COMPARING APPROACHES FOR ESTIMATING RESIDENTIAL PV SYSTEM ORIENTATIONS

Johan Lindahl[a], Joakim Munkhammar[b] and Gustav Öhgren[a]
[a]Becquerel Sweden AB, Staffansvägen 14, SE-74142 Knivsta, Sweden, johan.lindahl@becquerelsweden.se
[b]Department of Civil and Industrial Engineering, Uppsala University, SE-751 21 Uppsala, Sweden,
joakim.munkhammar@angstrom.uu.se

ABSTRACT: Accurate knowledge of photovoltaic (PV) system orientations—tilt and azimuth—is crucial for forecasting generation, assessing self-consumption, and managing distributed solar power. Yet, orientation data are often missing or unreliable in PV databases. This study compares three approaches for determining the orientation of residential PV systems in Sweden: (i) self-reported values from the national capital subsidy program, (ii) statistical orientation distributions based on PV penetration, and (iii) a LiDAR-based method using aerial imagery and linear regression. Validation was performed against a manually derived ground truth dataset of 104 PV polygons across three municipalities. Self-reported orientations proved highly inaccurate, primarily because the application form allowed only a single entry for a system's azimuth and tilt and lacked clear guidance on azimuth conventions, compounded by imprecise or careless estimations from applicants. In contrast, the LiDAR-based method closely matched the ground truth, with R^2 values of 0.96 for azimuth and 0.82 for tilt and captured the diversity of real-world orientations more accurately than the statistical model. These results demonstrate that LiDAR-based remote sensing provides a reliable, scalable, and high-accuracy approach for determining PV system orientations, supporting improved generation forecasts, grid operation planning, and the integration of distributed solar power at high penetration levels.
Keywords: Photovoltaics, Orientation methods, Azimuth, Tilt, Remote Sensing

1 INTRODUCTION

Solar power is growing rapidly worldwide. In 2024, photovoltaic (PV) systems made up 75% of all newly installed power generation capacity globally, adding an estimated 600 GW [1]. Within the European Union, solar now accounts for 14% of total electricity generation, of which about half is distributed PV systems [1].

Integrating large shares of PV power into the power grid is challenging due to weather variability [2], which affects system stability and requires effective management [3], [4]. As distributed PV adoption grows, accurately forecasting both generation and self-consumption [5] at local and regional levels becomes crucial for optimizing grid operation and energy management [6], [7].

The orientation, i.e., tilt and azimuth, of a PV system highly influences the PV power output profile. Although installed capacity is often reported in PV databases, tilt and azimuth are most often not reported [8]. For instance, a global study by Killinger et al. [9] found that among 14.8% of the world's installed PV capacity analyzed (as of 2017), tilt and azimuth data were available for only 1.7%.

To address this gap, several approaches have been proposed: (1) collect the azimuth and tilt through manual self-reporting — usually connect to different subsidy schemes, (2) assuming optimal tilt and azimuth for maximum power generation at each location [10], [11], (3) estimating a single representative orientation by minimizing discrepancies between simulated and observed regional PV power output [7], (4) using statistical models based on existing PV system orientations [12] or rooftop solar potential [13], (5) applying machine learning or parameterization models to infer orientation from reported PV generation [14], (6) employing remote sensing techniques, such as aerial image analysis combined with 3D building data [15] or calibrating tilt based on nearby systems in non-profit PV databases [16] and (7) using LiDAR data and linear regression to estimate the orientation [17], [18], [19], [20].

This study presents a novel comparison of three of these approaches that has been applied in Sweden for assessing the orientation of PV systems. The first approach evaluates orientation data reported by PV system owners or installers to the Swedish direct capital subsidy program registry [21], [22]. The second approach applies to the model by Ramadhani et al. [13], which estimates statistical orientation distributions based on PV penetration levels. The third approach leverages the method developed by Lingfors et al. [17], which uses Light Detection and Ranging (LiDAR) data to derive tilt and azimuth angles of solar energy systems.

The aim of this study is to evaluate and compare the three distinct approaches for assessing the orientation of residential PV systems, focusing on their accuracy and practical applicability.

2 DATA

2.1 Data from the Swedish Capital Subsidy Program

In mid-2009, a subsidy program was launched in Sweden in which actors could apply for direct capital support for PV installations [21]. The scheme remained in place until 2021, though it was revised several times — for example, lowering support levels in response to declining technology prices and rising market demand [22].

Applications for the capital subsidy had to include project location, applicant details (e.g., address and contact information), planned start and completion dates, and a project description. Moreover, the description needed to specify the type of PV system, whether it would be grid-connected, the estimated total rated power of the modules (kW), and the planned installation site. In the application form, applicants could also enter the azimuth and tilt of the system in two dedicated fields, following the exact instruction in Swedish directly translated to:

"Orientation – azimuth and tilt must be specified as degree values, with azimuth ranging from 0–360 and tilt from 0–90."

This information was recorded and stored in the subsidy program database by the Swedish Energy Agency. Although many applications omitted orientation data, a considerable share did include it. In this study, those entries represent approach (1) described in the introduction

— i.e., the collection of azimuth and tilt through manual self-reporting.

In this study, azimuth is defined such that −90° corresponds to East, 0° to South, and 90° to West and consequently the azimuth data from the direct capital subsidy database was recalculated to match this definition.

3 METHODOLOGY

3.1 Manually Created Ground Truth Dataset

To create a ground truth dataset, residential PV systems first need to be located. For this purpose, the *Alfrödull* [23] remote sensing pipeline was applied to the latest aerial orthophotos from the Swedish Land Survey to identify residential PV installations within the Swedish municipalities of Fagersta (May 2024), Falun (June 2024) and Karlshamn (May 2024).

In the first step of the *Alfrödull* pipeline, a Convolutional Neural Network (CNN) deep-learning model, developed in [24], is used to detect PV systems in aerial imagery across the target area. When applied in Sweden, the CNN algorithm has been shown to accurately identify over 95% of all existing solar energy systems at the municipal level [24].

In the second step, polygons of the identified PV systems are generated, either manually or preferably using the U-net segmentation method described in [25]. These polygons provide the coordinates and orthographic (top-down) area of each PV module cluster. This enables the extraction of LiDAR data for the rooftops hosting the PV systems, which is then used to calculate tilt and azimuth, as further described in section 3.2 and in detail in [17]. A residential PV system containing two PV polygons, generated using the above-mentioned remote sensing pipeline, is illustrated in Figure 1.

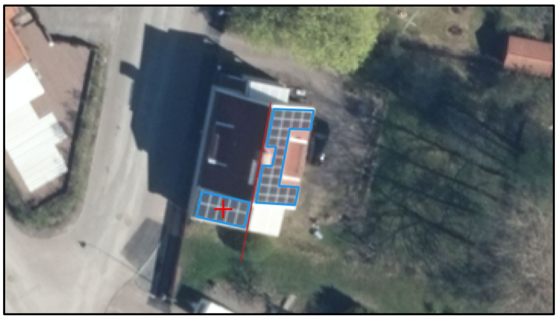

Figure 1. Orthophoto showing a single-family residential house with two distinct clusters of PV modules, each outlined by blue polygons. © Lantmäteriet, 2024. The dark red line indicates the azimuth measurement performed in QGIS for the polygon marked with a bright red cross.

Two different methodologies were applied to manually determine the orientation of the generated residential PV polygons in the ground truth dataset. Azimuth was obtained through manual measurements in QGIS using the "Azimuth Measurement" plugin, as illustrated in Figure 1. Tilt was determined by extracting Google Street View images of the building containing the PV polygon, selecting views taken at an estimated ~85–95° angle relative to the measured azimuth of the polygon, i.e. close to perpendicular. An online angle measurement tool was then applied to these images, as illustrated in Figure 2.

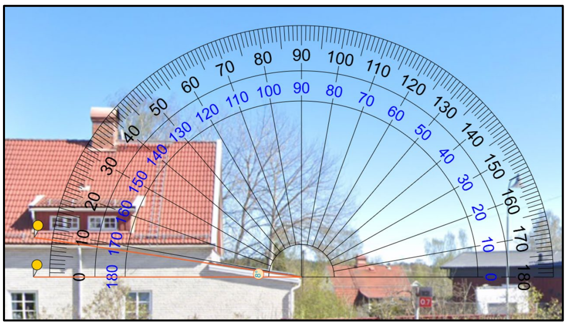

Figure 2. Google Street View image taken approximately perpendicular to the PV polygon marked with a red cross in Figure 1. Image © Google, 2024. The overlay illustrates the online angle measurement tool and the manually derived tilt of the PV system.

The manual derivation of both azimuth and tilt angles is subject to uncertainties. Firstly, the Swedish orthophotos are not orthorectified orthophotos — i.e., they are projected relative to terrain without accounting for building heights — which means that some buildings appear tilted or skewed. Moreover, it is not always possible to obtain Google Street View images that are perfectly perpendicular to the roof being measured. Both factors contribute to slightly distorted measurements. In addition, there is the risk of human error when performing measurements based solely on screenshotted street view images. Overall, the authors estimate that the accuracy of manual tilt and azimuth determinations for the manual ground truth is within ±3° for both angles.

While manually derived azimuths could be collected for all 4,817 residential PV polygons across the three municipalities, only 1,012 of these had a corresponding subsidy application, and just 249 of those applications included both tilt and azimuth values provided by the applicant. The availability of Google Street View images taken close to perpendicular to the polygons' azimuth further reduced our dataset to 104 polygons.

When analyzing the database of the Swedish subsidy scheme, a major drawback was identified: applicants could only enter a single azimuth and tilt value in the submission form. For PV systems consisting of multiple clusters of modules with different orientations, this restriction inevitably introduces errors in the reported orientation. For example, in Figure 1, the marked southernmost smaller PV polygon has a tilt of approximately 8° and an azimuth of 13°, whereas the adjacent larger polygon has a tilt of 45° and an azimuth of –79°. The applicant for this system reported a tilt of 45° and an azimuth of –90°, which reasonably represents the larger polygon but completely misrepresents the smaller one. To assess how widespread this application form and database design limitation was, the authors classified systems with more than one polygon into *primary* and *secondary polygons*, assigning the *primary* status to the cluster whose tilt and azimuth values were closest to the established ground truth. In the example of Figure 1, the larger east-facing polygon was designated as the *primary polygon*, while the smaller southernmost polygon was considered secondary. If two polygons are installed on the same roof facet, both are assigned as *primary* or *secondary*.

3.2 The LiDAR Data and Linear Regression Method

As demonstrated by Lingfors et al. [17], the tilt and azimuth of a solar energy system can be derived from

LiDAR point cloud data. Since Swedish orthophotos are not orthorectified orthophotos PV polygons may be spatially misaligned. To overcome this, Lingfors et al. [17] developed and described in detail an orthorectification procedure to align PV polygons with the LiDAR data.

The results in [17] showed that for most PV polygons detected by the *Alfrödull* remote sensing pipeline in municipal scans, the available Swedish LiDAR data contained a sufficient number of points to enable reliable orientation estimation. However, in cases where LiDAR data could not be used with confidence, a set of special cases and rules was defined to assign a plausible and realistic orientation. These include:

1. **Insufficient LiDAR points** → no regression performed; tilt fixed at 26°[1], azimuth set to the southernmost long edge the polygon.
2. **Ground-mounted systems (>2 m from buildings)** → tilt set to 30°, azimuth set to the southernmost long edge of the polygon, panels assumed to be 0.75 m above ground.
3. **Vertical systems** (not identifiable from aerial images) → added manually with tilt 90°, azimuth set to the southernmost long edge of the polygon.
4. **Flat-roof systems (tilt <5°)** → tilt adjusted to 10°[2], azimuth set to the southernmost long edge of the polygon.
5. **Flat-roof systems oriented east–west** → treated as bi-directional by splitting the PV polygon. Tilt adjusted to 10°, two different azimuths set based on the southernmost long edge of the polygon with 180° difference.
6. **Unrealistically high tilts from LiDAR** (caused by outdated or misleading data) → treated as in case 1.

These special cases are also relevant for some of the PV systems included in this study.

In [17], the method was evaluated on 3,500 Swedish solar energy systems using a manually derived ground truth azimuth dataset, following the same procedure described in Section 3.1. For 91–95% of the systems, the model accurately estimated the azimuth within the stated uncertainty margin of ±3° [17].

3.3 The Statistical Model Based on PV Penetration

The Ramadhani et al model [13], can be used to estimate the distribution of residential PV tilt and azimuth angles based on penetration level — defined in this study as the number of residential PV systems divided by the total number of residential buildings. Flat roofs (0° in tilt) are first excluded, and then the tilt and azimuth distributions of non-zero tilt roofs are estimated according to the stochastic variables:

$$\mathbb{X}_{Tilt} \sim \mathcal{N}[\mu(x), \sigma(x)],$$

for tilt and:

$$\mathbb{X}_{Azimuth} \sim \mathcal{U}[-180(x), 180(x)],$$

for azimuth. Here $x \in [0, 1]$ is the penetration level defined as the number of roofs with PV installations divided by total amount of roofs. In [13], the parameters were determined to be $\mu(x) = -0.078x + 26.295$ and $\sigma(x) = 0.028x + 6.429$. The complete Ramadhani et al model is then based on using these distributions with an addition of a degenerate distribution at zero for the tilt distribution for the flat roofs, see [13], [26] for more detailed information.

It should be mentioned that the model settings were determined in [13] based on a PV model with the hypothesis of best-roofs-first (in terms of yearly solar energy yield) for PV installations.

In this study the Ramadhani statistical model is tested on the three municipalities of Knivsta, Uppvidinge and Falun and compared to a model which fits normal distributions to both tilt and azimuth using the Matlab function "fitdist". According to a remote sensing-based study on PV penetration levels [27] Knivsta reached a residential penetration level of $x = 0.0375$ in 2023, while Uppvidinge reported $x = 0.0231$ in 2022 and Falun $x = 0.0232$ in 2020.

To assess the goodness-of-fit between the produced distributions of the model and the distribution of the observations, the Kolmogorov-Smirnov (K-S) test statistic was used. Formally, the K-S test statistic $K \in [0, 1]$ is based on the maximum deviation between two distributions [28]:

$$K = \max_{x} |F_1(x) - F_1(x)|,$$

where F_1 and F_2 are cumulative density functions (CDFs) of the probability distributions to be tested. The K-S test statistic is a negatively oriented score such that lower values mean higher goodness-of-fit.

3.4 Model evaluation

The accuracy of two methods, the LiDAR-based method and database-derived parameters, in estimating tilt and azimuth of PV systems was evaluated using the metrics shown in Table 1.

Table 1. Evaluation metrics and their definitions.

Evaluation metric	Definition
Coefficient of determination	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - x_i)^2}{\sum_{i=1}^{n}(y_i - \bar{x})^2}$
Root Mean Square Error	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2}$
Mean Absolute Error	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - x_i \rvert$

In the evaluation metric formulas presented in Table 1, y_i represents the estimated value of a datapoint, x_i the corresponded measured value, $\bar{x}$ the overall mean of the measured values, n the number of observations. The variables thus refer to either tilt or azimuth.

[1]Based on a sample of 100 PV systems with correctly identified azimuths, the mean tilt was 26°, which appears to be the most common roof inclination for single-family houses in Sweden [17], confirmed in Figure 4 in this study.

[2]The prevailing practice in Sweden for PV installations on large commercial and industrial (C&I) flat roofs is to mount the modules in rows with a tilt of 10° [29].

4. RESULTS AND DISCUSSION

4.1 Comparing Manual Self-Reporting and the LiDAR based Method with the Manually Created Ground Truth Dataset

As described in Section 3.1, a complete-case total data set of 104 PV system polygons in the three municipalities could be created for which applicants for the Swedish capital subsidy program had filled in both azimuth and tilt estimations, and for which it also was possible to manually derive both azimuth and tilt values from the orthophotos and Google Street view images, respectively. In addition, the LiDAR data and linear regression method of [17] could also generate both azimuth and tilt values based on either the sufficient LiDAR data or the six special cases specified in section 3.2.

When comparing the results of the three methods it became evident that many applicants for the capital subsidy used inconsistent definitions. The application form instructs that azimuth values should range from 0° to 360°, and when azimuth data from the subsidy database was analyzed, the authors concluded that different applicants appear to have used either of the two conventions:

- Clockwise from North (0°–360°): East = 90°, South = 180°, West = 270°, North = 0°.
- Clockwise from South (0°–360°): East = 270°, South = 0°, West = 90°, North = 180°.

To enable a more fair comparison of the methods, the authors therefore adopted for each PV system a generous interpretation of which of the two azimuth convention the applicant likely have used — to bring them as close as possible to the ground truth — and then recalculated that value to the definition used in this study, i.e., that −90° corresponds to the East, 0° to the South, and 90° to the West.

Figure 3 shows the distribution of azimuth angles for each PV polygon derived by the three methods (after the generous interpretation modification of the subsidy scheme data) while Figure 4 presents the corresponding distribution of tilt angles.

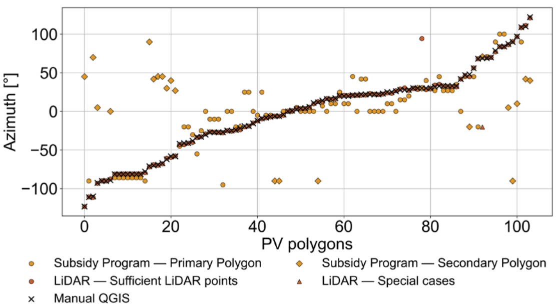

Figure 3. Distribution of azimuth angles derived by three different methods for each PV polygon. Yellow markers show values reported in the Swedish direct capital subsidy program (circles = *primary polygons*, squares = *secondary polygons*). Orange markers represent results from the LiDAR-based method (circles = sufficient LiDAR points, triangles = special-case estimations). Black crosses indicate manually measured azimuth angles from QGIS.

As shown in Figure 3, there is a strong correlation between the manually derived ground truth and the LiDAR-based methodology for azimuth angles: only 2 out of 104 PV polygons (1.9%) differed by more than ±3°. In contrast, applications submitted under the capital subsidy program showed much poorer accuracy, with 86 polygons (82.7%) deviating by more than ±3° from the ground truth. Not only were errors more frequent in the self-reported data, but the magnitude of the deviations was also often substantial, as seen in Figure 3.

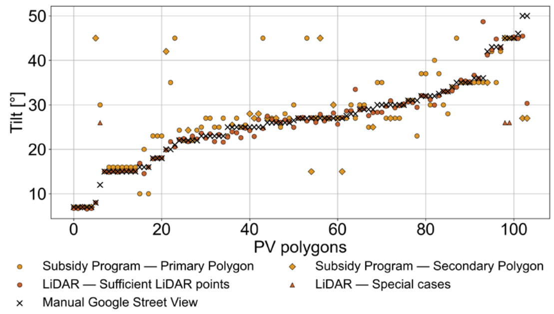

Figure 4. Distribution of tilt angles derived by three different methods for each PV polygon. Yellow markers show values reported in the Swedish direct capital subsidy program (circles = *primary polygons*, squares = *secondary polygons*). Orange markers represent results from the LiDAR-based method (circles = sufficient LiDAR points, triangles = special-case estimations). Black crosses indicate manually measured tilts from Google Street View.

Figure 4 shows a similar pattern for tilt angles. Here too, the LiDAR-based method aligns closely with the ground truth: only 7 out of 104 PV polygons (6.7%) differed by more than ±3°. By comparison, 35 polygons (33.7%) from the subsidy database deviated by more than ±3°. Again, the self-reported data show both more errors and larger discrepancies.

Furthermore, the accuracy of the subsidy database and the LiDAR-based method in estimating tilt and azimuth as compared to the ground truth dataset is summarized in and Table 2 (azimuth) Table 3 (tilt), showing R^2, MAE, and RMSE.

Table 2. Accuracy of the LiDAR-based method and the self-reporting method of the Swedish subsidy database as compared to the manually derived ground truth dataset in estimating azimuth, illustrated by R^2, MAE, and RMSE.

Case / Method	R^2	MAE [°]	RMSE [°]
LiDAR	0.96	2.17	10.82
Database	0.06	31.38	52.55

Table 3. Accuracy of the LiDAR-based method and the self-reporting method of the Swedish subsidy database as compared to the manually derived ground truth dataset in estimating tilt, illustrated by R^2, MAE, and RMSE.

Case / Method	R^2	MAE [°]	RMSE [°]
LiDAR	0.82	1.49	3.91
Database	0.30	4.21	7.76

When interpreting the results in Tables 2 and 3, it should be noted that the ground truth dataset is not absolute, as the manual measurements of azimuth and tilt angles are subject to uncertainties, estimated by the authors to be within ±3°. Nevertheless, the trends observed in Figures 3 and 4 are corroborated by the statistical measures presented in Tables 2 and 3. The LiDAR-based method demonstrates superior accuracy compared to the database-derived values. For instance, considering R^2, the LiDAR-based method achieves 0.96 for azimuth and 0.82 for tilt, whereas the database-derived parameters yield 0.3 and 0.06, respectively.

As mentioned earlier, a major source of discrepancy in both tilt and azimuth between the direct capital subsidy data and the ground truth or LiDAR-based method is that applicants could only enter a single value for each parameter in the submission form. Figures 3 and 4 show, however, that this limitation does not fully explain the deviations: 21 *primary* polygons differed by more than ±10° in azimuth, and 6 by more than ±10° in tilt. When accuracy was assessed using only the *primary polygon* values (as defined in Section 3.1), the results were R^2, MAE, and RMSE values of 0.57, 3.27º, and 5.73º for tilt, and 0.81, 13.92º and 20.14º for azimuth. Even under this favorable assumption, the database accuracy remains substantially lower than that of the LiDAR-based method, most likely because applicants were imprecise or careless when estimating roof orientations.

For the LiDAR-based method, 3 of the 104 polygons could not be reliably assessed due to insufficient LiDAR points (special case 1). In these cases, a fixed tilt of 26° was assigned and the azimuth was set perpendicular to the southernmost long edge of the polygon. This simplification introduced substantial errors, including one case where the southernmost edge did not align with the module cluster's lower boundary, leading to an azimuth offset of –90°. The second major azimuth error came from a non-special-case polygon that deviated by +64°, properly due to noise in the LiDAR data. All three special-case polygons also appeared among the seven polygons with tilt deviations greater than ±3°, with actual errors of –19°, –19°, and +14°. The remaining four non-special-case polygons showed tilt deviations of –19.7°, –4.6°, +5.5°, and +12.7°.

Overall, the LiDAR-based method demonstrated a stronger fit for azimuth than for tilt, as indicated by the higher R^2 value. At the same time, higher MAE and RMSE for azimuth suggest that occasional large errors can occur, as seen in the two polygons with –90° and +64° deviations.

4.2 Comparing the Statistical Model and the LiDAR based Method

The formal K–S test statistics for the agreement between modeled and observed distributions are presented in Table 4 for both the Ramadhani et al. model and the fitted normal distributions, separately for tilt and azimuth. For tilt, the K–S statistics are similar for the two models, indicating comparable performance. In contrast, for azimuth there is a pronounced difference: the Ramadhani et al. model departs much more strongly from the observed distribution than the fitted normal model. Keeping in mind that the K-S test statistic score is a negatively oriented score, the normal distribution model outperforms the Ramadhani model for the azimuth angle distribution estimation.

Table 4. Kolmogorov-Smirnov test statistics for Ramadhani model / normal distribution.

Variable	Knivsta	Falun	Uppvidinge
Tilt	0.19/0.16	0.14/0.22	0.19/0.15
Azimuth	0.49/0.07	0.50/0.07	0.46/0.06

This comparison is not entirely a fair comparison from a modelling and general perspective, since the normal distribution models were in fact trained on the data sets given, and not another data set and generalized as instructed for the Ramadhani model case. However, the mean value for the fitted normal distributions ranged from 3.87 to 4.48 and the standard deviation from 52.17 to 61.83, which shows a close universal fit to these distributions for these data sets, albeit both for Swedish towns and for similar low penetration levels. A proposal could be that the distribution of the azimuth is in fact universal regardless of penetration level and, based on the mean value of this data, that μ=4.263 and σ=52.335 are suggested parameters for this distribution. Investigating the universality of this proposition is left for future work.

It should also be mentioned that while the Ramadhani model as a statistical method is testable on the distributions of orientations of an aggregate of buildings in this study, it lacks the detailed system-to-system comparison features that are inherent to the self-reporting and LiDAR-based methods, such as the results in Section 4.1.

Moreover, since the majority of applications in the direct capital subsidy scheme lack reported orientation values, a direct comparison of general distributions is also not feasible.

However, as the LiDAR-based method generates both azimuth and tilt angles for all polygons of all PV systems within a scanned area, it enables a comparison of the overall distribution of orientations. Figure 5 presents such a comparison of the azimuth distributions obtained from the Ramadhani statistical model [13] and the LiDAR-based method [17] for all actual residential PV systems in the three Swedish municipalities of Knivsta, Uppvidinge and Falun in 2022.

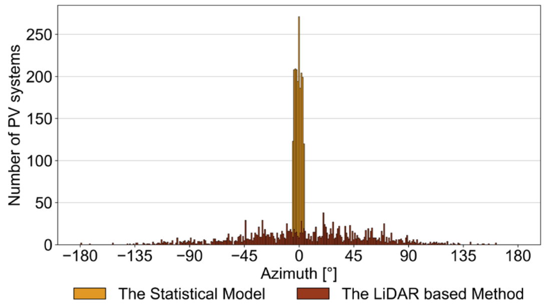

Figure 5. Histograms of the azimuth distribution from random samples using the Ramadhani et al. [13] model, which estimates orientation based on PV penetration levels, and the Lingfors et al. [17] method, which derives tilt and azimuth angles using LiDAR data, for 1922 actual residential PV systems in three Swedish municipalities in 2022.

As shown in Figure 5, the Ramadhani statistical model — assuminh that the "best" roofs are utilized first — produces a much narrower distribution of azimuth angles than the LiDAR-based method, which explain the high K-S test statistic values. While this result was expected from the probability distribution fit estimates, the difference is substantial in system azimuth orientation diversity. The LiDAR-based method is highly accurate in deriving azimuth angles for PV polygons, both in our dataset (where it produced only 2 erroneous azimuths out of 104 polygons) and in previous work [17], which demonstrated accurate azimuth estimation within a ±3° uncertainty margin for 91–95% of systems in these three municipalities. We therefore conclude that the assumption underlying the statistical model — that PV systems are installed on the best-oriented residential roofs — does not reflect reality and leads to erroneous estimates of the actual orientation distribution of residential PV systems.

5 CONCLUSIONS

This study compared three approaches for assessing the orientation of residential PV systems in Sweden: (i) self-reported values from the national capital subsidy program [21], (ii) statistical orientation distributions based on PV penetration levels [13], and (iii) a LiDAR data and linear regression method [17].

The results demonstrate that self-reported values are highly unreliable. During the analyses, two major design flaws were identified in the application form of the Swedish direct capital subsidy scheme. Firstly, the form provided only a single field for azimuth and tilt angles. In practice, many PV systems consist of multiple clusters of modules installed with different orientations. Restricting applicants to one entry per system has therefore proven to be a significant flaw, leading to substantial errors in capturing the true orientation of PV systems. Second, the instructions for reporting azimuth angles were insufficient. It became evident that applicants had used at least two different conventions — where an azimuth of 0° was taken to represent either north or south.

Since it is impossible to know which azimuth convention was applied in each individual case without a ground truth for comparison, the entire subsidy database becomes unreliable for analyzing azimuth data in isolation. Combined with the restriction of allowing only one orientation value per system, these design flaws render the Swedish capital subsidy program database practically unusable for orientation statistics. This serves as an important lesson for the design of future self-reporting systems.

However, because a ground truth was established in this study, it was possible to filter out these form-related issues and assess the accuracy of the self-reporting method. The results show that substantial errors arise not only from database design, but also from imprecise or careless estimations by applicants.

By contrast, the LiDAR-based approach [17] showed strong agreement with the manually derived ground truth, for both tilt and azimuth. Although occasional large errors occurred in special cases with insufficient LiDAR data or noise, overall performance was far superior to the database values. This suggests that LiDAR offers a robust and scalable means to derive system-level orientations at high accuracy.

The statistical model by Ramadhani et al. [13] provided some insight into aggregate distributions but failed to capture the diversity of real-world orientations. Its underlying assumption — that PV installations preferentially occupy the “best” roofs — led to a much narrower azimuth distribution than observed with LiDAR and it is concluded that the statistical approach cannot substitute for detailed system-level information.

In conclusion, LiDAR-based methods currently provide the most accurate and comprehensive option for determining PV system orientations at scale, while self-reported databases and statistical models are of limited reliability. For researchers and actors such as grid operators and aggregators, the adoption of remote sensing approaches based on LiDAR-derived system orientations will be essential for improving PV generation forecasts, enabling more effective network planning, and supporting the transition to high shares of distributed solar power.

6 ACKNOWLEDGEMENTS

The author gratefully acknowledges financial support from the Swedish Energy Agency (Project number P2023-00440). The Agency had no role in the study’s design, execution, or interpretation.

7 REFERENCES

[1] IEA PVPS task 1 *et al.*, “Trends in Photovoltaic Applications — 2024,” 2024.

[2] R. Luthander, D. Lingfors, and J. Widén, “Large-scale integration of photovoltaic power in a distribution grid using power curtailment and energy storage,” *Solar Energy*, vol. 155, pp. 1319–1325, 2017, doi: 10.1016/j.solener.2017.07.083.

[3] N. Etherden and M. H. J. Bollen, “Increasing the hosting capacity of distribution networks by curtailment of renewable energy resources,” in *Proceedings ot the 2011 IEEE PES Trondheim PowerTech*, IEEE, 2011, pp. 1–7. doi: 10.1109/PTC.2011.6019292.

[4] G. Barchi, M. Pierro, and D. Moser, “The impact of photovoltaic power estimation modeling on distribution grid voltages,” in *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, 2021, pp. 1–6.

[5] R. Luthander, J. Widén, D. Nilsson, and J. Palm, “Photovoltaic self-consumption in buildings: A review,” *Appl Energy*, vol. 142, pp. 80–94, 2015, doi: 10.1016/j.apenergy.2014.12.028.

[6] M. Pierro *et al.*, “Photovoltaic generation forecast for power transmission scheduling: A real case study,” *Solar Energy*, vol. 174, no. October, pp. 976–990, 2018, doi: 10.1016/j.solener.2018.09.054.

[7] M. Pierro *et al.*, “Impact of PV / Wind Forecast Accuracy and National Transmission Grid Reinforcement on the Italian Electric System,” 2022.

[8] Å. L. Sørensen, J. Hole, D. Bjerkehagen, and H. T. Walnum, “From customers to prosumers: PV systems impact on residential load profiles, peak power, and coincidence,” in *CIRED 2025 Conference*, 2025, pp. 1–5.

[9] S. Killinger *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*, vol. 173, no. August, pp. 1087–1106, 2018, doi: 10.1016/j.solener.2018.08.051.

[10] E. Hartvigsson, M. Odenberger, P. Chen, and E. Nyholm, “Estimating national and local low-voltage grid capacity for residential solar photovoltaic in Sweden, UK and Germany,” *Renew Energy*, vol. 171, pp. 915–926, 2021, doi: 10.1016/j.renene.2021.02.073.

[11] D. Lingfors and J. Widén, “Development and

validation of a wide-area model of hourly aggregate solar power generation," *Energy*, vol. 102, pp. 559–566, 2016, doi: 10.1016/j.energy.2016.02.085.

[12] Y. M. Saint-Drenan, G. H. Good, M. Braun, and T. Freisinger, "Analysis of the uncertainty in the estimates of regional PV power generation evaluated with the upscaling method," *Solar Energy*, vol. 135, pp. 536–550, 2016.

[13] U. H. Ramadhani, D. Lingfors, J. Munkhammar, and J. Widén, "On the properties of residential rooftop azimuth and tilt uncertainties for photovoltaic power generation modeling and hosting capacity analysis," *Solar Energy Advances*, vol. 3, no. February, p. 100036, 2023, doi: 10.1016/j.seja.2023.100036.

[14] S. Killinger, N. Engerer, and B. Müller, "QCPV: A quality control algorithm for distributed photovoltaic array power output," *Solar Energy*, vol. 143, pp. 120–131, 2017.

[15] K. Mayer *et al.*, "3D-PV-Locator: Large-scale detection of rooftop-mounted photovoltaic systems in 3D," *Appl Energy*, vol. 310, no. December 2021, p. 118469, 2022, doi: 10.1016/j.apenergy.2021.118469.

[16] G. Kasmi, L. Dubus, P. Blanc, and Y. M. Saint-Drenan, "Towards unsupervised assessment with open-source data of the accuracy of deep learning-based distributed PV mapping," in *CEUR Workshop Proceedings*, 2022.

[17] D. Lingfors, R. Johansson, and J. Lindahl, "Deriving the orientation of existing solar energy systems from LiDAR data at scale," *Solar Energy*, vol. 291, no. 113344, 2025, doi: 10.1016/j.solener.2025.113344.

[18] J. Martín-Jiménez, S. Del Pozo, M. Sánchez-Aparicio, and S. Lagüela, "Multi-scale roof characterization from LiDAR data and aerial orthoimagery: Automatic computation of building photovoltaic capacity," *Autom Constr*, vol. 109, no. September 2019, p. 102965, 2020, doi: 10.1016/j.autcon.2019.102965.

[19] B. Tian, R. C. G. M. Loonen, R. Valckenborg, and J. L. M. Hensen, "A fully automated urban PV parameterization framework for improved estimation of energy production profiles," 2025. doi: https://doi.org/10.48550/arXiv.2505.19876.

[20] D. Lingfors, J. Bright, N. Engerer, J. Ahlberg, S. Killinger, and J. Widén, "Comparing the capability of low- and high-resolution LiDAR data with application to solar resource assessment, roof type classification and shading analysis," *Appl Energy*, vol. 205, no. June, pp. 1216–1230, 2017, doi: 10.1016/j.apenergy.2017.08.045.

[21] Sveriges Riksdag, *Svensk författningssamling — Förordning (2009:689) om statligt stöd till solceller*. Sweden, 2009.

[22] H. Rydehell, B. Lantz, I. Mignon, and J. Lindahl, "The impact of solar PV subsidies on investment over time — the case of Sweden," *Energy Econ*, vol. 133, no. April, p. 107552, 2024, doi: 10.1016/j.eneco.2024.107552.

[23] L. Molin, S. Ericson, D. Lingfors, J. Munkhammar, and J. Lindahl, "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," in *40th European Photovoltaic Solar Energy Conference and Exhibition (EUPVSEC)*, 2023, pp. 020527–001. doi: http://dx.doi.org/10.4229/EUPVSEC2023/5DV.2.10.

[24] J. Lindahl, R. Johansson, and D. Lingfors, "Mapping of decentralised photovoltaic and solar thermal systems by remote sensing aerial imagery and deep machine learning for statistic generation," *Energy and AI*, vol. 14, p. 100300, 2023, doi: 10.1016/j.egyai.2023.100300.

[25] Â. Frimane, R. Johansson, J. Munkhammar, D. Lingfors, and J. Lindahl, "Identifying small decentralized solar systems in aerial images using deep learning," *Solar Energy*, vol. 262, 2023, doi: 10.1016/j.solener.2023.111822.

[26] U. H. Ramadhani, F. Johari, O. Lindberg, and J. Munkhammar, "A city-level assessment of residential PV hosting capacity for low-voltage distribution systems considering rooftop data and uncertainties," *Appl Energy*, vol. 371, no. February, p. 123715, 2024, doi: 10.1016/j.apenergy.2024.123715.

[27] E. Ekstrand and F. Hermodsson, "Assessing Rooftop Solar Energy Adoption — The remaining potential across market segments and the impact of socio-economic factors," Master Thesis, Uppsala University, 2024.

[28] W. J. Conover, *Practical Nonparametric Statistics*, Third Edit. New York: John Wiley & Sons, 1999.

[29] S. Liljeroth, "A techno-economic study of commercial and industrial PV systems in Sweden," Master Thesis, Uppsala University, 2025.