VALIDATION OF A PV GENERATION MODEL FOR SIMULATION OF WIDE AREA AGGREGATED DISTRIBUTED PV POWER GENERATION THAT TAKES INDIVIDUAL SYSTEMS LOCATION AND ORIENTATION INTO ACCOUNT

Lisa Molin, Sara Ericson, David Lingfors[1], Joakim Munkhammar[1] and Johan Lindahl[2]
[1] Uppsala University
[2] Becquerel Sweden AB
[2] Staffansvägen 14, SE-74142 Knivsta, +46768511773, johan@becquerelsweden.se

ABSTRACT: Understanding the photovoltaic (PV) power generation's temporal and spatial patterns is vital for grid balancing. This study aims to validate a simulation model for historical decentralized PV power generation, extending it to encompass the unique orientation of all PV systems within the Swedish municipality Knivsta. In a previous research project, a Convolutional Neural Network exhibited a 95% accuracy of identifying PV systems within Knivsta. In this project, using Light Detection and Ranging data, the orientation and area of detected PV systems was estimated. By combining this information with local weather and irradiance data, historical PV power generation was simulated. The regression analysis demonstrates strong correspondence between simulated and measured hourly generation for six reference systems, with coefficients of determination between 0.69–0.83. This study derives generic module parameters based on installation year and an average DC-to-AC ratio, enabling municipal-level simulations. Simulations for 2022, considering one scenario with optimal orientation for all PV systems and one scenario with derived real-condition orientations, reveal a smoothing effect in the daily pattern of aggregated PV generation, if considering real orientations. At the peak hour, power generation was found to be 10% lower when considering individual orientations compared to assuming optimal orientation across all facilities.
Keywords: Photovoltaics, LiDAR, System orientation, Power simulation, Grid balance

1 INTRODUCTION

Solar photovoltaic (PV) installations have experienced rapid growth over the past decade [1]. Projections from the International Energy Agency indicate that this upward trajectory will continue, with annual installations consistently setting new records [2]. Consequently, this trend is expected to drive further expansion in small-scale and decentralized electricity generation. For instance, in Sweden, distributed PV capacity reached 1 453 MW by the end of 2021, constituting a remarkable 92% of the total installed PV capacity [3].

The decentralized nature of distributed PV presents challenges in accurately measuring and predicting power generation. A significant portion of PV-generated power is self-consumed, leaving only grid-fed electricity subject to measurement by distribution system operators (DSOs) [4]. Consequently, the total aggregated PV electricity generation remains unknown in Sweden, as well as in most other countries globally. This lack of insight into the timing and quantity of PV-generated power poses significant challenges for DSOs in terms of grid balancing and capacity planning [5]. Consequently, there is a need for methods to forecast and estimate expected electricity generation from distributed PV systems [6].

A grid-based simulation model for the Nordic countries designed by [7], was tailored in a Swedish research project [8], to calculate the average hourly power output per installed DC capacity for PV systems. This model necessitates technical specifications of system components, including installed direct current (DC) and alternating current (AC) capacities, as well as the orientation (tilt and azimuth) of the PV modules. The model requires technical information of the system components, the installed DC and AC capacities, and the orientation of the PV modules. It also applies local hourly weather and irradiance data, with a spatial resolution of 2.5×2.5 km, which requires the geographical location of the PV system [9,10].

Recent research have yielded effective methods for remotely determining the precise geographical positions of PV and solar thermal panels, employing aerial imagery and machine learning techniques [11–17]. In one study, succeeding this work, a convolutional neural network (CNN) trained on orthophotos demonstrated a remarkable accuracy, achieving a 95% success rate in detecting PV systems within a Swedish municipality [16].

A significant challenge in modelling distributed PV power generation lies in the often unknown orientation of solar panels [6, 12]. Nonetheless, studies have showed the potential of using light detection and ranging (LiDAR) data to estimate the orientation of PV systems [18–20]. In a Swedish study [20], PV system orientations were assessed by computing the orientations of roof facets, using LiDAR data provided by the Swedish Land Survey. According to [19], LiDAR-derived orientations were found to be more accurate than reported data from the system owners, which often contain substantial uncertainties.

The combination of aerial imagery, LiDAR data and machine learning techniques enables remote geographical positioning and orientation estimations of distributed PV systems. When coupled with the previously mentioned simulation model [8] for computing the average hourly power output, this approach makes it possible to simulate the electricity generation from PV systems with known technical characteristics. To our knowledge, estimating the aggregated PV generation of distributed systems in a large geographic area using this approach, has so far not been done. Further, this methodology lacks validity. Therefore, in this study we implemented this method to simulate the hourly power generation of six reference PV systems and validated the result, by comparing the simulated PV power generation to the measured generation.

The underlaying motivation of the study is that predicting aggregated power generation, where the individual orientations are considered, can offer benefits for DSOs. As they lack the information of the orientation of PV systems, the Swedish DSOs traditionally operate on the assumption that all PV systems simultaneously can deliver their nameplate AC capacities, to ensure the adequacy of grid dimensions. This scenario is unlikely to

10.4229/EUPVSEC2023/5DV.2.10
020527-001

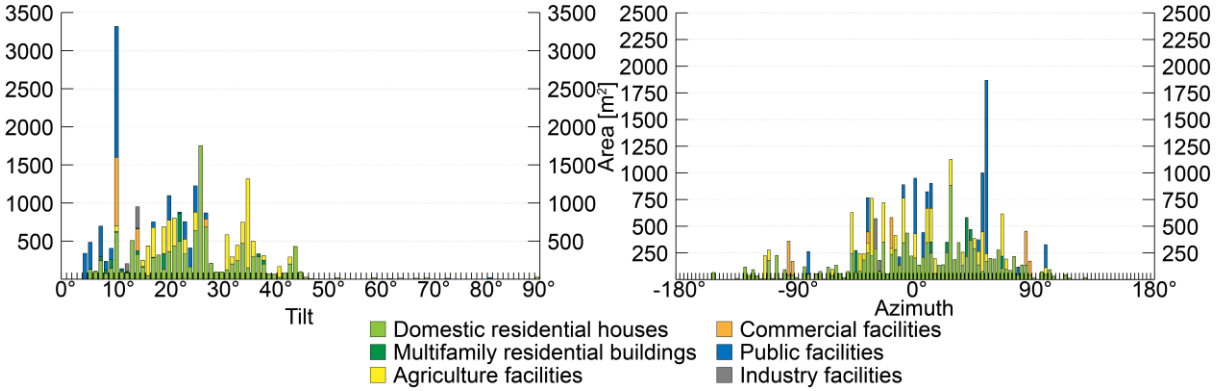

Figure 1. The tilt and the azimuth of all PV systems in Knivsta Municipality, identified and classified in the study of [10] and where the orientations have been calculated by the LiDAR methodology presented in [11].

occur since PV systems are typically installed with the same orientation as the roof facet [16,18], resulting in a wide spread of orientations, which therefore are not optimally oriented [21]. Based on the methodology-approach described earlier, the orientations across the Swedish Municipality of Knivsta were assessed and are illustrated in Figure 1. The broad spectrum of orientations within the municipality serves as compelling evidence that PV systems are improbable to generate their installed AC capacity in a synchronized manner. In addition to the primary goal of validating the PV power simulation model, this study also aspires to offer an initial quantitative assessment of the magnitude of the smoothing effect resulting from the wide spread of orientations of PV systems in the Swedish Municipality of Knivsta.

2 METHOD

This section presents the model for simulating PV power generation, the validation of the model, the data used in this study and assumptions made when simulating on a municipal scale.

2.1 Simulating PV power generation

2.1.1 PV power model

The average hourly electricity generation from a PV system per kW installed DC capacity P_{DC}/kW_{DC} (W/kW_{DC}) is calculated with a modified version of Eq. (1) from Campana et. Al [7] by;

$$P_{DC/kW_{DC}} = G_{g,t}\left[1+\frac{\mu}{\eta}(T_a - T_{STC}) + \frac{\mu}{\eta}\frac{9.5}{5.7+3.8v}\frac{(NOCT-20)}{800}(1-\eta)G_{g,t}\right], \quad (1)$$

where η is the efficiency of the PV module at standard test conditions (STC) (%), μ is the temperature coefficient of the maximum power efficiency (%/°C), T_a is the ambient temperature (°C), T_{STC} is the temperature during STC (25°C), v is the wind speed (m/s), NOCT is the nominal operating cell temperature (°C) and $G_{g,t}$, is the global tilted irradiance (W/m^2).

The inverter was simulated with the Steca Grid 300 efficiency curve [7]. To handle potential curtailment from the inverter, the model uses a minimizing function to generate the hourly AC power output per installed kW DC capacity, $P_{AC/kWDC}$, at every hour expressed by;

$$P_{AC/kW_{DC}} = min\left(P_{DC/kW_{DC}}, \frac{1000}{P_{max,DC}/P_{max,AC}}\right), \quad (2)$$

where $P_{max,DC}$ is the installed DC capacity and $P_{max,AC}$ is the installed AC capacity.

To generate the final total hourly power output from a PV system, in kilowatt instead of watt, P_{AC} (kW), the power generation per installed kW_{DC} needs to be multiplied with the nameplate capacity of the system, using the following equation:

$$P_{AC} = \frac{P_{AC/kW_{DC}} \times P_{max,DC}}{1000}. \quad (3)$$

Calculating the global tilted irradiance $G_{g,t}$, involves the both a transposition model and an absorption model. The transposition model utilized in the study, known as Perez1990, is developed for Nordic countries and publicly available in the PV_LIB [22]. The absorption model takes the irradiance output from the transposition model, the tilt of the PV modules, the solar elevation and the angle of incidence, to output the global tilted irradiance $G_{g,t}$. Spectral losses and losses due to snow, soiling and shadows are neglected in the model [7].

2.1.2 Weather and irradiance data

The weather and irradiance data employed in this research is derived from the models STRÅNG and MESAN, developed by the Swedish Meteorological and Hydrological Institute (SMHI). The STRÅNG model produces (global horizontal and direct normal) irradiance data for all Nordic countries, with a spatial resolution of 2.5×2.5 km, on an hourly basis [9]. The MESAN model generates hourly meteorological data, such as the ambient temperature and wind speed, with the same spatial resolution as STRÅNG [10].

2.2 Model validation

The validation was executed in two steps through a comparison between the measured PV power generation and the computed generation, derived from the PV power model presented in Section 2.1.1. The first validation step compares the measured generation to simulated results with orientations provided by the reference system owners. In contrast, the second validation step employs individual LiDAR-derived orientations to all systems. This scenario aims to approximate real conditions and the resulting aggregated PV power generation. The time period for the simulations were limited to the years of

10.4229/EUPVSEC2023/5DV.2.10
020527-002

2018−2022, due to the availability of high-resolution weather data and LiDAR data since 2018 [10,23].

2.2.1 Reference data of PV systems

Hourly measurements of electricity generation were provided by CheckWatt AB for six PV systems (PVS1–PVS6), situated in the municipalities scanned in [16]. The reference systems have a letter indicating their respective municipality; “U” is Uppvidinge, “F” is Falun and “K” is Knivsta.

The gathered information about the reference systems were the coordinates (WGS84), the orientation (tilt and azimuth), installed DC and AC capacity, module and inverter model and the year of installation. If the systems consisted of two or more parts with different orientations, the installed DC capacity and orientations of each subsystem were reported, as well as the configuration of the inverter(s). Details on size and timing of system expansion during the studied years were also taken into consideration. The information from the reference system owners was cross-checked with the Swedish direct capital subsidy, system data reported to the local grid owners, the size and number of modules, as well as manually estimated azimuth from arial images. The characteristics of the reference systems are displayed in Table 1 below.

Table I: Characteristics of the reference systems.

System	Installed DC capacity [kW]	Installed AC capacity [kW]	Number of sub-systems
PVS1-U	10.1	9.2	1
PVS2-U	11.4	17.0	1
PVS3-F	17.2	15.0	1
PVS4-F	19.5	16.4	3
PVS5-K	8.1	8.2	2
PVS6-K	20.9	10.0	4

2.2.2 LiDAR data

In the second validation step, the orientations of the reference systems were derived from LiDAR data. It consists of laser pulses which gives spatial 3D-information about surfaces [19,24]. The LiDAR dataset employed in this research was gathered from SMHI in March 2018, with a resolution of 2.5×2.5 km [23]. The data was also utilized when simulating the aggregated PV power generation for Knivsta Municipality, which is described in a following section.

2.2.3 Regression analysis

For this study, the coefficient of determination (R^2), the mean absolute error (MAE), the mean squared error (MSE), the root mean squared error of residuals (RMSE) and the mean bias error (MBE) were calculated for all simulations of reference systems PVS1–PVS6. Scatter plots for each reference facility were plotted and color-coded based on the time of day and time of year. The time intervals were chosen by applying different intervals and their effect on appearing patterns.

2.3 Simulating on a municipal level

To assess the smoothing effect of the wide spread of PV system orientations, the aggregated PV power generation was simulated for all identified PV systems within Knivsta. The generation was simulated for every hour of 2022 for two different scenarios. The first scenario assumes that all systems are having the optimal orientation for Knivsta, which posits their ability to generate their nameplate AC capacity simultaneously. The optimal orientation for the period of 2018–2021 was calculated by the grid-based optimization tool developed in [8], which takes local weather patterns into account, and resulted in a tilt of 45.62° and an azimuth of 174.24° for Knivsta. The second scenario implements individual orientations approximated using LiDAR data. A schematic view of the data and models used for simulating the PV power generation on a municipal level is illustrated in Figure 2.

2.3.1 Generic module parameters and DC-to-AC ratio

To simulate the aggregated PV generation for distributed systems on a municipal scale, with unknown technical properties, generic module parameters were estimated based on the year of installation. In a previous project, module data sheets of generic PV modules available on the Swedish PV market were collected for years 2014−2019 [25]. Complementing this database with additional data sheets collected for this study from Swedish retailers, resulted in a database of 200 different data sheets for monocrystalline and polycrystalline silicone PV modules available on the Swedish market on different occasions between 2015−2020.

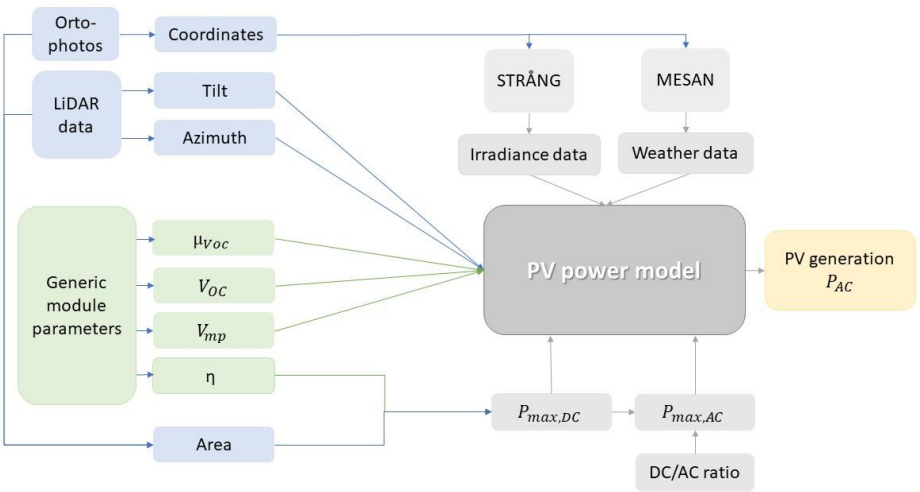

Figure 2. A schematic outline of the models and data for computing the PV power generation for an identified PV system.

10.4229/EUPVSEC2023/5DV.2.10
020527-003

From this database, an annual average of the following module parameters was derived for the years of 2015−2020; (1) the efficiency η, (2) the installed DC capacity, (3) the nominal voltage V_{mp}, (4) the open circuit voltage V_{oc} and (5) the temperature coefficient of the open circuit voltage μ_{Voc}. Additionally, the NOCT was assumed to be 45° for all modules. Consequently, this established an average and generic module representing each year between 2015 and 2020.

The installed DC capacity of each system was estimated based on the generic module efficiency and the LiDAR-corrected PV polygon area. Furthermore, an average DC-to-AC ratio was computed to determine the inverter curtailment and resulting hourly AC power output. A ratio 1.08 was derived from the average DC-to-AC ratio of 115 PV systems installed on single-family houses in Sweden collected for a study in 2020 [26]. The assumption was made that all PV systems had the same DC-to-AC ratio.

3 RESULTS

This section presents the validation of the PV power model. This is followed by the results from simulating the aggregated PV generation of distributed systems in Knivsta Municipality, using the PV power model.

3.1 Model validation

The total simulated and measured electricity production for the reference facilities is presented in Table II and III. As revealed by the tables, it is apparent that the model exhibits a consistent tendency to overestimate PV generation. Furthermore, it is important to emphasize the substantial variability in the percentage variances observed across the reference systems.

Table II: The measured and simulated PV power, using orientations provided by the reference system owners, and the percentual difference compared to the measured power.

System	Simulated PV power [MWh]	Measured PV power [MWh]	Difference [%]
PVS1-U	41.0	39.0	5.1
PVS2-U	43.9	40.1	9.5
PVS3-F	57.0	46.6	22.3
PVS4-F	67.7	65.5	3.4
PVS5-K	41.2	31.3	31.6
PVS6-K	86.3	64.8	33.3
Total	337.1	287.3	17.3

Table III: The measured and simulated PV power, using approximated orientations from LiDAR data, and the percentual difference compared to the measured power.

System	Simulated PV power [MWh]	Measured PV power [MWh]	Difference [%]
PVS1-U	40.9	39.0	4.7
PVS2-U	42.8	40.1	6.7
PVS3-F	53.5	46.6	15.0
PVS4-F	70.3	65.5	7.4
PVS5-K	41.3	31.3	31.8
PVS6-K	86.7	64.8	33.9
Total	335.6	287.3	16.8

Figure 3 provides scatter plots for one of the reference systems, PVS2-U, showcasing the results for both validation steps.

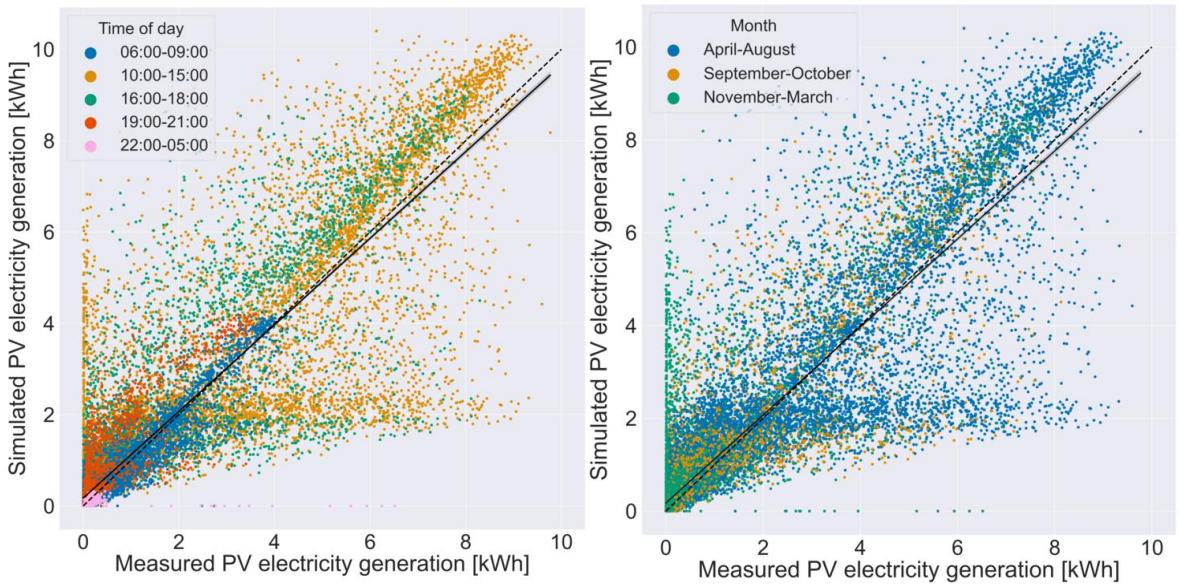

Figure 3a. Scatter plot for validation step 1 for PVS2-U, using orientations provided by the reference system owners. The data points are divided by colour, showcasing different time-intervals of the day on the left and for monthly time intervals on the right. The fitted regression is represented by a continuous line and the dotted line indicates a perfect model.

10.4229/EUPVSEC2023/5DV.2.10
020527-004

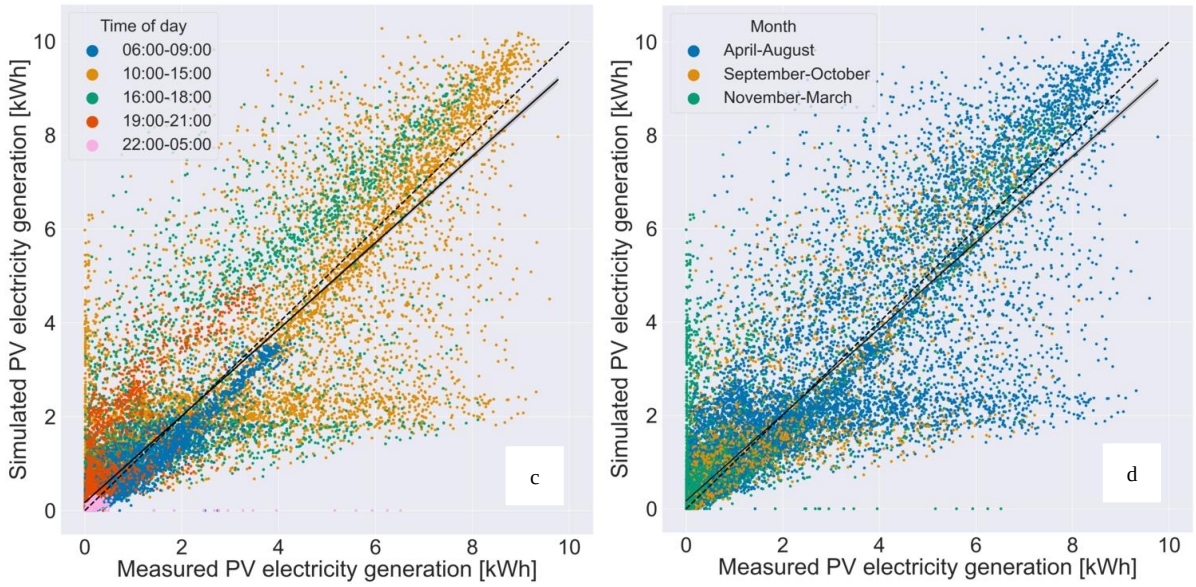

Figure 3b. Scatter plots for validation step 2 for PVS2-U, using individual orientations from LiDAR data. The data points are divided by colour, showcasing different time-intervals of the day on the left and for monthly time intervals on the right. The fitted regression is represented by a continuous line and the dotted line indicates a perfect model.

Several clusters of data points are apparent in Figure 3, at different times of the day and year, as denoted by the color-coded representations. These clusters were identified in several of the reference systems and are conceptually depicted in Figure 4, and are subject to further exploration in the subsequent section. Comprehensive scatter plots for all facilities are provided in Appendix A.

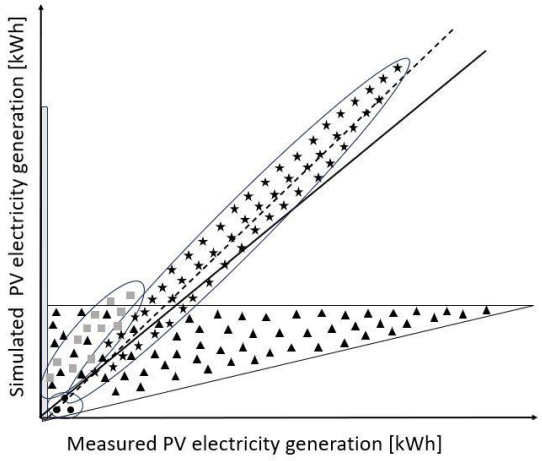

Figure 4. A graphic representation of reoccurring patterns in the reference system scatter plots.

3.1.1 Daily and yearly variation patterns

There is one cluster (stars in Figure 4) where the data points closely align with the perfect fit, showing that the simulated production values correspond well to the measured data. Within this cluster, it appears to be a slight overestimation of the simulated PV generation compared to the measured data. All daytime hours (6 am to 9 pm) occur in this cluster. However, a substantial portion of the data points is represented by blue dots in Figure 3b and 3d, indicating their predominant occurrence within the period of April to August.

A noticeable cluster or "tail" is evident, where the simulated PV electricity generation is approximately 2 kWh (triangles in Figure 4), representing instances where the model underestimates the PV generation compared to the measured data. This cluster has an upper limit, commonly close to 2−3 kWh, see Appendix A. It encompasses all hours of the day and months of the year, with the highest production levels on the right side, aligning with midday and summer months, and gradually tapering towards lower values on the left, which coincide with night-time and the winter months. This consistent behaviour of the PV power estimations, across all months and hours, could potentially be attributed to systematic faults, such as measurement inaccuracies in the MESAN and STRÅNG data.

An overestimation of the simulated PV electricity generation is displayed during the evening hours of 7−9 pm, as depicted by the red dots in Figure 3 (squares in Figure 4). This model overestimation was also observed when analysing the difference between the measured and simulated values over the day, as there is a peak in the percentual difference in-between 8−9 pm. A similar peak is discernible in the morning, around 6 am, although its magnitude is relatively lower. Additionally, within the same cluster, the model appears to overestimate electricity generation in the late afternoon hours 4−6 pm, as indicated by the green dots in Figure 3. In contrast, the PV power model consistently underestimates production during night-time and in the early morning hours (cluster of circles in Figure 4). This pattern is observed across all scatter plots in Appendix A.

Notably, there are certain hours along the y-axis where the power model predicts PV power generation despite the absence of measured production (the gray rectangle on the y-axis in Figure 4). These events are primarily observed during the months of November to March, evident in the various scatter plots featured in Appendix A. This pattern might be attributed to the losses caused by snow cover, as it is not accounted for in the model.

Lastly, there are some measured production hours on the x-axis, mostly during night-time, where the model predicts zero power generation. These instances are infrequent and generally minor, with only a few exceptions. PVS6-K breaks this pattern, exhibiting notable

10.4229/EUPVSEC2023/5DV.2.10
020527-005

positive measured values on the x-axis, which are likely due to errors in the measurements.

3.1.2 Regression analysis

The simulated PV power, using orientations supplied by the reference system owners, corresponds well with the measured PV generation with coefficients of determination between 0.70−0.82. The RMSE, which characterizes the typical dispersion of data points around the regression line, varies from 0.90−3.41 kWh. Every reference facility displays a positive MBE, aligning with the model's tendency to overestimate total production. All statistical measurements for the two validation steps are presented in Table IV and V.

Table IV: *Statistical measurements for validation step 1, employing orientations from reference system owners. For R^2, red is the lowest value, green is the highest value and yellow is the 50th percentile (median). For the other metrics, green is the highest value, and red is the lowest value. N denotes the number of hours simulated.*

PVS	N	R^2	RMSE [kWh]	MSE [kWh]	SSR [kWh]	MAE [kWh]	MBE [kWh]
1-U	33 912	0.77	0.90	4.44	30656	0.44	0.06
2-U	33 720	0.78	1.02	5.00	34554	0.47	0.11
3-F	30 806	0.70	3.41	11.52	105008	0.84	0.34
4-F	42 280	0.73	2.31	8.40	97764	0.71	0.05
5-K	43 796	0.77	0.52	2.24	22951	0.35	0.23
6-K	40 188	0.82	1.73	9.45	69620	0.78	0.54

Table V: *Statistical measurements for validation step 2, employing LiDAR-derived orientations. For R^2, red is the lowest value, green is the highest value and yellow is the 50th percentile (median). For the other metrics, green is the highest value, and red is the lowest value. N denotes the number of hours simulated.*

PVS	N	R^2	RMSE [kWh]	MSE [kWh]	SSR [kWh]	MAE [kWh]	MBE [kWh]
1-U	33 912	0.79	0.92	4.39	31306	0.45	0.05
2-U	33 720	0.79	0.99	4.75	33278	0.47	0.08
3-F	30 806	0.69	3.30	10.59	101670	0.80	0.23
4-F	42 280	0.74	2.7	9.06	100308	0.71	0.11
5-K	43 796	0.77	0.52	2.24	22761	0.34	0.23
6-K	40 188	0.83	1.67	9.59	67166	0.77	0.55

Table IV and V demonstrates that there are varying results among the reference systems, with certain systems demonstrating a stronger alignment between the simulated and measured PV power generation in comparison to others.

When simulating with LiDAR-derived orientations, the PV power generation corresponds slightly better to the measured values, yielding coefficients of determination ranging from 0.69−0.83 and RMSE values from 0.52−3.30 kWh. Other measurements of correlation produce similar outcomes to those observed with orientations provided by the facility owners, although the LiDAR-derived orientation marginally enhances these metrics as well. Nevertheless, this improvement in the relationship between the simulated and measured PV power is not significant. Unlike verification of the azimuth, which can easily be cross-checked using aerial imagery, verifying the tilt angle is more challenging. To test the sensitivity of the results when varying the angle, the angle was adjusted for each reference system. Altering the LiDAR-derived tilt by +20° angle led to a cumulative difference in simulated power ranging from -4.8−3.9% and altering the tilt by -20° led to a difference of -7.5−3.7%.

When examining the variations in the correlation between modelled and measured PV power generation over different times of the year, the average coefficient of determination for each month is summarized in Table VI.

Table VI: The R^2-values each month for the reference systems PVS1−PVS6.

PVS	1-U	2-U	3-F	4-F	5-K	6-K
January	0.54	0.40	0.51	0.38	0.35	0.42
February	0.61	0.49	0.59	0.63	0.51	0.67
March	0.87	0.87	0.75	0.72	0.62	0.85
April	0.86	0.86	0.78	0.78	0.82	0.85
May	0.76	0.76	0.73	0.73	0.81	0.78
June	0.79	0.80	0.74	0.75	0.85	0.77
July	0.74	0.72	0.72	0.70	0.66	0.79
August	0.78	0.76	0.70	0.74	0.57	0.66
September	0.78	0.78	0.73	0.75	0.83	0.80
October	0.75	0.75	0.67	0.66	0.77	0.82
November	0.43	0.41	0.52	0.51	0.67	0.70
December	0.34	0.25	0.29	0.22	0.28	0.36

The R^2-values in Table VI are considerably lower for the winter months, especially between November–February. Also, the relative difference in electricity generation was calculated to be largest during the same months. On the other hand, the total generated power and the absolute difference is lower during the winter months, compared to the rest of the year.

3.2 Simulating the PV generation in Knivsta Municipality

The generic module parameters that were derived for the relevant years of installation (2015–2020) are presented in Table VII.

Table VII: Generic module parameters for silicon modules in the years of 2015–2020.

Year	2015	2016	2017	2018	2019	2020
η [%]	16.3	17.0	17.8	18.3	18.9	20.5
V_{mp} [V]	32.2	31.8	32.7	33.3	35.5	38.2
V_{oc} [V]	39.5	38.9	40.3	40.4	43.0	45.5
μ_{Voc} [mV/°C]	-123	-122	-124	-117	-127	-124
Power [W]	265	281	298	309	329	416
Modules	39	24	24	44	30	39
Manu-facturers	14	15	17	12	9	10

A trend of rising η, V_{mp} $V_{oc,}$ and nominal power with increasing years is evident and expected. Exceptions from this trend, such as a decreasing V_{oc} year 2016, may be attributed to limited data that year. Therefore, the number of manufacturers and modules used as a basis for calculating the average technical parameter are also presented in Table VII.

Since the generic module parameters are estimations from limited data, a sensitivity analysis was conducted. The three parameters η, V_{mp} and μ_{Voc}, were adjusted to their minimum and maximum values within the dataset to test the weakness of these estimations. Altering the nominal voltage and the temperature coefficient had a minimal impact on the simulated aggregated PV power generation, compared to varying the efficiency, which significantly influenced the results.

3.2.1 Comparing the two simulation scenarios

The aggregated PV power generation of Knivsta

10.4229/EUPVSEC2023/5DV.2.10
020527-006

Municipality 2022 was calculated to 4 601 MWh, when simulating with individual LiDAR-derived orientations of the distributed systems. Simulating with all systems having the optimal orientation (of 45.62° tilt and 174.24° azimuth) resulted in a aggregated power generation of 5 200 MWh. The daily sum of the PV power generation for both simulations and the difference between them are visualized in Figure 5.

The difference in PV power between the two cases is relatively small during summer months (May–August), whereas it becomes considerably lager for the remainder of the year, except for the period from November to January, during which overall production is lower.

Given the significance of power generation data for grid sizing, a key aspect of this research is studying the day with the highest hourly power generation. In Figure 6, the days with the highest hourly PV generation for each simulation are presented. The peak PV generation in the scenario with individual orientations, of 3.72 MWh, occurs at 1 pm on May 18th. In the other scenario, where all systems are optimally oriented, the peak power generation of 4.14 MWh is achieved at 12–1 pm on March 30th. The simulation considering individual orientations is generating 3.28 MWh this same hour. Notably, during the hours of peak PV generation for both cases, the scenario with optimal orientations reaches the estimated aggregated AC capacity of 4.14 MWh. This scenario also reaches its peak another time, at 12 pm on April 5th, though this is not included in the figure. In contrast to the scenario implementing an optimal orientation, the scenario with individual orientations never generates the aggregated AC capacity, which is illustrated in Figure 6. This can be explained by that the orientations for PV systems in Knivsta Municipality are not optimally oriented, as evidenced by the dispersion in Figure 1.

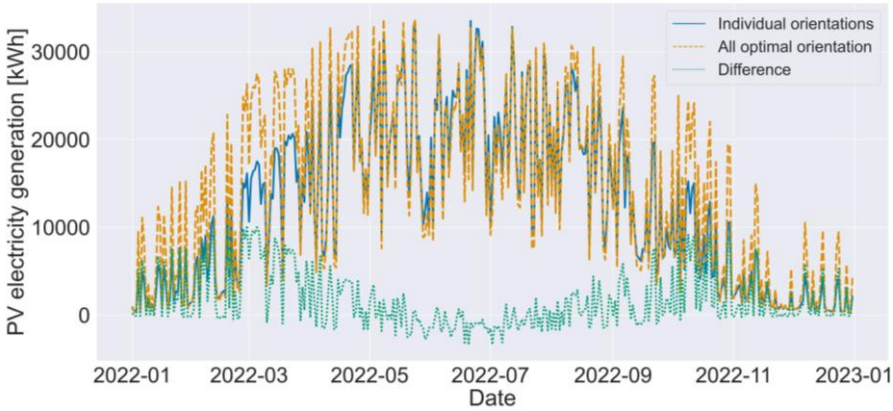

Figure 5. The daily aggregated PV power generation for Knivsta Municipality 2022, simulated for two different scenarios, and the daily difference in generation between the scenarios.

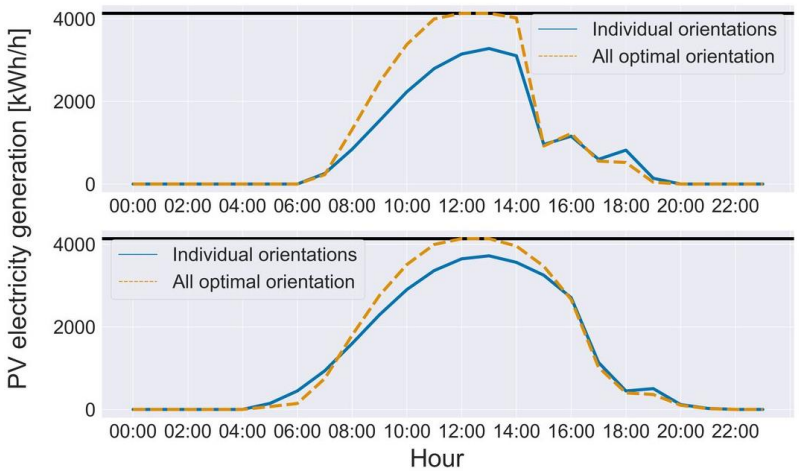

Figure 6. The aggregated PV power generation for two days with the highest generation hours 2022. The black line is the estimated aggregated AC capacity for Knivsta Municipality. The lower graph shows the peak hour with individual orientations on 1 pm May 18th and the graph above shows March 30th, where all optimal orientations achieve the aggregated AC capacity.

10.4229/EUPVSEC2023/5DV.2.10
020527-007

Another important aspect for comprehending PV generation characteristics is the average production curve. Figure 7 presents the average simulated power generation across all hours of the day for both scenarios. During morning and evening periods, the case of individual orientations exhibits higher simulated PV power generation. Conversely, the remaining hours of the day show greater production with all systems optimally oriented. Consequently, there are intersections occurring roughly at 8 am and 5 pm.

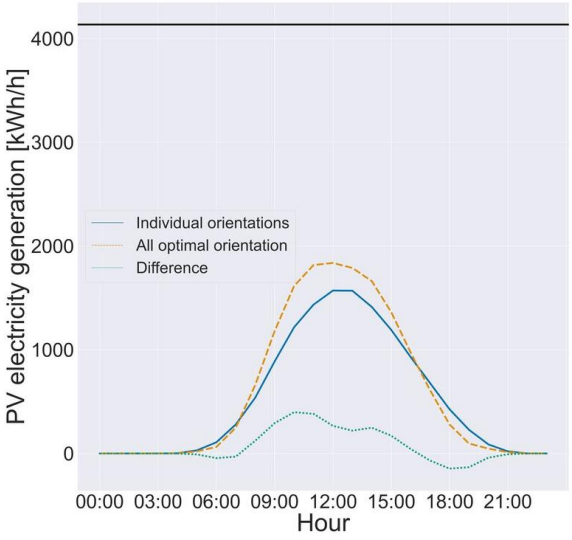

Figure 7. The hourly average PV electricity generation curve simulated for 2022. The black line is the approximated aggregated AC capacity of 4.14 MWh for Knivsta Municipality.

The peak mean production is observed at noon in both simulation scenarios, reaching 1.57 MWh for systems having individual orientations and 1.87 MWh for optimally oriented systems. It is important to note, as depicted in Figure 7, that neither of the average production curves approaches the estimated aggregated AC capacity for Knivsta Municipality. This average is with respect to all seasons, including the winter months characterized by low PV production.

4 DISCUSSION

It is essential to clarify that in simulating the aggregated PV power generation at the municipal level, certain assumptions were made, introducing uncertainties in the study's findings. Firstly, each polygon within Knivsta Municipality was modelled to have an individual inverter and a fixed DC-to-AC ratio, due to the lack of information on the actual configuration. In practice, some polygons may constitute subsystems situated within the same facility, potentially sharing a common inverter. Secondly, the generic module parameters for each year between 2015−2020 were based on the average from 200 modules on the Swedish market (ranging between 24−39 modules for each year). A more comprehensive dataset would be beneficial to refine the approximations of generic parameters. The importance of accurately estimating the module efficiency was highlighted in a sensitivity analysis, since varying this parameter significantly influenced the results. Additionally, the DC-to-AC ratio which governs the AC electricity generation, is currently sourced from 115 private residential PV systems. Expanding the database to include broader system capacities and different typical DC-to-AC ratios of different market segments would probably improve DC-to-AC ratio estimations.

Furthermore, it would be appropriate to compare the current inverter model of the Steca Grid 300 efficiency curve with other alternatives.

Another know drawback of the simulations are the inaccuracies in the weather and irradiance data sourced from the MESAN and STRÅNG models, which significantly impact the precision of the results. Though the average error statistics are not published for the studied years, as a reference, during the period from 1999−2009 the hourly RMSE for global radiation and the direct normal radiation from STRÅNG was 30% and 57%, respectively [27].

The initial finding in this study was that the modelled PV power corresponds well to the measured PV power generation. The results were marginally better when employing orientations obtained from LiDAR. This observation aligns with prior research [19], affirming a higher reliability of LiDAR data in orientation estimation compared to manual estimations of the PV system owners.

The model tends to overestimate PV power generation, especially in the late afternoon. This is possibly due to temperature differences between assumed and actual cell temperatures. The model employs air temperature data from MESAN, while cell temperatures tend to be higher in the afternoon, due to heat accumulation within modules and the roof. Consequently, the simulation might overestimate PV power generation when using a temperature coefficient based on lower cell temperatures. Likewise, the opposite situation applies to the morning discrepancy, wherein measured production exceeds the simulated production. Another explanation for the overestimation could be unaccounted shadowing effects from the surroundings. One study revealed that shading led to a 35% reduction in the anticipated global maximum power point for monocrystalline modules and a 33% decline for polycrystalline modules [28].

Further, considerably larger percentage errors of the simulated production compared to the recorded production, were noted during months of winter, accompanied by reduced R^2-values. One plausible explanation could be the influence of snow cover losses. A recent study estimated these losses to vary between 0−20% of the annual PV power generation yield in Sweden when using an average tilt angle of 25° [29], which differs from the mean tilt angle of the reference systems at 31°. Notably, the current PV power simulation model does not account for snow losses, potentially leading to an overestimation of PV electricity generation.

Since the largest absolute error between modelled and measured production occurs in summer daylight hours, shading-related losses exert a significant impact on simulation outcomes. Snow-related losses, on the other hand, have a comparatively smaller effect due to lower absolute magnitudes of simulation errors in winter. Therefore, the implementation of a shadow loss model is more crucial for enhancing result accuracy.

5 CONCLUSION

A model for simulating the historic PV power generation from individual systems was validated, employing LiDAR-derived orientation, as well as local

10.4229/EUPVSEC2023/5DV.2.10
020527-008

weather and irradiance data generated by the models STRÅNG and MESAN. Simulating with orientations provided by the reference system owners resulted in R^2-values between 0.70–0.84. Using orientations from LiDAR-data resulted in overall slightly improved metrics, indicating that LiDAR-data can be used to adequately estimate unknown orientations and may even generate more accurate results. Hence, the results from the validation suggest that the PV power model is valuable for estimating hourly PV power generation and that LiDAR data accurately estimates the orientations. Currently, the PV model neglects both shading and snow losses. Integrating these losses might decrease overestimations of the PV power model, thus enhancing accuracy.

The PV electricity generation model was further developed and used to estimate the aggregated distributed PV power generation in Knivsta Municipality. Simulating the aggregated distributed power generation considering unique LiDAR-derived orientations, never reached the total AC capacity installed in Knivsta. Furthermore, it was shown that the installed AC capacity is not generated in a synchronized manner. Instead, there is a noticeable smoothing effect in the daily pattern of the aggregated PV generation when implementing individual orientations, compared to one optimal orientation. This is explained by the wide spectrum of orientations in Knivsta. An initial quantification of this effect is that the power generation was 10% lower at the peak hour 2022, when considering individual orientations compared to assuming a uniform optimal orientation. This discovery can free grid capacity for further deployment of distributed PV systems.

6 ACKNOWLEDGEMENTS

This work was financed by the Swedish Energy Agency through the research project P2023-00440, and in addition supported by Solar Electricity Research Center Sweden (SOLVE) and the Swedish strategic research programme StandUp for Energy.

7 REFERENCES

[1] IRENA. Renewable capacity statistics 2021 Internaional Renewable Energy Agency (IRENA). Abu Dhabi: 2021.

[2] International Energy Agency. Renewable Energy Market Update - Outlook for 2021 and 2022. 2021.

[3] Lindahl J, Oller Westerberg A. National Survey Report of PV Power Applications in Sweden 2021. 2022.

[4] Luthander R, Widén J, Nilsson D, Palm J. Photovoltaic self-consumption in buildings: A review. Applied Energy 2015;142:80–94. https://doi.org/10.1016/j.apenergy.2014.12.028.

[5] Widén J, Shepero M, Munkhammar J. Probabilistic Load Flow for Power Grids With High PV Penetrations Using Copula-Based Modeling of Spatially Correlated Solar Irradiance. IEEE Journal of Photovoltaics 2017;7:1740–5. https://doi.org/10.1109/JPHOTOV.2017.2749004.

[6] Ramadhani UH, Lingfors D, Munkhammar J, Widén J. On the properties of residential rooftop azimuth and tilt uncertainties for photovoltaic power generation modeling and hosting capacity analysis. Solar Energy Advances 2023;3:100036. https://doi.org/10.1016/j.seja.2023.100036.

[7] Campana PE, Landelius T, Andersson S, Lundström L, Nordlander E, He T, et al. A gridded optimization model for photovoltaic applications. Solar Energy 2020;202:465–84. https://doi.org/10.1016/j.solener.2020.03.076.

[8] Lindahl J, Campana PE, Hansen MK. Beräkningsverktyg för optimering av solcellsparker. n.d.

[9] SMHI, Naturvårdsverket, Strålsäkerhetsmyndigheten. STRÅNG 2017. https://strang.smhi.se/ (accessed January 20, 2023).

[10] SMHI. Analysmodell (MESAN) | SMHI 2017. https://www.smhi.se/data/oppna-data/meteorologiska-data/analysmodell-mesan-1.30445 (accessed January 20, 2023).

[11] Kausika BB, Nijmeijer D, Reimerink I, Brouwer P, Liem V. GeoAI for detection of solar photovoltaic installations in the Netherlands. Energy and AI 2021;6:100111. https://doi.org/10.1016/j.egyai.2021.100111.

[12] Yu J, Wang Z, Majumdar A, Rajagopal R. DeepSolar: A Machine Learning Framework to Efficiently Construct a Solar Deployment Database in the United States. Joule 2018;2:2605–17. https://doi.org/10.1016/j.joule.2018.11.021.

[13] Kruitwagen L, Story KT, Friedrich J, Byers L, Skillman S, Hepburn C. A global inventory of photovoltaic solar energy generating units. Nature 2021;598:604–10. https://doi.org/10.1038/s41586-021-03957-7.

[14] Costa MVCV da, Carvalho OLF de, Orlandi AG, Hirata I, Albuquerque AO de, Silva FV e, et al. Remote Sensing for Monitoring Photovoltaic Solar Plants in Brazil Using Deep Semantic Segmentation. Energies 2021;14:2960. https://doi.org/10.3390/en14102960.

[15] Hu W, Bradbury K, Malof JM, Li B, Huang B, Streltsov A, et al. What you get is not always what you see—pitfalls in solar array assessment using overhead imagery. Applied Energy 2022;327:120143. https://doi.org/10.1016/j.apenergy.2022.120143.

[16] Lindahl J, Johansson R, Lingfors D. Mapping of decentralised photovoltaic and solar thermal systems by remote sensing aerial imagery and deep machine learning for statistic generation. Energy and AI 2023;14:100300.

[17] Ren S, Hu W, Bradbury K, Harrison-Atlas D, Malaguzzi Valeri L, Murray B, et al. Automated Extraction of Energy Systems Information from Remotely Sensed Data: A Review and Analysis. Applied Energy 2022;326:119876. https://doi.org/10.1016/j.apenergy.2022.119876.

[18] Killinger S, Lingfors D, Saint-Drenan Y-M, Moraitis P, van Sark W, Taylor J, et al. On the search for representative characteristics of PV systems: Data collection and analysis of PV system azimuth, tilt, capacity, yield and shading. Solar Energy 2018;173:1087–106. https://doi.org/10.1016/j.solener.2018.08.051.

[19] Lingfors D, Killinger S, Engerer NA, Widén J, Bright JM. Identification of PV system shading using a LiDAR-based solar resource assessment model: An evaluation and cross-validation. Solar Energy 2018;159:157–72. https://doi.org/10.1016/j.solener.2017.10.061.

[20] Lingfors D, Bright JM, Engerer NA, Ahlberg J, Killinger S, Widén J. Comparing the capability of low- and high-resolution LiDAR data with application to solar resource assessment, roof type classification and shading analysis. Applied Energy 2017;205:1216–30. https://doi.org/10.1016/j.apenergy.2017.08.045.

[21] Sadeghian H, Wang Z. A novel impact-

assessment framework for distributed PV installations in low-voltage secondary networks. Renewable Energy 2020;147:2179–94. https://doi.org/10.1016/j.renene.2019.09.117.
[22] Sandia National Laboratories. PV Performance Modeling Collaborative | PV_LIB Toolbox 2023. https://pvpmc.sandia.gov/applications/pv_lib-toolbox/ (accessed January 27, 2023).
[23] Lantmäteriet. Laserdata Nedladdning, skog 2023.
[24] Lindahl J. Automatisk kartläggning av solpaneler och generering av solelsprognoser från flygbilder och maskininlärning. n.d.
[25] Gustavsson U, Rosenqvist L. Further Development of Njord, a statistical Instrument for Estimating International Installed Photovoltaic Capacities. master thesis. Linköping University, 2021.
[26] Oller Westerberg A, Lindahl J. Kostnadsstruktur för svenska villasystem. Stockholm: n.d.
[27] Swedish Meteorological and Hydrological Institute. STRÅNG validation n.d. https://strang.smhi.se/validation/validation.html (accessed June 14, 2023).
[28] Bimenyimana S, Asemota GNO, Kemunto MC, Li L. Shading effects in photovoltaic modules: Simulation and experimental results. 2017 2nd International Conference on Power and Renewable Energy (ICPRE), 2017, p. 904–9. https://doi.org/10.1109/ICPRE.2017.8390665.
[29] van Noord M, Landelius T, Andersson S. Snow-Induced PV Loss Modeling Using Production-Data Inferred PV System Models. Energies 2021;14:1574. https://doi.org/10.3390/en14061574.

Appendix A

Appendix displaying scatter plots of the remaining five PV systems (PVS1U, PVS3F, PVS4F, PVS5K and PVS6K) that were used in this study for the validation of the previously mentioned PV power model.

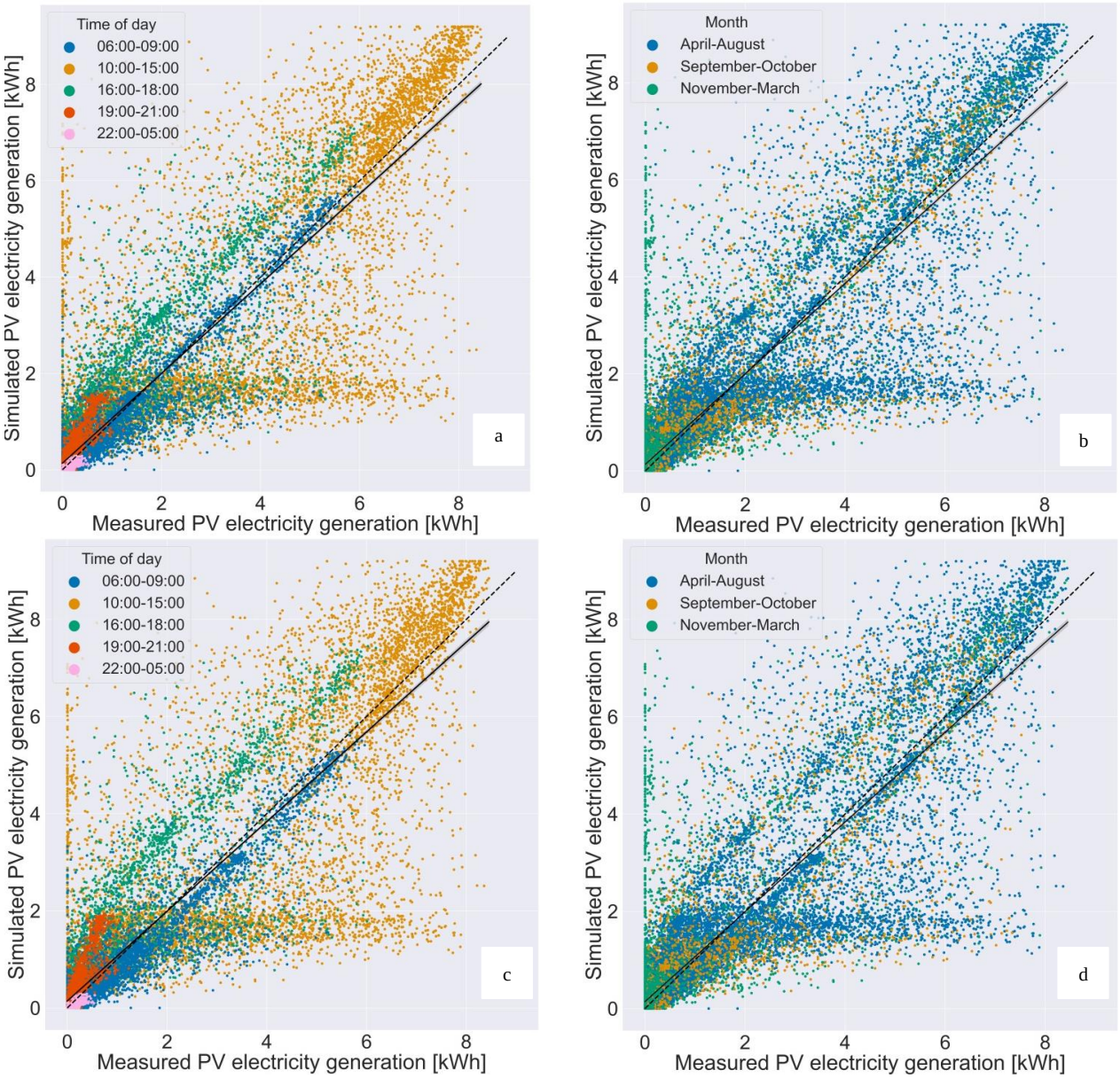

Figure 8. In (a) and (b), scatter plot for validation step 1, and in (c) and (d) validation step 2 for PVS1-U, both steps using known module- and inverter parameters disclosed by the system owners. The data points are divided using colour for time-intervals of the day in (a) and (c), and for monthly time intervals in (b) and (d). The fitted regression is represented by a continuous line and the dotted line indicates a perfect fit.

10.4229/EUPVSEC2023/5DV.2.10
020527-011

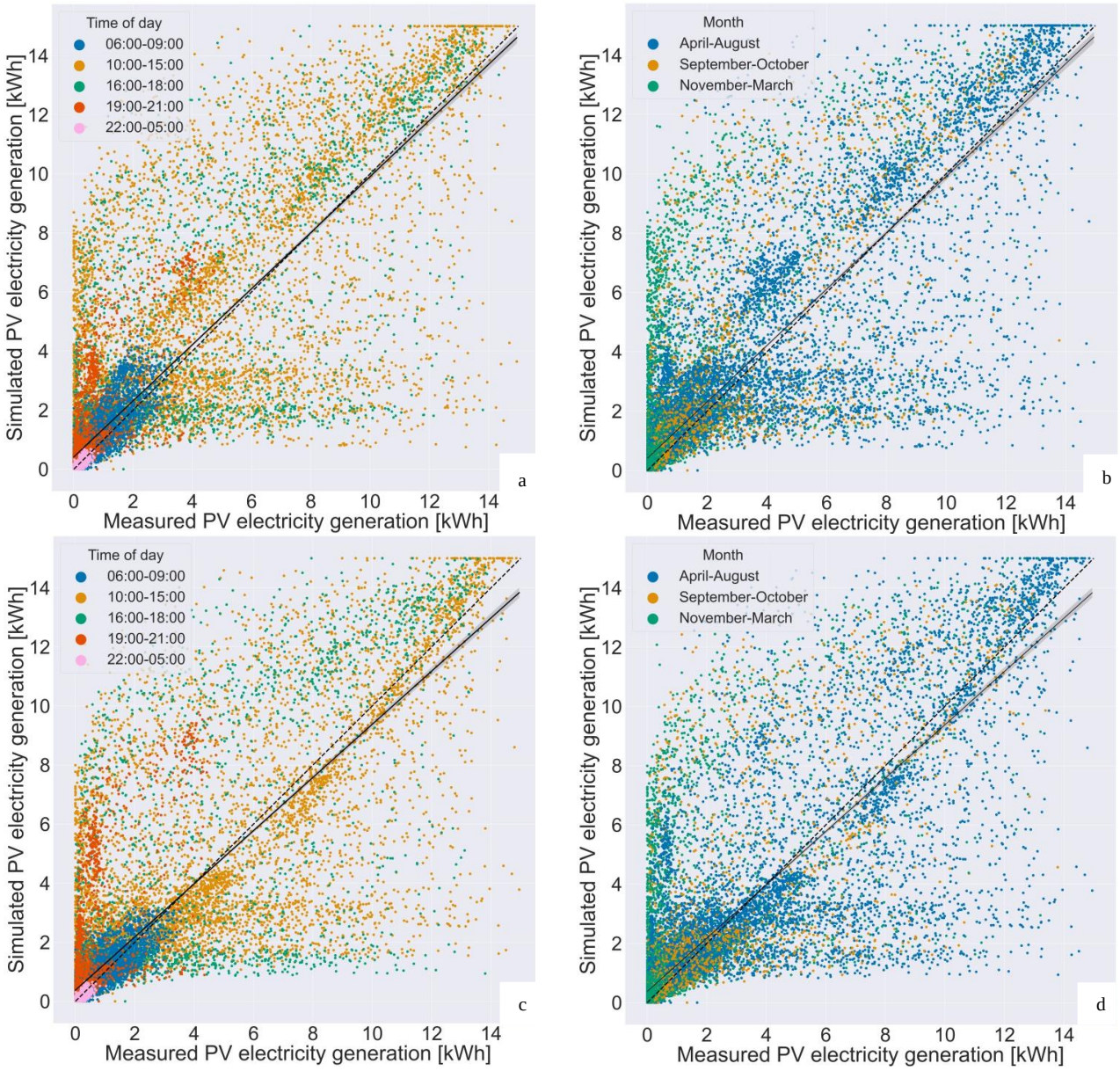

Figure 9. In (a) and (b), scatter plot for validation step 1, and in (c) and (d) validation step 2 for PVS3-F, both steps using known module- and inverter parameters disclosed by the system owners. The data points are divided using colour for time-intervals of the day in (a) and (c), and for monthly time intervals in (b) and (d). The fitted regression is represented by a continuous line and the dotted line indicates a perfect fit.

10.4229/EUPVSEC2023/5DV.2.10
020527-012

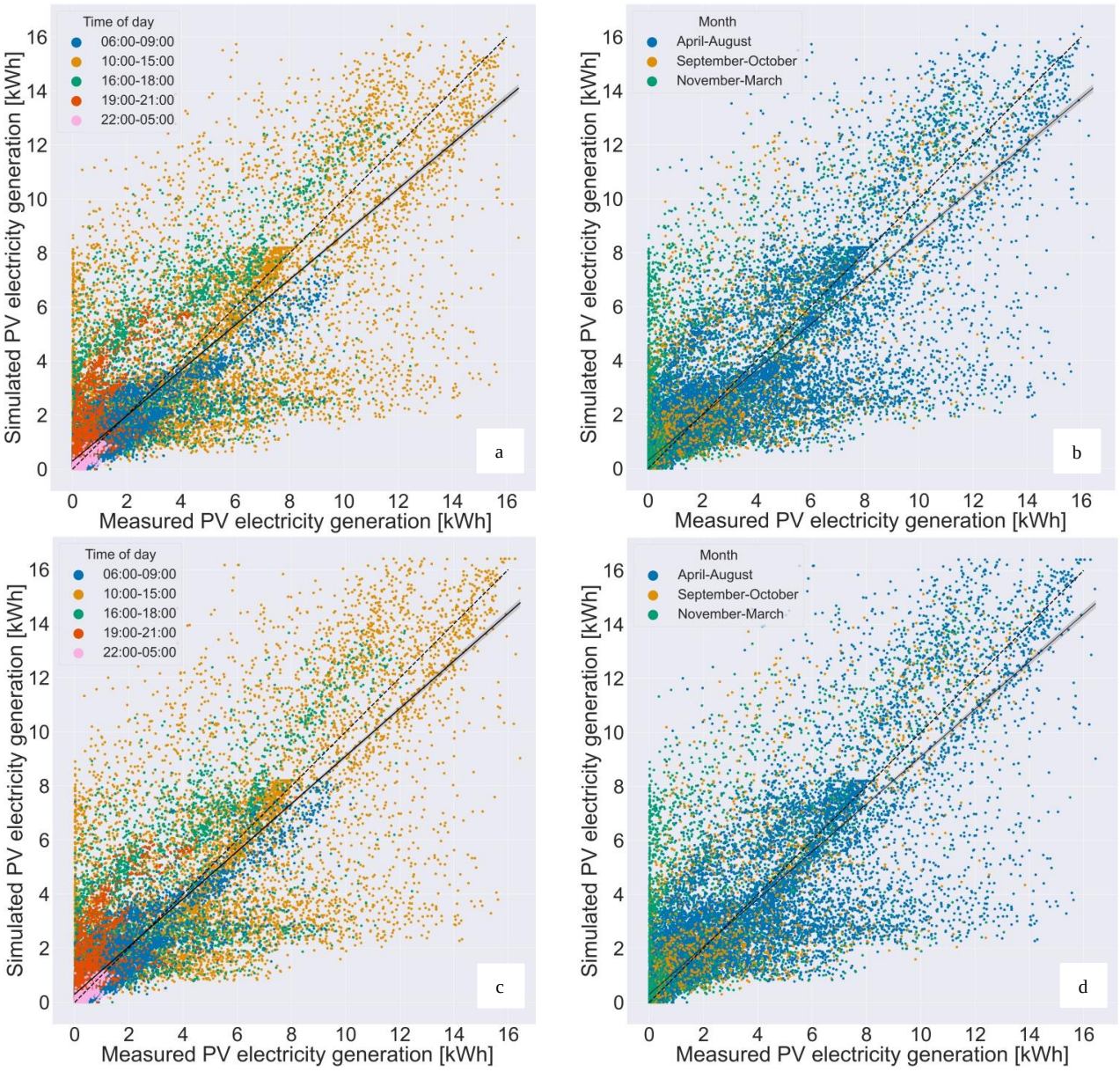

Figure 10. In (a) and (b), scatter plot for validation step 1, and in (c) and (d) validation step 2 for PVS4-F, both steps using known module- and inverter parameters disclosed by the system owners. The data points are divided using colour for time-intervals of the day in (a) and (c), and for monthly time intervals in (b) and (d). The fitted regression is represented by a continuous line and the dotted line indicates a perfect fit.

10.4229/EUPVSEC2023/5DV.2.10
020527-013

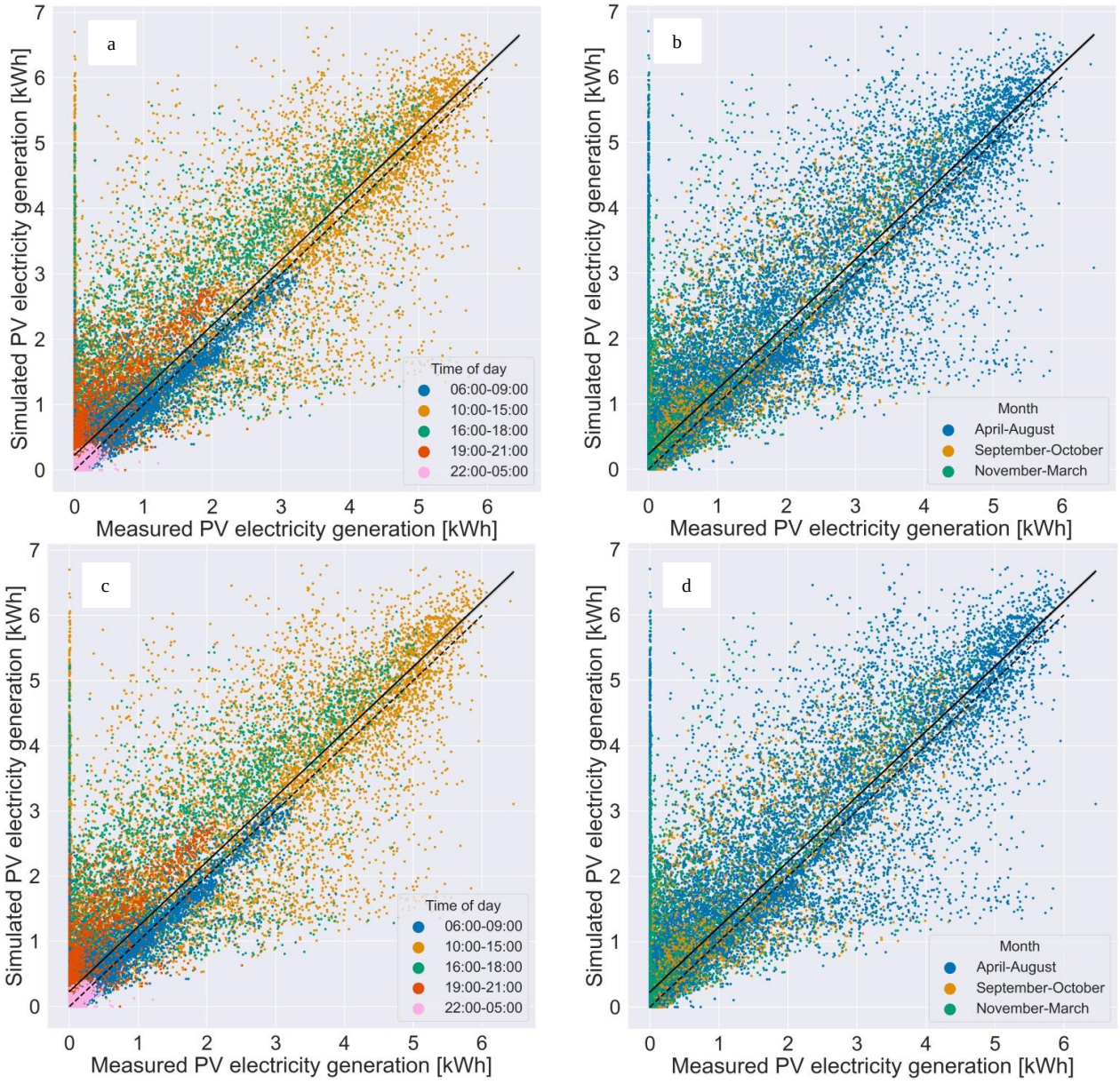

Figure 11. In (a) and (b), scatter plot for validation step 1, and in (c) and (d) validation step 2 for PVS5-K, both steps using known module- and inverter parameters disclosed by the system owners. The data points are divided using colour for time-intervals of the day in (a) and (c), and for monthly time intervals in (b) and (d). The fitted regression is represented by a continuous line and the dotted line indicates a perfect fit.

10.4229/EUPVSEC2023/5DV.2.10
020527-014

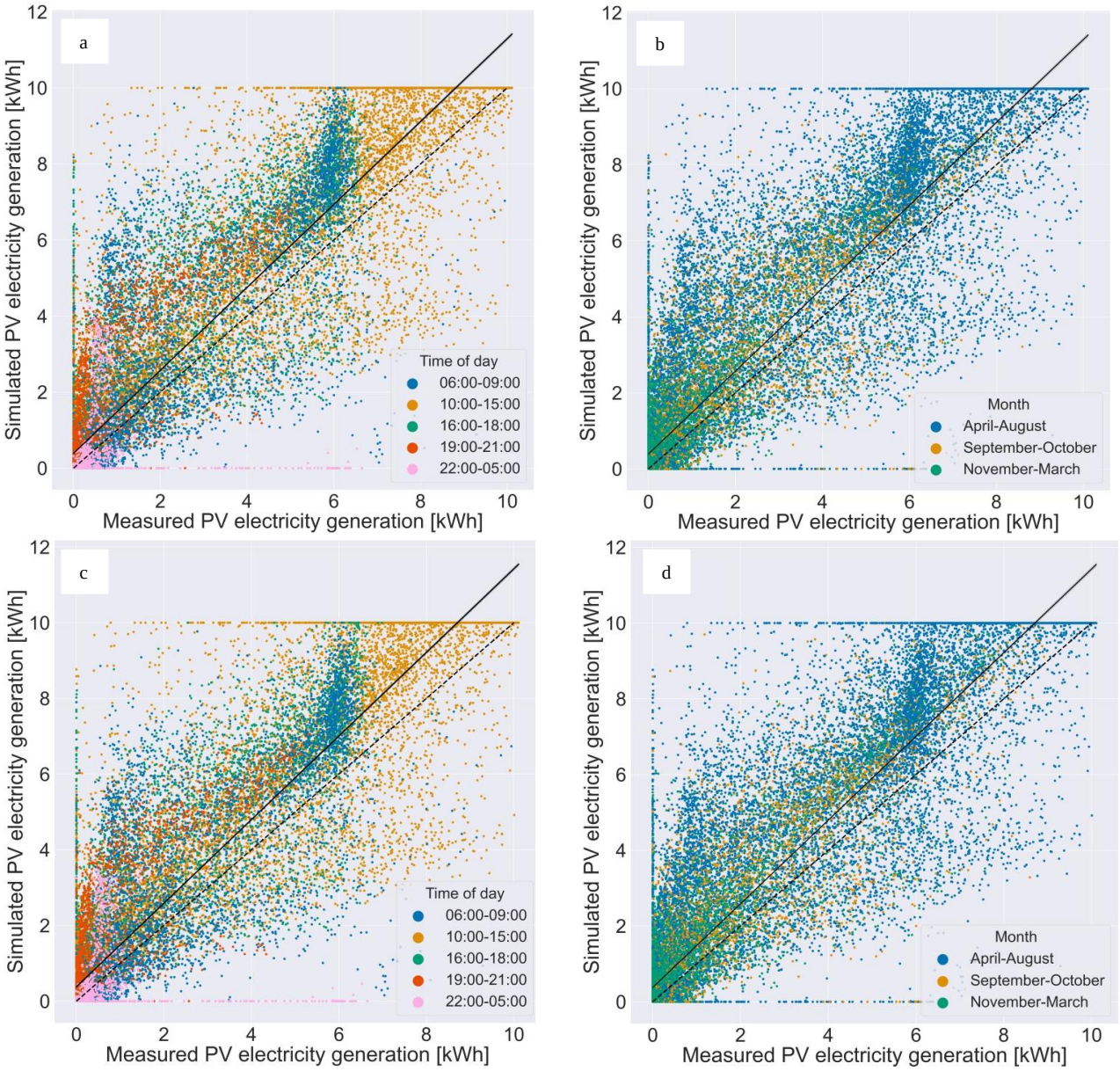

Figure 12. In (a) and (b), scatter plot for validation step 1, and in (c) and (d) validation step 2 for PVS6-K, both steps using known module- and inverter parameters disclosed by the system owners. The data points are divided using colour for time-intervals of the day in (a) and (c), and for monthly time intervals in (b) and (d). The fitted regression is represented by a continuous line and the dotted line indicates a perfect fit.

10.4229/EUPVSEC2023/5DV.2.10
020527-015