



Article

Analysis of Energy Transition Impact on the Low-Voltage Network Using Stochastic Load and Generation Models [†]

Raoul Bernards ^{1,*} , Werner van Westering ^{2,3} , Johan Morren ⁴ and Han Slootweg ⁴¹ Enexis Netbeheer B.V., Magistratenlaan 116, 5223 MB 's-Hertogenbosch, The Netherlands;² Delft Center of Systems & Control, Delft University of Technology, Delft Mekelweg 2, 2628 CD Delft, The Netherlands; w.h.p.vanwestering@tudelft.nl³ Alliander N.V., Dijkgraaf 4, P.O. Box 50, 6921 RL Duiven, The Netherlands⁴ Electrical Energy Systems, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands; j.morren@tue.nl (J.M.); j.g.slootweg@tue.nl (H.S.)

* Correspondence: raoul.bernards@enexis.nl

[†] This paper is an extended version of our paper published in CIRED 2019 Conference, Madrid, Spain, 3 June 2019, <http://dx.doi.org/10.34890/720>.

Received: 9 October 2020; Accepted: 16 November 2020; Published: 21 November 2020



Abstract: The energy transition poses a challenge for the electricity distribution network design as new energy technologies cause increasing and uncertain network loads. Traditional static load models cannot cope with the stochastic nature of this new technology adoption. Furthermore, traditional nonlinear power methods have difficulty evaluating very large networks with millions of cables, because they are computationally expensive. This paper proposes a method which uses copulas for modeling the uncertainty of technology adoption and load profiles, and combines it with a fast linear load flow model. The copulas are able to accurately model the stochastic behavior of solar irradiance, load measurements, and mobility data, converting them into electricity load profiles. The linear load flow model has better scalability and stability compared to traditional load flow models. The models are applied to a case study which uses a real-world dataset consisting of a realistic technology adoption scenario and a low-voltage network with millions of cables, which considers both voltage and current problems. Results show that risk profiles can be generated for all cables in the network, resulting in a valuable map for the district network operator as to where to focus their efforts.

Keywords: energy transition; copula stochastic modeling; low voltage distribution network

1. Introduction

The energy transition poses a challenge for electricity distribution network design as new energy technologies cause increasing and uncertain network loads. Distributed energy generation becomes more common as Photovoltaics (PV) get cheaper, shifting electricity production from predictable large plants to consumers [1]. Additionally, consumers are electrifying their transportation and heating needs with electric vehicles (EV) and heat pumps (HP). Each of these techniques has a significant impact on the residential electricity consumption [2].

Distribution network operators (DNO), which are responsible for maintaining the medium and low-voltage (LV) networks, are trying to determine the exact impact of these technologies. This is especially challenging on a low voltage level, given that the technology adoption of individual customers is hard to predict. Furthermore, because of their vast size, the impact on urban low voltage networks is hard to simulate.

For predicting the increasing and uncertain loads, this paper proposes a method which uses copulas for modeling the uncertainty of technology adoption and load profiles. For determining the impact on the low-voltage network, a fast linear load flow model is described. The copulas and load flow models are combined in a case study considering a large real-world low-voltage network. Results show that risk profiles can be generated for all cables in the network, resulting in a valuable map for the district network operator as to where to focus their efforts.

1.1. Related Work

Traditionally, the impact of load on low-voltage networks is determined by evaluating a single deterministic peak load. This peak load is determined via a Velandar approximation [3] which accounts for peak load covariance between consumers. While this approach is straightforward and requires relatively little data, it cannot cope with the uncertain and irregular load profiles caused by the adoption of new energy technologies [2]. In more recent literature, probabilistic approaches are the norm for assessing low-voltage networks. With a probabilistic peak load, it becomes possible to make risk-based decisions and consider multiple planning alternatives [4]. Considering residential load, many probabilistic load models are designed from the bottom up [5–7], starting at individual appliance level and their operational states, and combining these into a specific model for each house. While accurate, the bottom-up design requires information on a household level, which is generally not available by DNOs. Furthermore these models have a high modeling intensity, i.e., they require many individual variables and parameters to be modeled correctly. This necessitates the use of additional assumptions and renders them impractical for application [8]. Conversely, spatial load forecasting methods based on historical growth data can be employed such as in [9]. However, this approach lacks the possibility to evaluate different future scenarios, due the absence of modeling the underlying drivers for changes in load.

Other work focuses on probabilistic simulation through random sampling from measured datasets, e.g., in [4,10,11]. Either electric load profiles are sampled directly or profiles of different variables (e.g., heat demand or solar irradiance) are sampled and converted to electric load. While these approaches allow for ease of use and low modeling intensity is required, they are very dependent on the actual values that occurred by chance in the used datasets.

The statistical modeling of measurement data related to electricity networks and energy technologies such as PV, EV, and HP has been discussed in, for example, [12–17], using various distribution functions. However, all these models show applicability only to one specific set and no mention is made of generalizability to other data.

Another major point in the statistical modeling of load profiles is dealing with the inherent correlation between time steps. This effect can be accounted for by copulas [18], which are the basis of this paper. Copulas can be used to model the correlation structure between different random variables, allowing for separate modeling of the correlation and the marginal distributions. Copulas have been successfully applied to modeling loads in power system studies [19–23], most often focusing on modeling wind power generation.

1.2. Contributions

This work adds to the existing research on load and generation modeling by extending and implementing a flexible modeling method that is able to model the statistical behavior of load profiles and is applicable to a variety of data. Existing methods to model stochastic load profiles are often specifically tailored to a single application, requiring information and data on a very low level to accurately model the behavior. In low-voltage network planning, such low level information is usually not available, making the models infeasible for use without relying on additional assumptions. Moreover they are difficult or impossible to adapt to different technologies, limiting their efficient use in large scale network planning which needs to include many different load components. Our method allows for an accurate and standardized way to model the electric load behavior of different connections

and technologies, and can therefore be incorporated in existing network planning processes with relative ease.

Furthermore, stochastic load models are created for residential load, PV, EV, and HP, yielding insight into their electrical behavior and the uncertainties herein. The profiles are generated by first modeling the statistical behavior present in different measurement datasets and subsequently modeling the electric load based on the statistical representation of the data. Additionally, we show that these profiles are accurately modeled using the proposed methodology. The models also include different correlations which are, specifically, temporal correlations for residential load and PV generation, correlations between travel distance and arrival and departure times for EV charging, and the combined correlation with time and temperature for HP load. We also show that the use of a single combination of probability distribution functions (Gaussian mixture models with t-copulas) can effectively model many different datasets.

The stochastic nature of the load models puts a higher computational burden on simulations using these models. To effectively apply the models in large-scale network simulations, a fast network analysis methodology is therefore discussed and applied in the final sections of this paper. This analysis methodology allows for fast and detailed analysis of large network areas (e.g., two to three million connections). We demonstrate that by using this analysis methodology it becomes feasible to apply more detailed load models also for large scale network evaluations, which might previously have been limited by unacceptable computation times.

1.3. Paper Outline

In the next section, first the general profile modeling methodology will be discussed. Then, we show the application, fitting, and output of an application of the methodology to different sets of measurement data. The methodology is applied to create individual profile models for current residential load, PV, EV, and HP. In the subsequent section, the models are applied in a large scale network analysis using a fast and efficient loadflow algorithm. Here the uncertain impacts of an energy transition scenario are shown on the real-world LV network area of a distribution grid operator. Finally, the paper concludes with several conclusions and recommendations regarding the proposed models and their application.

2. Stochastic Profile Modeling

2.1. Methodology

The modeling of the load and generation profiles is based on a method proposed by the authors in [24]. In this work, the original method is extended with several modifications to allow the generation of load profiles based on measurements of different variables.

The core of the methodology is a two-stage approach to the modeling of stochastic correlated variables. First, the marginal distributions of the data variables are modeled using Gaussian mixture models (GMMs), while the correlation structure of the variables is modeled separately using copula functions. This detachment of the marginals and the correlation allows for a flexible modeling method, which can be effectively applied to a variety of data.

The different steps in the modeling methodology are detailed in Figure 1. In the first step, the measurement data are structured and preprocessed to remove missing values. Then, the data are clustered by quarter hour of the day, per day type (week or weekend), and per month. Consequently, GMMs are fit to model the marginal distribution of each quarter hour. The data are then transformed to uniform distributions using the CDF transform as in [25], and a copula is fit to the multivariate uniform distribution. To generate profiles first samples are generated from the copula functions, which are then converted back to the original scale of the marginal distributions using the inverse CDF transform.

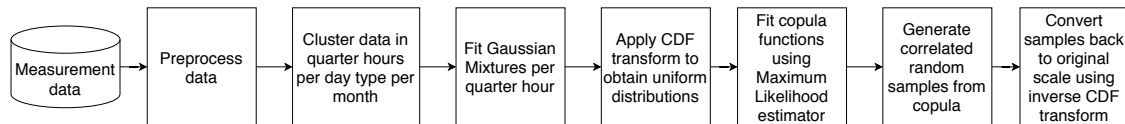


Figure 1. Diagram of the steps for generating the correlated load models.

Accuracy of the stochastic models is evaluated by the Continuous Rank Probability Score (CRPS), which is an often used score in ensemble forecasts [26]. It compares the difference between a forecast CDF and the empirical CDF. Additionally, the Root Mean Square Percentage Error (RMSPE) [27] is used to evaluate average output profiles to the measured average to quantify the level forecast error in the output model. A relative measure is used to judge the fit to the shape of the profile, as the absolute values between (15-min) clusters may vary significantly.

More details about the modeling technique can be found in [24]. In the next paragraphs, the individual steps will be explained more in-depth by applying the method to several datasets. The used datasets contain electric residential load profiles, solar irradiance measurements, mobility data, and heat pump electric load measurements.

2.2. Residential Load

First, we consider the modeling of the base residential load, which is the current load of households without any PV, EV, and/or HP system installed. The dataset used for this is an open data set of smart meter data of residential loads in 2013 [28]. This dataset contains recorded 15-min average energy consumption values for 80 households for a whole year. The households in this dataset vary from apartments to detached houses, with a gas connection for heating and cooking, and a wide range in yearly electricity consumption varying from 1000 to over 6500 kWh. Moreover, the households also do not have any energy transition technologies already installed.

2.2.1. Data and Preprocessing

To analyze and model the statistical behavior the measurements are first clustered into groups, based on time of day, day type (week or weekend day) and month of the year. This partitioning is based on the seasonality of the load and time of use behavior of residential consumers [7,29]. Each cluster now consists of all measurements for all households for all week or weekend days in a specific month at a specific 15-min time interval (e.g., the time interval of 12:00–12:15 h for all weekend days in June). This results in $12 \text{ months} \times 2 \text{ day types} \times 96 \text{ quarters} = 2304$ clusters with each around 1700 and 700 measurements for week and weekend days, respectively. Figure 2 illustrates the behavior of the residential load over time and the stochastic variation at specific time instants. The mean profiles in Figure 2a show a clear daily cycle, with a peak at 18:00 h of approximately 0.9 kW in the winter and 0.5 kW in summer. The wide confidence interval around the mean indicates a large spread in individual load values, which is displayed in more detail in Figure 2b,c, which shows the distribution of the load at two chosen 15-min intervals.

2.2.2. Model Fitting

First, GMMs are fit to the measurement data to obtain the marginal distributions of each individual quarter of an hour. The optimal number of components of each GMM is selected by evaluating at what number the Bayesian Information Criterion (BIC) is minimal. Figure 3a shows the evaluation of the BIC for GMMs with an increasing number of components, for several quarter of an hours in January. For all clusters, the optimal number of components varies between 3 and 7 components. Figure 3b show the fit of four of the mixture models to their respective data by comparing the probability density function (PDF) of the GMM to the measured data, indicating a good fit and showing no signs of overfitting. CRPS values of the different clusters are also low, 0.01 kW on average, ranging from 0.002 to 0.02 kW.

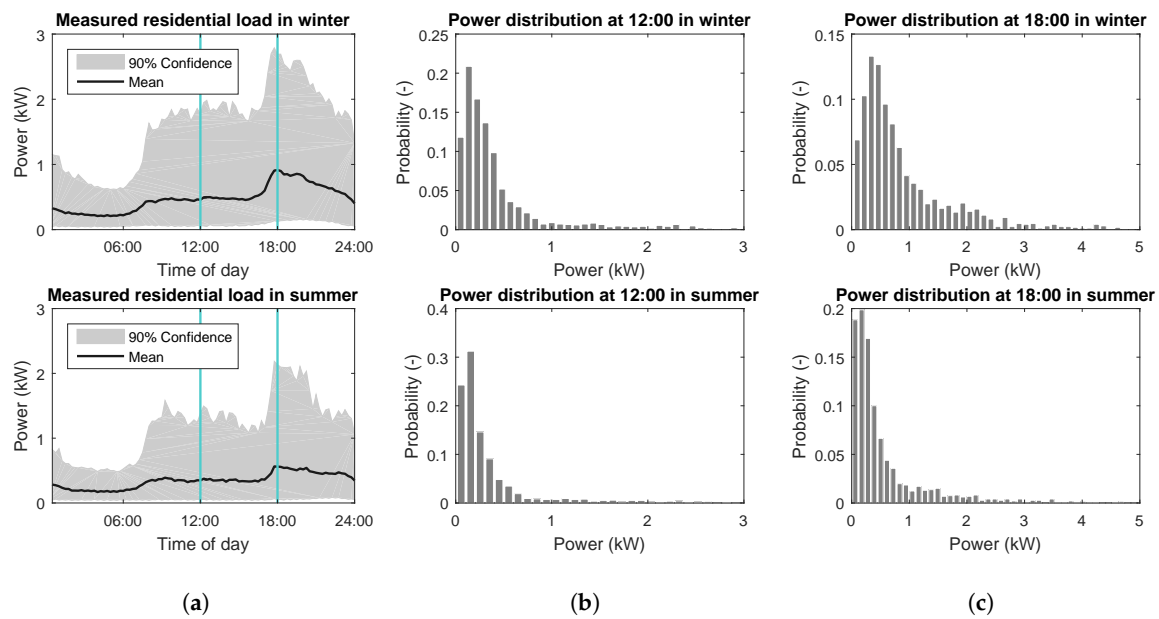


Figure 2. Load behavior in winter and summer. (a) Mean profiles and 90% confidence interval around the mean. (b,c) Stochastic behavior of load at specific time intervals.

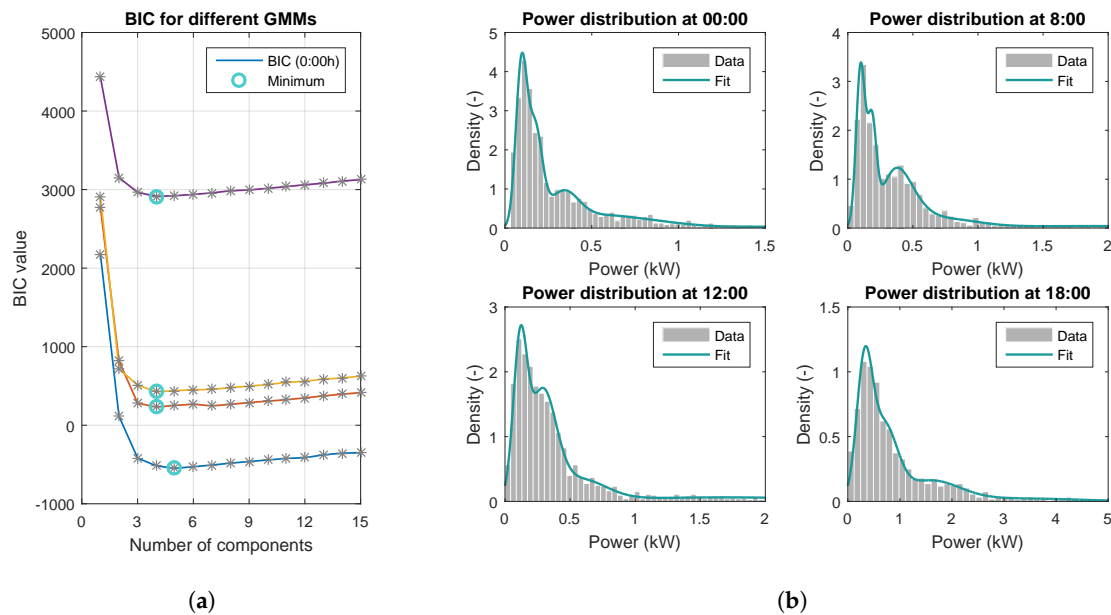


Figure 3. Fitting of mixture models to measurement data at four time intervals. (a) The Bayesian Information Criterion (BIC) values of the Gaussian mixture models (GMMs) for an increasing number of components. (b) PDFs of data and model fit using the optimal number of components.

Next, a transformation is applied to transform the data to (approximate) uniform distributions with a range between $[0, 1]$ by using the probability integral transform [18]. Copulas are fitted for modeling the dependence structure for all points in time for a each day-month date-type (e.g., a weekday in January). Different families of copulas [18,30] were considered and their “goodness of fit” was tested based on the log-likelihood of the transformed data points. For all studied cases, the t-copula produced the best overall results. Subsequently, these copula functions can be used to generate daily load profiles.

2.2.3. Profile Generation and Evaluation

By using the copulas, correlated samples are generated. The samples are obtained from day profile with a 15-min interval resulting in 96 correlated time steps per day. The generated samples are normalized on a $[0, 1]$ range and to transform the samples back to their original scale (representing their actual power values present in the data) an inverse CDF transformation with the marginal distributions was applied. Given that the CDF of mixture distributions cannot be obtained analytically, we opted for a two-step numerical method. First, the marginal CDFs were generated with small discrete steps. Subsequently, a line search was performed resulting in the correctly scaled power values.

Profiles can be now generated which represent the statistical behavior and correlations present in the original dataset. Figure 4 shows a comparison of measured and generated profiles for 30 and 300 random household profiles for four days in January. By visual inspection, the model appears to represent the time behavior and peak load well. The probability density distribution of the load for a full year is shown in Figure 5, indicating a proper capture of the stochastic yearly behavior by the model. Further, the correlation factors between quarters of an hour of the measured and modeled profiles are compared in Figure 6, again demonstrating a good representation of these correlations by the model with a root mean squared error (RMSE) of 0.05.

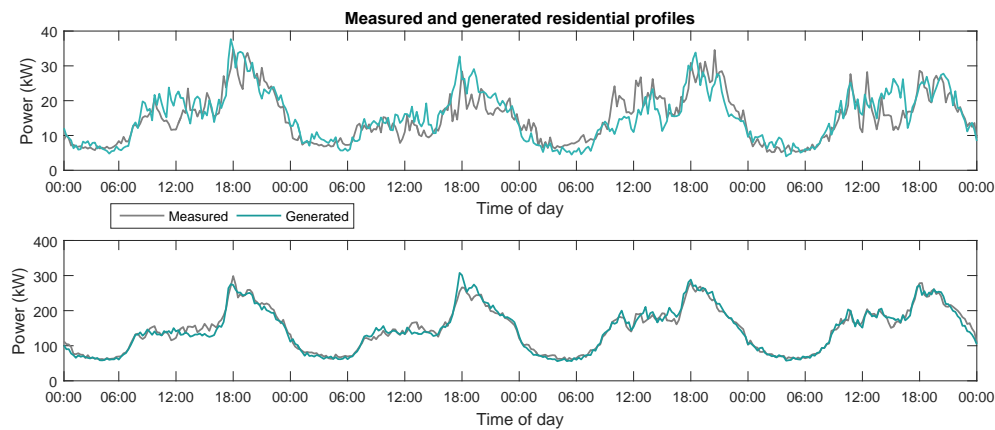


Figure 4. Comparison of measured and generated profiles for two week and weekend days in January for 30 (top) and 300 (bottom) households.

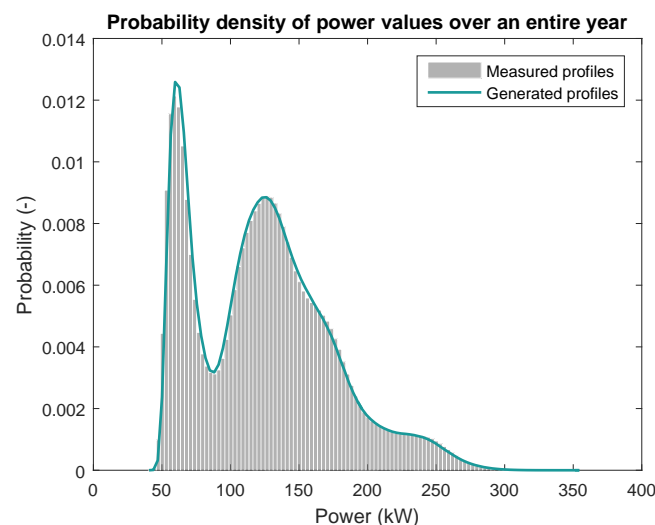


Figure 5. Probability density of year profiles generated comparing actual measured values and the the generated day profiles.

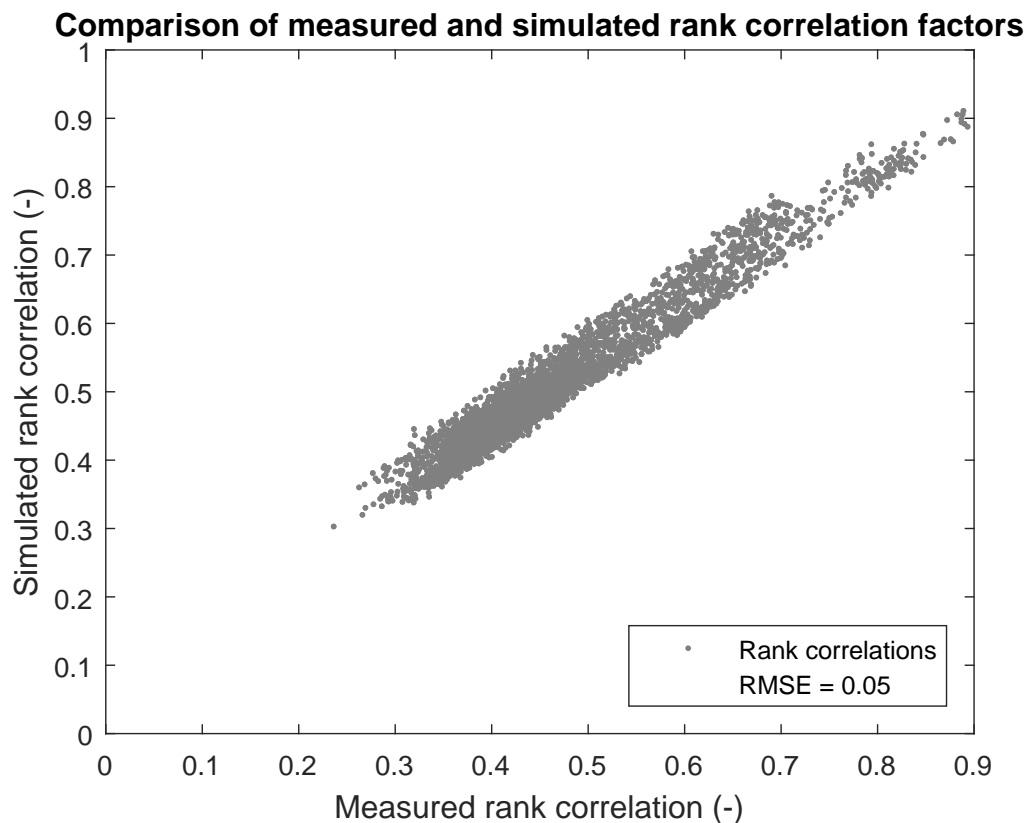


Figure 6. Plot showing the correlation between the actual measured values and the simulated values.

Due to the flexibility of the GMMs and copulas, profiles can be generated for PV, EV, and HP technologies in a similar fashion. However, for each of these cases some technology specific adaptations are required to achieve proper load or generation profiles. The next sections will focus on these extensions of the proposed method for each technology.

2.3. Photovoltaics

The power output of a PV system depends on the effective solar irradiance on the surface of the panels. This is affected by several factors such as the actual irradiance levels, the changing angle of the sun over the day (and year), and the orientation and inclination angle of the panels [31]. To model the PV output, we first model the irradiance profiles and subsequently convert the irradiance values to an instantaneous power output.

2.3.1. Data and Model Fit

The modeling of the irradiance profiles is based on measurements of the solar irradiance for a period of 10 years from four weather stations in the Netherlands, obtained from the openly available meteorological data of the Dutch Royal Meteorological Institute (KNMI) [32]. The data is clustered per 15-min interval of the day for each specific month (the distinction between week and weekend days is obviously not made), resulting in $12 \text{ months} \times 96 \text{ quarters} = 1152$ clusters with each around 1300 measurements. Figure 7 shows the fit of six mixture models to their respective data clusters for morning, noon, and evening clusters in January and June.

2.3.2. Model Fitting

Next, we again first model the marginal distributions of each individual quarter of an hour. Because there is no or negligible irradiance during night we only fit models on quarters between

(nautical) dawn and dusk (Nautical dawn and dusk are the moments when the sun is 12 degrees below the horizon, around this time the irradiance may become large enough for PV installations to start producing.). Evaluation of the optimal number of components showed the optimal performance for 2 to 4 components, depending on the time of day and year. Figure 7 shows the fit of six mixture models to their respective data clusters for morning, noon, and evening clusters in January and June. The densities of the fitted mixture models show to track the stochastic behavior of varying levels of irradiance well, with an average CRPS of 0.011 kW/m² for summer days and 0.050 for winter days.

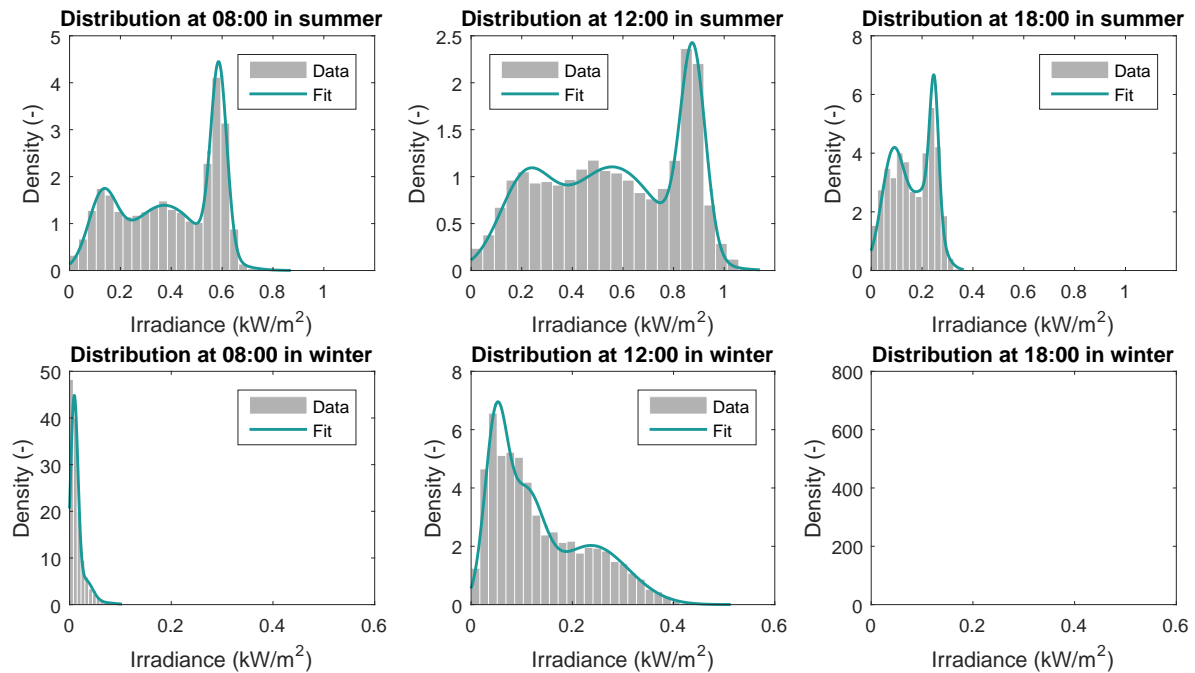


Figure 7. PDFs of data and fitted mixture models for irradiance measurements at three time intervals in summer and winter.

To model the correlation between solar irradiance at different time intervals of the same day, copula functions are then fit to the transformed distributions of the data. The correlation matrices of the solar irradiance are constructed for the quarters between dawn and dusk for each month and the remaining correlation values are fixed at 0. Analyzing different model copula function fits showed the t-copula to also provide the best fit for solar irradiance data.

2.3.3. Profile Generation and Evaluation

By sampling values from the copula functions and transforming these back using the GMMs, stochastic irradiance profiles can be generated. Figure 8 shows a comparison of averaged measured and generated profiles of 30 and 300 random two day profiles in summer and winter, indicating the model closely represents the measured profiles. The respective RMSPE values of the summer and winter profiles are 2.9% and 5.3%.

To obtain profiles of the power output of a PV system the global irradiance profiles are converted to the irradiance on an inclined surface using the calculation methods from [31,33,34].

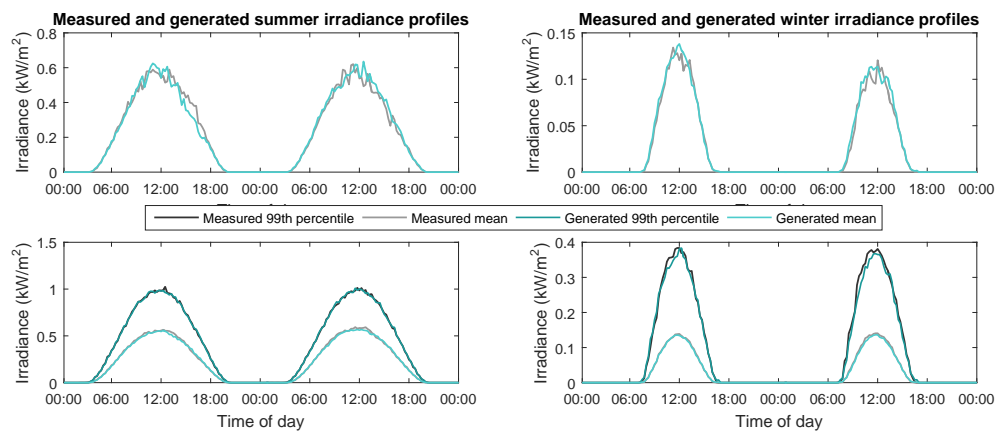


Figure 8. Comparison of measured and generated profiles of 30 (**top**) and 300 (**bottom**) random two day profiles in summer and winter.

2.4. Electric Vehicles

To model electric vehicle's charging profiles a modified approach is used, as rather than fitting the marginal distributions directly to measured charging profiles, we model arrival times and trip distances and convert to electric load profiles afterwards. This has several reasons: First, because data on mobility is openly available for many countries, this way of building up the models is more generically applicable than those that rely only on measured charging profiles which might not be available or not representative enough. Second, in absence of some smart charging scheme, the charging profiles of electric vehicles usually take the form of block shaped curves with an event-based character, i.e., either charging at full power or not charging. A stochastic representation of these profiles would look like a combination of Dirac (delta) functions at zero and full charging power, which are not well representable by a Gaussian mixture with non-zero variances. Third, the charging profiles would change with varying charging rates or when implementing smart charging to shift charging time and alter power consumption levels, this can more easily be integrated in this way.

2.4.1. Data and Model Fit

The load profiles of electric vehicles are based on surveyed mobility data from a Dutch Mobility research [35]. The dataset contains over 30.000 journeys by car, made by over 11,000 individual persons. From this dataset, the arrival and departure times, and distances traveled are extracted and modeled. To implement a (from a network perspective) worst-case situation, we order the data such that all persons are assumed to charge their car battery after the last homebound trip and the total energy needed is based on their cumulative distance traveled that day. Hereby it is assumed that the battery capacity in each car is sufficient for the total distance driven that day. The surveyed and modeled distributions of the departure and arrival times, and the distance traveled are shown in Figure 9.

2.4.2. Model Fitting

The surveyed and modeled distributions of the departure and arrival times, and the distance traveled are shown in Figure 9. The plots again show a good tracking of the marginal stochastic behavior of the different variables, which is further supported by the low Continuous Ranked Probability Score (CRPS) values of 0.11 and 0.08 min, and 10.2 meters for, respectively, departure time, arrival time, and distance traveled.

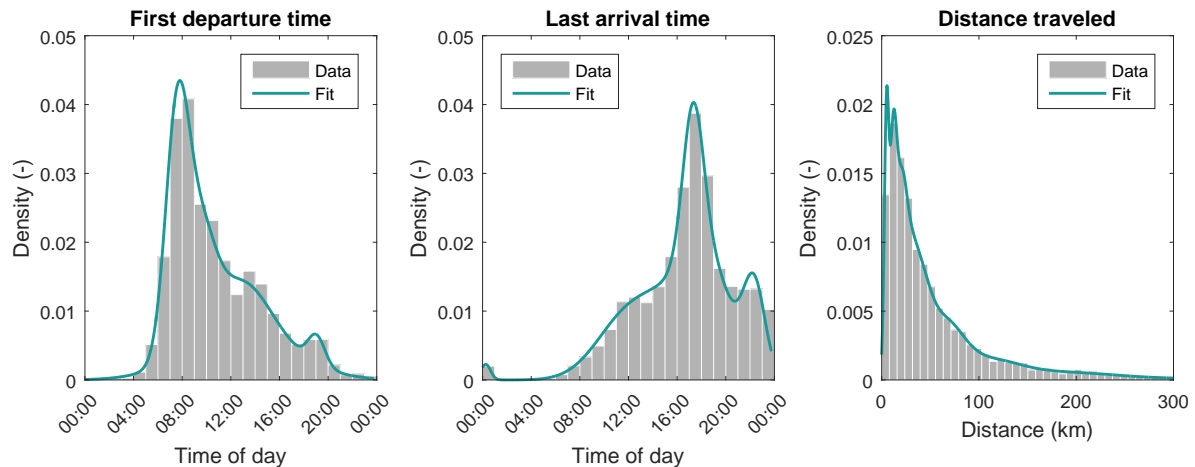


Figure 9. PDFs of data and fitted mixture models for first departure time, last home arrival time, and cumulative distance traveled.

2.4.3. Profile Generation and Evaluation

By sampling from the fitted copula functions and applying the previously described transformation, combinations of departure time, arrival time, and travel distance are generated. Using the average efficiency of currently available electric vehicles, the power consumption while driving is assumed to be 6 km/kWh [36]. Given the available information on traveled distance, the total required energy each trip is calculated. These energies are then converted in profiles by assuming a charge rate of 3.6 kW, which corresponds to the single phase residential charge power. With the final assumption that people start charging their car at full power at the time they arrive at home, representative charging profiles have been generated. Figure 10 shows individual and aggregated profiles for residential EV charging. The average charging duration is roughly 2.5 h, which aligns with the average travel distance of 54 km. The bottom graph shows a comparison of the average simulated charging profiles to the average of measured profiles from a dataset provided by ElaadNL. This dataset contains measurements EV charging at 428 (semi-)public charging stations in the Netherlands for a full year in 2017. From this dataset, the charging stations which are located in residential areas and that did not show charging peak in the morning (to exclude charging when arriving at work) were selected for comparison; furthermore, the set was filtered to include only day profiles with at least one charging session. Overall, the general shape of the profiles is similar; however, the simulated profile has a lower peak, which also occurs a bit earlier in the day compared to the measured profile. The differences may be related to the fact that the measured profiles are from (semi-)public charging stations, and not from actual residential consumer connections, this might cause a different charging behavior at individual stations. For instance, as the charging stations are in the public space multiple cars may connect on a single day, for usually shorter charging duration, and some stations are less regularly used than would be expected for a private residential charging station. Moreover, the driving behavior of the average EV driver in the recent past would not necessarily resemble the the average overall driving behavior of the current Dutch car fleet, e.g., the difference in peak suggests that on average current EV drivers arrive at home a bit later than the overall average driver.

2.5. Heat Pumps

For modeling the electrical loads of heat pumps, measurements from the Your Energy Moment project [37] was used. Because the heating demand is also strongly influenced by the outside temperature, the interdependency between electric load, time of day, and temperature will be modeled to obtain realistic profiles.

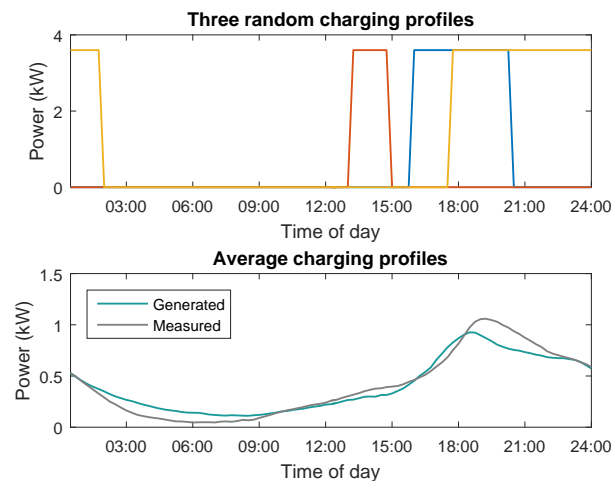


Figure 10. Individual and aggregated profiles for charging of electric vehicles at a charging rate of 3.6 kW.

2.5.1. Data and Preprocessing

The used measurements consist of 15-min intervals of the electric load of 37 heat pumps during one year, along with the temperature at each time interval. The heat pumps used in the pilot project had a two-step compressor with two power levels: 1 kW for regular heating needs and 1.8 kW for tap water. For households with higher or lower heating demand, we assume the installed power of the heat pump will be sized accordingly, thus the overall usage profile remains identical only the actual power consumption changes. The data are again clustered per month, day type, and quarter of an hour of the day.

2.5.2. Model Fitting

In Figure 11, the PDF of the electric load of a heat pump is shown for evening time in winter. Three distinct peaks can be discerned, signifying the three general states of operation of the heat pump: off, room heating/cooling, and tap water heating. During the off-state, a small amount of power is still being used to keep the water flowing through the system. The other two peaks occur around the compressor steps of the heat pump, the variance being caused by the change of operation state during the measured 15-min intervals.

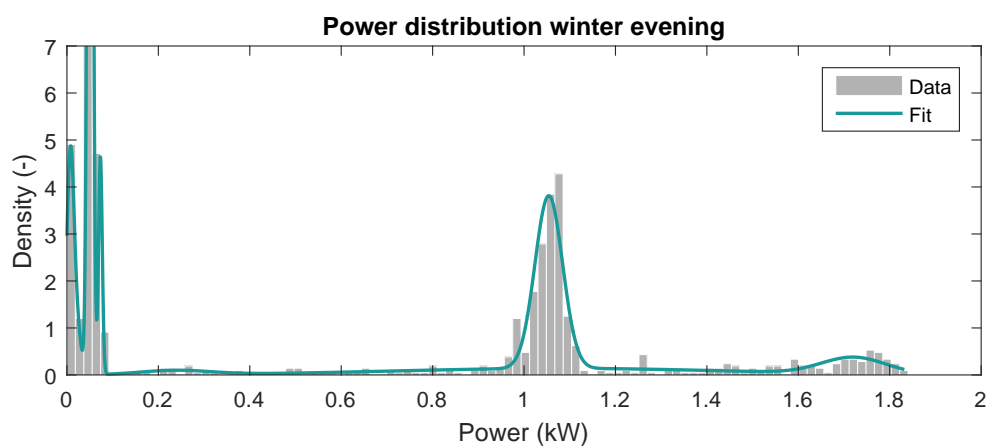


Figure 11. Measured and fitted PDF of heat pump electric load during winter evening, showing three peaks at the levels of power usage for the different operating states.

Once again, copula functions are used to model the correlations present in the dataset. For the heat pump case, besides the autocorrelations of the load itself, also the correlation between load and outside temperature will be taken into account. Figure 12 shows the correlation matrices of the load vs. load (Figure 12a) and the temperature vs. load (Figure 12b). The autocorrelation of the load is characterized by high correlation factors for subsequent time intervals, degrading for time intervals further away. The correlation between load and temperature is negative (higher temperature equals lower load), with the largest correlations in late night and early morning, and late evening time intervals, indicated by the darker shading in the figure. Notably, the correlation of the load with the temperature during the afternoon is much lower, and for time intervals further apart virtually nonexistent. The correlation factors between load and temperature appear relatively low. This is mainly caused by two reasons. First, the fact that only for relatively low temperatures (<5 degrees Celsius) the correlation is decidedly higher. For moderate and high temperatures, the correlation quickly drops. As the correlation plot shows the average correlations over the entire year, the moderate temperatures make up a large portion of the dataset, resulting in a low overall correlation. Second, The electric power use of the heat pumps is partially used for tap water heating. This type of heat demand has a much lower correlation with outdoor temperature and is much more dependent on the time of day.

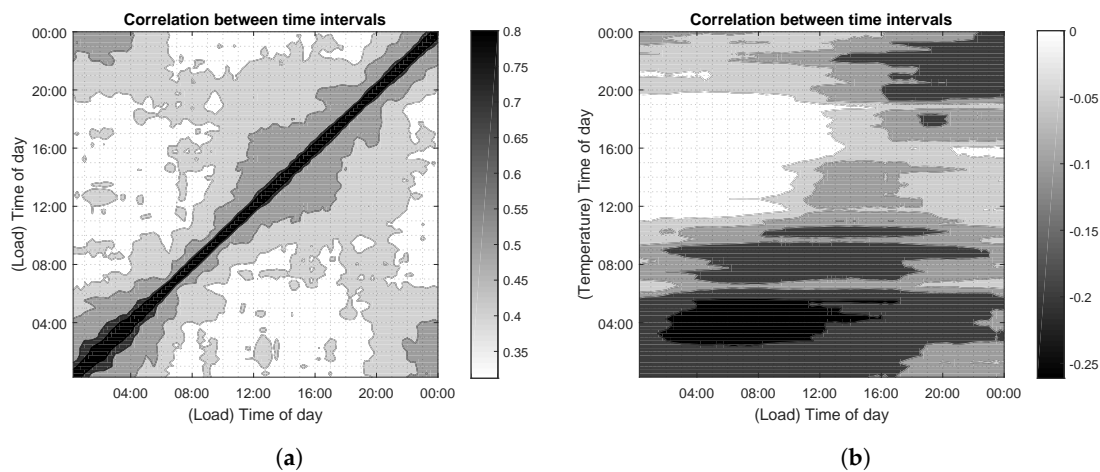


Figure 12. Correlation of power consumption at different time intervals, with darker shading indicating a larger magnitude of the correlation factor. (a) Autocorrelation of electric load and (b) correlation of load with temperature.

2.5.3. Profile Generation and Evaluation

By sampling from the copula functions random load–temperature profiles can be generated. Figure 13 shows the average load and temperature profiles generated with the models compared to the measured averages. The model shows to provide a good representation of the measurements, with an RMSPE of 2.7% for the winter load profile and 5.1% for the summer profile. The average aggregated peak load occurs in the early morning and is approximately 0.65 kW.

To apply these stochastic models for large scale network evaluations, they will be incorporated in a linear network model as described in the next section.

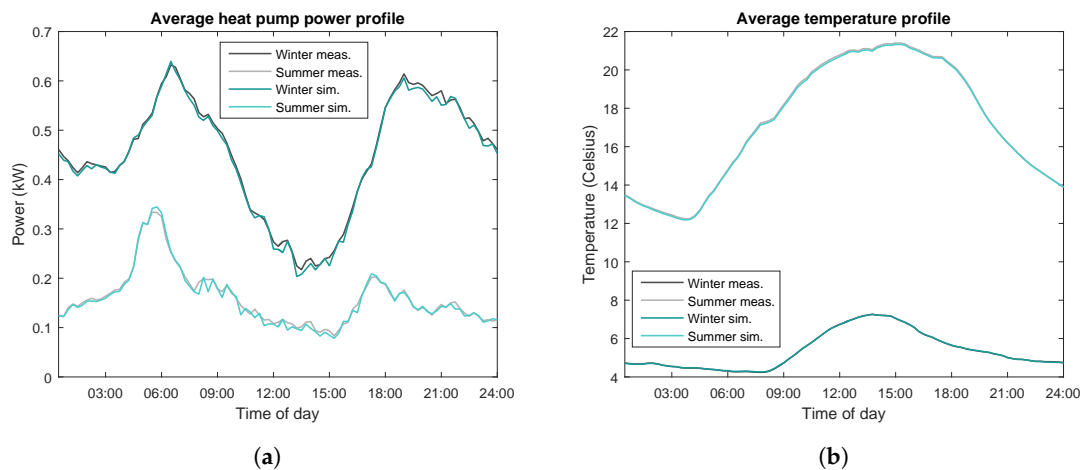


Figure 13. (a) Average heat pump load profile during winter and summer. (b) corresponding average temperature profiles.

3. Evaluating a Large Low Voltage Electricity Network

This section introduces a linear network model which is suitable to evaluate very large low voltage distribution networks [38,39]. Contrary to the models presented in the previous section, the proposed load flow approach is deterministic. Therefore, a Monte Carlo approach is applied to simulate the impact of the stochastic peak loads. The general load flow modeling approach is displayed in Figure 14 and is described more thoroughly in [38,39], where also its accuracy is assessed on a test bed. Additionally, high-efficiency solvers for this load flow formulation are considered in [40,41].

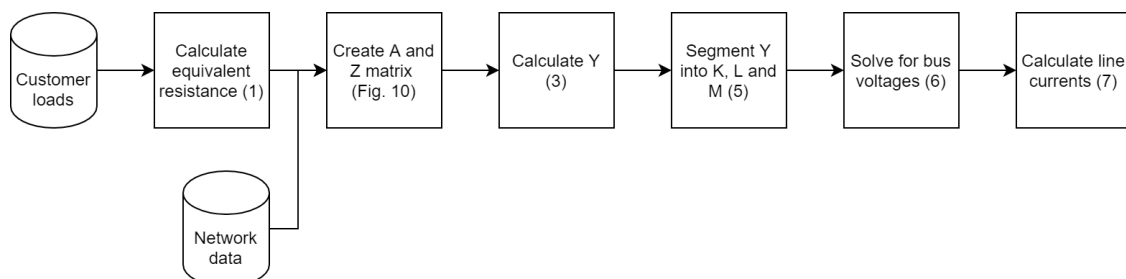


Figure 14. Diagram of the steps for solving the linear load flow problem.

In calculating power flows in distribution networks, it is sufficient to use a steady state electrical model. The most common way to calculate these properties is by using a load flow model [42,43]. The consumer load is traditionally modeled by combining constant power, constant current, and constant impedance load models [43]. The constant power model (i.e., the network power load is independent from the network state) is most commonly used. However, while using a constant power load the load flow model is nonlinear, which makes simulating large networks challenging due to increasing computational time and convergence problems [41].

To create a load flow model which is capable of modeling very large networks, we only use a constant impedance model. This results in a fast linear model, which is still sufficiently accurate [38,40] and over 8 times faster than a nonlinear approach [40].

An example constant impedance network is displayed in Figure 15. This network consists of two impedances which represent the consumer loads and a swing bus which represents the connection to the medium voltage grid. The customers are modeled as an impedance which is directly connected to the ground. No current enters/leaves the network at the cable joints and the influence of the transformer is neglected.

To create the constant impedance model as in Figure 15, all load is converted to an equivalent impedance:

$$Z_{eq} = U_{n,ref}^2 / P_{user} \quad \forall n \in \mathcal{N} \quad (1)$$

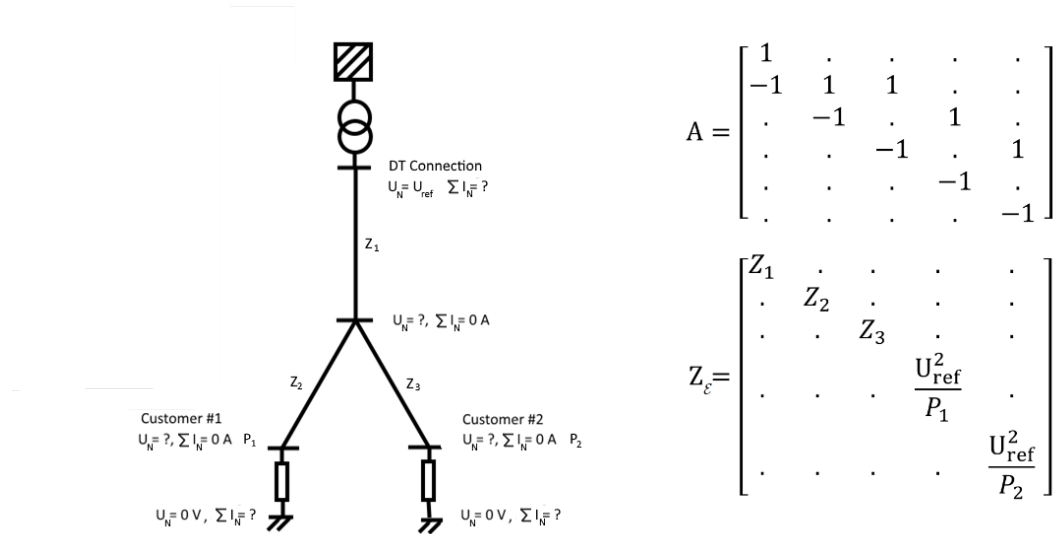


Figure 15. An example low-voltage distribution network (left) and its corresponding A and Z_{ϵ} matrices (right).

In the formula above, Z_{eq} is the equivalent impedance representing the load, P_{user} is the consumer load, n is an index which corresponds to the bus to which the consumer is connected, and $U_{n,ref}$ is a reference voltage at the location of the consumer. Given that the actual voltage at the consumer's location is not exactly known, the reference voltage is set to the nominal voltage. This assumption introduces a small but acceptable error [38,40]. The consumer load was assumed to have a power factor of 1.

The grid is described by a graph. The graph G is defined as $G = (\mathcal{N}, \mathcal{E})$. Here, \mathcal{N} represent the network buses and \mathcal{E} are the network edges/lines. The modeling objective is to calculate the nodal voltages $U_{\mathcal{N}}$ and the cable currents $I_{\mathcal{E}}$.

The starting point is Ohm's law, with which the nodal voltages and bus currents can be determined:

$$I_{\mathcal{N}} = \bar{Y} U_{\mathcal{N}} \quad (2)$$

$I_{\mathcal{N}}$ is the bus current, i.e., the current that enters/leaves the network and \bar{Y} is the "admittance matrix", which can be obtained by [44]

$$\bar{Y} = A Z_{\epsilon}^{-1} A' \quad (3)$$

A is the network directional connection matrix. Here, a matrix row represents a network bus and a column represents a network line. Each line has no more and no less than one starting and one end point. These points are denoted by "1" and -1 , respectively. Because of the quadratic nature of (3), the interchanging the minus sign between buses will result in the same admittance matrix \bar{Y} . Z_{ϵ} is a diagonal matrix with a diagonal containing the inverse of the impedance of each line (\mathcal{E}).

As neither $I_{\mathcal{E}}$ and $U_{\mathcal{N}}$ are completely known, solving (2) requires some extra steps. One way to efficiently solve this problem is to segment the matrices and solving two separate subproblems. If one sorts $I_{\mathcal{N}}$, \bar{Y} , and $U_{\mathcal{N}}$ so that all swing and ground buses are $\in U_1$, then the problem can be written as

$$I_{\mathcal{N}} = \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}, \bar{Y} = \begin{bmatrix} K & L \\ L' & M \end{bmatrix}, U_{\mathcal{N}} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \quad (4)$$

At regular network buses, I_2 is equal to $\bar{0}$ as no currents enter/leave the network. U_1 , which contains all the voltages at the swing/ground buses, is completely known. This results in the following load flow problem,

$$\begin{bmatrix} I_1 \\ \bar{0} \end{bmatrix} = \begin{bmatrix} K & L \\ L' & M \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \quad (5)$$

A practical way to obtain U_2 is

$$MU_2 = -L'U_1 \quad (6)$$

This equation can be defined as the standard matrix equation $Ax = B$, which can be solved by many of-the-shelf solvers. This formulation is efficient enough to be solve networks with millions of cables as will be demonstrated in Section 4. It is possible to increase the load flow speed with the numerical techniques described in [40].

Once all voltages are known, the cable currents can be obtained using

$$I_{\mathcal{C}} = Z_E A' U_N \quad (7)$$

A major advantage of the described linear approach is that it is susceptible to convergence problems which can occur by using nonlinear models [43,45]. These unfeasible solutions tend to happen more often in large networks with high loads. Given that network matrices with over 22 million cables always contain some errors, and loads increase due to the energy transition, applying a linear model is a logical choice.

4. Case Study: Congestion in a Large Real-World Low-Voltage Power Grid

The objective of this case study is to find the number of overloaded cables and voltage problems in a large LV grid, using the stochastic peak loads generated from the copula models. An diagram giving an overview of the case study is presented in Figure 16. An overload problem is defined as a peak load current which is higher than the rated capacity of the corresponding cable. A voltage problem is defined as a voltage drop of more than 4.5% of the nominal voltage on the LV network, as in accordance with DNO policy.

For simulating the impact of new energy technologies on the distribution network, the network of the Dutch DNO Alliander was studied. Alliander DNO is the largest DNO of the Netherlands serving three million consumers. It is responsible for maintaining the Medium Voltage and Low-Voltage distribution network. The LV network which is considered in this paper has more than 80,000 km of cables and over 38,000 MV/LV transformers. The network spans roughly one-third of the Dutch distribution network and is described with 22 million cable segments. The low-voltage network operates on 400 V/230 V and is modeled as a single-phase balanced network.

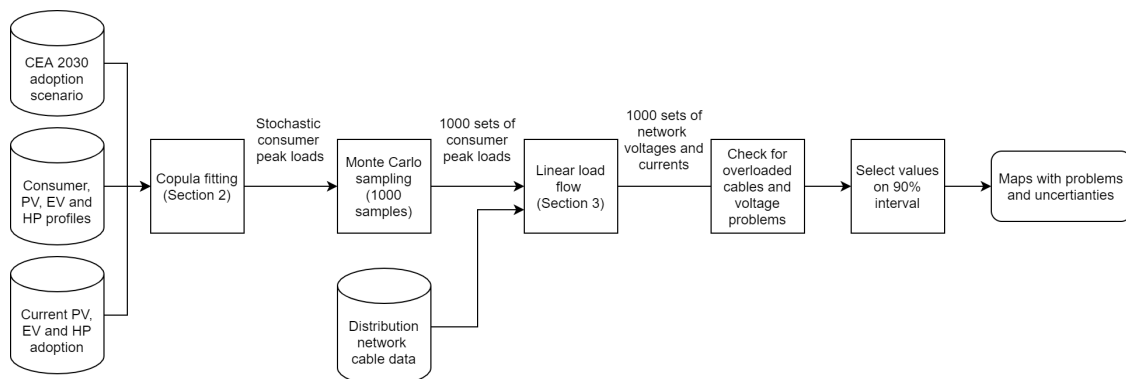


Figure 16. Diagram giving an overview of the steps performed in the case study.

The Dutch power grid is very similar to many other European/US grids, but has a few distinguishing properties, which make this work more relevant. In contrast to most other countries, almost all electricity cables in the Netherlands are underground. This is mainly motivated by the fact that the Netherlands generally has a soft, non-rocky soil. While it greatly increases the network's resistance to stormy weather, it has as downside that it is very expensive to replace cables, making the planning methods of this paper indispensable.

Furthermore, as the Netherlands is very densely populated, the low-voltage part of the power grid is very large. An average MV/LV transformer has over 100 consumer connections, a 120–630 kVA power rating, and over 2.5 km over cable. This is a major difference to, for example, a rural US network, where a MV/LV transformer often serves less than five consumers has a power rating of less than 50 kVA and a few hundreds of meter cable. Because of the vast size of the studied LV grid, this case study is of great value for DNOs which operate in densely populated areas.

4.1. General Approach and Scenario Description

The scenario used in this study is the one provided by the Climate and Energy Agreement (CEA) [46], which is widely accepted as very plausible and is basis for the energy policy of the Dutch government.

The year of 2030 was selected for which the CEA provides adoption numbers for the PV, EV, and HP technologies. According to the CEA scenario, in 2030 the Alliander area contains 4.9 GW of extra solar power generation, 1.2 million extra EVs, and 2.8 million HPs compared to the current situation. This means that 40% of all low voltage electricity demand will be produced by local solar power, 39% of all cars will be an EV and 37% of the consumers will have a HP. The current adaptation of PV, EV, and HP is 10%, 3%, and below 3%, respectively.

In accordance to the diagram in Figure 16, the numbers of the CEA scenario are applied to the copula models described in Section 2 resulting in stochastic load profiles for every consumer. These load profiles were sampled using a Monte Carlo approach, resulting in a set of customer peak loads. Each set of peak loads was evaluated using the linear load flow model from Section 4, following the diagram presented in Figure 14.

4.2. Simulation Results

The copula model was implemented Matlab and the load flow model was implemented in the R programming language. By using sparse matrices and a QR decomposition method calling BLAS/LAPACK routines, the equations of the LV grid with 22 million cables were solved within 30 s using a single processor core of an Intel i7-6820HQ processor.

A thousand Monte Carlo iterations have been performed on the stochastic consumer peak loads resulting from the models from Section 2. For every iteration, all the currents and voltages in the low voltage network have been calculated. The cable currents were compared to the rated power of each cable and the voltages were compared to the allowed voltage range. This resulted in a thousand sets with voltage and current problems and their location.

The total number of overloaded cables and voltage problems did not differ much between scenarios. The number of overloaded cables was found to be 3–4% of the total number of LV cables. This may seem like a low percentage, but as the LV of Alliander network spans over 80,000 km of underground cable, the cost implications are significant.

The number of voltage problems was calculated to be 10–12%. This is a large number and voltage problems are generally cost-intensive to solve. However, while this predicts a large future problem from a policy view (i.e., the voltages are outside the allowed design specifications), it will likely not be noticed by most consumers. Most household appliances work well within a larger voltage range than the specified 4.5% difference from the nominal voltage, with PV inverters being the exception. It remains to be seen how much cable replacement will be driven by voltage problems.

Figure 17 shows the geographical distribution of the load and the voltage and current problems. It can be observed that some areas in the Netherlands are expected to have much more voltage

problems than others. The west of the Netherlands has the oldest network and is the most densely populated. It is not surprising that most voltage problems are expected in this area. Because Monte Carlo methods can produce outliers, the top and bottom 5% of the voltages have been filtered from the results, creating a confidence interval of 90%.

In an attempt to show the local uncertainty between Monte Carlo scenarios, Figure 18 shows the difference between the highest and lowest voltage within the 90% confidence interval. Darker areas have more variability in the end voltage of the consumers on average. While this is an imperfect measure of the variability between voltages, because not every area has the same number of consumers, it gives an indication in which areas the effect of the new technologies is less certain. This can be used by the DNO to focus their data collection.

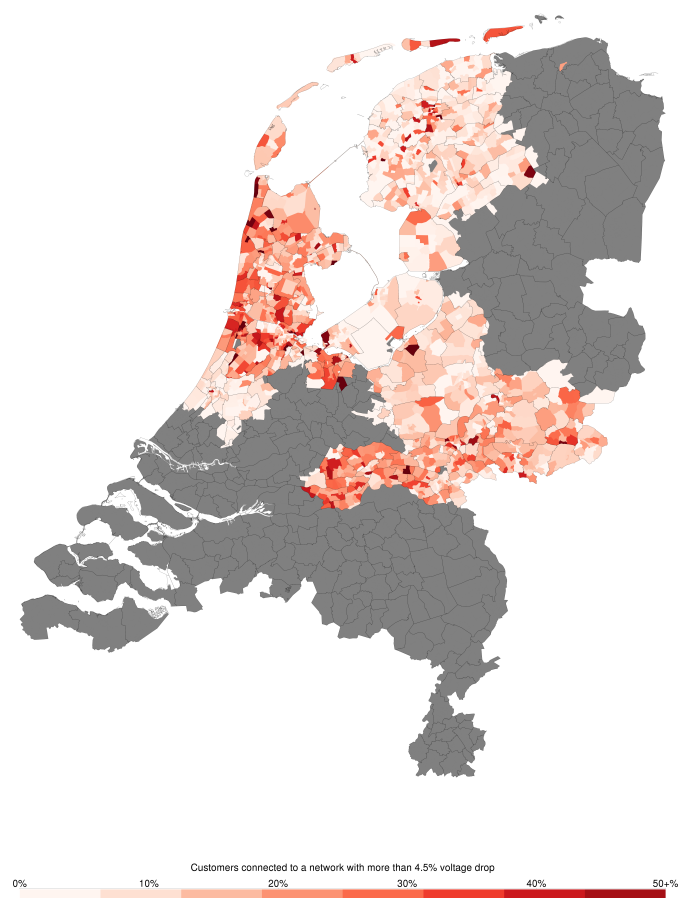


Figure 17. Geographic distribution of voltage problems in 2030 within 90% confidence. The gray area is The Netherlands in which Alliander operates the colored area. The total number of consumers with voltage problems in this figure is 11%.

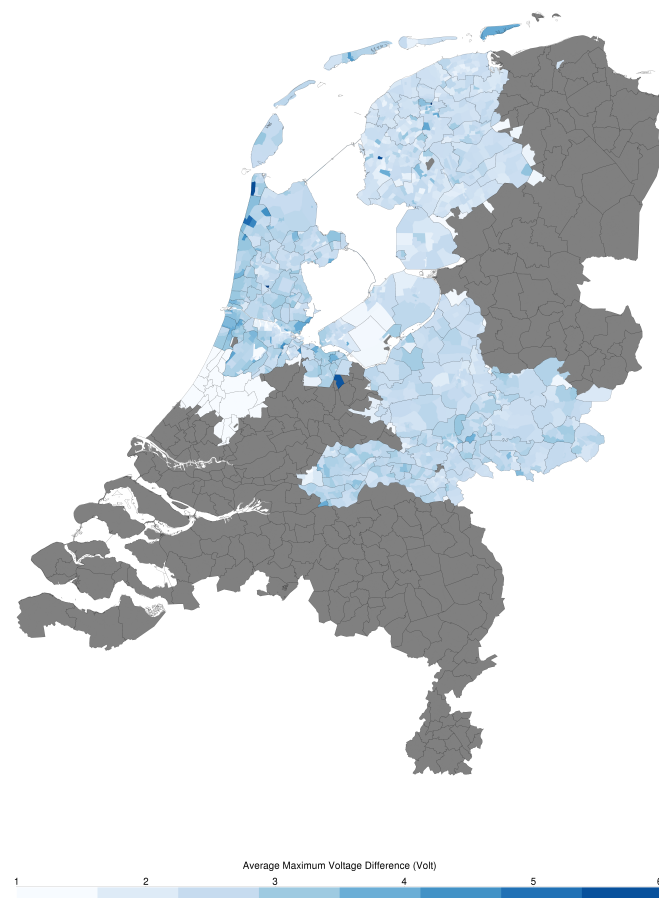


Figure 18. The difference between the highest and lowest voltage within the 90% confidence interval. It is a measure for the local variability between scenarios.

5. Conclusions

This paper presented a general framework for modeling the effects of energy transition on the distribution network. Its potential is demonstrated by applying it to a real-world network.

The proposed modeling approach using Gaussian mixture models and copula functions is shown to be both accurate and flexible, and is therefore an effective method to represent the stochastic behavior and correlations of a variety of variables.

The advantages of this methodology over existing methods are that multiple different technologies or datasets can be modeled using the same approach with minor variations, and a low modeling intensity is required given an adequate dataset. This also allows the methodology to be automated, for ease of use in practical network planning situations. Furthermore, due to the parametric nature of the used models, they can be scaled and altered to be adapted to local variations if required. Last, the separate stochastic representation of individual technologies allows for a risk based analysis of the change in electric load under various future scenarios. A limitation of the load modeling approach is the heavy reliance on available (measurement) data and that the relatively smooth copula functions are limited in describing very volatile correlation structures with.

A linear approach to modeling large low voltage networks has been described and applied to a very large network of Alliander DNO. The model was sufficiently fast to evaluate a the network and provided valuable insights in the impact of the new technology on the grid. Due to the higher

electricity demand, 3–4% of the cable segments were overloaded and 10–12% consumers are expected to have voltage problems based on the current grid.

Given the size of the studied LV network, the number of problems is expected to pose a serious challenge for the DNO in the near future. However, using the geographical overviews from this paper, the DNO has a valuable tool as to where to focus their efforts.

It is the authors' ambition to expand the modeled network to other network voltage levels (medium and high voltage) and create an integral model for the entire Dutch electricity network. The theory presented in this paper makes it possible to predict problems in the distribution network, which help in maintaining and securing the distribution grid for the next decades.

Author Contributions: All the authors contributed equally in all aspects of this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Rijksdienst voor Ondernemend Nederland | TKI Urban Energy.

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

References

1. U.S. Energy Information Administration. *International Energy Outlook 2016*; U.S. Energy Information Administration: Washington, DC, USA, 2016; Volume 0484, pp. 1–2.
2. Veldman, E.; Gibescu, M.; Slootweg, H.J.G.; Kling, W.L. Scenario-based modelling of future residential electricity demands and assessing their impact on distribution grids. *Energy Policy* **2013**, *56*, 233–247. [\[CrossRef\]](#)
3. Velander, S. Methods of operational analysis applied to distribution of electric power. *Tek. Tidskrif* **1952**, *82*, 293–299.
4. Navarro-Espinosa, A.; Ochoa, L.F. Probabilistic Impact Assessment of Low Carbon Technologies in LV Distribution Systems. *IEEE Trans. Power Syst.* **2016**, *31*, 2192–2203. [\[CrossRef\]](#)
5. Dickert, J.; Schegner, P. A Time Series Probabilistic Synthetic Load Curve Model for RESIDENTIAL Customers. In Proceedings of the 2011 IEEE PES PowerTech, Trondheim, Norway, 19–23 June 2011; pp. 1–6. [\[CrossRef\]](#)
6. Collin, A.J.; Tsagarakis, G.; Kiprakis, A.E. Development of Low-Voltage Load Models for the Residential Load Sector. *IEEE Trans. Power Syst.* **2014**, *29*, 2180–2188. [\[CrossRef\]](#)
7. Grandjean, A.; Adnot, J.; Binet, G. A review and an analysis of the residential electric load curve models. *Renew. Sustain. Energy Rev.* **2012**, *16*, 6539–6565. [\[CrossRef\]](#)
8. Labeeuw, W.; Deconinck, G. Residential electrical load model based on mixture model clustering and markov models. *IEEE Trans. Ind. Inform.* **2013**, *9*, 1561–1569. [\[CrossRef\]](#)
9. Evangelopoulos, V.; Karafotis, P.; Georgilakis, P. Probabilistic Spatial Load Forecasting Based on Hierarchical Trending Method. *Energies* **2020**, *13*, 4643. [\[CrossRef\]](#)
10. Navarro-Espinosa, A.; Mancarella, P. Probabilistic modeling and assessment of the impact of electric heat pumps on low voltage distribution networks. *Appl. Energy* **2014**, *127*, 249–266. [\[CrossRef\]](#)
11. Veldman, E.; Verzijlbergh, R.A. Distribution grid impacts of smart electric vehicle charging from different perspectives. *IEEE Trans. Smart Grid* **2015**, *6*, 333–342. [\[CrossRef\]](#)
12. Lo Franco, F.; Ricco, M.; Mandrioli, R.; Grandi, G. Electric Vehicle Aggregate Power Flow Prediction and Smart Charging System for Distributed Renewable Energy Self-Consumption Optimization. *Energies* **2020**, *13*, 5003. [\[CrossRef\]](#)
13. Lojowska, A.; Kurowicka, D.; Papaefthymiou, G.; Van Der Sluis, L. Stochastic modeling of power demand due to EVs using copula. *IEEE Trans. Power Syst.* **2012**, *27*, 1960–1968. [\[CrossRef\]](#)
14. Pashajavid, E.; Aliakbar Golkar, M. Copula-based multivariate stochastic modeling of load demand due to plug-in electric vehicles in order to be integrated in distribution system planning. In Proceedings of the CIRED 22nd International Conference on Electricity Distribution, Stockholm, Sweden, 10–13 June 2013; Volume 0333, pp. 10–13.
15. Widén, J.; Wäckelgård, E. A high-resolution stochastic model of domestic activity patterns and electricity demand. *Appl. Energy* **2010**, *87*, 1880–1892. [\[CrossRef\]](#)

16. Mu, Y.; Wu, J.; Jenkins, N.; Jia, H.; Wang, C. A Spatial-Temporal model for grid impact analysis of plug-in electric vehicles. *Appl. Energy* **2014**, *114*, 456–465. [\[CrossRef\]](#)
17. Chen, S.; Zhang, H.; Guan, J.; Rao, Z. Agent-based modeling and simulation of stochastic heat pump usage behavior in residential communities. In *Building Simulation*; Springer: Berlin, Germany, 2020; pp. 1–19.
18. Nelsen, R.B. *An Introduction to Copulas*; Springer: Berlin, Germany, 2013; Volume 53, p. 276. [\[CrossRef\]](#)
19. Wang, H.; Zou, B. Probabilistic computational model for correlated wind speed, solar irradiation, and load using Bayesian network. *IEEE Access* **2020**, *8*, 51653–51663. [\[CrossRef\]](#)
20. Philippe, W.P.J.; Zhang, S.; Eftekharij, S.; Ghosh, P.K.; Varshney, P.K. Mixed Copula-Based Uncertainty Modeling of Hourly Wind Farm Production for Power System Operational Planning Studies. *IEEE Access* **2020**, *8*, 138569–138583. [\[CrossRef\]](#)
21. Cui, M.; Krishnan, V.; Hodge, B.M.; Zhang, J. A copula-based conditional probabilistic forecast model for wind power ramps. *IEEE Trans. Smart Grid* **2018**, *10*, 3870–3882. [\[CrossRef\]](#)
22. Bina, M.T.; Ahmadi, D. Aggregate domestic demand modelling for the next day direct load control applications. *IET Gener. Transm. Distrib.* **2014**, *8*, 1306–1317. [\[CrossRef\]](#)
23. Zhao, Y.; Liu, Q.; Kuang, J.; Xie, K.; Du, W. Modeling multivariate dependence by nonparametric pair-copula construction in composite system reliability evaluation. *Int. J. Electr. Power Energy Syst.* **2020**, *124*, 106373. [\[CrossRef\]](#)
24. Bernards, R.; Morren, J.; Slootweg, J.G. Statistical Modelling of Load Profiles Incorporating Correlations using Copula. In Proceedings of the 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), Torino, Italy, 26–29 September 2017.
25. Papaefthymiou, G.; Kurowicka, D. Using copulas for modeling stochastic dependence in power system uncertainty analysis. *IEEE Trans. Power Syst.* **2009**, *24*, 40–49. [\[CrossRef\]](#)
26. Zamo, M.; Naveau, P. Estimation of the continuous ranked probability score with limited information and applications to ensemble weather forecasts. *Math. Geosci.* **2018**, *50*, 209–234. [\[CrossRef\]](#)
27. Göçken, M.; Özçalıcı, M.; Boru, A.; Dosdoğru, A.T. Integrating metaheuristics and artificial neural networks for improved stock price prediction. *Expert Syst. Appl.* **2016**, *44*, 320–331. [\[CrossRef\]](#)
28. Liander N.V. *Smart Meter Data*; Liander N.V., Duiven, The Netherlands, 2013. Available online: <https://www.liander.nl/over-liander/innovatie/open-data/data> (accessed on 9 October 2020).
29. Stephen, B.; Mutanen, A.J.; Galloway, S.; Burt, G.; Jarventausta, P. Enhanced load profiling for residential network customers. *IEEE Trans. Power Deliv.* **2014**, *29*, 88–96. [\[CrossRef\]](#)
30. Ciapessoni, E.; Cirio, D.; Pitto, A.; Massucco, S.; Silvestro, F. A novel approach to account for uncertainty and correlations in probabilistic power flow. In Proceedings of the IEEE PES Innovative Smart Grid Technologies Conference Europe, Istanbul, Turkey, 12–15 October 2015; pp. 1–6. [\[CrossRef\]](#)
31. Ye, G.; Nijhuis, M.; Cuk, V.; Cobben, J. Stochastic Residential Harmonic Source Modeling for Grid Impact Studies. *Energies* **2017**, *10*, 372. [\[CrossRef\]](#)
32. KNMI. *Daggegevens Van Het Weer in Nederland*; KNMI: De Bilt, The Netherlands, 2017.
33. Benganem, M. Optimization of tilt angle for solar panel: Case study for Madinah, Saudi Arabia. *Appl. Energy* **2011**, *88*, 1427–1433. [\[CrossRef\]](#)
34. Stine, W.; Harrigan, H. *Power from the Sun*; John Wiley & Sons: Hoboken, NJ, USA, 1986.
35. CBS. *Onderzoek Verplaatsingen in Nederland 2014*; Centraal Bureau voor de Statistiek: Den Haag, The Netherlands, 2015. [\[CrossRef\]](#)
36. Office of Energy Efficiency and Renewable Energy. *Fuel Economy Guide 2017*; Technical Report; U.S. Department of Energy: Washington, DC, USA, 2017.
37. Klaassen, E.A.M.; Asare-Bediako, B.; De Koning, C.P.; Frunt, J.; Slootweg, J.G. Assessment of an algorithm to utilize heat pump flexibility-theory and practice. In Proceedings of the 2015 IEEE Eindhoven PowerTech, Eindhoven, The Netherlands, 29 June–2 July 2015. [\[CrossRef\]](#)
38. Van Westering, W.; Hellendoorn, H. Low voltage power grid congestion reduction using a community battery: Design principles, control and experimental validation. *Int. J. Electr. Power Energy Syst.* **2020**, *114*, 105349. [\[CrossRef\]](#)
39. van Westering, W.; Droste, B.; Hellendoorn, H. Combined Medium Voltage and Low Voltage Simulation to Accurately Determine the Location of Voltage Problems in Large Grids. In Proceedings of the 25th International Conference on Electricity Distribution, Madrid, Spain, 3–6 June 2019.

40. Sereeter, B.; van Westering, W.; Vuik, C.; Witteveen, C. Linear power flow method improved with numerical analysis techniques applied to a very large network. *Energies* **2019**, *12*, 4078. [[CrossRef](#)]
41. Sereeter, B.; Vuik, C.; Witteveen, C. Mathematical Formulations and Algorithms for Fast and Robust Power System Simulations. Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 2020.
42. Le Nguyen, H. Newton-Raphson method in complex form [power system load flow analysis]. *IEEE Trans. Power Syst.* **1997**, *12*, 1355–1359. [[CrossRef](#)]
43. Kersting, W.H. *Distribution System Modeling and Analysis*; CRC Press: Boca Raton, FL, USA, 2001.
44. Kirtley, J. *6.061 Introduction to Power Systems Class Notes Chapter 5 Introduction to Load Flow*; MIT Open Courseware: Cambridge, MA, USA, 2018.
45. Sarmiento, J.E.; Alvez, C.A.; de Nadai, N.B.; Zambroni de Souza, A.C.; Carreno, E.M.; Ribeiro, P.F. A Complex-Valued Three-Phase Load Flow for Radial Networks: High-Performance and Low-Voltage Solution Capability. *IEEE Trans. Power Syst.* **2019**, *34*, 3241–3249. [[CrossRef](#)]
46. Hekkenberg, M.; Koelemeijer, R. *Analyse van Het Voorstel Voor Hoofdlijnen van Het Klimaatakkoord*; PBL Netherlands Environmental Assessment Agency: Den Haag, The Netherlands, 2018.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).