energies

MDPI

Article

Probabilistic Methodology for Calculating PV Hosting Capacity in LV Networks Using Actual Building Roof Data

Miha Grabner [1,2,*], Andrej Souvent [2], Nermin Suljanović [2], Andrej Košir [1] and Boštjan Blažič [1]

[1] Faculty of Electrical Engineering, University of Ljubljana, Tržaška c. 25, 1000 Ljubljana, Slovenia; andrej.kosir@fe.uni-lj.si (A.K.); bostjan.blazic@fe.uni-lj.si (B.B.)

[2] Electric Power Research Institute Milan Vidmar, Hajdrihova 2, 1000 Ljubljana, Slovenia; andrej.souvent@eimv.si (A.S.); nermin.suljanovic@eimv.si (N.S.)

* Correspondence: miha.grabner@eimv.si

Received: 01 August 2019 date; Accepted: 22 October 2019; Published: 25 October 2019

Abstract: There has been an increasing trend of integrating photovoltaic power plants (PVs). One of the important challenges for distribution system operators is to evaluate the total installed power of a PV that a particular network can host (or PV hosting capacity) while keeping voltage and element constraints within required limits. The major drawback of the existing methods for calculating PV hosting capacity is that they use the same installed power of the PV systems for all simulated PVs, as these methods do not use external data sources about building roofs. As a consequence, this has a significant impact on the final accuracy of the results. This paper presents a probabilistic methodology for calculating the PV hosting capacity in low voltage (LV) networks. The main contribution of this paper is the improved modeling of PV generation using actual building roof data when calculating the PV hosting capacity, as every building is treated according to its actual solar potential. Monte Carlo simulations with incorporated stochastic consumption and PV generation models are utilized for load flow calculations of the actual LV network. The simulation results presented in this paper prove that the proposed methodology increases the accuracy of the final PV hosting capacity calculations.

Keywords: LV networks; hosting capacity; Monte Carlo; PV

1. Introduction

There has been an increasing trend of installed photovoltaic power plants (PVs) in the European Union (EU) due to various national subsidies and net-metering schemes. Solar energy has great potential, mainly due to unutilized building roof surfaces. High distributed generation penetration, however, can cause overloads and voltage violations. Therefore, it is important to analyze different network development scenarios in advance [1–3].

Regulation in many countries promotes new smart grid technologies that enable higher PV penetration in the network, so assessment of the most appropriate technology for a particular network has become crucial. Calculating network hosting capacity is important for the comparison of benefits introduced by different smart grid technologies [4–6].

PV hosting capacity is a special case of network hosting capacity, where different PV penetration scenarios are analyzed. The PV hosting capacity can be interpreted as the maximum installed power of PVs (or a maximum number of consumers with PVs) that a particular network can host while still keeping voltage and element constraints within the required limits. A major benefit of the PV hosting capacity concept is the clear criteria for the PV penetration assessment, which makes the PV hosting capacity concept specific, measurable, and practical [6]. In order to statistically assess the PV hosting

capacity, various scenarios must be calculated for different PV locations and penetrations. This can provide a statistical description of the possible levels of PV penetration while voltages and element loadings are still within the required range.

Many papers have dealt with this kind of problem. The obtained results are usually presented as a network violation probability, depending on the number of consumers with installed PVs or total nominal power of PVs in the network. The paper [7] described low voltage (LV) network planning based on the Monte Carlo approach, taking into account different random distributed generation (DG) locations. The probabilities of voltage levels and the maximum PV penetration limits were assessed. The study in reference [8] proposed a similar approach, where the PV hosting capacity was evaluated considering the probability of voltage limit violation occurrences at the customer level, according to different PV power factors. The approach developed by the Electric Power Research Institute (EPRI) [9] also proposed a stochastic approach when creating PV deployment scenarios. The stochastic nature of the analysis takes into account the uncertainty in the size and location of the future PV systems [9]. In reference [10], the work proposed a Monte Carlo-based technique to assess the impacts of different PV penetrations on LV networks, in order to estimate their corresponding hosting capabilities. Similarly, the authors in reference [11] proposed a probabilistic methodology for evaluating the impacts of PVs, electric heat pumps, micro combined heat and power units, and electrical vehicles on network operation states. The authors in reference [12] studied how to improve PV hosting capacity with a battery storage system. They developed a method to select the optimal size of the battery and converter unit, as well as the optimal placement of an LV-central battery storage system. PV hosting capacity increase potential for different local Volt/var control strategies and the associated grid losses for various types of LV feeders have been analyzed and compared in reference [13]. The authors in reference [14] proposed and demonstrated an integrated planning model that allowed distribution PV network hosting capacity constraints to be included in long-term grid planning. They demonstrated the impact of integrating a large amount of distributed small-scale storage batteries for use as hosting capacity enhancement in the distribution network in order to quantify its benefit in terms of overall system costs. In reference [15], a probabilistic-based framework to determine the maximum integration limits of DGs considering the voltage rise and voltage deviation constraints was proposed. This framework required the use of the hosting capacity model, which was formulated as a non-linear optimization problem. The authors in reference [16] performed extensive sensitivity studies to quantify the effects of several factors on the PV hosting capacity. The effects of number of customers with PV generators, PV power factor, voltage magnitude on the medium voltage system, load level, and conductor impedances were investigated. The authors in reference [17] addressed the applicability of PV hosting capacity, studying the methodologies of harmonic voltage distortion caused by PV. The voltage rise due to harmonic injection was analyzed and discussed, with the aim of validating the discussed model and also putting forward recommendations for connecting PV generators across other network systems. The authors in reference [18] studied the impact of single-phase PVs on PV hosting capacity by utilizing different percentages of PVs per phase and, consequently, studying the voltage unbalance. Similar approaches to those mentioned above were also presented in references [19–23]. Finally, comprehensive literature reviews have been provided in References [24,25].

The above overviewed methods, however, did not leverage building roof data when randomly allocating PVs in the network. Therefore, when performing simulations, buildings that had high PV potential and buildings that did not have PV potential were treated equally. As a consequence, the final results of PV hosting capacity were significantly impacted, which has also been shown in reference [7]. The method proposed in this paper utilizes roof data and outperforms these other methods, which is shown in the use-case section.

In comparison with other approaches, the proposed method extends the current input data with actual building roof data. This has become possible with the availability of new data sources. With 3D-point cloud data derived from light detection and ranging (LIDAR) and land register geographic information system (GIS) data, it is possible to form 3D roof models. Further, the solar irradiance for every roof can be determined by using actual tilts and orientations of particular roof surfaces. This

enables the calculation of PV system nominal power and stochastic generation for every roof from solar irradiance measurements.

The main contribution of this paper is the improved modeling of PV generation using actual building roof data when calculating the PV hosting capacity, as every building is treated according to its actual solar potential.

The structure of the paper is as follows: in Section 2, the proposed methodology is thoroughly described. In Section 3, the proposed methodology is demonstrated in a real LV network, using actual LIDAR and GIS data of an analyzed village. Finally, the discussion and conclusion are given in Sections 4 and 5, respectively.

2. Methodology

The PV hosting capacity is the maximum installed power of PVs that a particular network can host while still keeping network constraints within the required limits. It depends on a location and installed power of PV systems. We cannot predict the location nor the installed power of future PVs, so the problem can be formulated as follows: given the LV network with N number of consumers, the PV hosting capacity can be evaluated by examining all possible combinations of locations and installed power of PV systems. Calculating load flow for all possible combinations is computationally almost impossible, so we leverage the Monte Carlo method for solving this problem.

The proposed methodology includes three steps. The first (and the most time-consuming) step is the acquisition and preprocessing of data from various sources, as shown in Figure 1. The second step is stochastic modeling of consumption and solar generation, which is used for probabilistic load flow calculations. The third and final step is the improved calculation of the PV hosting capacity using actual building roof data and leveraging Monte Carlo simulations, which is also the major contribution of this paper.

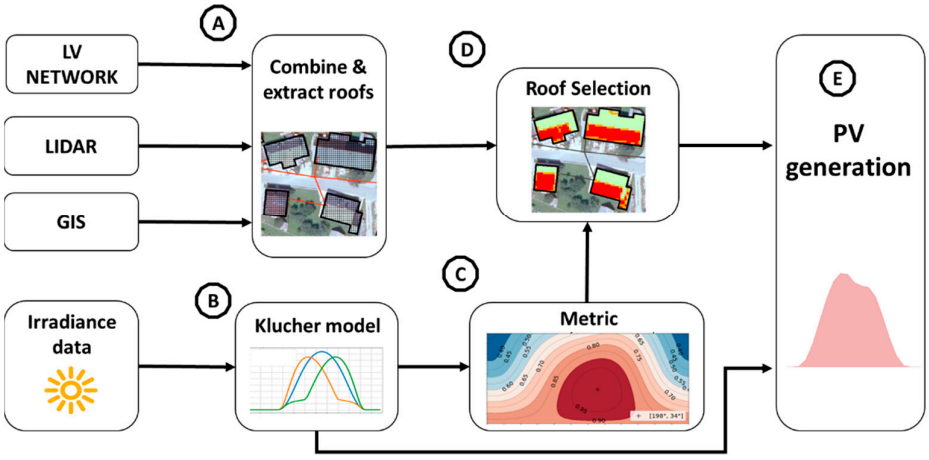

Figure 1. Photovoltaic power plant (PV) generation modeling.

2.1. Input Data Preparation

Input data preparation is an important part of the proposed methodology. The required input data sources are

- electrical network data;
- measurements from smart meters (not essential);
- LIDAR;
- GIS (land register); and
- solar irradiance data.

Network data are collected and pre-processed for a selected middle voltage (MV) feeder and corresponding LV grid. Low voltage consumers are modeled using smart metering data. In the absence of such data, stochastic models for similar areas can be used [7].

For modeling the solar irradiance on a building roofs, a digital elevation model (DEM) has to be obtained from 3D point cloud data, where the 3D point cloud data can be derived from LIDAR. LIDAR provides unstructured 3D point cloud data consisting of millions of points with x, y, and z components. LIDAR data preprocessing has been described, in detail, in reference [26]. Further, the tilt and orientation have to be calculated for every grid cell and converted into a vector polygon format. GIS data from the land register is used for extracting building roofs from LIDAR and linking the roofs to the LV network. The output of the LIDAR and GIS data preparation are polygons of building roofs, with tilt and orientation information for every grid cell of 1 m × 1 m (step A in Figure 1) and are further represented in a matrix form.

Next, all possible combinations of tilt and orientation using 5° step are calculated, where tilt goes from 0° to 90° and orientation goes from 0° to 180°. A final number of combinations is denoted as n, which is 648. Finally, $\boldsymbol{A}$ is area matrix of roof surfaces in m^2 for every tilt-orientation combination and for every building roof polygon, where r is the number of roofs in LV network. The following matrix represents tilt-orientation information for every roof:

$$\boldsymbol{A} = \begin{bmatrix} a_{1,1} & \dots & a_{1,r} \\ \vdots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,r} \end{bmatrix}, where\ \boldsymbol{A} \in \boldsymbol{R}^{nxr} \quad (1)$$

For the stochastic modeling of PV generation, hourly measurements of global horizontal irradiance (GHI) and diffuse horizontal irradiance (DHI) have to be obtained from the nearest weather station. The two input solar irradiance time-series vectors (which will be used in the later analysis) are defined as follows:

- $\boldsymbol{I_m}\ \left[\frac{W}{m^2}\right]$ represents the GHI measurements; and
- $\boldsymbol{I_{m,\ d}}\ \left[\frac{W}{m^2}\right]$ represents the DHI measurements, for the last few years and with an hourly resolution.

Next, these two vectors have to be transposed and stacked vertically in order to form a solar irradiance measurements dataset matrix $\boldsymbol{X}$.

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{I_m^T} \\ \boldsymbol{I_{m,\ d}^T} \end{bmatrix}, where\ \boldsymbol{X} \in \boldsymbol{R}^{2xm} \quad (2)$$

In the equation above, m denotes the number of observations.

GHI and DHI measurements, together with tilt and orientation information for every roof, are essential for modeling PV solar generation in further calculations and will be described in the next section.

2.2. Consumption and PV Generation Modeling

Deterministic load flow analyses are not appropriate for modeling the LV network operating conditions due to the stochastic nature of consumers and PVs [3,27,28]. Instead, probabilistic load flow analyses must be carried out, taking into account the stochastic modeling of demand and generation. The proposed approach relies on the already developed stochastic load models derived from smart metering data in Slovenia, which have been described in Reference [7].

The most important step of the proposed methodology is the stochastic PV generation modeling of the building roofs, having different tilts and orientations.

A detailed PV generation modeling is important for two reasons,

- selecting roofs with high PV potential; and
- accurate modeling of stochastic PV generation on different roof surfaces.

First, the variables that will be used in the further analysis are presented in Table 1.

Table 1. Solar irradiance variables.

Variable name	Unit	Explanation
Irradiance	$I\ \left[\frac{W}{m^2}\right]$	radiant flux received by a surface per unit area;
Radiant exposure	$H\ \left[\frac{Wh}{m^2}\right]$	radiant energy received by a surface per unit area.

2.2.1. Solar Irradiance on Tilted and Oriented Surfaces

Different buildings have differently oriented and tilted roofs; thus, this has to be incorporated when modeling solar irradiance. Solar irradiance measurements from weather stations usually include global and diffuse solar irradiance data on a horizontal surface [29,30]. Therefore, there are various models which are able to predict irradiance on a tilted and oriented surface (hereafter denoted by I_{T}) from measured GHI and DHI. In order to calculate solar irradiance on a particular roof surface from $\boldsymbol{I_m}$ and $\boldsymbol{I_{m,\ d}}$ weather station measurements, the well-known model developed by Klucher [29] was be used. The Klucher model is defined by the following equation [29,30]:

$$I_{\mathrm{T}} = I_{\mathrm{h,b}}R_{\mathrm{b}} + I_{\mathrm{h,d}}\left(\frac{1+\ \cos\beta}{2}\right)\left[1+\ F'\sin^3\left(\frac{\beta}{2}\right)\right]\ \left[1+F'\cos^2\theta\,\sin^3\theta_{\mathrm{z}}\right] + I_{\mathrm{h}}\rho\left(\frac{1-\ \cos\beta}{2}\right) \quad (3)$$

where I_{T} denotes solar irradiance on a tilted surface, $I_{\mathrm{h,b}}$ denotes the direct-normal component of solar irradiance on the horizontal surface, R_{b} denotes a variable geometric factor which is the ratio of tilted and horizontal solar beam irradiance, $I_{\mathrm{h,d}}$ denotes the global diffuse horizontal solar irradiance, β denotes surface tilt angle from horizon, θ denotes incident angle of the surface, θ_{z} denotes the zenith angle, I_{h} denotes global horizontal solar irradiance, and ρ denotes the hemispherical–hemispherical ground reflectance.

The Klucher model is crucial in the proposed methodology and can be understood as a function mapping GHI and DHI weather station measurement to global irradiance (GI) on a desired tilted and oriented surface (step B in Figure 1).

2.2.2. Selecting Roofs with High PV Potential

Different roofs are characterized with different surfaces and, consequently, different solar potentials. This section describes how to extract the most suitable surfaces for every roof (e.g., north-facing surfaces for a particular roof will not be used in a model).

First, mean annual irradiance vectors $\boldsymbol{I_{A,m}}$ and $\boldsymbol{I_{A,m,d}}$ are calculated from $\boldsymbol{I_m}$ and $\boldsymbol{I_{m,d}}$, transposed and stacked vertically.

$$\boldsymbol{X_A} = \begin{bmatrix} \boldsymbol{I_{A,\ m}^T} \\ \boldsymbol{I_{A,\ m,\ d}^T} \end{bmatrix}, where\ \boldsymbol{X_A}\ \in\ \boldsymbol{R}^{2x8760} \quad (4)$$

The resulting matrix $\boldsymbol{X_A}$ has 8760 columns, as analyzed data have hourly resolution and columns in a matrix $\boldsymbol{X_A}$ represent annual data (365 days × 24 h).

Next, the Klucher model is applied to $\boldsymbol{X_A}$ for all possible combinations of tilt and orientation using 5° step (n combinations).

$$\boldsymbol{I_A} = \begin{bmatrix} i_{A,\ 1,1} & \cdots & i_{A,\ 1,8760} \\ \vdots & \vdots & \vdots \\ i_{A,\ n,1} & \cdots & i_{A,\ n,8760} \end{bmatrix}, where\ \boldsymbol{I_A}\ \in\ \boldsymbol{R}^{nx8760} \quad (5)$$

In the last step, values in a matrix $\boldsymbol{I_A}$ are summarized for every row separately to obtain annual solar radiant exposure for every tilt-orientation combination. The result is represented as a column vector $\boldsymbol{H}$.

$$\boldsymbol{H} = \sum_{j=1}^{8760} \boldsymbol{I}_{\boldsymbol{A},\ :\boldsymbol{j}}, where\ \boldsymbol{H} \in \boldsymbol{R}^{nx1} \quad (6)$$

Further, vector $\boldsymbol{H}$ is normalized.

$$\boldsymbol{H_{norm}} = \frac{\boldsymbol{H}}{\boldsymbol{max(H)}} \quad (7)$$

Next, $\boldsymbol{H_{norm}}$ is further changed in such way that only surfaces having normalized annual solar radiant exposure greater than 0.8 are selected, which means that in further simulations PV systems will be placed only on the most suitable roof surfaces. This vector represents a metric of the PV potential for every surface (n combinations) depending on how much annual solar radiant exposure a particular surface receives.

It is not optimal to cover the whole roof surface with PVs, and in some cases, a roof does not have any potential. It turns out that using a threshold of 0.8 leads to the installed power of PV panels being similar to those actually used in our country. This means that only roof surfaces with tilt-orientation combinations greater than 80% of the optimal annual radiant exposure are selected for PV placement in our further calculations.

Next, Boolean column vector $\boldsymbol{H_{metric}}$ is calculated from $\boldsymbol{H_{norm}}$ having n rows for all tilt-orientation combinations and equals to 1 for all suitable tilt-orientation combinations and to zero for other surfaces. The resulting vector $\boldsymbol{H_{metric}}$ serves as a metric for selecting only the most suitable roof surfaces in further analysis (step C and D in Figure 1).

After selecting roofs with high PV potential, stochastic PV generation for various weather scenarios on different roofs must be developed.

2.2.3. Stochastic PV Generation Modeling using Actual Roofs Surfaces

This section presents the development of stochastic PV generation models for differently oriented and tilted roofs and for various weather scenarios (step E in Figure 1). The main contribution of this paper is the improved modeling of PV generation using actual building roof data, as described below.

PV generation is calculated from the solar irradiance for every surface of a particular roof. This brings a new challenge, as solar irradiance and, consequently, daily profiles are different for every tilt and orientation of the roof. Consequently, PV stochastic modeling in such cases is more complex, as the stochastic generation should be modeled for every tilt-orientation combination separately and, then, aggregated for every roof.

The following solar irradiance matrix $\boldsymbol{I}$ is calculated from $\boldsymbol{X}$ using the Klucher model for all n tilt-orientation combinations and only for summer days (the highest solar irradiance in Slovenia is in summer), where s denotes the number of summer days in a dataset $\boldsymbol{X}$

$$\boldsymbol{I} = \begin{bmatrix} i_{1,1} & \dots & i_{1,s} \\ \vdots & \vdots & \vdots \\ i_{n,1} & \dots & i_{n,s} \end{bmatrix}, where\ \boldsymbol{I} \in \boldsymbol{R}^{nxs} \quad (8)$$

where $i_{1,1}$ is the GI on a tilt 0° and an orientation 0° for first summer solar irradiance measurement, whereas $i_{n,s}$ is the GI on a tilt 90° and an orientation 180° calculated from the last summer solar irradiance measurement.

PV generation for a chosen hour in the Monte Carlo simulations is acquired as follows:

1. Randomly choose one column for a previously chosen hour from matrix $\boldsymbol{I}$. Here we utilize random sampling with replacement from a finite population, which means that one column can be selected more than once. The result is a vector $\boldsymbol{D} \in \boldsymbol{R}^{nx1}$ holding irradiance for every tilt-orientation combination on a particular day and represents one weather scenario.
2. Select roofs with high PV potential using $\boldsymbol{H_{metric}}$ and utilize PV systems (μ_{PV}) and inverter efficiency ($\mu_{inverter}$) to derive PV generation for all tilt-orientation combinations from vector $\boldsymbol{D}$ ($\boldsymbol{P}_g$ has unit [W/m^2])

$$\boldsymbol{P_g} = \mu_{PV} \cdot \mu_{inverter} \cdot \boldsymbol{H_{metric}} \cdot \boldsymbol{D}, where\ \boldsymbol{P_g} \in \boldsymbol{R}^{nx1} \quad (9)$$

PV systems' and inverters' efficiencies of 15% and 95% are utilized in the calculations to obtain actual PV generations.

3. Let matrix $\boldsymbol{B}$ represent a subset of matrix $\boldsymbol{A}$, where $\boldsymbol{B}$ holds only the selected columns (randomly chosen roofs during PV hosting capacity calculations) and the total number of chosen roofs is denoted as i. For every column j in a matrix $\boldsymbol{B}$ (every selected roof), we calculate

$$\boldsymbol{P_{g,\ roof,\ j}} = \sum_{i=1}^{n} \left(\boldsymbol{B}_{i,j} \cdot \boldsymbol{P_g}\right), where\ \boldsymbol{P_{g,\ roof,\ j}} \in \boldsymbol{R} \quad (10)$$

4. Which means that PV generation is summarized for all surfaces in a particular roof to create PV generation on a particular day.
   After calculating PV generation for every roof in step 3, the results are stacked vertically to create a PV generation vector.

$$\boldsymbol{P_{g,\ PV}} = \begin{bmatrix} \boldsymbol{P_{g,\ roof,\ 1}} \\ \vdots \\ \boldsymbol{P_{g,\ roof,\ i}} \end{bmatrix}, where\ \boldsymbol{P_{g,\ PV}} \in \boldsymbol{R}^{ix1} \quad (11)$$

5. Output of the procedure in step 4 is PV generated power for every chosen roof on a particular day, which is further used in load flow calculations.

2.3. Calculating PV Hosting Capacity Using the Monte Carlo Method

This section explains how the final load flow calculations are performed using the Monte Carlo method in order to calculate the PV hosting capacity.

Probabilistic load flow analyses with the Monte Carlo method have been widely used for simulating LV networks operating states, considering various DG types [31–36]. In order to perform analyses for different seasons, day types, or hours, sequential Monte Carlo has to be used and stochastic models have to be made for every calculated sequence separately. The results of the Monte Carlo load flow simulations are computed based on repeated random sampling and statistical analysis.

Probabilistic load flow calculations using Monte Carlo methods actually perform deterministic load flow calculations for a large number of times using different combinations of nodal load and generation values [3]. In further load flow calculations, the load, PV generation and the location of consumers with PVs are represented as random variables. The result of the proposed algorithm is a set C of possible PV hosting capacity values. Further, the probability distribution for a set C can be analyzed and interpreted.

The proposed algorithm workflow for calculating the PV hosting capacity using actual roof surfaces is shown in Figure 2. The term "network violations" in Figure 2 refers to the following constraints: The voltage limits considered are in accordance with the SIST EN 50160 standard, which defines voltage limits for MV and LV networks within ±10% of the nominal voltage (U_n). The maximally allowed voltage at MV and LV levels were 110% of U_n. The maximally allowed cable loading was set to 75% nominal apparent power (S_n; due to cable installation factors) and 100% S_n for the transformer.

The proposed algorithm workflow shown in Figure 2, for every hour, is as follows (numbers in the list below refer to Figure 2):

1 Choose real number K which denotes the number of iterations and equals to a size of a final set C.

2 Set $i = 0$, where i denotes the total number of chosen roofs and create empty matrix $\boldsymbol{B}$, which later holds the information about the roof data.

3 Randomly select one column from a matrix $\boldsymbol{A}$ holding roofs data and add a selected column to $\boldsymbol{B}$. Here, we utilize random sampling without replacement, which means that every roof

(location) has the same chance of being chosen and that every roof is chosen only once (until the network violation occur and all the locations are reset).

4 Generate stochastic load (described in Section 2.2) for all consumers and form a load matrix $\boldsymbol{P}_l$ of size r (number of roofs and consumers).

$$\boldsymbol{P_l} = \begin{bmatrix} \boldsymbol{P}_{l,\ consumer\ 1} \\ \vdots \\ \boldsymbol{P}_{l,\ consumer\ r} \end{bmatrix}, where\ \boldsymbol{P_l} \in\ \boldsymbol{R}^{rx1} \tag{12}$$

5 Generate stochastic PV generation (described in Section 2.2.3) for all chosen roofs and form a PV generation matrix $\boldsymbol{P}_{g,PV}$ using roof matrix $\boldsymbol{B}$ from step 3.

6 Calculate load flow. Here, we solve the following equations:

$$P_{Gk} - P_{Dk} = V_k \sum_{j=1}^{N} V_j \left[G_{kj} \cos\left(\delta_k - \delta_j\right) + B_{kj} \sin(\delta_k - \delta_j) \right]; \quad k = 1,2, \ldots N$$

$$Q_{Gk} - Q_{Dk} = V_k \sum_{j=1}^{N} V_j \left[G_{kj} \sin\left(\delta_k - \delta_j\right) - B_{kj} \cos(\delta_k - \delta_j) \right] \tag{13}$$

$$G_{kj} + jB_{kj} = (k, j)\ element\ of\ the\ bus\ admittance\ matrix$$

where P_{Gk} is active power generation, P_{Dk} is active power demand, Q_{Gk} is reactive power generation, Q_{Dk} is reactive power demand, V_k is voltage, and δ_k is voltage angle for a particular node k in a network. V_j is voltage and δ_j is voltage angle for a node j which is close to the node k. In our case, P_{Gk}, P_{Dk}, Q_{Gk}, and Q_{Dk} are random variables representing consumer demand and PV generation. P_{Dk} equals to P_l (calculated in step 7), $Q_{Dk} = 0.33\ P_l$, P_{Gk} for every bus equals to $\boldsymbol{P}_{g,PV}$ (calculated in step 8) and $Q_{Gk} = 0$.

7 Check network violations (voltage limits and element loadings). If there are no violations, return to a step 3. If there is at least one violation, go to the next step.

8 Aggregate the nominal power of all PVs in the network without the last one (there were no violations until the last PV was added). If $P_{g,j}$ denotes installed power of j-th added PV system in the network and C_k is one value in a set C of PV hosting capacity values then C_k is calculated as follows:

$$C_k = \sum_{j=1}^{i-1} P_{g,j} \tag{14}$$

9 Add C_k to a set C which holds PV hosting capacities.

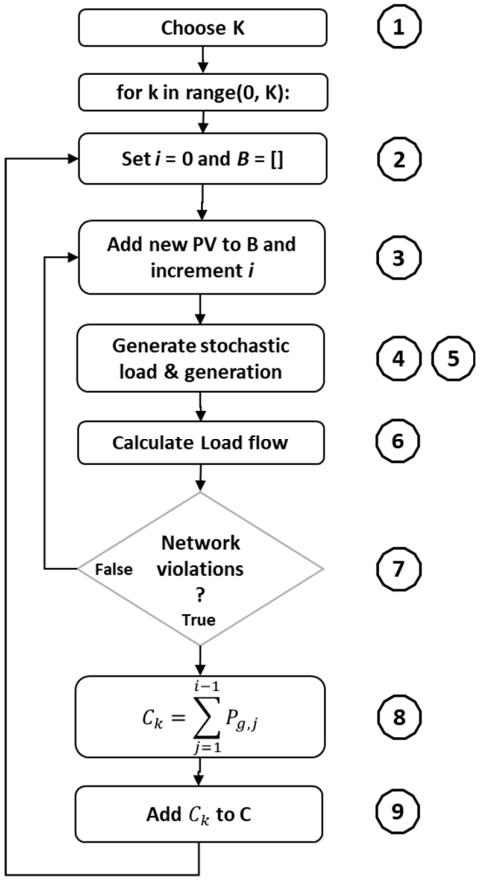

Figure 2. PV hosting capacity algorithm workflow.

3. Case Study

3.1. Data Preparation

The proposed methodology was validated on a real LV network using LIDAR and land register GIS data of actual buildings. The first step is network modeling. The chosen LV network was supplied by a 250 kVA MV/LV transformer, having 7 feeders and 93 consumers. The main feeder lines were Al 70 mm^2 cables, whereas the side branches were mainly Al 35 mm^2 cables.

The output of the electrical network LIDAR and GIS data preparation was polygons of the building roofs, along with tilt and orientation information for every 1 m x 1 m grid cell connected to the LV grid, as described in Section 2.1.

3.2. Solar Potential Results for an Analyzed Village

After the appropriate network and roof modeling, solar radiant exposure for every roof was calculated. This enabled assessment of the solar potential of the analyzed rural network (village) and selection of the roofs that had high PV potential, which will be used in further analyses as described in Section 2.2. As described in Sections 2.2.2, the normalized column vector $\boldsymbol{H}_{norm}$ was calculated. Here, the vector $\boldsymbol{H}_{norm}$ is reshaped to a matrix having orientation values on x-axis and tilt values on y-axis, so that the resulting matrix for a chosen area can be plotted (shown in in Figure 3). In order to receive as much radiant exposure as possible, it is important to have a slightly west oriented PV panel. The optimal position around the analyzed weather station was at azimuth 198 ° and tilt 34 °. In order for the surface to receive at least 90% of the maximal yearly radiant exposure (red color), there were a wide range of tilt-orientation (azimuth) combinations available. The authors in reference [37] performed a national study of solar energy in Slovenia and came up with similar results.

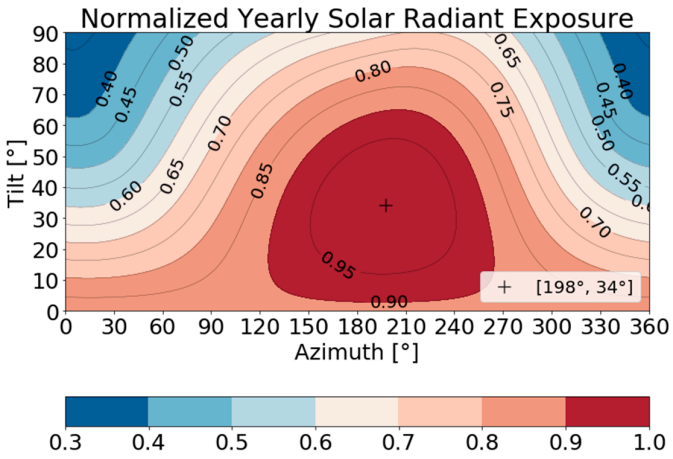

Figure 3. Normalized yearly solar radiant exposure on a tilted and oriented surface for an analyzed area, which serves as a basis for selecting appropriate roof surfaces, depending on their solar potential.

The solar potential for the analyzed village for a part of the chosen LV network is presented in Figure 4. Every color represents the normalized yearly radiant exposure that a particular polygon receives, depending on an optimal yearly radiant exposure, as presented in Figure 3. Red denotes that a particular surface has at least 80% of the optimal yearly radiance exposure; thus, those were the most suitable roof surfaces (mostly south-facing parts of the roofs), followed by the orange, which was within 60–80%. Finally, the light green color indicates the surfaces which were the least suitable for PV placement. As defined by the methodology proposed in Section 2.2.2., PV panels were placed only on the roof surfaces that were most suitable for PV placement and colored with red in Figure 4.

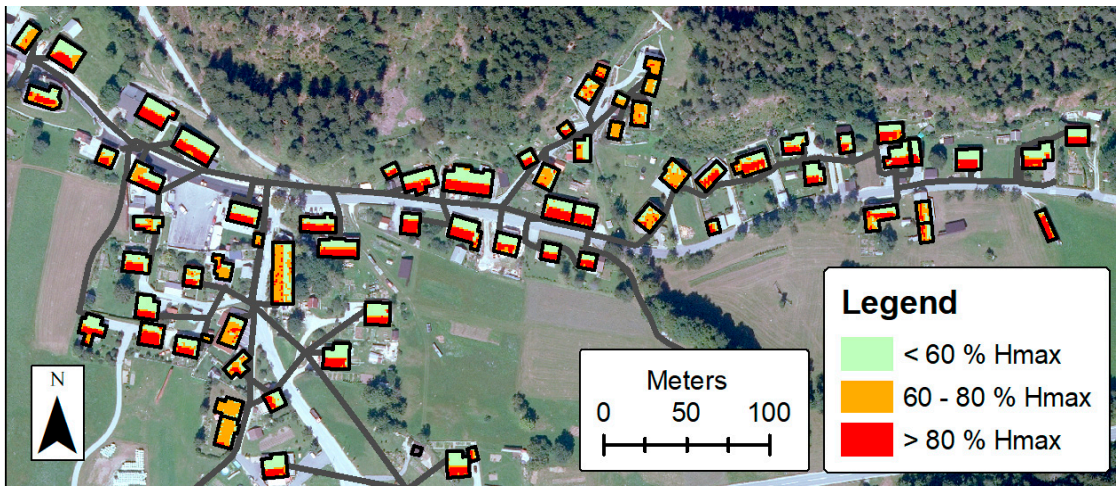

Figure 4. Solar irradiance on a building roofs enables assessment of the solar potential for an analyzed village.

3.3. PV Hosting Capacity Results

The last step in the proposed methodology for calculating the PV hosting capacity was to perform the load flow simulations using the Monte Carlo method, which enabled assessment of the PV hosting capacity. The result of the proposed method is a set *C* of PV hosting capacity values. Here, we utilized kernel density estimation of a set *C*, which is a non-parametric way to estimate the probability density function of a random variable, to plot the results in Figure 5.

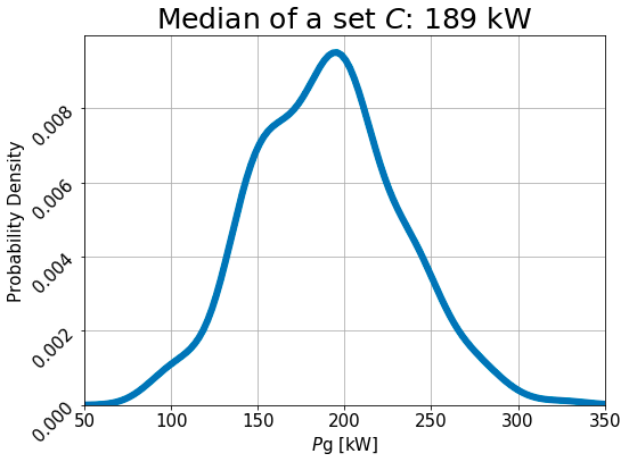

Figure 5. Results of calculating PV hosting capacity with the proposed methodology. Figure shows kernel density estimation of a set *C* of PV hosting capacity values calculated using the Monte Carlo method.

Since it is easier to examine the quantiles of the sample from a cumulative histogram, we show it in Figure 6. The label "actual roofs" refers to our proposed methodology using actual building roof surfaces. The median PV hosting capacity is 189 kW. The values on the left part of the graph are buildings that are far from the transformer station and have a greater impact on the network voltages due to lesser short circuit apparent power. On the other hand, the values on the right of the graph are the buildings near the transformer station.

Finally, the proposed methodology was compared with existing ones using the same PV installed power for every roof in a model. Since other methods do not utilize LIDAR and GIS data for modeling PV generation, those models cannot distinguish between different roofs. Therefore, other methods usually utilize the most often installed power of PV systems such as 10 kW or 15 kW. The blue line shows results when installed power of PVs is 15 kW, whereas the green shows results when installed power is 10 kW. There are two major shortcomings of these approaches. First, all roofs are treated equally, which means that roofs having high PV potential and roofs without PV potential are modeled the same. Second, the chosen PV installed power such as 10 kW or 15 kW significantly impacts the final results, as shown in Figure 6 (median hosting capacity when using 10 kW PVs is almost 50% higher than median hosting capacity when using 15 kW).

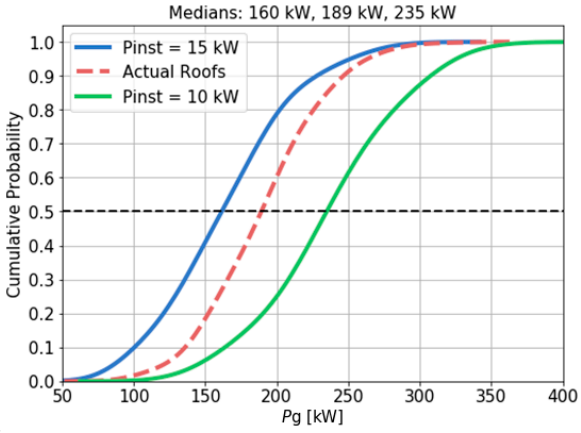

Figure 6. Comparison of the proposed PV hosting capacity methodology using actual roof surfaces (dashed red line) with existing ones, that use the same PV installed power per roof across all buildings (blue and green line).

As described in the methodology, the algorithm randomly chose one roof at a time and performed load flow analysis sequentially until a network violation occurred. This allows presenting the simulation results depending on the total PV installed power in the network, as shown in Figure 7.

The maximal voltage in the LV network for every simulation is represented with a circle, which is colored depending on violations. Green represents simulations without network violations, and red represents the simulations where violations occurred. The sizes of the red circles are correlated to the number of violations. It can be seen that violations started to occur when the total installed power of PV was around 100 kW.

All the load flow calculations were performed using the PandaPower Python library [38].

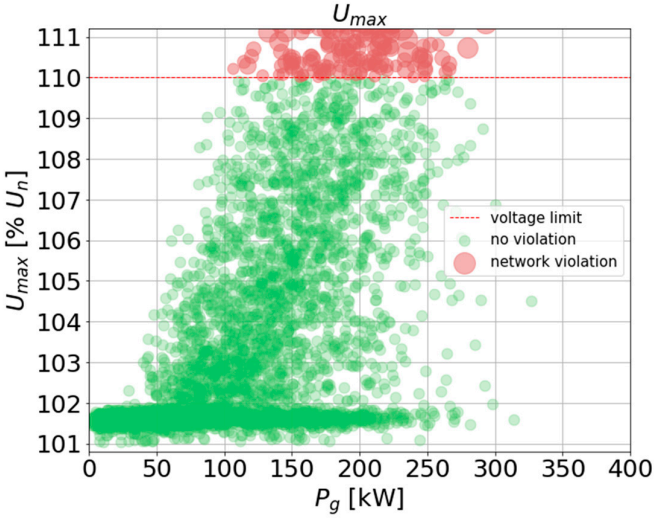

Figure 7. U_{max} depending on total PV installed power in the network for all simulations.

4. Discussion

It was confirmed that the chosen PV installed power significantly impacted the final results of the PV hosting capacity in our case study. It can be seen (from the blue line in Figure 6) that, when the chosen PV installed power of all PVs in the network was 15 kW, the median PV hosting capacity was 160 kW. When the PV installed power was 10 kW, the median PV hosting capacity was 235 kW (the green line in Figure 6). Furthermore, it can be seen that the difference was significant (50%) and that the actual PV hosting capacity results (the dashed red line in Figure 6) lay somewhere in between.

Choosing the single most appropriate PV installed power for modeling all the PVs in the network is hard as, in reality, every house will have a different PV panel. The major drawback of existing methods is that they use the same installed power for all simulated PVs and, consequently, the results can vary a lot, as shown in Figure 6; this problem has already been addressed in reference [7].

This is the case because distribution system operators do not possess any additional data about building roofs, and also as LIDAR data, which are essential for modeling building roofs, are not available in all countries. Additionally, combining electrical LV network data with other data sources is very hard and time-consuming, as it requires proper data science skills from preprocessing to connecting all the data sources.

Our approach uses actual roof surfaces in load flow calculations, which consequently results in the improved accuracy of calculating the PV hosting capacity. Additionally, using actual roof surfaces in the calculations complicates stochastic PV generation modeling, due to the fact that generation depends on the tilt and orientation of a PV system, as opposed to other approaches in the literature where stochastic PV generation modeling is straightforward (where the same orientation and tilt is used for all PVs).

5. Conclusions

In this paper, we presented a probabilistic methodology for calculating the PV hosting capacity using actual building roof data, which was achieved by combining electrical network, LIDAR, and GIS data. Stochastic PV generation models were developed for differently oriented and tilted roofs and for various weather scenarios. The proposed methodology was thoroughly described and confirmed in a case study.

The major drawback of existing PV hosting capacity approaches is that they use the same installed power of PV systems for all simulated PVs, as these methods do not use external data sources about building roofs. Consequently, all roofs are treated equally, which means that roofs having high PV potential and roofs without PV potential are modeled the same. Another problem with existing methods is that a chosen PV installed power significantly impacts the final result, which was also confirmed in our case study. The major motivation behind our paper was to improve PV generation modeling by using external data sources.

The main contribution of this paper is the improved modeling of PV generation using actual building roof data when calculating PV hosting capacity, as every building is treated according to its actual solar potential. It was confirmed that using actual building roof data is important for accurate calculation of PV hosting capacity in our case study by comparing our approach with existing ones. Our approach uses actual roof data in simulations, which consequently results in improved accuracy of final PV hosting capacity calculations.

Author Contributions: Conceptualization, M.G., A.S., N.S., B.B.; Methodology, M.G., B.B.; Software, M.G.; Validation, M.G.; Formal Analysis M.G.; Writing - original draft preparation, M.G.; Writing – review and editing, M.G., A.S., N.S., A.K., B.B.; Visualization, M.G.; Supervision, N.S., N.S, A.K., B.B.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bletterie, B.; Kadam, S.; Renner, H. On the classification of low voltage feeders for network planning and hosting capacity studies. *Energies* **2018**, *11*, 651.
2. Georgilakis, P.S.; Hatziargyriou, N.D. A review of power distribution planning in the modern power systems era: Models, methods and future research. *Electr. Power Syst. Res.* **2015**, *121*, 89–100.
3. Chen, P.; Chen, Z.; Bak-Jensen, B. Probabilistic load flow: A review. In Proceedings of the 2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, Nanjing, China, 6–9 April 2008; pp. 1586–1591.
4. Arshad, A.; Püvi, V.; Lehtonen, M. Monte Carlo-based comprehensive assessment of PV hosting capacity and energy storage impact in realistic finnish low-voltage networks. *Energies* **2018**, *11*, 1467.
5. Wang S., Chen S., Ge L., and Wu L. Distributed Generation Hosting Capacity Evaluation for Distribution Systems Considering the Robust Optimal Operation of OLTC and SVC, *IEEE Trans. Sustainable Energy.* **2016,** *7*, 1111–1123.
6. Bollen, M.H.; Hassan, F. *Integration of Distributed Generation in the Power System*; John Wiley & Sons: Hoboken, NJ, USA, 2011; Volume 80.
7. Kolenc, M.; Papič, I.; Blažič, B. Assessment of maximum distributed generation penetration levels in low voltage networks using a probabilistic approach. *Int. J. Electr. Power Energy Syst.* **2015**, *64*, 505–515.
8. Ballanti, A.; Pilo, F.; Navarro-Espinosa, A.; Ochoa, L.F. Assessing the benefits of PV var absorption on the hosting capacity of LV feeders. In Proceedings of the IEEE PES ISGT Europe 2013, Copenhagen, Denmark, 6–9 October 2013; pp. 1–5.
9. Smith, J.; Rylander, M. Stochastic analysis to determine feeder hosting capacity for distributed solar PV. *Electr. Power Res. Inst. Palo Alto CA Tech. Rep.* **2012**, *1026640*, 0885–8950.
10. Navarro, A.; Ochoa, L.F.; Randles, D. Monte Carlo-based assessment of PV impacts on real UK low voltage networks. In Proceedings of the 2013 IEEE Power & Energy Society General Meeting, Vancouver, BC, Canada, 21–25 July 2013; pp. 1–5.

11. Navarro-Espinosa, A.; Ochoa, L.F. Probabilistic impact assessment of low carbon technologies in LV distribution systems. *IEEE Trans. Power Syst.* **2015**, *31*, 2192–2203.
12. Divshali, P.H.; Soder, L. Improving Hosting Capacity of Rooftop PVs by Quadratic Control of an LV-Central BSS. *IEEE Trans. Smart Grid* **2019**, *10*, 919–927.
13. Schultis, D.-L. Comparison of Local Volt/var Control Strategies for PV Hosting Capacity Enhancement of Low Voltage Feeders. *Energies* **2019**, *12*, 1560.
14. Saad, S.N.M.; van der Weijde, A.H. Evaluating the Potential of Hosting Capacity Enhancement Using Integrated Grid Planning modeling Methods. *Energies* **2019**, *12*, 3610.
15. Abad, M.S.S.; Ma, J.; Zhang, D.; Ahmadyar, A.S.; Marzooghi, H. Probabilistic Assessment of Hosting Capacity in Radial Distribution Systems. *IEEE Trans. Sustain. Energy* **2018**, *9*, 1935–1947.
16. Torquato, R.; Salles, D.; Pereira, C.O.; Meira, P.C.M.; Freitas, W. A Comprehensive Assessment of PV Hosting Capacity on Low-Voltage Distribution Systems. *IEEE Trans. Power Deliv.* **2018**, *33*, 1002–1012.
17. Oliveira, T.; Carvalho, P.; Ribeiro, P.; Bonatto, B. PV Hosting Capacity Dependence on Harmonic Voltage Distortion in Low-Voltage Grids: Model Validation with Experimental Data. *Energies* **2018**, *11*, 465.
18. Arshad, A.; Püvi, V.; Lehtonen, M. Monte Carlo-Based Comprehensive Assessment of PV Hosting Capacity and Energy Storage Impact in Realistic Finnish Low-Voltage Networks. *Energies* **2018**, *11*, 1467.
19. Ding, F.; Mather, B. On Distributed PV Hosting Capacity Estimation, Sensitivity Study, and Improvement. *IEEE Trans. Sustain. Energy* **2017**, *8*, 1010–1020.
20. Al-Saadi, H.; Zivanovic, R.; Al-Sarawi, S.F. Probabilistic Hosting Capacity for Active Distribution Networks. *IEEE Trans. Ind. Inf.* **2017**, *13*, 2519–2532.
21. Palacios-Garcia, E.; Moreno-Muñoz, A.; Santiago, I.; Moreno-Garcia, I.; Milanés-Montero, M. PV Hosting Capacity Analysis and Enhancement Using High Resolution Stochastic Modeling. *Energies* **2017**, *10*, 1488.
22. Ding, F.; Mather, B.; Gotseff, P. Technologies to increase PV hosting capacity in distribution feeders. In Proceedings of the 2016 IEEE Power and Energy Society General Meeting (PESGM), Boston, MA, USA, 2016; pp. 1–5.
23. Dubey, A.; Santoso, S.; Maitra, A. Understanding photovoltaic hosting capacity of distribution circuits. In Proceedings of the 2015 IEEE Power & Energy Society General Meeting, Denver, CO, USA, 2015; pp. 1–5.
24. Mulenga, E.; Bollen, M.H.J.; Etherden, N. A review of hosting capacity quantification methods for photovoltaics in low-voltage distribution grids. *Int. J. Electr. Power Energy Syst.* **2020**, *115*, 105445.
25. Ismael, S.M.; Aleem, S.H.A.; Abdelaziz, A.Y.; Zobaa, A.F. State-of-the-art of hosting capacity in modern power systems with distributed generation. *Renew. Energy* **2019**, *130*, 1002–1020.
26. Freitas, S.; Catita, C.; Redweik, P.; Brito, M.C. Modelling solar potential in the urban environment: State-of-the-art review. *Renew. Sustain. Energy Rev.* **2015**, *41*, 915–931.
27. Conti, S.; Raiti, S. Probabilistic load flow using Monte Carlo techniques for distribution networks with photovoltaic generators. *Sol. Energy* **2007**, *81*, 1473–1481.
28. Navarro-Espinosa, A.; Ochoa, L.F. Increasing the PV hosting capacity of LV networks: OLTC-fitted transformers vs. reinforcements. In Proceedings of the 2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 17–20 February 2015; pp. 1–5.
29. Klucher, T.M. Evaluation of models to predict insolation on tilted surfaces. *Sol. Energy* **1979**, *23*, 111–114.
30. Loutzenhiser, P.G.; Manz, H.; Felsmann, C.; Strachan, P.A.; Frank, T.; Maxwell, G.M. Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation. *Sol. Energy* **2007**, *81*, 254–267.
31. El-Khattam, W.; Hegazy, Y.; Salama, M. Investigating distributed generation systems performance using Monte Carlo simulation. *IEEE Trans. Power Syst.* **2006**, *21*, 524–532.
32. Torquato, R.; Shi, Q.; Xu, W.; Freitas, W. A Monte Carlo simulation platform for studying low voltage residential networks. *IEEE Trans. Smart Grid* **2014**, *5*, 2766–2776.
33. Vallée, F.; Klonari, V.; Lisiecki, T.; Durieux, O.; Moiny, F.; Lobry, J. Development of a probabilistic tool using Monte Carlo simulation and smart meters measurements for the long term analysis of low voltage distribution grids with photovoltaic generation. *Int. J. Electr. Power Energy Syst.* **2013**, *53*, 468–477.
34. Ruiz-Rodriguez, F.; Hernandez, J.; Jurado, F. Probabilistic load flow for radial distribution networks with photovoltaic generators. *IET Renew. Power Gener.* **2012**, *6*, 110–121.
35. Widén, J.; Wäckelgård, E.; Paatero, J.; Lund, P. Impacts of distributed photovoltaics on network voltages: Stochastic simulations of three Swedish low-voltage distribution grids. *Electr. Power Syst. Res.* **2010**, *80*, 1562–1571.

36. Grabner, M.; Bregar, Z.; Ivanjko, Š.; Valenčič, L. Improved model for the spatial load forecasting of the Slovenian distribution network. *CIRED Open Access Proc. J.* **2017**, *2017*, 2354–2357.
37. Kastelec, D.; Rakovec, J.; Zakšek, K.; Medved, S. *Sončna Energija v Sloveniji*; Založba ZRC, ZRC SAZU: Ljubljana, Slovenia, 2007.
38. Thurner, L.; Scheidler, A.; Schäfer, F.; Menke, J.-H.; Dollichon, J.; Meier, F.; Minecke, S.; Braun, M. Pandapower—An open-source python tool for convenient modeling, analysis, and optimization of electric power systems. *IEEE Trans. Power Syst.* **2018**, *33*, 6510–6521.