Estimation of Residential PV Power Generation Using Panel Azimuth Information

Gussan Mufti, Markos Asprou and Christos Panayiotou
KIOS Research and Innovation Center of Excellence and Department of Electrical and Computer Engineering
University of Cyprus
Nicosia, Cyprus
{mufti.gussan, asprou.markos, christop}@ucy.ac.cy

***Abstract*— The high integration of residential Photovoltaics (PVs) to the low voltage distribution network, poses several challenges to the operation of the power systems today. The main reason is that roof-top solar PV sites are usually unobservable by the distribution system operators (DSOs) and consequently their generation is not measured. Therefore, the accurate knowledge regarding the real time generation of the residential PVs will become an imperative need. In this paper, a methodology for estimating the power generated by all the residential PVs of a certain area using the available information from some observable residential PVs is proposed. Considering the high impact of the solar panel orientation on the PV generation, the proposed method clusters together the residential PVs with the same orientation. The results show that there is a maximum reduction in the mean absolute error by 62% in the summer season for 25 degree panel azimuth in comparison to the case that the generation of the unobservable PVs are estimated without using the orientation of the solar panels. The methodology was validated using real data from residential PVs.**

Keywords—Big data, clustering, PV power estimation, residential PVs, solar panel azimuth.

NOMENCLATURE

β	Slope of the solar panel
γ_i	Panel azimuth of the i^{th} panel
GM_i^{γ}	Geometric Mean (GM) of panels azimuth γ for residential prosumer i
$MAE_{NP(e,t,\gamma)}$	Mean Absolute Error (MAE) of estimated normalized power of the unobservable prosumer for γ degree panel azimuth at time t
$MAE_{NP(e,t)}$	MAE of the base case estimation of the unobservable prosumer without considering panel azimuth at time t
NP_i	Normalized Power (NP) of prosumer i
N_{γ}	Number of observable prosumers belonging to a panel azimuth cluster γ
$NP_{e,t,i}$	Estimated normalized power of prosumer i at time t
$NP_{e,t,GM_i^{\gamma}}$	Estimated normalized power of the unobservable prosumer i for a specific representative panel azimuth GM_i^{γ}
$NP_{true,t,\gamma\ i}$	True normalized power of prosumer i at time t belonging to a panel azimuth cluster γ
P_i	Instantaneous power produced by the solar panels of prosumer i
P_p	Peak capacity of the installed solar panels of the prosumer at Standard temperature Conditions (STC)
$w_{i,x}$	Weights that are given in Table I for four cases (A, B, C, D)

I. INTRODUCTION

A. Motivation and Background

The penetration of residential PVs in the distribution grid is getting higher today since thousands of electricity consumers aims to become energy neutral using renewable energy sources. This is in line with the attempt of the European Commission to promote the green energy, therefore the residential PV generation will be a considerable amount in the energy mix of a country in the near future. In reality, the generation produced by the residential PVs is hardly known by the operators and essentially is a source of uncertainty to the grid operation. In theory, one may estimate the total PV generation by measuring every single solar generation site. However, considering the enormous growth of small scale and roof top PV sites [1], building an infrastructure to collect and continuously monitor all these small-scale sites is not a practical nor a cost effective solution.

B. Relevant Literature

Considering the increase in the share of the residential PVs generation in the distribution grid several research works are conducted for the estimation of the total energy generated by these residential PVs. Some researchers tried to develop methodologies that aggregate the output power of the residential PVs in a specific region [2, 3]. In [4], the authors proposed k-means clustering and principal component analysis to estimate the total power generation of all the sites based on a selected subset of the sites. However, since the efficiency and the orientation of the solar panels were not considered, the estimations were accurate for the sunny days and not accurate on the cloudy days of the year. A load forecasting method using extreme weather conditions was proposed in [5] to estimate the residential PVs installed on the feeder level. However, an optimal temperature threshold was decided to select the number of days for the training model. In [6], the authors proposed a methodology to estimate the solar power generated by the residential PVs using only publicly available data set. The availability of the results to only specific months of the year is a limitation to

This work was supported in part by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 957739 (OneNet), and in part by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 739551 (KIOS CoE - TEAMING) and from the Republic of Cyprus through the Deputy Ministry of Research, Innovation and Digital Policy.

978-1-6654-0557-7/22/$31.00 ©2022 IEEE

the proposed methodology. In [7], the authors provided a publicly available data for the solar PV installations in United Kingdom. In [8, 9], an equivalent virtual PV prosumer was modeled using historical data of individual residential PV prosumers. In [10], an upscaling method was proposed in which the PV sites were clustered based upon the similar power generation pattern. A neural network was then used to estimate the power generation of all the clusters. However, the authors suggested that better estimations were possible if an individual neural network for each cluster was used. Some data driven techniques were also implemented in [9, 11, 12].According to the literature review, there are still challenges in the accurate estimation of the residential PV generation. First, data outliers are significant factors in affecting the estimation results in data driven methods; methodology to specify the outliers, can improve the estimation model. Second, in an area, only a limited number of PV prosumers can be accessible and the large variances in the individual PV prosumers introduce challenges in correctly estimating the PV generation at the aggregated level.

C. *Contributions and Organization*

These challenges are considered in this research work, proposing a methodology that considers the panel azimuth information for each residential prosumer (having PVs). Therefore, the contributions of this paper are the following:

- In this paper, two estimation schemes are proposed to estimate the PV generation of the prosumers that their PV generation is not measured, hereafter unobservable prosumers. Both schemes manage to improve the estimation of the PV generation of the unobservable prosumers. The estimation schemes consider that the orientation of the installed solar panels is known and the PV power generated by the panels for the specific prosumers is measured through smart inverters, hereafter observable prosumers.
- The panel azimuth information is used after the preprocessing step, which not only clusters all the prosumers having same panel azimuth, but also helps in identifying the outlier prosumer within the grouped prosumers. Thus, using panel azimuth of prosumers as static information serves as a double check for outliers.
- Real PV generation data for a whole year having different panel azimuths were used to test the developed schemes.

The remaining of the paper is organized as follows. Section II introduces the proposed methodology and the framework that can be followed to estimate the PV generation of unobservable residential prosumers. The validation results for the proposed method are shown in Section III, while the paper concludes in Section IV.

II. METHODOLOGY

In this section the proposed framework for estimating the power generation of the residential PVs using the available information from some observable prosumers is shown. The estimation framework first clusters the prosumers based upon the orientation of their solar panels as described in Section II.A and then the normalized power of the unobservable prosumers is estimated using the method described in Section II.B. Fig. 1. explains the methodology using a flow chart. The preprocessing of the data set was carried out for those prosumers whose data was available for a year. The prosumers which had non reporting data either due to the faulty sensors of the system or due to electrical faults were discarded. In the next step the geometric mean of panel azimuth and normalized power were calculated for each prosumer and the clustering of the prosumers was done for the two estimation schemes. The first scheme is based on panel azimuth clustering, and the second scheme is based on panel azimuth and distance clustering. The two schemes are then compared using performance indicators.

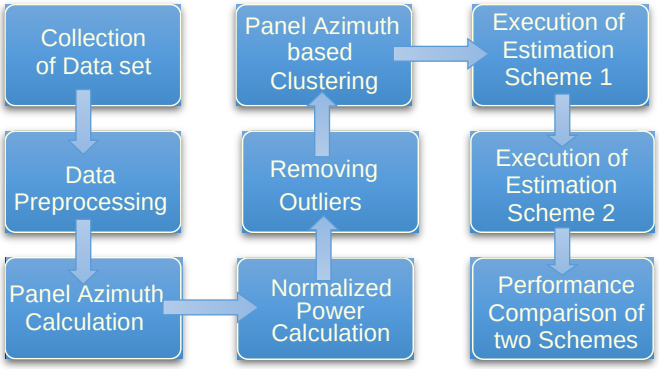

Fig. 1. Flowchart for the Methodology

A. *Clustering of Prosumers*

The power generated by the solar panels is mathematically defined by the two angles: β, the slope of the solar panel and γ, the panel azimuth [13]. The domestic prosumers have fixed (non-tracking) solar panels and the power generated by them mainly depends upon the azimuth of the panels. For this paper γ, which is the solar panel angle relative to the direction facing south, was taken -180° for east, 0° for south and +180° for west[14, 15]. The variable γ for each solar panel installed in residential prosumer is assumed to be known after the installation of the PVs since it is a static information. A representative panel azimuth for a residential prosumer i can be calculated by using the Geometric Mean (GM) of γ of each panel as shown below,

$$GM_i^{\gamma} = \sqrt[n]{\|\gamma_1\| \times \|\gamma_2\| \times \cdots \|\gamma_n\|} \qquad (1)$$

where n is the total number of solar panels installed by the domestic prosumer and γ_i is the panel azimuth of the i^{th} panel. The geometric mean was chosen in this paper instead of the arithmetic mean due to the large variations in the individual panel azimuth of each prosumer. In particular, the large variations tend to affect the arithmetic mean since it is dominated by numbers on the larger scale whereas the geometric mean handles the large variations, due to its multiplicative nature. Thus a better representation of the net panel azimuth of the prosumer was achieved through the geometric mean [16, 17]. The net panel azimuth of the prosumers was rounded off to the nearest 5 degrees in order to make coherent clusters. All the prosumers of an area with the same geometric mean panel azimuth were clustered to the same category.

B. *Estimation of the Normalized Power for the Unobservable Prosumers in a Cluster*

After the clustering of the prosumers according to their net geometric mean panel azimuth, the estimation of the

normalized power of the unobservable prosumers follows. To have a common reference for all the prosumers (since they have different installed capacity), the Normalized Power (*NP*) of each prosumer was calculated as,

$$NP_i = \frac{P_i}{P_p} \tag{2}$$

where P_i represents the instantaneous values of the power produced by the solar panels for the i^{th} prosumer and the P_p is the peak capacity of the installed solar panels of the prosumer. The P_p of each prosumer is taken from manufacturer's data sheet under the STC (Standard Temperature Conditions). In order to estimate the normalized power of the PVs for the unobservable prosumers, two estimation schemes are proposed in this paper. The first one (Scheme 1) considers only the panel azimuth of the prosumers, while the second one (Scheme 2) considers both distance (of the unobservable prosumer from the observable one) and the panel azimuth to estimate the normalized power. Both Schemes are compared to the base case estimation, which does not consider any panel azimuth information for the PV generation estimation of the unobservable prosumers.

1) Base Case Estimation

In this scheme the estimation of the PV generation for the unobservable prosumers is done without considering the role of the panel azimuth. Therefore, the normalized power of the unobservable prosumer was estimated by averaging the normalized power of the observable prosumers as,

$$NP_{e,t,i} = \frac{1}{N_x} \sum_{x \in \text{region}} NP_{i,x} \tag{3}$$

where N_x is number of customers in region *x.* Since all the observable prosumers that were in the region x were included for the estimation of the PV generation of the unobservable prosumer, the estimated normalized power ($NP_{e,t,i}$) is simply the average value of all the observable prosumers normalized powers (in the region) at time *t*.

2) Scheme 1-Estimation using the Panel azimuth information

In this scheme the estimation of the normalized power of the unobservable prosumer was done by considering the panel azimuth information only. Within each cluster, a prosumer's normalized power curve was estimated using the normalized power curve of the observable prosumers belonging to the same cluster as,

$$NP_{e,t,GM_i^\gamma} = \frac{1}{N_\gamma} \sum_{x \in GM_1^\gamma, \cdots GM_n^\gamma} (w_{i,x} \times NP_{i,x}) \tag{4}$$

where N_γ is number of observable prosumers belonging to a panel azimuth cluster γ and NP_{e,t,GM_i^γ} represents the estimated normalized power of the unobservable prosumer *i* for a specific representative panel azimuth GM_i^γ. $NP_{i,x}$ represents the normalized power of the observable prosumer belonging to panel azimuth clusters $GM_1^\gamma, GM_2^\gamma, \ldots, GM_n^\gamma$ and can be calculated by using (2). $w_{i,x}$ represents the weights that are given in Table I for four cases (A, B, C, D). Case A gives equal weights to the prosumers that are having 5 degrees and more than 5 degrees difference of net panel azimuth from the cluster of unobservable prosumer. For Case B the weights were redistributed such that there was a decrease in the weights of the prosumer lying in the same cluster as that of unobservable prosumer and an increase in the weights of the prosumer having 5-degree difference from the cluster of the unobservable prosumer. In Case C the role of the prosumers having more than 5-degrees difference of the panel azimuth from the unobservable prosumer was decreased in the estimation while it was nullified in Case D. It should be noted that all the observable prosumers were considered in the estimation procedure except for Case D in which the observable prosumers that have more than 5 degrees difference of net panel azimuth from the cluster of the unobservable prosumer were not considered.

TABLE I. WEIGHTAGE SCHEMES FOR SCHEME 1

Cases	***Same Panel azimuth as unobservable Prosumer***	***(+, -) 5 degrees from the Panel azimuth of unobservable Prosumer***	***More than (+, -) 5 degrees from the panel azimuth of the unobservable prosumer***
Case A	0.80	0.10	0.10
Case B	0.75	0.15	0.10
Case C	0.80	0.15	0.05
Case D	0.80	0.20	0

3) Scheme 2-Estimation based on the Panel azimuth and the distance

In this scheme the distance of the unobservable prosumer from the observable prosumer was also taken into the consideration along with the panel azimuth of the unobservable prosumer for the estimation purposes. For the estimation, (5) was used as well, however the weights ($w_{i,x}$) that were used to multiply the normalized power of the observable prosumers are also based on the distances between the observable and unobservable prosumers as shown in Table II. Case A gives equal weights to far away prosumers and nearby prosumers that have the same azimuth, whereas for Case B the weights are increased for nearby prosumers. In Case C the weights of the far away prosumers are reduced and their effect in the estimation is nullified in Case D.

TABLE II. WEIGHTAGE SCHEMES FOR SCHME 2

Cases	***Closest to unobservable Prosumer having the same panel azimuth***	***Second and third closest to unobservable prosumer having the same panel azimuth***	***Remaining prosumers in the panel azimuth cluster as the unobservable prosumer***
Case A	0.80	0.10	0.10
Case B	0.75	0.15	0.10
Case C	0.80	0.15	0.05
Case D	0.80	0.20	0

C. Performance Metric for the Comparison

In order to assess the performance of the estimation schemes, the percentage reduction of the Mean Absolute Error (*MAE*) is used. This metric is calculated by first calculating the MAE between the estimated normalized power ($NP_{e,t,\gamma\ i}$) of an unobservable prosumer and the true normalized power (since it is actually known).

$$MAE = \sum_{t=1}^{n} \left| NP_{e,t,\gamma\ i} - NP_{true,t,\gamma\ i} \right| = \sum_{t=1}^{n} |e_i| \tag{6}$$

The Percentage Reduction of the MAE (*PRMAE*) of the estimated normalized power of the unobservable prosumer ($NP_{e,t,25}$) in relation to the MAE of the base case estimation is then calculated as,

$$PRMAE = \frac{MAE_{NP(e,t,\gamma)} - MAE_{NP(e,t)}}{MAE_{NP(e,t)}} \times 100 \qquad (7)$$

III. Results

The proposed estimation schemes were developed using Python version 3.6 on a standard PC with an Intel Core i5-2430M CPU running at 2.40 GHz with 12 GB of RAM. Section III.A describes the data set that was used in the estimation schemes. Section III.B presents the simulation results for prosumers having 25 degree panel azimuth. Section III.C concludes the results section by presenting the simulation results for prosumers having 10, 20, 40 and 50 degrees panel azimuths to show the applicability of the estimation schemes for other panel azimuth angles.

A. Description of the Data Set

To verify the effectiveness of the estimation schemes a real data set is used that includes real power generation from residential solar prosumers of Nicosia (35°10′N 33°22′E), Cyprus. The real values were recorded from September 2020 to September 2021. The total number of solar residential prosumers in this data set is 110. The power production values from observable prosumers (P_i) were recorded every 15 minutes and they were accessed using the solar edge website [18], while the measurements were preprocessed [19, 20]. In order to generalize the estimation results, data for the four seasons of the year were used. It should be noted that although the normalized power of all the prosumers in each cluster is known, for evaluation purposes one prosumer in the cluster is considered as unobservable and the other as observables.

B. Simulation Results for Percentage Reduction in MAE for 25 Degree Panel azimuth

To show that the estimations using Scheme 1 and Scheme 2 are effective, for all the four seasons of the year, two single days of the year are chosen for the summer and autumn season, while an extended time frame of 15 days are selected to represent the spring and winter season. In particular, 1st of November 2020 is selected to represent autumn season, 1st of July 2021 to represent summer season, first 15 days of February 2021 to represent winter season and first 15 days of April 2021 as spring season. The four cases (A, B, C, D) are compared with the base case estimation that is calculated using (3).

1) Percentage Reduction of MAE for a Single Day of the Season

Fig. 2 illustrates the percentage reduction of the MAE for the autumn season. In general, both Scheme 1 and Scheme 2 present a considerable reduction of the MAE, indicating that the information regarding the panel azimuth helps the estimation of the normalized power of the unobservable power. Moreover, the simulation results show that Scheme 2 is effective in most of the cases for reducing the percentage of MAE for the autumn season. As shown in Fig. 2, the maximum reduction in MAE is shown by Scheme 2 for prosumer 2. For prosumer 3, Cases C and D are showing better results for Scheme 1 and Cases A and B of Scheme 2 results in larger reduction in MAE. In general, both schemes result in reduction of MAE, indicating the importance of considering panel azimuth in the estimation procedure.

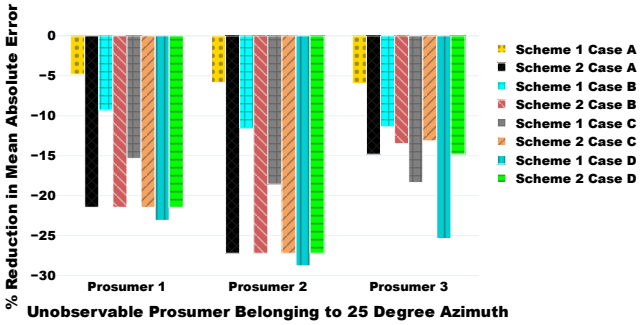

Fig. 2. Percentage Reduction for Simulation on 1st November 2020

For the summer season, the percentage reduction in the MAE for each prosumer is higher in comparison to autumn season. More specifically, as shown in Fig. 3 the percentage reduction in the MAE for prosumer 4 is 62%, which is the maximum one compared to the case when no panel azimuth information was used for the estimation of the normalized power. Specifically in the summer season, this maximum percentage reduction in the estimation of the normalized power can be helpful to reduce the uncertainty when PV curtailment might be needed for maintaining system operation in proper limits.

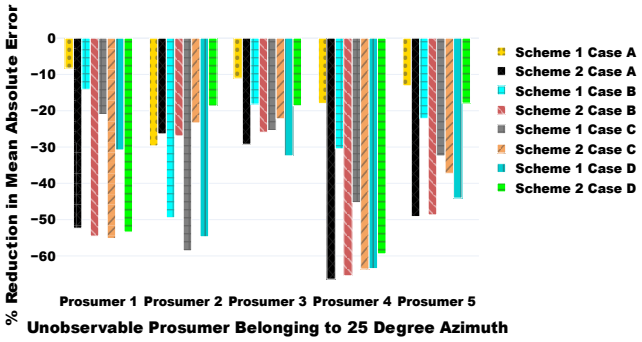

Fig. 3. Percentage Reduction for Simulation on 1st July 2021

Fig 4 shows the estimation curves for the autumn season which was the most volatile case among the 25-degree panel azimuth prosumers. There were three prosumers in the cluster for that day and prosumer 2 was taken as an unobservable prosumer. The estimated curves show that the base case using no panel azimuth information and Scheme 1 cases using panel azimuth information only estimates the curves to be lower than the real values of the prosumer 2. On the other hand, Scheme 2 using both the panel azimuth and distance information was effective in capturing the real values of prosumer 2 even on a cloudy day (as shown by the dip around 13:00). This also shows that even with a small number of prosumers the distance between the observable and unobservable prosumer can have significant reduction on the estimation error. Although there were variations in the PV generation on this cloudy day, all the cases of Scheme 2 follow the trend of the real values of the prosumer.

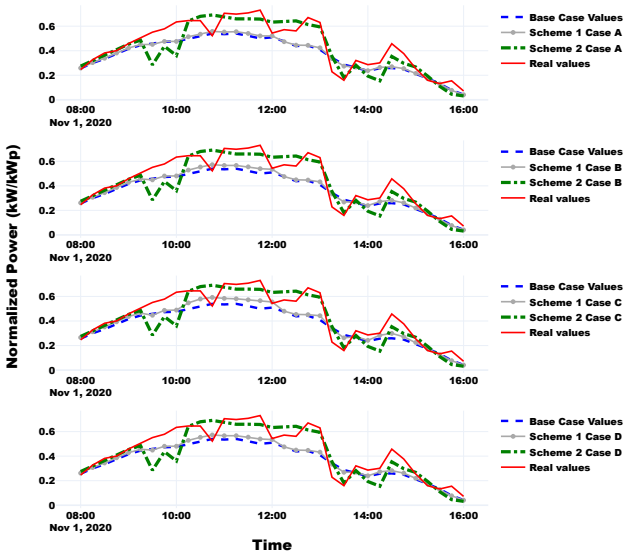

Fig. 4. Normalized Power Estimations for Prosumer 2 on 1st November 2020

2) Percentage Reduction of MAE for 15 Days of the Season

To show that the estimation scheme is generalized for an extended timeframe the first 15 days of February are selected for the winter season and first 15 days of April are selected for the spring season. Fig. 5 and Fig. 6 shows the results of the simulations for the winter and spring season. For the winter season, Scheme 2 results in greater percentage reduction in the MAE in comparison to Scheme 1. The maximum reduction is observed for prosumer 2 with 52% decrease for Case C weightages and the minimum reduction is recorded for prosumer 3 (around 18%) for Case A.

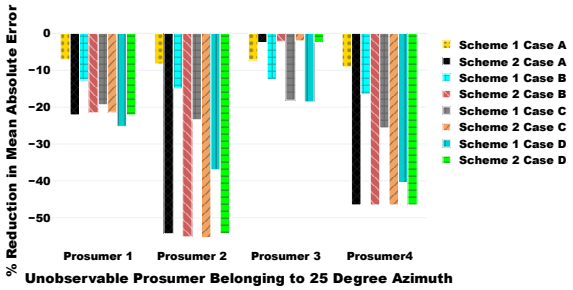

Fig. 5. Percentage Reduction for Simulation 1st February till 15th February

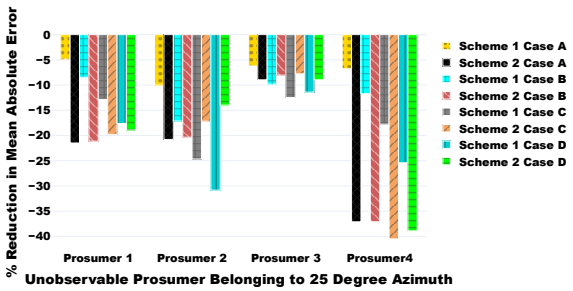

Fig. 6. Percentage Reduction for Simulation 1st April till 15th April

Fig. 6 illustrates the percentage reduction of the MAE for the spring season. In general, both Scheme 1 and Scheme 2 present a considerable reduction of the MAE, indicating that the information regarding the panel azimuth helps in the estimation of the normalized power of the unobservable power. Moreover, the simulation results show that Scheme 2 (considering both distance and azimuth) is more effective in most of the cases for reducing the percentage of MAE for the spring season. The maximum reduction is observed in prosumer 4 which is almost the same for the four cases A, B, C and D of Scheme 2. For prosumers 1 and 4, Scheme 2 demonstrates a larger percentage reduction of MAE for all the cases whereas for prosumer 2 and 3 Cases A and B of Scheme 2 are better and Case C and D of Scheme 1 show larger percentage reduction of MAE. Based on the estimation results from the four seasons, it can be concluded that the panel azimuth information plays a critical role in the reduction of MAE for the 25-degree azimuth prosumers.

C. Simulation Results for Percentage Reduction in MAE of 10,20,40 and 50 Degree Panel azimuths

To better illustrate the effect of azimuth information in the estimation of the PV generation of unobservable prosumers, four different panel azimuths other than 25-degree panel azimuth are selected. The panel azimuths used for the simulations are 10-, 20-, 40- and 50-degree panel azimuths. More specifically, in Figs. 7-10 each prosumer representing a different panel azimuth to show that both proposed Schemes provides better results than the base case estimation. The prosumer which has the maximum percentage reduction in the MAE for each season of the year, is selected as the representative prosumer for each panel azimuth. The real values of the PV generation of each prosumer are taken from the data set as explained in Section III.A. For these simulations, 24th of December 2020 is selected to represent winter season, 4th of September 2021 to represent autumn season, 4th April 2021 as spring season and 4th July 2021 to represent summer season.

Fig. 7 and Fig. 8 show the percentage reduction for the prosumers belonging to different panel azimuths of Nicosia for winter and autumn seasons respectively. For the winter season, Scheme 2 results in greater percentage reduction in the MAE in comparison to Scheme 1. The maximum reduction is observed for prosumer 3 belonging to 40-degree panel azimuth with 40% decrease for Case C weights and the minimum reduction is recorded for prosumer 4 (around 22%) belonging to 50-degree panel azimuth for Case B weights.

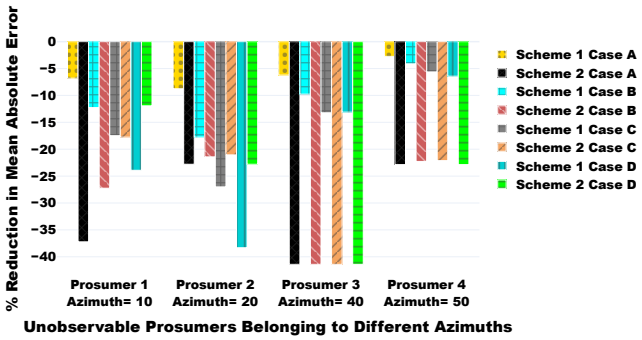

Fig. 7. Percentage Reduction for Simulation on 24th December 2020

For the autumn season, the trend is retained, and the maximum reduction is observed for prosumer 2 belonging to 20-degree panel azimuth with 60% decrease for Case C weights and the minimum reduction is recorded for prosumer 3 which shows 23% reduction. The spring and summer seasons are shown in Fig. 9 and Fig. 10 respectively. Based on the estimation results from the four

seasons, it can be concluded that the Scheme 2 results in greater percentage reduction in the MAE in comparison to Scheme 1 for all the prosumers having different panel azimuths.

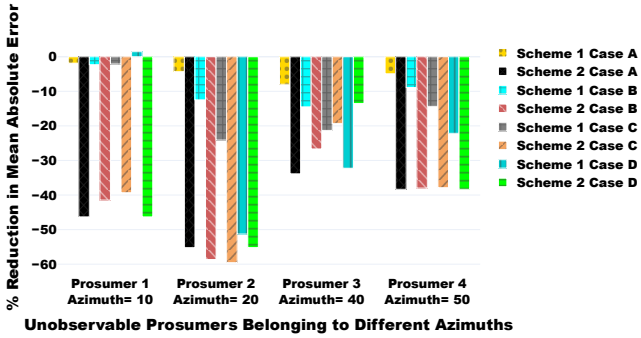

Fig. 8. Percentage Reduction for Simulation on 4th September 2021

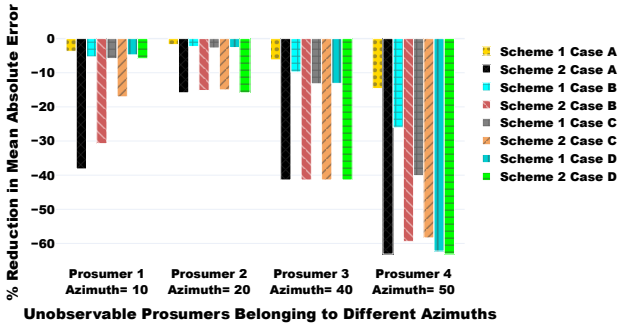

Fig. 9. Percentage Reduction for Simulation on 4th April 2021

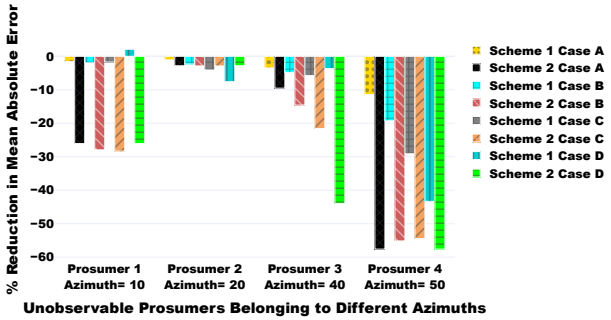

Fig. 10. Percentage Reduction for Simulation on 4th July 2021

IV. Conclusion

This paper deals with the estimation of the residential PV power generation that is not measured by any measurement device. Two estimation schemes are proposed in this work that take into account the panel azimuth for estimating the PV generation. The results of the case studies that are carried out show the importance of panel azimuth information when estimating the normalized power for the residential prosumers. Furthermore, it is shown that the large variances that occur in the large fleet of the residential PV prosumers can be captured by employing the estimation schemes. In particular, the estimation scheme based on panel azimuth information and distance (Scheme 2) significantly improves the estimation of the normalized power for the unobservable prosumers for all the seasons of the year. Future work includes the use of machine learning techniques that would be useful in the estimation of the normalized power of each prosumer by considering past data sets and learning parameters that may be able to tackle the abrupt changes that occur on the cloudy days of autumn and winter.

Acknowledgment

The authors would like to thank Eletoyia Photovoltaics for providing access to the dataset of solar prosumers.

References

[1] M. Salvia *et al.*, "Climate mitigation in the Mediterranean Europe: An assessment of regional and city-level plans," *Journal of Environmental Management*, vol. 295, p. 113146, 2021.

[2] T. B. P. Nguyen, Y.-K. Wu, and M.-H. Pham, "A Novel Data-Driven Method to Estimate Invisible Solar Power Generation: A Case Study in Taiwan," in *2022 IEEE/IAS 58th Industrial and Commercial Power Systems Technical Conference (I&CPS)*, 2022, pp. 1-10: IEEE.

[3] Y.-K. Wu, Y.-H. Lai, C.-L. Huang, N. T. B. Phuong, and W.-S. Tan, "Artificial Intelligence Applications in estimating invisible solar power generation," *Energies*, vol. 15, no. 4, p. 1312, 2022.

[4] H. Shaker, H. Zareipour, and D. Wood, "A data-driven approach for estimating the power generation of invisible solar sites," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2466-2476, 2015.

[5] M. Sun, C. Feng, and J. Zhang, "Factoring behind-the-meter solar into load forecasting: Case studies under extreme weather," in *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2020, pp. 1-5: IEEE.

[6] H. Shaker, H. Zareipour, and D. Wood, "Estimating power generation of invisible solar sites using publicly available data," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2456-2465, 2016.

[7] D. Stowell *et al.*, "A harmonised, high-coverage, open dataset of solar photovoltaic installations in the UK," *Scientific Data*, vol. 7, no. 1, pp. 1-15, 2020.

[8] Y. Wang, N. Zhang, Q. Chen, D. S. Kirschen, P. Li, and Q. Xia, "Data-driven probabilistic net load forecasting with high penetration of behind-the-meter PV," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3255-3264, 2017.

[9] J. M. Bright, S. Killinger, D. Lingfors, and N. A. Engerer, "Improved satellite-derived PV power nowcasting using real-time power data from reference PV systems," *Solar Energy*, vol. 168, pp. 118-139, 2018.

[10] M. Pierro *et al.*, "Data-driven upscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data," *Solar Energy*, vol. 158, pp. 1026-1038, 2017.

[11] K. Li, F. Wang, Z. Mi, M. Fotuhi-Firuzabad, N. Duić, and T. Wang, "Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation," *Applied Energy*, vol. 253, p. 113595, 2019.

[12] H. Shaker, D. Manfre, and H. Zareipour, "Forecasting the aggregated output of a large fleet of small behind-the-meter solar photovoltaic sites," *Renewable Energy*, vol. 147, pp. 1861-1869, 2020.

[13] J. Braun and J. Mitchell, "Solar geometry for fixed and tracking surfaces," *Solar Energy*, vol. 31, no. 5, pp. 439-444, 1983.

[14] A. Karafil, H. Ozbay, M. Kesler, and H. Parmaksiz, "Calculation of optimum fixed tilt angle of PV panels depending on solar angles and comparison of the results with experimental study conducted in summer in Bilecik, Turkey," in *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, 2015, pp. 971-976: IEEE.

[15] J. A. Duffie and W. A. Beckman, *Solar engineering of thermal processes*. John Wiley & Sons, 2013.

[16] R. Graziani and P. Veronese, "How to compute a mean? The Chisini approach and its applications," *The American Statistician*, vol. 63, no. 1, pp. 33-36, 2009.

[17] D. McAlister, "XIII. The law of the geometric mean," *Proceedings of the Royal Society of London*, vol. 29, no. 196-199, pp. 367-376, 1879.

[18] M. Holm, "Solaredge technologies, inc. equity research-capabilities far beyond inverter technology," 2021.

[19] D. Kuhlman, *A python book: Beginning python, advanced python, and python exercises*. Dave Kuhlman Lutz, 2009.

[20] C. R. Harris *et al.*, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357-362, 2020.