using forecasted weather data, the future power of the system can be predicted [8], which enables effective energy management, aiding in grid integration of solar energy and balancing electricity supply and demand [9]. Typically, the weather data can be forecasted from numerical weather prediction (NWP) [10], sky cameras [11], or satellite images [12]. Both application scenarios of power modeling play a significant role in

* Corresponding author.
E-mail address: ajain@lbl.gov (A. Jain).

https://doi.org/10.1016/j.renene.2024.121493
Received 26 June 2024; Received in revised form 23 September 2024; Accepted 26 September 2024
Available online 30 September 2024

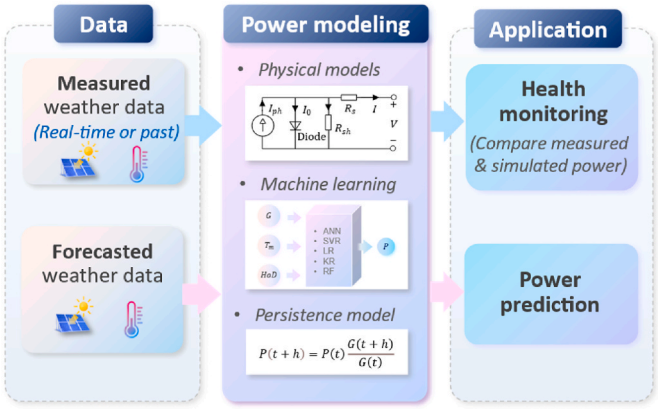

Fig. 1. Two major application scenarios of power modeling: health monitoring and power prediction, which leverage real-time measured or forecasted weather data, respectively. Both applications play a significant role in enhancing the efficiency and reliability of solar energy systems.

optimizing the efficiency and reliability of solar energy systems [1,13].

Common power modeling methods can be broadly categorized into physical models [14], machine learning models [15], or persistence models [16] (as presented in Fig. 1). Persistence models are widely used for power prediction, operating on the assumption that the power output at a given time closely resembles the power output at the same time on the previous day [17]. The performance of persistence models highly depends on the temporal correlation of current and past data, which limits their power modeling accuracy [3]. However, due to their simplicity, persistence models are frequently used as benchmark models for power modeling [18].

Machine learning (ML) models are popular techniques for power modeling, especially in the area of power prediction [19]. These models analyze historical data to identify patterns and correlations without needing detailed knowledge of PV systems [20]. Popular ML models applied in power prediction include Artificial neural network (ANN) [21,22], Long short-term memory (LSTM) [23,24], Support vector regression (SVR) [25,26], Random Forest (RF) [27,28], and Linear regression (LR) [29]. Several review works have systematically summarized these ML models for power prediction [18–20]. Despite their advantages, the major drawbacks of ML models include the requirement of large amount training data, complexity in hyperparameter-tuning, over-fitting problems, and lack of interpretability [32].

Physical equivalent circuit modeling is a common and the most basic method for power modeling [18,33]. The typical physical models are the single-, double-, or three-diode models, distinguished by the number of parameters used to characterize the model [34]. A review of these different physical models can be found in Refs. [14,35]. Note that, for physical models, the input weather data generally refers to the ground weather data, such as plane-of-array irradiance (G_{poa}) and module temperature. Thus, under the power prediction scenario, the forecasted global weather data (such as *GHI* and ambient temperature) needs to be translated into these ground weather data [36] to perform physical-model-based power modeling [37].

The major challenge for physical-model-based power modeling lies in the accurate estimation of model parameters [38]. Relying on nominal datasheet information for parameter estimation presents a significant limitation [36], as it fails to accurately reflect the current health of PV systems, especially those affected by years of degradation [39]. Using field-measured current-voltage characteristics (I-V curves) enables the acquisition of high-precision physical parameters [34]. However, the characterization of field I-V curves necessitates additional measurement equipment and will interrupt the operation of the system. Thus, the I-V curves of the entire PV array are not readily available, especially for large-scale PV systems [34]. While it is feasible to characterize a single reference PV module installed near the array, the health status of this reference module may not represent that of the entire PV array [40]. These findings emphasize the complexity and limitations in the parameter estimation of physical models for power modeling, especially for degraded PV systems.

Seen in this light, this paper proposes a powering modeling method based on a dynamically updated physical model (PV-Pro), a tool we previously proposed to extract the single-diode model parameters from the basic production data of the PV system [41,42]. By leveraging the recent production data, PV-Pro can rebuild a physical model that reflects the current degradation status of the PV system. Along with the operation of the system, this physical model will be dynamically updated. Using weather data as input, the model can then achieve a precise power modeling for the PV system.

The contribution of this paper is then reflected in the following points: A dynamically-updated physical model for power modeling is proposed, which leverages the basic production data without requiring I-V characterization; This method is suitable for power modeling of degraded PV systems, with robust performance against variations in degradation levels and seasonal impacts; This method is fully-interpretable and free of hyperparameter tuning compared to machine learning models; This method is also applicable on newly-installed systems with a limited amount of production data; The proposed method is coded into an open-source Python-based tool.

The remainder of the paper is organized as: Section 2 outlines the comprehensive methodology, encompassing details on the power modeling techniques and error metrics. Section 3 presents the power modeling conversion performance using synthetic datasets, where different types and severity of degradation are addressed. Section 4 evaluates the performance of four field PV systems. The effects of seasons and over/under-estimation are specifically analyzed. Section 5 summarizes the pros and cons of the evaluated power modeling techniques. Finally, Section 6 concludes the paper.

2. Methodology

This section provides a comprehensive picture of the methodology of power modeling. Section 2.1 introduces the power modeling techniques, including the proposed dynamically-updated physical model, alongside other widely-used methods for comparison analysis (traditional physical model, smart persistence, and machine learning models). Section 2.2 describes the error metrics to quantify the power modeling performance.

2.1. Power modeling techniques

2.1.1. Dynamically-updated physical model (PV-Pro)

The dynamically-updated physical modeling method is based on the equivalent electrical circuit model of the PV system. This approach differs from traditional physical models in the dynamical or periodical update of the model parameters. Typically, updating the parameters of field PV modules would require sophisticated I-V characterization, which, as noted in the Introduction, is impractical for large-scale in-field PV systems. To address this, we adopt PV-Pro, a tool we previously proposed [41]. It can dynamically update model parameters using only basic production and environmental data, eliminating the need for additional measurements such as I-V curves [34]. In this work, by adjusting the input data and the workflow, PV-Pro can be applied for power modeling, as depicted in Fig. 2.

Specifically, the workflow of PV-Pro for power modeling is illustrated in Fig. 3. The input includes the recent production data (DC voltage V_{DC} and current I_{DC}) and weather data (irradiance and module temperature) collected up to the day before power modeling (*Day N* – 1). The length of historical data (L) could vary from days to months, with a default setting of 7 days. Then, the raw input data is preprocessed to ensure the data quality and consistency with two major operations: daylight saving time (DST) correction and outlier removal. The DST

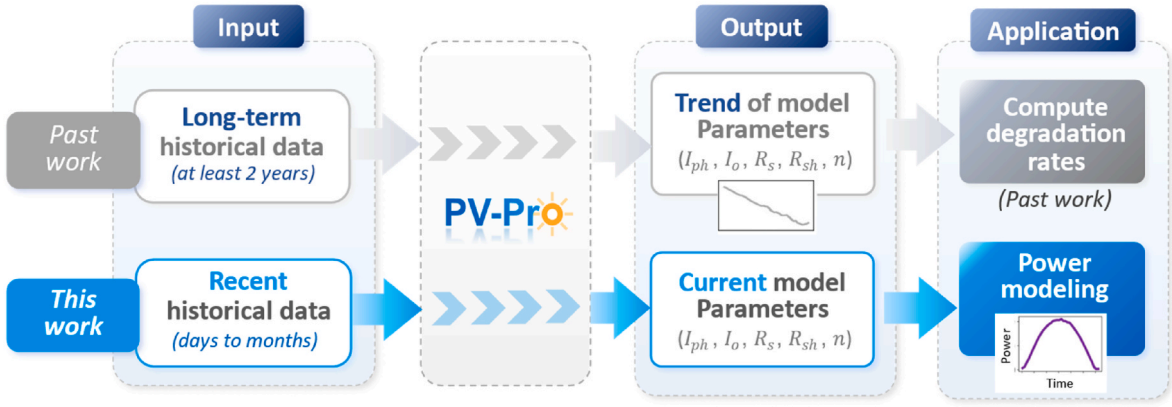

Fig. 2. Comparison of past and current work using PV-Pro, where the major difference lies in the length of input data and the application. Past work requires long-term historical data (>2 years) to extract the evolution trend and degradation rates of SDM parameters. The work in this paper leverages the recent data of the PV system to rebuild a model reflecting the current health status of the PV system, which then enables precise power modeling.

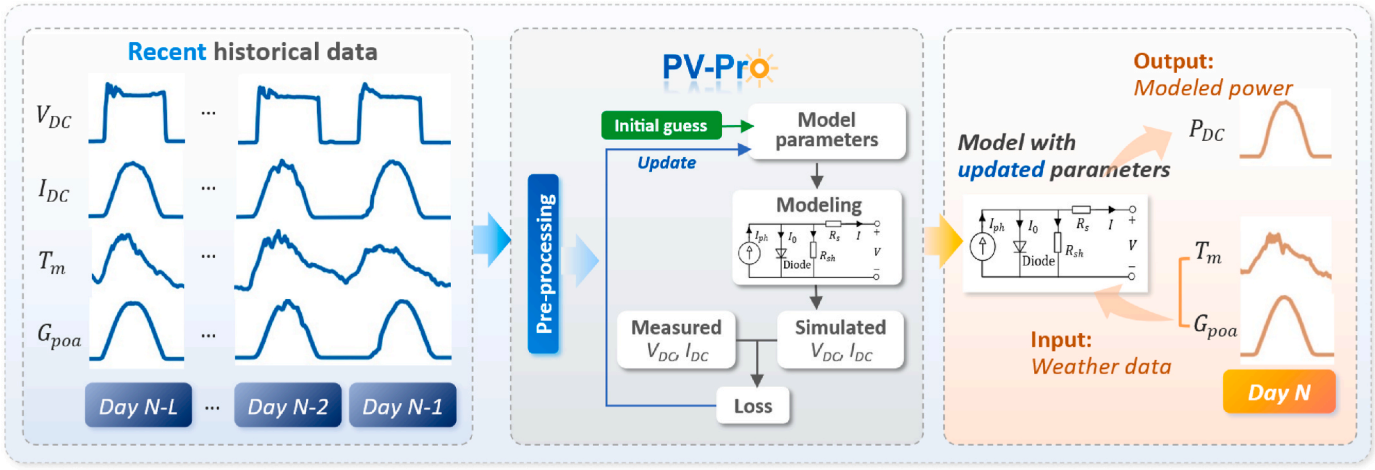

Fig. 3. Flowchart illustrating power modeling using PV-Pro. Historical production and weather data are leveraged, where the length of data can vary from days to months. After pre-processing of data, PV-Pro fits the model parameters by minimizing the loss between measured and simulated V_{DC} and I_{DC}. Then, this model with updated parameters can output precise modeled power using the real-time/forecasted weather data as input.

shifts are adjusted using the Solar-data-tools [43]. The outliers are identified by performing a linear regression of the DC current as a function of G_{poa} and of the DC voltage by module temperature [41].

For the base physical model, we employ the widely-used DeSoto singe-diode model (SDM) [44], as which includes five primary parameters, *i.e.*, the photocurrent (I_{ph}), saturation current (I_o), series resistance (R_s), shunt resistance (R_{sh}), and the diode factor (n). These parameters under different irradiance (G) and cell temperature (T_c) are expressed from (2) to (5) based on the values at the reference condition (I_{ph_ref}, I_{0_ref}, R_{sh_ref}, R_{s_ref}, n_{ref}). T_c can be calculated from the module temperature (T_m) via Sandia Array Performance Model [45]. Note that the R_{sh} in the DeSoto model is proportional to the inverse irradiance. This may cause problems in the parameter extraction, for example, R_{sh} could become unbounded as irradiance decreases to 0. Thus, we add a constant (G_{sh_extra}) as shown in (1).

$$I = I_{ph} - I_0\left[\exp\left(\frac{V+IR_s}{nN_S k_B T/q}\right) - 1\right] - (V+IR_s)\left(\frac{1}{R_{sh}} + G_{sh\_extra}\right) \tag{1}$$

$$I_{ph} = \frac{G}{G_{ref}}\left[I_{ph\_ref} + \alpha_{I_{sc}}\left(T_c - T_{c\_ref}\right)\right] \tag{2}$$

$$I_0 = I_{0\_ref}\left[\frac{T_c}{T_{c\_ref}}\right]^3 \exp\left[\frac{1}{k}\left(\frac{E_{g\_ref}}{T_{c\_ref}} - \frac{E_g}{T}\right)\right] \tag{3}$$

$$E_g = E_{g\_ref}\left[1 - dE_g dT\left(T_c - T_{c\_ref}\right)\right] \tag{4}$$

$$R_{sh} = R_{sh\_ref}\frac{G_{ref}}{G} \tag{5}$$

where,

- N_S: Number of cells connected in series
- k_B: Boltzmann constant
- q: Electron's charge
- E_g, E_{g_ref}: Material bandgap/at reference condition
- $\alpha_{I_{sc}}$: temperature coefficient of I_{sc}.
- $dE_g dT$: temperature coefficient of bandgap energy

After an initial guess of the five model parameters based on the module datasheet, PV-Pro models the PV system and obtains the simulated V_{DC} and I_{DC}. The L2 loss [46] is then calculated between the measured and simulated V_{DC} and I_{DC}. Using this loss, L-BFGS-B solver [47] updates the model parameters iteratively until the joint loss is minimized or the maximum iterations are reached. This whole process is repeated periodically to obtain a physical model with updated parameters. The update frequency depends on the user's settings and can be as daily, weekly, monthly, etc. In this work, the frequency is set as daily. The obtained dynamic physical model parameters can closely model the

PV system's actual health condition. In this way, using the measured or predicted weather data on *Day N*, the accuracy of the corresponding modeled power could be improved, especially for field-degraded PV systems.

2.1.2. Nominal physical model

For benchmarking purposes, the traditional and commonly used power modeling method, which relies on the physical model with nominal parameters, will also be addressed in this work. The base physical model is similarly set as DeSoto singe-diode model [44]. Typically, the nominal model parameters are sourced from the module manufacture datasheet. Given that the datasheet does not explicitly provide the model parameters, the 'pvlib.ivtools.sdm.fit_desoto' [45] function is used for this extraction. This approach is named the nominal physical model. This benchmarking allows a comparison between the dynamically updated model against a model with constant nominal parameters.

2.1.3. Persistence model

The persistence model is also a straightforward and commonly employed benchmark model for PV power modeling, especially in power prediction. It operates under a simple assumption that the power output at a given time closely resembles the power output at the same time on the previous day. The traditional persistence model [17] does not consider the factor of irradiance, which leads to its limited accuracy. In this work, we adopt a refined version known as the Smart Persistence model [16] (expressed in (4)), which improves accuracy by adjusting the power output using the ratio of current to historical irradiance.

$$P(t) = P(t-h)\frac{G(t)}{G(t-h)} \tag{6}$$

where, G refers to the plane-of-array irradiance; h is the time gap between current and past measurement, typically set to 24 h.

2.1.4. Machine learning models

Machine learning (ML) models are increasingly popular methods for power modeling, which directly map the weather data to output power without requiring detailed knowledge of the PV system. Drawing from literature research [30,31], this study employs five common machine learning models: Multilayer Perceptron (MLP) of Artificial neural network (ANN), Random Forest (RF), Support vector regression (SVR), Kernel Ridge (KR), and Linear regression (LR). Input features for the models include the plane-of-array irradiance (G_{poa}), module temperature (T_m), and hour of day (*HoD*) with the output as the power, as illustrated in Fig. 4. The data frequency depends on field measurements, which can vary from minute to sub-hour intervals.

The ML models are trained and validated using the data (length of L days) collected up to the day before power modeling ($Day\ N-1$), where 80 % for training and 20 % for validation. Similar to PV-Pro, the update frequency of ML models is set daily as well. The performance of machine learning models primarily hinges on the hyperparameters. Based on a comprehensive review of machine learning models for power modeling/prediction [31], a list of the hyperparameters for fine-tuning is outlined in Table 1.

The grid search method is used to systematically explore diverse combinations of these hyperparameters. Note that, different from machine learning applications in the literature, we also consider the length of historical data (used for training and validation) as a 'hyperparameter', recognizing that each model has an optimal amount of data for training. We systematically vary the length of historical data from 3 days to 3 months and calculate the power modeling error to determine the most suitable historical data length for each model. Further details can be found in Section An of Supplementary Information (SI).

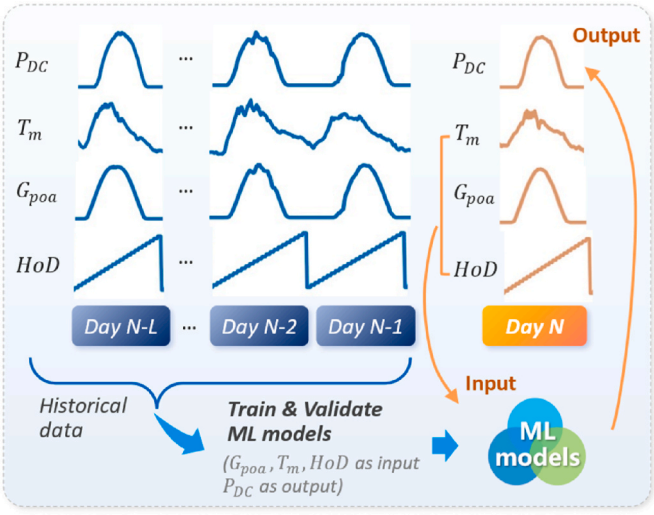

Fig. 4. Flowchart of machine learning (ML) models for power modeling. Historical data are used to train and validate the ML models, where the inputs include irradiance (G_{poa}), module temperature (T_m), and hour of the day (*HoD*) with the DC power (P_{DC}) as output. The ML models are updated daily. Then, using the real-time or forecasted G_{poa}, T_m, and *HoD* data, the ML models output the power. Note that the ML models generally require a large amount of historical data for training, which may not be available for newly-installed PV systems.

Table 1
Machine learning models and hyperparameters.

Model	Category	Hyperparameters
ANN (MLP)	Neural network	• hidden layer sizes: (5), (10), (50), (5,5), (10,10), (5,5,5), (100) • alpha: 1e-4, 1e-3, 1e-2, 1e-1 • activation function: ReLU
SVR	Support vector model	• kernel: 'rbf' • epsilon: 0.5, 0.25, 0.2, 0.1 • C: 1e-3, 1e-2, 0.1, 0.5, 1
LR	Linear model	• normalize: True or false • fit_intercept: True or false
KR	Kernel ridge model	• kernel: 'polynomial', 'rbf', 'linear' • alpha: 1e-3, 1e-2, 1e-1
RF	Ensemble method	• k_estimators: 50, 100, 250, 300 • max_depth: none, 5, 8 • min_sample_leaf: 0.01, 1, 10, 100

2.2. Error metrics

To quantify the power modeling error, we adopt two common metrics, *i.e.*, normalized mean absolute error (*nMAE*) and normalized bias error (*nBE*) [31], as described from (7)-(8). Both metrics are normalized by the nominal capacity of the PV system.

$$nMAE = \frac{\sum_{i=1}^{N}\left|P_{mod,i} - P_{meas,i}\right|}{P_{nominal}} \tag{7}$$

$$nBE = \frac{P_{mod,i} - P_{meas,i}}{P_{nominal}} \tag{8}$$

where, P_{meas} and P_{mod} are the measured and modeled power, respectively. $P_{nominal}$ refers to the nominal power of the PV system. *nMAE* reflects the total imbalance between the estimated and the actual power. *nBE* indicates the over- or under-estimation of the modeled power.

3. Case studies: Power modeling using synthetic datasets

The power modeling performance of the proposed methods and

other candidate techniques is first evaluated via generated synthetic datasets. The benefits of using synthetic datasets lie in the ability to individually address and control different types and severity of degradation of PV systems. Section 3.1 presents the generation of the synthetic datasets. Section 3.2 evaluates the performance using an example dataset of a degraded PV system. Section 3.3 examines the impact of various types and severities of degradation on the power modeling performance.

3.1. Generation of synthetic datasets of degraded PV systems

To generate synthetic datasets, we simulate the output of an 11.8 kW PV system. The system contains 50 mono-c-Si modules (Sharp_NU_U235F2), arranged in 5 strings with each string of 10 modules. The nominal parameters of the module, sourced from the datasheet, are listed in Table 2, where the single-diode model (SDM) parameters are estimated by using the 'pvlib.ivtools.sdm.fit_desoto' function [45]. The weather data are retrieved from the NSRDB database (version 3.0.1) [48] at 37° 53′ 24″N 122° 15′ 36″W. The module back sheet temperature (T_m) is estimated from the irradiance, ambient temperature, and wind speed using the temperature translation method [49]. The simulation of the PV system's output is carried out using the 'pvlib.pvsystem.single-diode' function [45].

To approximate a field-degraded PV system, we introduce specific changes to the single-diode model (SDM) parameters, reflecting typical field degradation patterns reported in the literature. Three primary SDM parameters are modified based on their nominal values (consistently over time): I_{ph} is decreased by 2 %, R_s is increased by 20 %, and R_{sh} is decreased by 20 %. Using these degraded SDM parameters and field weather data, a synthetic dataset of a degraded PV system is generated for analysis.

3.2. Power modeling performance

Using the generated synthetic dataset, we evaluate the proposed dynamically-updated physical model (PV-Pro) and other candidate power modeling techniques. For PV-Pro and the five machine learning models, the models are updated daily based on the previous days' data. The daily power modeling performance is tracked over a one-year period and the results are displayed in Fig. 5 (a). To provide a more detailed view, examples of modeled and reference power on a clear and cloudy day are illustrated in Fig. 5(b) and (c), respectively. A summary of the annual power error is given in Fig. 5 (d).

It is shown in Fig. 5 (a) that PV-Pro consistently shows lower error rates throughout the year compared to other methods. This can also be seen from the summarized annual error in Fig. 5 (d). Compared to the best of the other candidate power modeling techniques, PV-Pro reduces the average error (*nMAE*) by 25.6 %, highlighting its effectiveness in improving the power modeling performance.

3.3. Impact of degradation levels

The synthetic dataset used in Section 3.2 addresses the degradation of three parameters at a certain severity at the same time (presented in Section 3.1). To independently evaluate the impact of each degradation type and severity on the performance, the degradation severity of I_{ph}, R_s, and R_{sh} are individually varied: I_{ph} is decreased by 0–5%, R_s is increased by 0–50 %, and R_{sh} is decreased by 0–50 %. Under each case, new synthetic datasets are generated, and all techniques are re-evaluated. The corresponding annual power modeling results are presented in Fig. 6.

Across all degradation scenarios, the nominal physical method shows a quasi-linear or exponential increase in error with the severity of degradation. In contrast, the data-driven methods, including PV-Pro and machine learning (ML) models, exhibit a stable performance overall, with minimal sensitivity to variations in degradation severity. Notably, PV-Pro consistently outperforms other power modeling techniques across all degradation scenarios. This suggests that PV-Pro's dynamic updating mechanism is particularly effective at maintaining accuracy despite changes in system conditions.

Table 2
Nominal module parameters (IV and SDM parameters) for simulated PV system.

IV parameters	Value	SDM parameters	Value
V_{mp_ref}	38.3 V	I_{ph_ref}	6.0 A
I_{mp_ref}	5.65 A	I_{o_ref}	1E-10 A
V_{oc_ref}	45.89 V	n_{ref}	1.2
I_{sc_ref}	6.0 A	R_{s_ref}	0.35 Ω
P_{mp_ref}	216 W	R_{sh_ref}	600 Ω

4. Case studies: Power modeling using field datasets

To evaluate the power modeling performance on real degraded PV systems, the study incorporates several field tests. Section 4.1 presents a detailed case study of a 271 kW ground PV system, where the annual power modeling error, impact of seasons, and over-/under-estimation rate of power are carefully examined. Following the same analysis path, Section 4.2 extends the analysis to three additional PV systems to evaluate the consistency and reliability of the proposed method across PV systems with different settings and system scales. Note that, the goal of this section is to evaluate the performance of power modeling techniques. Thus to avoid introducing additional uncertainty caused by the forecasting of weather data, the weather data used for the techniques are based on the field measurement, i.e., plane-of-array irradiance (G_{poa}) and module temperature, which provides a more accurate record of the real environmental condition of the PV modules.

4.1. Case study of a 271 kW field PV system

The NIST ground PV system (Fig. 7) is selected for study, which contains 1152 mono-c-Si 235 W PV modules (Sharp_NU_U235F2), yielding a rated DC output of 271 kW [50]. This system is located in Maryland, USA (39.132170, −77.213990) with the climate zone of 'DH' based on Köppen-Geiger-Photovoltaic climate classification [51] and 'T5:H4' according to the PV Climate Zone (PVCZ) method [52]. The operation and environmental data are continuously recorded since 2015. The plane-of-array irradiance is measured by a reference cell and the module temperature by a probe attached to the back sheet of the PV module [50]. We selected the data in 2018 to study the annual power modeling error, which represents a period after a 3-year operation for the PV system.

4.1.1. Annual power modeling performance

The proposed dynamical physical model (PVPro) is applied for this field dataset, and its performance is compared with other candidate techniques, including persistence, nominal, and five machine learning models (presented in Section 2.1). The modeled power error spanning an entire year in 2018 is presented in Fig. 8 (a), where examples of modeled power on a clear and cloudy day are illustrated in Fig. 8(b) and (c), respectively. A summary of the annual power error is given in Fig. 8 (d).

From the year-long performance presented in Fig. 8 (a) and (d), we can notice the persistence model displays larger fluctuation and higher error. However, it may be also noted that the persistence model can achieve low error rates (<1 %), even outperforming PVPro on certain days, such as in Fig. 8 (b). This variability is because the persistence model's performance is highly dependent on the similarity of power shape between the current and the previous day. If the two consecutive days have similar weather conditions, such as being both clear, the persistence model can provide a good estimation of the power. Conversely, if the weather conditions differ significantly, its performance will suffer considerably. That explains the large fluctuation of the

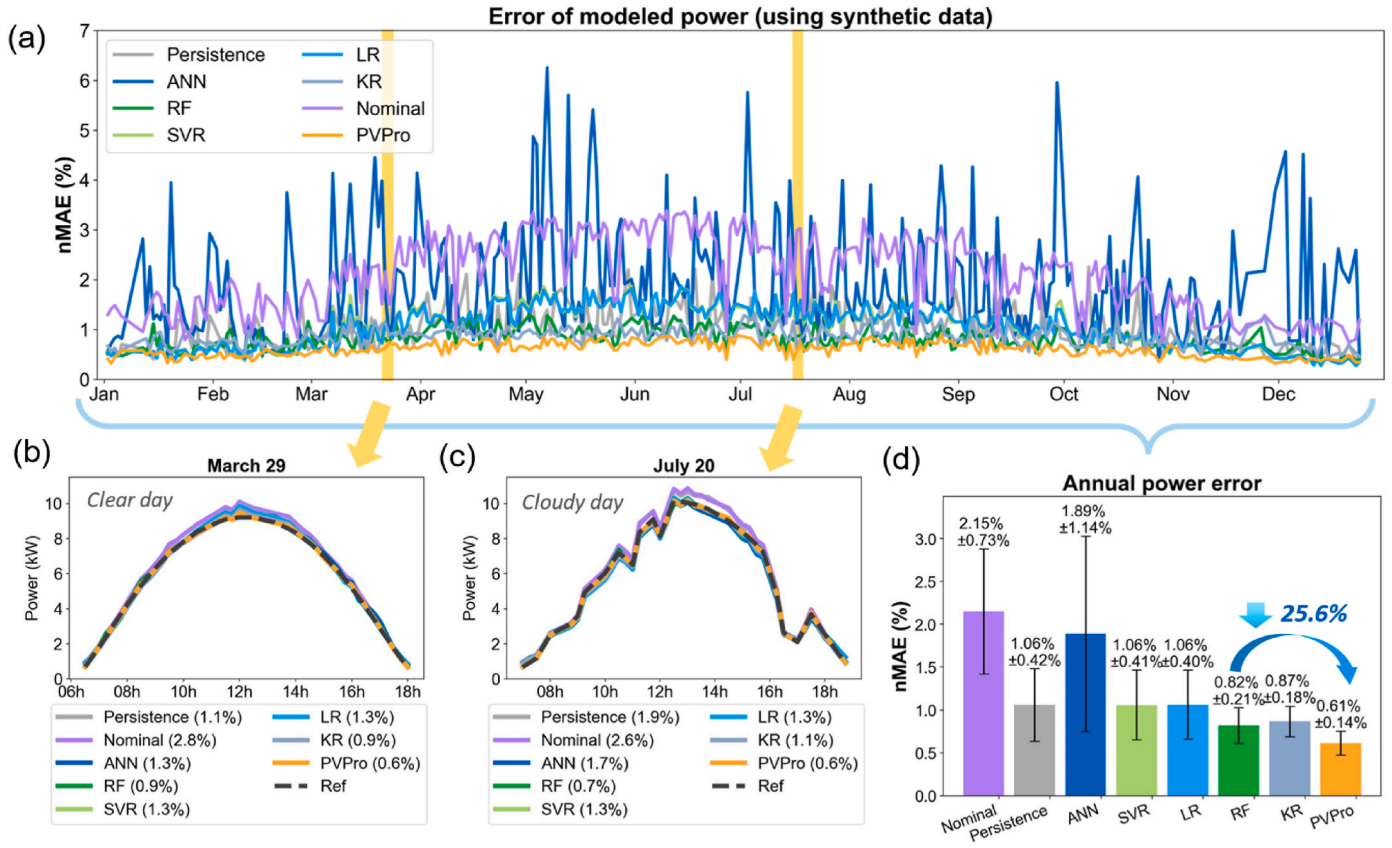

Fig. 5. (a) Error of modeled power (*nMAE*) using synthetic data set over a year. Example of modeled power on a clear day (b) and a cloudy day (c) (*nMAE* is presented for each technique in the legend). (d) The summarized annual error of modeled power (*nMAE*) using synthetic data (the values above each bar refer to the "mean ± std".).

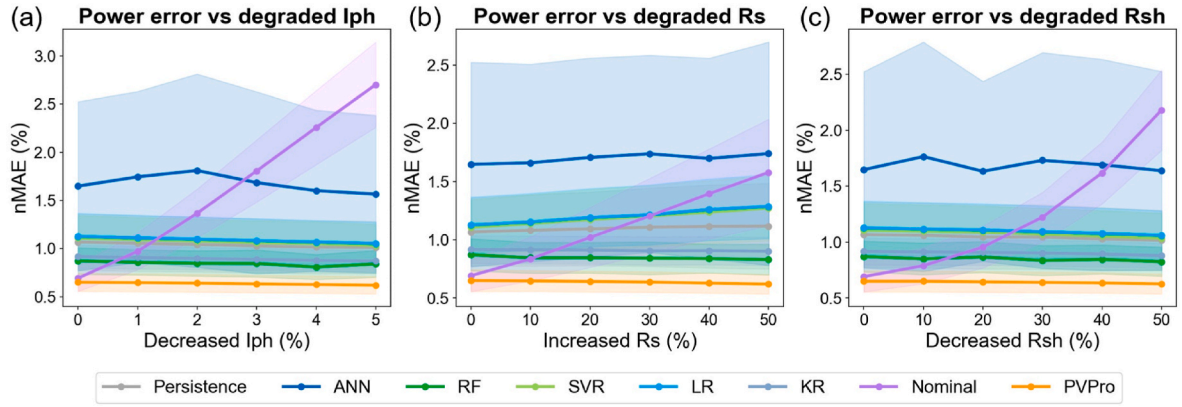

Fig. 6. Error of annual modeled power (*nMAE*) using synthetic datasets as a function of decreased I_{ph} (a), increased R_s (b), and decreased R_{sh} (c). The error of the nominal physical method increases quasi-linearly with the increasing severity of degradation. The data-driven methods (PV-Pro and ML models) exhibit an overall stable performance across various degradation levels. PV-Pro consistently outperforms other power modeling techniques, underscoring its robustness in accurate power modeling even under varying degrees of system degradation.

error of the persistence model across an entire year. For the nominal physical model, the system's mild degradation after just three years of operation results in acceptable modeling performance (1.68 %). Regarding the machine learning models, except for the ANN, the SVR, RF, and KR models exhibit similar performance (~1.6 %). Overall, PVPro outperforms all the candidate models by demonstrating a lower and more stable modeling error, with a year-averaged *nMAE* of 1.23 %, reducing the error by 21.2 % compared to the best-performing alternative technique (KR), which highlights the effectiveness of PVPro on the field PV power modeling.

4.1.2. *Over- and under-estimation analysis*

For grid-connected PV systems, another crucial angle to evaluate the power modeling is the frequency of severe over- or under-estimation of the output power, which greatly risks the reliability issues of the power system. Notably, over-estimation of power output is particularly problematic for system operators, as it can complicate the rapid deployment of backup power units and the implementation of load reduction measures [17]. Thus, we examine these conditions by quantifying the error using the normalized bias error (*nBE*). The distribution of *nBE* of predicted power in 2018 is depicted in Fig. 9, where the occurrence

Fig. 7. NIST-ground array with 1152 PV modules for power modeling test.

frequency (named as density in Fig. 9) for severe over-estimation (when *nBE* >10 % or *nBE* >20 %) is also listed.

It is shown in Fig. 9 that persistence and ANN models introduce a higher frequency of severe overestimation (ratio of *nBE* >10 % higher than 4 % and ratio of *nBE* >20 % higher than 1 %). Comparatively, PVPro achieves the lowest frequency of severe overestimation, underscoring its robust capability to mitigate instances of significant overestimation of the PV system's power.

4.1.3. Effect of seasons

To evaluate the seasonal impact on the power modeling, we partitioned the year-long power modeling error (*nMAE*) in Fig. 8 and the overestimation ratio in Fig. 9 (*nBE* >10 %) into four seasons and calculated the average error, as detailed in Fig. 10. The season windows are defined as follows: Spring (March 1 to May 31), Summer (June 1 to August 31), Fall (September 1 to November 30), and Winter (December 1 to February 28).

The power error of each method presents seasonal fluctuations in Fig. 8. Typically, the power error and overestimation ratio are lower during summer and higher during winter. This trend is especially evident in the power overestimation ratio (*nBE* >10 %), which drops to approximately 0.03 % during summer. This may be due to the higher and more stable irradiance in summer, which improves the accuracy of the measurements of environmental conditions and system modeling, and vice versa for winter. Notably, PV-Pro surpasses other techniques consistently across all seasons.

4.2. Application to multiple PV systems

In this section, we extend the power modeling evaluation to three additional field PV systems in the U.S. selected from the PVDAQ data lake [53]. These PV systems differ in capacity scales, operating years, and climate zones as mapped in Fig. 11, where System 2, 3, and 4 are the ones to be analyzed and System 1 refers to the NIST ground system analyzed in Section 4.1. The climate zone is determined by the PV Climate Zone (PVCZ) method [52], which distinguishes locations based on climate stressors more relevant to PV degradation.

Following the same power modeling path applied for the NIST ground system (System 1), the power modeling performance for Systems 2 to 4 (over one year) is accordingly analyzed. Similarly, PVPro and the ML models are updated daily, where the procedures are detailed in Figs. 3 and 4, respectively. The power modeling error and the power overestimation ratio (*nBE* >10 %) are summarized with System 1's

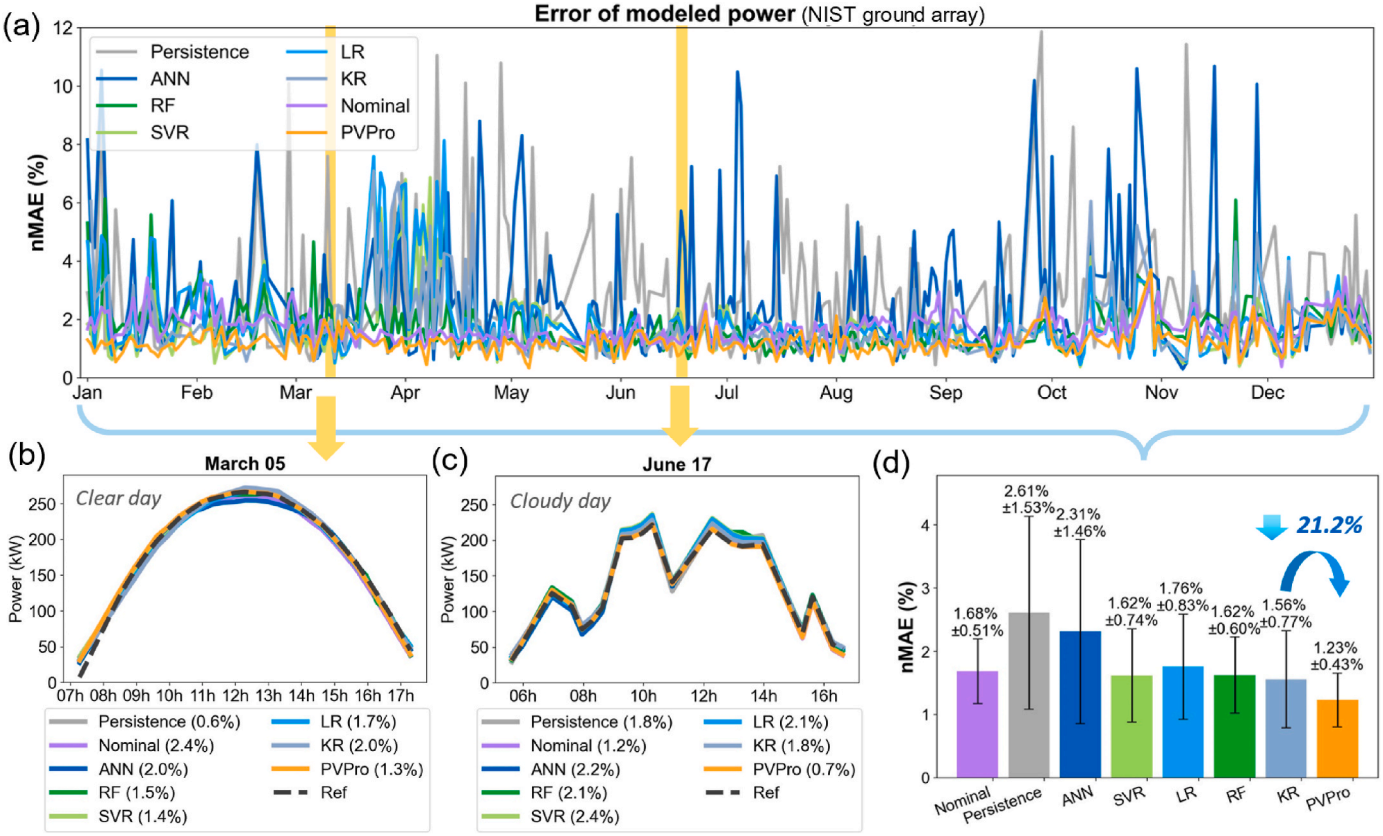

Fig. 8. (a) Error of modeled power (*nMAE*) using NIST ground dataset in 2018. Example of modeled power on a clear day (b) and a cloudy day (c) (*nMAE* is presented for each technique in the legend). (d) The summarized annual error of modeled power (*nMAE*) (the values above each bar refer to the "mean ± std".) PVPro exhibits lower and more stable errors with a decrease of 21.2 % compared to the best of other techniques (*i.e.*, KR).

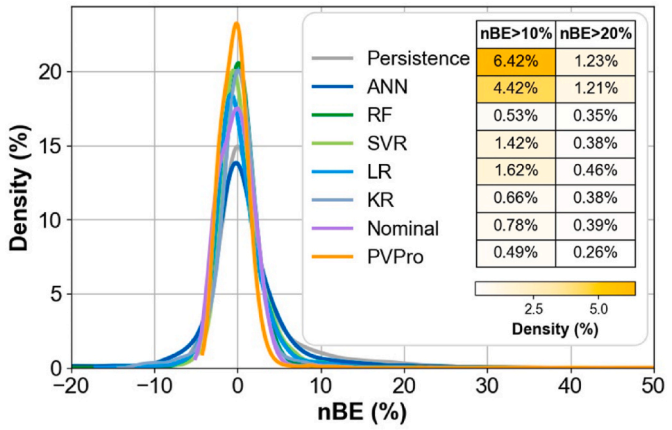

Fig. 9. Distribution of *nBE* of modeled power in 2018 using different techniques. PV-Pro has a lower frequency (<0.5 %) of significant power overestimation (*nBE* >10 % or 20 %) compared to other techniques, effectively minimizing the adverse impacts on the power system.

results in Fig. 12. The '*mean* ± *std*' of the four values is marked to show the discrepancy of performance across different systems. Note that for System 4, the module information is unknown. Thus, the nominal physical model cannot be applied to System 4.

From the results of the four PV systems, it is shown that the model performance varies from site to site. For example, the *nMAE* of the nominal physical model is below 3 % for System 1 and 3 but increases to 6.8 % in System 2. A similar trend can be observed also from *nBE* ratio in Fig. 12 (b). This discrepancy is due to the significant degradation of System 2 (after operation of 16 years), rendering the nominal parameters unfit for the actual system's condition. The *nBE* ratio, which highlights the frequency of severe power overestimation (*nBE* >10 %), shows a greater variation across the four systems compared to the *nMAE* metric. The smart persistence model, which relies on the similarity of weather conditions between two consecutive days, also exhibits a distinct performance across different systems.

In contrast, the data-driven models, such as machine learning and PV-Pro, exhibit an overall less pronounced difference between systems. This is attributed to their ability to adapt and learn from historical production data, allowing them to better capture the system's evolving degradation status. Notably, PV-Pro exhibits robust performance with lower power errors (average *nMAE* = 1.4 %, *nBE* = 1.0 %), reducing the error by 17.6 % on *nMAE* and 34 % on *nBE* compared to the best-performing alternative model (*i.e.*, KR). This highlights PV-Pro's capability for accurate power modeling across diverse degraded field PV systems.

5. Discussion

The nominal physical model method is a simple and basic approach to power modeling. However, as the PV system degrades over years of operation, the nominal parameters no longer reflect the current state of the system, leading to significant over-estimation of output power. This over-estimation is particularly problematic for grid-connected PV systems, because it necessitates the rapid deployment of backup power units and the reduction of load.

The persistence model is commonly used as a benchmark for assessing power modeling performance. However, in the literature, many studies employ the *naïve* type of persistence model [54], which simply assumes the future power output will be the same as the past observation. Consequently, this *naïve* model generally leads to relatively poor performance (average *nMAE* = 15 % as shown in Fig. S2 of SI). It is noteworthy that this naïve persistence model could be seamlessly replaced by the smart persistence model (presented in Section 2.1.3) without extra effort. This smart persistence model, considering the impact of irradiance, significantly enhances the accuracy of power predictions (average *nMAE* = 2.5 %), as illustrated in Fig. S2 of SI. We encourage the use of this updated persistence model as the benchmark for future research.

Machine learning (ML) models remain popular research in PV power modeling, especially for power prediction, due to their ability to adaptively learn from historical data without requiring detailed knowledge of the PV system. In this research, the selection of ML models and the fine-tuning of hyperparameters were guided by the successful experience reported in the literature [31]. However, across the application of four PV systems with distinct sizes, climate zones, and module technologies, the optimal machine learning model varied from site to site, as presented in Fig. 12. Therefore, we recommend that future researchers interested in machine learning for power modeling not rely solely on a single model recommended in the literature. Instead, we suggest exploring different candidates to identify the most suitable model for the specific PV system. Drawing from this research and literature studies, potential

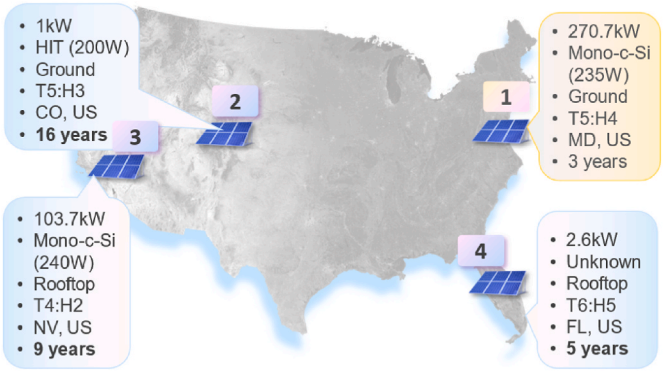

Fig. 11. Location of the PV systems in the U.S. for evaluation. System 1 is already analyzed in Section 4.1. System 2, 3, and 4 will be addressed. The metadata of the PV system is listed, including the capacity, module type, mounting type, climate zone, location, and duration of operation.

(a)

	Spring	Summer	Fall	Winter
Persistence	2.85	2.58	2.45	2.54
ANN	2.47	2.13	2.37	2.32
RF	1.62	1.26	1.73	2.03
SVR	1.81	1.57	1.47	1.63
LR	1.8	1.56	1.62	2.14
KR	1.71	1.3	1.57	1.76
Nominal	1.45	1.57	1.74	2.08
PVPro	1.16	1.12	1.34	1.35

nMAE (%): 2.5, 2.0, 1.5

(b)

	Spring	Summer	Fall	Winter
Persistence	9.99	5.91	4.37	4.65
ANN	8.65	2.98	2.91	3.87
RF	1.09	0.25	1.12	1.01
SVR	3.91	0.21	1.12	1.32
LR	4.89	0.17	1.06	0.7
KR	1.58	0.12	0.79	0.62
Nominal	0.89	0.08	1.19	1.7
PVPro	0.62	0.03	0.79	0.65

nBE>10% (%): 8, 6, 4, 2

Fig. 10. Errors of the estimated power (*nMAE*) and the power overestimation ratio (*nBE* >10 %) under the four seasons in 2018. Overall, the summer is associated with lower power errors, whereas the winter and spring tend to display higher errors. PV-Pro outperforms other techniques across all seasons from both metrics.

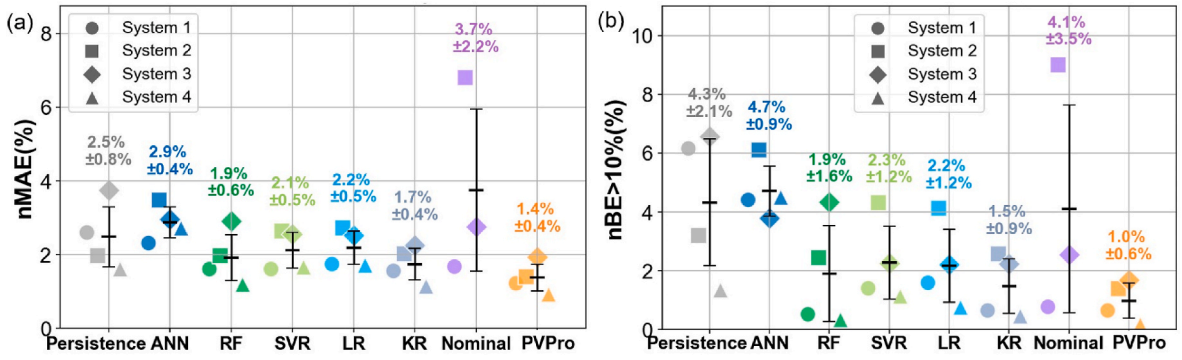

Fig. 12. Error of modeled power (*nMAE*) and power overestimation ratio (*nBE* >10 %) of four PV systems. The '*mean* ± *std*' of the four systems is marked above for each technique. PV-Pro consistently demonstrates robust performance and lower power error and overestimation ratio across different PV systems.

models to evaluate may include kernel ridge (KR), support vector regression (SVR), and random forest (RF).

For the data-driven models, like machine learning and PVPro (fusion of statistical and physical models), the availability of data is crucial. In this research, each machine learning is applied with its optimal length of historical data for training, as discussed in Section 2.1.4. Some models, like ANN, may require up to 60 days of training data for optimal performance. Here, we test an extreme case by providing these models with short-length data (3 days) and present their performance in Fig. S3 in SI. Interestingly, the results indicate that PV-Pro can still achieve a low power modeling error (*nMAE* = 1.26 %) even with this limited training data. This highlights the suitability of PVPro for application in newly-installed PV systems.

The training time for machine learning and PVPro models, using their optimal length of training data, is within 2 s on the Apple MacBook Pro with an M1 16G chip. This rapid training time fully supports a frequent update (like daily) of model parameters for power prediction.

The power modeling using PVPro is fundamentally a hybridization of statistical (data-driven) and physical methods. The leverage of historical data serves to reconstruct the physical model of the PV system. The power modeling, achieved through equivalent-circuit modeling, ensures that the output power adheres to the physical rules of PV cells and is fully interpretable. As PVPro rebuilds the model by fitting the production data, the challenge mainly lies in the quality of the data and fitting process. For future work, the pre-processing step will be enhanced to improve the data quality, especially in the identification of the operation conditions (inverter on MPP or clipping), removal of outliers, and use of clear-sky data. The current fitting process requires setting lower and upper bounds for the SDM parameters, with the default settings being constant and covering a wide range. An algorithm will be developed to adjust these ranges based on measurements and previously extracted parameters to further improve the fitting process. In essence, the model rebuilt by PVPro, which mirrors the current degradation status of the PV system, not only enables precise real-time health monitoring for operational and maintenance purposes but also holds promise for accurate and robust power prediction of degraded PV systems.

6. Conclusion

This paper presents a dynamically-updated physical model technique (PVPro) for power modeling for degraded PV systems. Using basic production and weather data PVPro dynamically reconstructs a precise physical model to reflect the current degradation status of the PV system without the need for additional characterization. PVPro is compared with the popular power modeling techniques in the literature. The results reveal that PVPro achieves an outstanding power prediction performance with the average *nMAE* = 1.4 % across four field PV systems, surpassing the best of other techniques with a reduction of error of 17.6 %. Additionally, PVPro demonstrates robustness across different seasons and degradation severity. Moreover, PVPro performs well with a limited amount of operational data (3 days), making it suitable for application in newly-installed PV systems. Future work will focus on the improvement of the pre-processing of data and the application of PVPro on more large-scale PV systems. This approach, developed as a Python-based open-source tool, holds promise for integration into real-time health monitoring and physical-model-based power prediction of degraded PV systems.

Data and code availability

The field PV datasets are from the PVDAQ data lake and National Institute of Standards and Technology (NIST): https://data.openei.org/submissions/4568, https://pvdata.nist.gov.

The proposed tool (PVPro) is coded as a Python-based package available at GitHub repository: https://github.com/DuraMAT/pvpro.

CRediT authorship contribution statement

Baojie Li: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Conceptualization. **Xin Chen:** Writing – review & editing, Conceptualization. **Anubhav Jain:** Writing – review & editing, Supervision.

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.

Acknowledgement

The project was primarily funded and intellectually led as part of the Durable Modules Consortium (DuraMAT), an Energy Materials Network Consortium funded under Agreement 32509 by the U.S. Department of Energy (DOE), Office of Energy Efficiency & Renewable Energy, Solar Energy Technologies Office (EERE, SETO). Lawrence Berkeley National Laboratory is funded by the DOE under award DE-AC02-05CH11231. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.renene.2024.121493.

References

[1] T. Ma, H. Yang, L. Lu, Solar photovoltaic system modeling and performance prediction, Renew. Sustain. Energy Rev. 36 (2014) 304–315, https://doi.org/10.1016/J.RSER.2014.04.057.

[2] S.R. Madeti, S.N. Singh, Monitoring system for photovoltaic plants: a review, Renew. Sustain. Energy Rev. 67 (2017) 1180–1207, https://doi.org/10.1016/J.RSER.2016.09.088.

[3] R. Ahmed, V. Sreeram, Y. Mishra, M.D. Arif, A review and evaluation of the state-of-the-art in PV solar power forecasting: techniques and optimization, Renew. Sustain. Energy Rev. 124 (2020) 109792, https://doi.org/10.1016/J.RSER.2020.109792.

[4] M. Kumar, P. Malik, R. Chandel, S.S. Chandel, Development of a novel solar PV module model for reliable power prediction under real outdoor conditions, Renew. Energy 217 (2023) 119224, https://doi.org/10.1016/J.RENENE.2023.119224.

[5] C. Brester, V. Kallio-Myers, A.V. Lindfors, M. Kolehmainen, H. Niska, Evaluating neural network models in site-specific solar PV forecasting using numerical weather prediction data and weather observations, Renew. Energy 207 (2023) 266–274, https://doi.org/10.1016/J.RENENE.2023.02.130.

[6] I.M. Moreno-Garcia, E.J. Palacios-Garcia, V. Pallares-Lopez, I. Santiago, M. J. Gonzalez-Redondo, M. Varo-Martinez, R.J. Real-Calvo, Real-time monitoring system for a utility-scale photovoltaic power plant, Sensors 16 (2016) 770, https://doi.org/10.3390/S16060770, 770 16 (2016).

[7] D.S. Pillai, N. Rajasekar, A comprehensive review on protection challenges and fault diagnosis in PV systems, Renew. Sustain. Energy Rev. 91 (2018) 18–40, https://doi.org/10.1016/j.rser.2018.03.082.

[8] C. Voyant, G. Notton, S. Kalogirou, M.L. Nivet, C. Paoli, F. Motte, A. Fouilloy, Machine learning methods for solar radiation forecasting: a review, Renew. Energy 105 (2017) 569–582, https://doi.org/10.1016/j.renene.2016.12.095.

[9] L. Visser, T. AlSkaif, W. van Sark, Operational day-ahead solar power forecasting for aggregated PV systems with a varying spatial distribution, Renew. Energy 183 (2022) 267–282, https://doi.org/10.1016/J.RENENE.2021.10.102.

[10] E. Sarmas, E. Spiliotis, E. Stamatopoulos, V. Marinakis, H. Doukas, Short-term photovoltaic power forecasting using meta-learning and numerical weather prediction independent Long Short-Term Memory models, Renew. Energy 216 (2023) 118997, https://doi.org/10.1016/J.RENENE.2023.118997.

[11] A. Niccolai, E. Ogliari, A. Nespoli, R. Zich, V. Vanetti, Very short-term forecast: different classification methods of the whole sky camera images for sudden PV power variations detection, Energies 15 (2022) 9433, https://doi.org/10.3390/EN15249433, 9433 15 (2022).

[12] Z. Si, M. Yang, Y. Yu, T. Ding, Photovoltaic power forecast based on satellite images considering effects of solar position, Appl. Energy 302 (2021) 117514, https://doi.org/10.1016/J.APENERGY.2021.117514.

[13] M. Pierro, D. Gentili, F.R. Liolli, C. Cornaro, D. Moser, A. Betti, M. Moschella, E. Collino, D. Ronzio, D. van der Meer, Progress in regional PV power forecasting: a sensitivity analysis on the Italian case study, Renew. Energy 189 (2022) 983–996, https://doi.org/10.1016/J.RENENE.2022.03.041.

[14] A. Dolara, S. Leva, G. Manzolini, Comparison of different physical models for PV power output prediction, Sol. Energy 119 (2015) 83–99, https://doi.org/10.1016/J.SOLENER.2015.06.017.

[15] E. Ogliari, A. Dolara, G. Manzolini, S. Leva, Physical and hybrid methods comparison for the day ahead PV output power forecast, Renew. Energy 113 (2017) 11–21, https://doi.org/10.1016/J.RENENE.2017.05.063.

[16] H.T.C. Pedro, D.P. Larson, C.F.M. Coimbra, A comprehensive dataset for the accelerated development and benchmarking of solar forecasting methods, J. Renew. Sustain. Energy 11 (2019), https://doi.org/10.1063/1.5094494.

[17] D.P. Larson, L. Nonnenmacher, C.F.M. Coimbra, Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest, Renew. Energy 91 (2016) 11–20, https://doi.org/10.1016/J.RENENE.2016.01.039.

[18] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F.J. Martinez-de-Pison, F. Antonanzas-Torres, Review of photovoltaic power forecasting, Sol. Energy 136 (2016) 78–111, https://doi.org/10.1016/J.SOLENER.2016.06.069.

[19] J. Gaboitaolelwe, A.M. Zungeru, A. Yahya, C.K. Lebekwe, D.N. Vinod, A.O. Salau, Machine learning based solar photovoltaic power forecasting: a review and comparison, IEEE Access 11 (2023) 40820–40845, https://doi.org/10.1109/ACCESS.2023.3270041.

[20] J. Tian, R. Ooka, D. Lee, Multi-scale solar radiation and photovoltaic power forecasting with machine learning algorithms in urban environment: a state-of-the-art review, J. Clean. Prod. 426 (2023) 139040, https://doi.org/10.1016/J.JCLEPRO.2023.139040.

[21] A. Mellit, A.M. Pavan, V. Lughi, Deep learning neural networks for short-term photovoltaic power forecasting, Renew. Energy 172 (2021) 276–288, https://doi.org/10.1016/J.RENENE.2021.02.166.

[22] K. Wang, X. Qi, H. Liu, A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network, Appl. Energy 251 (2019) 113315, https://doi.org/10.1016/J.APENERGY.2019.113315.

[23] T. Limouni, R. Yaagoubi, K. Bouziane, K. Guissi, E.H. Baali, Accurate one step and multistep forecasting of very short-term PV power using LSTM-TCN model, Renew. Energy 205 (2023) 1010–1024, https://doi.org/10.1016/J.RENENE.2023.01.118.

[24] F. Wang, Z. Xuan, Z. Zhen, K. Li, T. Wang, M. Shi, A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework, Energy Convers. Manag. 212 (2020) 112766, https://doi.org/10.1016/J.ENCONMAN.2020.112766.

[25] J. Wang, R. Ran, Z. Song, J. Sun, Short-term photovoltaic power generation forecasting based on environmental factors and GA-SVM, J. Electr. Eng. Technol. 12 (2017) 64–71, https://doi.org/10.5370/JEET.2017.12.1.064.

[26] B. Wolff, J. Kühnert, E. Lorenz, O. Kramer, D. Heinemann, Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data, Sol. Energy 135 (2016) 197–208, https://doi.org/10.1016/J.SOLENER.2016.05.051.

[27] D. Liu, K. Sun, Random forest solar power forecast based on classification optimization, Energy 187 (2019) 115940, https://doi.org/10.1016/J.ENERGY.2019.115940.

[28] M. Massaoudi, I. Chihi, L. Sidhom, M. Trabelsi, S.S. Refaat, F.S. Oueslati, Enhanced random forest model for robust short-term photovoltaic power forecasting using weather measurements, Energies 14 (2021) 3992, https://doi.org/10.3390/EN14133992, 3992 14 (2021).

[29] M. AlShafeey, C. Csáki, Evaluating neural network and linear regression photovoltaic power forecasting models based on different input methods, Energy Rep. 7 (2021) 7601–7614, https://doi.org/10.1016/J.EGYR.2021.10.125.

[30] M.J. Mayer, Benefits of physical and machine learning hybridization for photovoltaic power forecasting, Renew. Sustain. Energy Rev. 168 (2022) 112772, https://doi.org/10.1016/J.RSER.2022.112772.

[31] D. Markovics, M.J. Mayer, Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction, Renew. Sustain. Energy Rev. 161 (2022) 112364, https://doi.org/10.1016/J.RSER.2022.112364.

[32] A. Alcañiz, D. Grzebyk, H. Ziar, O. Isabella, Trends and gaps in photovoltaic power forecasting with machine learning, Energy Rep. 9 (2023) 447–471, https://doi.org/10.1016/J.EGYR.2022.11.208.

[33] M.A. Hameed, I. Kaaya, M. Al-Jbori, Q. Matti, R. Scheer, R. Gottschalg, Analysis and prediction of the performance and reliability of PV modules installed in harsh climates: case study Iraq, Renew. Energy 228 (2024) 120577, https://doi.org/10.1016/J.RENENE.2024.120577.

[34] A.M. Humada, M. Hojabri, S. Mekhilef, H.M. Hamada, Solar cell parameters extraction based on single and double-diode models: a review, Renew. Sustain. Energy Rev. 56 (2016) 494–509, https://doi.org/10.1016/J.RSER.2015.11.051.

[35] E. Ogliari, A. Dolara, G. Manzolini, S. Leva, Physical and hybrid methods comparison for the day ahead PV output power forecast, Renew. Energy 113 (2017) 11–21, https://doi.org/10.1016/J.RENENE.2017.05.063.

[36] M.J. Mayer, G. Gróf, Extensive comparison of physical models for photovoltaic power forecasting, Appl. Energy 283 (2021) 116239, https://doi.org/10.1016/J.APENERGY.2020.116239.

[37] W. Wang, D. Yang, N. Huang, C. Lyu, G. Zhang, X. Han, Irradiance-to-power conversion based on physical model chain: an application on the optimal configuration of multi-energy microgrid in cold climate, Renew. Sustain. Energy Rev. 161 (2022) 112356, https://doi.org/10.1016/J.RSER.2022.112356.

[38] A.R. Jordehi, Parameter estimation of solar photovoltaic (PV) cells: a review, Renew. Sustain. Energy Rev. 61 (2016) 354–371, https://doi.org/10.1016/J.RSER.2016.03.049.

[39] A.M. Humada, S.Y. Darweesh, K.G. Mohammed, M. Kamil, S.F. Mohammed, N. K. Kasim, T.A. Tahseen, O.I. Awad, S. Mekhilef, Modeling of PV system and parameter extraction based on experimental data: review and investigation, Sol. Energy 199 (2020) 742–760, https://doi.org/10.1016/J.SOLENER.2020.02.068.

[40] B. Li, D. Diallo, A. Migan-Dubois, C. Delpha, Performance evaluation of IEC 60891: 2021 procedures for correcting I–V curves of photovoltaic modules under healthy and faulty conditions, Prog. Photovoltaics Res. Appl. (2022), https://doi.org/10.1002/PIP.3652.

[41] B. Li, T. Karin, B.E. Meyers, X. Chen, D.C. Jordan, C.W. Hansen, B.H. King, M. G. Deceglie, A. Jain, Determining circuit model parameters from operation data for PV system degradation analysis: pvpro, Sol. Energy 254 (2023) 168–181, https://doi.org/10.1016/J.SOLENER.2023.03.011.

[42] X. Sun, R.V.K. Chavali, M.A. Alam, Real-time monitoring and diagnosis of photovoltaic system degradation only using maximum power point—the Suns-Vmp method, Prog. Photovoltaics Res. Appl. 27 (2019) 55–66, https://doi.org/10.1002/PIP.3043.

[43] B.E. Meyers, E. Apostolaki-Iosifidou, L.T. Schelhas, Solar Data Tools: Automatic Solar Data Processing Pipeline, Conference Record of the IEEE Photovoltaic Specialists Conference 2020-June, 2020, https://doi.org/10.1109/PVSC45281.2020.9300847, 0655–0656.

[44] W. De Soto, S.A. Klein, W.A. Beckman, Improvement and validation of a model for photovoltaic array performance, Sol. Energy 80 (2006) 78–88, https://doi.org/10.1016/J.SOLENER.2005.06.010.

[45] W.F. Holmgren, C.W. Hansen, M.A. Mikofski, Pvlib python: a python package for modeling solar energy systems, J. Open Source Softw. 3 (2018) 884, https://doi.org/10.21105/JOSS.00884.

[46] Z. Allen-Zhu, Y. Li, Z. Song, A Convergence Theory for Deep Learning via Over-parameterization, 36th International Conference on Machine Learning, ICML, 2019, pp. 362–372, https://doi.org/10.48550/arxiv.1811.03962, 2019-June (2018).

[47] D.C. Liu, J. Nocedal, On the limited memory BFGS method for large scale optimization, Math. Program. 45 (1 45) (1989) 503–528, https://doi.org/10.1007/BF01589116 (1989).

[48] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, J. Shelby, The national solar radiation data base (NSRDB), Renew. Sustain. Energy Rev. 89 (2018) 51–60, https://doi.org/10.1016/J.RSER.2018.03.003.

[49] D.L. King, W.E. Boyson, J.A. Kratochvill, Photovoltaic array performance model, SAND2004-3535 Report, https://www.osti.gov/servlets/purl/919131, 2004. (Accessed 8 September 2024).

[50] M. Boyd, T. Chen, B. Dougherty, NIST Campus Photovoltaic (PV) Arrays and Weather Station Data Sets, National Institute of Standards and Technology. U.S.

Department of Commerce, Washington, D.C., 2017. https://doi.org/10.18434/M3S67G. (Accessed 8 September 2024).
[51] J. Ascencio-Vásquez, K. Brecl, M. Topič, Methodology of Köppen-Geiger-Photovoltaic climate classification and implications to worldwide mapping of PV system performance, Sol. Energy 191 (2019) 672–685, https://doi.org/10.1016/J.SOLENER.2019.08.072.
[52] T. Karin, C.B. Jones, A. Jain, Photovoltaic degradation climate zones. Conference Record of the IEEE Photovoltaic Specialists Conference, 2019, pp. 687–694, https://doi.org/10.1109/PVSC40753.2019.8980831.
[53] NREL, Photovoltaic data acquisition (PVDAQ) public datasets. https://data.openei.org/submissions/4568, 2021. (Accessed 26 September 2024).
[54] C. Monteiro, T. Santos, L.A. Fernandez-Jimenez, I.J. Ramirez-Rosado, M. S. Terreros-Olarte, Short-term power forecasting model for photovoltaic plants based on historical similarity, Energies 6 (2013) 2624–2643, https://doi.org/10.3390/EN6052624, 2624-2643 6 (2013).