CHALMERS
UNIVERSITY OF TECHNOLOGY

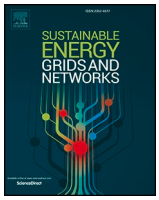
An open data-based model for generating a synthetic low-voltage grid to estimate hosting capacity

Sustainable Energy, Grids and Networks 39 (2024) 101483

Contents lists available at ScienceDirect

Sustainable Energy, Grids and Networks

journal homepage: www.elsevier.com/locate/segan



An open data-based model for generating a synthetic low-voltage grid to estimate hosting capacity

Therese Lundblad *, Maria Taljegard, Niclas Mattsson, Elias Hartvigsson, Filip Johnsson

Division of Energy Technology, Chalmers University of Technology, Gothenburg 412 96, Sweden

ARTICLE INFO

Keywords:
Synthetic grid representation
Hosting capacity
Electric vehicles
Home charging
Low-voltage grid
Electricity grid

ABSTRACT

This study develops and applies an open data-based reference electricity grid analysis (REGAL) model designed to create a synthetic representation of a low-voltage (LV) grid for a country-size geographic area. The model enables large-scale grid simulation in which new loads, such as electric vehicle charging, can be added to estimate their impacts on the current LV grid. The modeling is carried out in three steps: (1) generation of a synthetic LV grid; (2) addition of residential loads, including electric vehicle charging; and (3) evaluating if the grid capacity is exceeded. The grid is generated by selecting transformers and cables so that the system can fulfill the current demand while meeting national regulations and standards for distribution grids, all at the lowest total cost. This paper presents the results of calibration and validation against real-world data for the predicted electricity demands and synthetic grid generated by the model. Different calibration values were explored, and the accuracy of the estimations of grid capacities was calibrated using proprietary real-world data from grid operators. For a region with multiple grid cells, an average deviation from real-world data of ±10 % was achieved. For an average area of 1 km^2, the error was 44.5 %, which means that the model is not suitable for analysis on this geographic level. However, the level of accuracy is deemed sufficient for initial estimations of hosting capacity for larger geographic areas, such as a region or a country, thereby enabling estimations of hosting capacity in new areas that lack publicly accessible grid capacities.

1. Introduction

Traditionally, distribution grids were built to distribute electricity to consumers with a more or less inflexible load. As more sectors are electrified, including the transport sector, and electricity production becomes increasingly decentralized and more volatile (from increasing shares of wind and solar power), greater stress is exerted on the local electricity grid and on its operational ability to connect production and demand in both space and time. A large share of the new loads imposed on the low-voltage (LV) electricity grid is linked to the electrification of passenger cars, in the form of charging of electric vehicles (EVs). There have been several studies on passenger EVs and their impacts on electricity systems (see reviews by Nazari-Heris et al. [1], Kumar et al. [2], and Nour et al. [3]). For example, Hedegaard et al. [4] have analyzed how large-scale implementation of EVs would influence the North European electricity systems. They have shown that enabling intelligent EV charging and discharging can support variable renewable electricity production, although they conclude that the effects of EVs on electricity systems will vary between different countries. The impacts of EV charging on electricity generation and storage have been investigated using energy systems modeling by Taljegard et al. [5], reaching the conclusion that controlled charging of EVs can reduce investments in peak power capacity, reduce the need for storage technologies, and stimulate increased shares of solar and wind power generation. However, few studies have included the capacities of distribution grids and how those grids are affected by new loads from EV charging and local electricity generation, such as through photovoltaics (PV). The flexibility in EV charging included in this type of study could influence the needed grid capacity, but not evaluate it.

Reviews of the methods and tools for estimating the numbers of EVs and amounts of solar PV that can be hosted by the capacities of current LV grids – hereinafter referred to as *hosting capacity* – have been presented by Umoh et al. [6] and Carmelito and Filho [7]. An additional review of the calculations for hosting capacity has been presented by

Abbreviations: APT, Apartment; DeSA, Demographic statistical area; LV, Low-voltage; EV, Electric vehicle; MV, Medium voltage; PV, Photovoltaic; SFD, Single-family dwelling.

* Corresponding author.
E-mail address: therese.lundblad@chalmers.se (T. Lundblad).

https://doi.org/10.1016/j.segan.2024.101483
Received 9 April 2024; Received in revised form 7 June 2024; Accepted 14 July 2024
Available online 24 July 2024

Abideen et al. [8]. All of these publications describe methodologies for the evaluation of hosting capacity and their strengths and weaknesses. The reviews have characterized previous studies based on the types of methodology for predicting how new loads influence electricity grids. Three types of methodologies for adding new loads have been identified: deterministic simulations; stochastic methodologies; and time-series methods [7]. Deterministic simulations use fixed values for the loads, typically the maximum loads, to generate a single value for the hosting capacity. Thus, deterministic simulations lack the possibility to consider uncertainty but have the benefits of simplicity, low data requirements, and a low demand for computational power, as compared to other methods [6,9]. Uncertainty has been addressed with stochastic methods, usually through Monte Carlo simulations, which consider the likelihood of uncertain variables and run a multitude of scenarios to determine how the uncertain variables affect the outcome. Drawbacks of this method are that the knowledge as to how the loads coincide in time is lost, and multiple uncertainties can lead to an unfeasible number of scenarios [9]. Carmelito and Filho [7] proposed and used a methodology that combines deterministic and stochastic methods to study EV charging, however, that does not allow for temporal resolution of the results. Time-series analysis uses actual measurements or generated data with a certain temporal resolution to identify when problems are likely to occur in the grid. However, this requires large amounts of input data and typically assumes that past data can also be representative of scenarios with the introduction of future loads, electricity generation, and storage technologies [6]. Luthander [10] has shown that when studying the hosting capacity of solar PV with a time-series analysis, the timestep chosen in the analysis can influence the accuracy of the results. Mulenga et al. [9] have pointed out that when evaluating the hosting capacity of solar PV, time-series analyses and stochastic analyses are more-relevant than deterministic studies due to the intermittency of PV generation. Since EV loads are also intermittent, albeit with partially predictable diurnal patterns, one could argue that this should also be true for EVs.

In the reviews of Umoh et al. [6] and Mulenga et al. [9], the evaluation criteria for the determination of hosting capacity have been addressed. Both studies have concluded that most studies of hosting capacities have used the same three performance indices: voltage magnitude; line or cable loading; and transformer loading. Some studies have also included other evaluation criteria, such as power loss and voltage unbalance [11,12,13,14].

Veldman et al. [15] have studied the impacts on distribution grids from the smart charging of EVs in an electricity grid of known topology in The Netherlands. The area that they cover in the study delivers electricity to 920,000 residential customers. The studied grid includes the medium voltage (MV) level (transformers and cables), as well as the transformers that convert between the MV and LV levels [15]. Charging profiles were achieved by running the driving data for conventional vehicles in an optimization model in which the introduced EV load can impact electricity prices. In doing so, they have assumed that vehicles can charge when not driving [15]. However, since only the transformer of the LV grid is considered, a simulation of the LV grid is not possible and other limiting factors in the LV grid can, therefore, not be investigated.

Luthander et al. [16] performed a study in which PVs and EV charging were added to a power system that included both the MV and LV grids in Sweden, covering an area with 5174 customers. They used actual grid capacities that were retrieved from the grid operator in the area. They studied voltage, current, and power fluctuations for two time periods of 2 weeks each, one in summertime and one in wintertime. For each time period, they added future PV generation and EV loads to the studied grid; the results were obtained separately for the winter and summer weeks as averages over a single day with a 1-hour time resolution [16]. How often the operational limits of the grid were exceeded could not be studied, as only a 2-week period was studied for each case. Furthermore, no information was provided as to how applicable the results might be to electricity grids other than the one studied.

The three previously described methods for the evaluation of hosting capacity assume that the capacity of the investigated grid is known. If the real grid capacities are unknown, an estimation of current grid capacities is needed. Today, it is often difficult to have access to grid capacities over a larger geographic area, since the operation of electricity grids, especially distribution grids, is typically split between many different actors, and national security restrictions limit access to the data. One option is to gain access to grid capacities for a smaller area, so as to perform a case study and subsequently extrapolate the results to a larger area [16]. Due to the lack of real grid capacities, some studies have used typical grids and extrapolates the results to larger areas [15]. An alternative method is to develop a synthetic grid based on data regarding the number of customers in an area and the grid design principles, and thereby analyze the hosting capacities for EVs and solar PVs [17,18,19].

Amme et al. [17] have combined local grid planning principles and GIS data, considering line congestion and voltage limitations, to generate a plausible, distributed representation of all the MV grids for Germany. Using this methodology, a deviation of approximately 10 % from real network data was achieved [17]. However, those authors did not consider the LV grid and its operational capacities.

Zhu et al. [12] have assessed the EV hosting capacities of two Australian MV and LV grids using stochastic time-series analysis. Zhu et al. [12] built the analysis on a model that was first presented by Nacmanson et al. [19]. They used the known capacities of the MV grid network and approximated the LV grid capacities based on data, which included the number of customers per transformer and the local design principles used for grid development in the area, such that they derived a synthetic LV grid representation [19]. The loads were simulated using a stochastic time-series analysis run with a 1-minute resolution over a 24-hour day [12]. The household load profiles with 1-minute resolution were created by interpolating data with a 30-minute time resolution [12]. They found that the hosting capacity needs to be estimated for different types of networks, and concluded that different components can be limiting factors [12]. No estimation was performed on a larger regional level (e.g., encompassing several MV grids), and no data on how accurate the estimations of LV grid capacities were presented in that study.

The abovementioned studies have provided valuable insights into how interactions between residential and EV loads can be modeled to evaluate using several methods the hosting capacities for EVs in the current LV grid. However, the methods used in those previous studies have the limitation that they are not able to consider the LV grids of larger geographic regions. Therefore, this work aims to create and validate a methodology for generating a synthetic LV grid for a large geographic area, such as a country, and then use this method to add future EV charging to residential loads. The model uses open data as inputs such as population density, dwelling distribution, and rules for grid design to generate an LV grid for Sweden. These types of data are usually more accessible than grid capacities. The developed method also suggests how to simulate the imposition of new loads on the generated grids in a way that takes into account the uncertainty as to where the loads might appear, as well as how to discern whether the operational limits of the LV grid have been exceeded. The study also presents the results of calibration and validation tests against real-world data for the predicted electricity demands and the synthetic grid generated by the model, and suggests calibration values that can be used for Sweden.

2. Method

In this study, a reference electricity grid analysis (REGAL) model is developed to generate a synthetic representation of the LV grid in Sweden based on population data and power system calculations. The REGAL model can be used to estimate hosting capacities in areas with unknown grid capacities. The present study builds on the LV grid model presented by Hartvigsson et al. [18], albeit with modifications designed

to serve new purposes and increase the level of accuracy when analyzing new research questions, including those related to hosting capacities for EVs. Examples of substantial changes include allowing for a high temporal resolution and analysis of EV charging using different charging strategies. Furthermore, the new version of the model uses several new datasets as inputs and has been validated and calibrated against large datasets of real-world grid data. This section is divided into two parts: (i) a description of the REGAL model and the different modeling steps (Section 2.1); and (ii) a description of the geographic input dataset that is used to run the REGAL model (Section 2.2). The description of the modeling in Section 2.1 also includes an outline of the method and datasets used for validation of the generated synthetic grid (Section 2.1.4).

2.1. The REGAL model

The REGAL model procedure is depicted in Fig. 1 and consists of the following modeling blocks: (1) generation of a synthetic least-cost LV grid with high geographic resolution that complies with national regulations; (2) grid simulation when adding EV charging to the residential load; and (3) evaluation of power system violations. Sections 2.1.1 to 2.1.3 describe these three modeling blocks in detail. The main input data to the different modeling blocks are indicated in white boxes in Fig. 1. Technical and economic input data are presented continuously through the model description and in Appendix A1, while the geographic input data are presented in detail in Section 2.2. The LV grid model in this study is developed for Sweden, whereby the country is divided into squares of $1\times1\,\mathrm{km}^2$, herein termed *grid cells*, resulting in 104,853 populated grid cells for Sweden. The LV grid is generated and simulated for each grid cell individually in the model. The REGAL model is written in Julia. The model code is available online under a permissive open-source license [20].

2.1.1. Synthetic low-voltage grid generation

The REGAL model generates a synthetic LV grid by selecting the number of transformers, cables, and feeders of different capacities for each grid cell. Only residential customers are considered in the model, both when the grid is dimensioned and during grid simulation. It is assumed that all households are connected to an LV grid. Fig. 2 shows an overview of the process used to generate the synthetic LV grid within a grid cell.

The REGAL model selects the number of transformers of different sizes in each grid cell using the total peak power of the grid cell. The peak power demand in a grid cell is estimated using Velander's formula in Eq. (1). Velander's formula represents an empirical relationship between peak power and annual energy use for Swedish demand patterns. It is used in the model because it was commonly used to dimension components when the major part of the Swedish LV grid was built. Velander's formula is still used by grid companies to dimension components during grid expansion in Sweden, although it is being replaced by other methods that rely on measured data. The peak power demand is used to estimate the demand for transformer capacity in the grid cell [see Eq. (2)]. A transformer margin, α, is applied to take into account that transformer capacities are typically over-dimensioned during grid expansion, to allow for future load increases. Transformers are available in discrete sizes and are subject to economics of scale, with decreasing cost per kVA for increasing transformer size. Eq. (3) is used to estimate how many transformers of a certain capacity will be needed to fulfill the transformer demand in the grid cell. One grid cell can have several transformers, although as symmetry within the grid cell is assumed, all transformers must be of the same size. Table 1 lists all the notations used in Eqs. (1)–(3). Data on the properties and costs of the components (transformers, feeders, and cables) are given in Appendix A1.

$$P_{T,\max} = k_{1,APT} \bullet N_{APT} \bullet E_{APT} + k_{2,APT}\sqrt{N_{APT} \bullet E_{APT}} \quad + k_{1,SFD} \bullet N_{SFD} \bullet E_{SFD,p} + k_{2,SFD}\sqrt{N_{SFD} \bullet E_{SFD,p}} \tag{1}$$

$$T = \alpha_p \bullet P_{T,\max} \tag{2}$$

$$T \le T_i \bullet N_{T,i} \quad i \in I \tag{3}$$

The average annual electricity demands used in Eq. (1) differ between single-family dwellings (SFD) and apartments (APT), as well as between urban and rural areas. In the model, it is assumed that the LV grids in rural electricity areas are dimensioned to supply more customers with electric heating for single-family dwellings, while in urban areas more of the customers in Sweden are connected to a district heating network. The annual electricity demands used in the model, which are taken from Zimmerman [21], are assumed to be 18,558 kWh for single-family dwellings in rural areas, 8416 kWh for single-family dwellings in cities, and 2404 kWh for apartments based on an average apartment with two adult inhabitants. The electricity demand for apartments is scaled by a factor of 1.25 to include electricity use in the common areas of the buildings, such as for elevators [22].

The geometry of the synthetic grid is constructed in the same way as presented by Hyvärinen [23], as illustrated in Fig. 3. A uniform distribution of transformers and customers within the grid cell is assumed, as the positions of the transformers and customers are unknown (Fig. 3). This means that the sizing of power lines is selected such that all customers are reached, and the capacity of each cable is selected to ensure that it meets the national standards regarding tripping time, i.e., the time before a protection device trips during a power fault. Furthermore, the power lines must have sufficient capacity to meet the peak power demand and to ensure voltage quality. A final check is done to ensure that the earthing impedance does not exceed the maximum values [24]. The estimations of the power demand in cables assume fuse sizes of 10 A in apartments 20 A in single-family dwellings. When there is more than one household, a coincidence factor for the cables is calculated. This is achieved by comparing the maximum power for all households, calculated using Eq. (1) for the whole load, to the sum of the maximum powers when using Eq. (1) for each household separately. The cable capacities for different segments are calculated iteratively until all of the abovementioned design criteria are met, following the process presented by Hartvigsson et al. [18].

In Fig. 3, the orange colored dots represent the customers who are supplied by the transformer, which in turn is indicated by a star. Each transformer can have up to four LV feeders, the cable for one of which is illustrated by the blue line in Fig. 3. Each cable that is connected to the feeder, called a branch (illustrated with yellow lines), is assumed to serve the same number of customers. MV feeders are not depicted in the

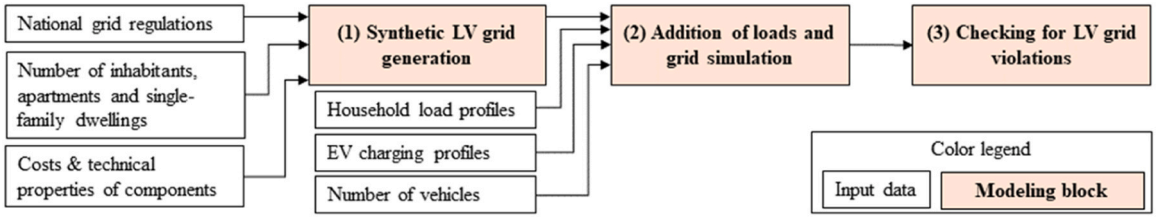


Fig. 1. Overview of the main modeling blocks and input data for the REGAL model.

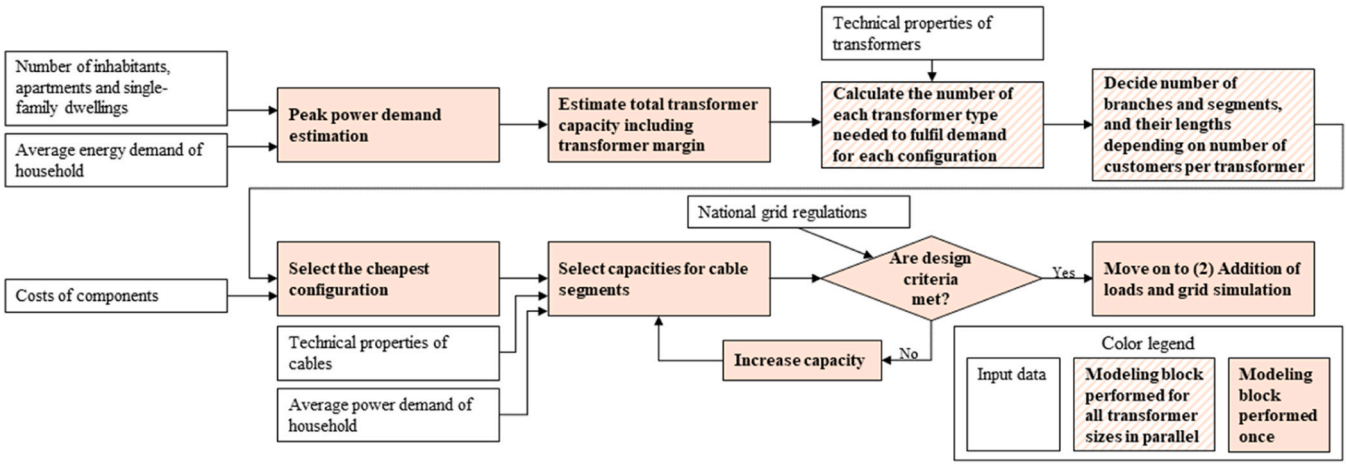


Fig. 2. Overview of the process used to generate the synthetic grid.

Table 1
Parameters and variables for Eqs. (1)–(3).

Notation	Description
$E_{SFD/APT}$	Annual energy consumption for a household of the specified type
I	Set of all possible transformer capacities
k_1, k_2	Velander's coefficients
$N_{SFD/APT}$	Number of households of the specified type
$N_{T,i}$	Number of transformers
p	Population density
$P_{T,max}$	Peak power demand imposed on the transformer
T	The total demand for transformer capacity
T_i	The capacity of transformer *i*
α_p	Transformer margin for the population density *p*

Table 2
Parameters and variables for Eq. (4).

Notation	Description
C_{tot}	The total cost of the synthetic grid
C_{LV}	Cost of low-voltage line per km
C_{MV}	Cost of medium-voltage line per km
$C_{T,i}$	Cost of transformer i
I	Set of all possible transformer capacities
l_{LV}	The total length of low-voltage lines
l_{MV}	Marginal length of medium-voltage lines
$N_{T,i}$	Number of transformers

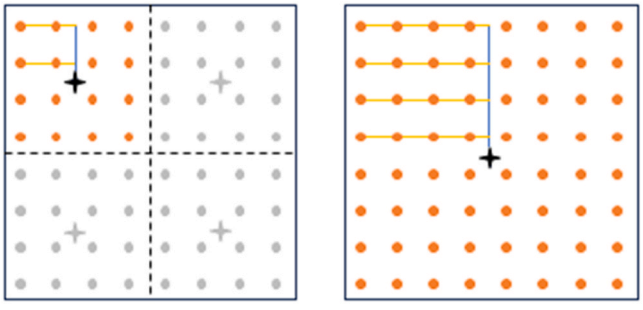


Fig. 3. Illustration of the placement scheme for the transformers (the black cross), cables (blue and yellow lines), and customers (orange dots) in a grid cell. This example includes 64 customers. Panels a and b illustrate two different configurations. Panel a assumes a lower capacity for each transformer compared to panel b and, thereby, more transformers. MV feeders are not included in the illustration.

figure. Multiple setups could reach all customers and supply the required transformer capacity, for example, either more transformers of a lower capacity with shorter LV lines and more MV feeders, or fewer transformers of a higher capacity with longer LV lines and fewer MV feeders. Fig. 3 includes two examples of layouts that can reach all of the customers: one with four transformers of lower capacity (left panel); and one with a single transformer of higher capacity (right panel). Cable dimensioning is performed for each possible layout based on the number of customers per transformer. Since multiple configurations can exist, the cost of each configuration that fulfills Eqs. (1)–(3) is calculated using the formula presented in Eq. (4). As seen in Eq. (4), the number of transformers, their capacities, and costs, as well as the cost for LV and MV lines and feeders used to connect the customers within the grid cell, are considered. Then, the system with the lowest cost is selected. The parameters and variables for Eq. (4) are provided in Table 4.

$$\min\left[C_{tot} = N_{T,i} \bullet C_{T,i} + l_{LV,i} \bullet C_{LV} + l_{MV,i} \bullet C_{MV}\right] \quad i \in I \tag{4}$$

2.1.2. Addition of loads and grid simulation

The second building block in the REGAL model involves the addition of residential and EV loads to the synthetic grid generated in the first block. According to the local regulations for the Swedish grid, the limit for deviations from the nominal voltage is evaluated over a 10-minute period, so the model used a 10-minute time resolution. Load profiles were randomly allocated to different households and repeatedly simulated using a Monte Carlo approach to consider the uncertainty as to where in the grid different loads could occur, while still being able to study when in time issues are likely to occur.

In the model, 35 different measured household load profiles with 10-minute time resolution were used: 20 for single-family dwellings and 15 for apartments. The load profiles included households with different heating technologies and were all gathered from a region that is centrally located in Sweden [21]. These profiles were shifted in time, up to 2 hours and additionally up to 1 full week, backwards or forwards in relation to the original profiles, so as to create at least 1120 different household load profiles of each type. The EV charging profiles are based on GPS driving data from 426 EVs [25]. Each EV is by default assumed to charge upon arrival at the home location. This is the most conservative assumption, since charging largely coincides with the afternoon peak load, while smart charging strategies can avoid this by shifting more of the charging to night-time. For more information about the charging profiles, see the paper of Taljegård et al. [26]. For the results shown in this paper, 100 % of the current vehicle fleet was assumed to be electrified. However, the model was developed to allow for the study of other fleet shares and other smart charging strategies. The maximum charging power for EVs was assumed to be 6.9 kW.

As there is no information about which customers the profiles represent best, the household and EV profiles could be allocated to any of the customers in the LV grid. Household and charging load profiles were randomly selected among those available, and then pre-aggregated into aggregated load profiles representing different numbers of single-family dwellings, apartments, and EVs. This pre-aggregation step was performed to decrease the computational time for the model. One model run of the REGAL model iterated over all 104,853 grid cells in Sweden and performed a number of simulations for each cell by selecting different random combinations of household and charging load profiles. The number of simulations used for the results presented in this work was 1500. When analyzing the impacts of specific variables, the same randomization of load profiles was used in the runs, which were compared to isolate the impact from the studied variable.

Due to the assumed symmetry of the loads and equipment within each cell, any violations that arise will occur in the longest cable coming from the transformer, so the calculations need only be performed for the longest cable in a grid cell. This assumption is based on the notion that a longer feeder will have a larger voltage drop, as it likely will have the highest number of customers. To enable the assessments of new loads and technologies in the LV grid, a simplified voltage-drop calculation, as given by Eq. (5), is used. It starts from a nominal voltage and then calculates the voltage drop over the transformer and cable segments up until point j along the LV feeder for a time-step t. The power demand for each segment is the sum of the power demands further down the LV feeder. A power factor of 0.9 is used for all household loads except EV charging, which is assumed to behave as a purely resistive load (with the notations given in Table 3).

$$U_{j,t} = U_N - \frac{R_{T,t} \bullet \left(P_{SFD,T,t} + P_{APT,T,t} + P_{EV,T,t}\right) - X_T \bullet pf_X \bullet \left(P_{SFD,T,t} + P_{APT,T,t}\right)}{U_N} - \sum_{j=1}^{J} \frac{\left(R_{line,j} \bullet \left(P_{SFD,j,t} + P_{APT,j,t} + P_{EV,j,t}\right) + X_{line,j} \bullet pf_X \bullet \left(P_{SFD,j,t} + P_{APT,j,t}\right)\right)}{U_N} t \in \tau \tag{5}$$

2.1.3. Evaluation of power system violations

The power system violations that are considered can be split into voltage variations outside the allowed limits and violations of thermal constraints, hereinafter referred to as *voltage* and *thermal* violations, respectively. Voltage violations occur when the voltage drops below or exceeds a fixed level, in this study set as a 5 % deviation from the nominal value to allow for some variation in the MV level and commercial loads before reaching the limit of 10 % commonly used in Sweden. The criteria for when a voltage violation occurs are given in Eq. (6) and Eq. (7), and the thermal violation criterion is expressed in Eq. (8).

Table 3
Parameters for Eq. (5).

Notation	Description
N_{EV}	Number of EVs per household
N_j	Number of customers supplied at point j
N_T	Number of customers supplied by the transformer
$P_{SFD/APT,t}$	Average load profile of households
P_{EV}	EV charging rate
pf_X	Relation between active and reactive power
$R_{line,j}$	Resistance of feeder at point j
R_T	Resistance of the transformer
t	Time-step
$U_{j,t}$	The voltage at point j
U_N	Nominal voltage
$X_{line,j}$	The reactance of the feeder at point j
X_T	Reactance of the transformer
$\lambda^s_{EV,t}$	Charging coincidence of EVs at percentile rank s
τ	Set of modeled time-steps

$$U_J \geq U_N \bullet 1.05 \tag{6}$$

$$U_J \leq U_N \bullet 0.95 \tag{7}$$

$$P_J \geq T_C \tag{8}$$

Here, T_C is the thermal capacity and P_J is the load from both households and EVs at point J. For each time-step and run, each of the criteria listed above for a power system violation is checked for all relevant points in the LV grid. If any of the criteria are met, they are recorded and the likelihood of a violation occurring for each time-step can be determined by calculating how many of the randomized vehicle and household combinations for that grid cell violate the criteria.

The share of the vehicle and household load profile combinations that results in power system violations is calculated for each time-step in each grid cell using Eq. (9). This share is then summarized over the time-series to derive the average number of time-steps with a violation of each type in a year for each grid cell. Over multiple iterations, the average number of violations will converge to the expected value (of the stochastic variable).

$$V_{x,tot} = \sum_{t=1}^{\tau} \frac{\sum_{s=1}^{S} V_{x,t,s}}{S} \tag{9}$$

where $V_{x,tot}$ is the average number of violations in a year of type x, $V_{x,t,s}$ is a binary variable describing whether or not a violation occurs in time-step t and for run s, and S is the total number of runs considered. The calculations are performed for each type of violation individually, as well as for all violations taken together.

2.1.4. Calibration and validation of the generated synthetic grid

The generated LV grid was first calibrated and then validated against real-world data supplied by several Swedish grid operators. The calibration focused on reducing the errors in the estimations of transformer capacities and the validation showed how the errors related to the initial electricity demand estimations and transformer capacities vary between different grid cells and groupings of grid cells. Detailed data on a grid cell level were provided by several grid operators, representing both large and small operators. The data cover 9477 out of 104,853 grid cells in Sweden, including grid cell data from 7 out of 21 regions and 81 out of 290 municipalities (for maps of the municipalities and regions in Sweden, see Statistics Sweden [27]). The data per grid cell consist of: (i) the annual electricity demands for residential and commercial customers; (ii) the number of transformers with different capacities; and (iii) the cable lengths. Since most of the grid cells contain both commercial and residential customers and the REGAL model only considers residential demands, allocations between residential and commercial customers were made according to the peak power demand. The annual electricity demands measured by the grid operators were used to calculate the peak power demands using Velander's formula [Eq. (1)] for household and residential customers, respectively. It was assumed that the residential customer's share of the total peak power demand, μ, in each grid cell was the same as their share of the total transformer capacity.

Some of the 9477 grid cells were excluded from the dataset based on the following criteria: (i) grid cells without registered inhabitants; and (ii) grid cells where the commercial customers accounted for more than 30 % of the peak power demand. These exclusions were made because the model only considers residential demands, and the empirical correlation between the transformer capacity data and the population data was significantly lower that level of commercial share of load. In total, 7382 out of the 9477 grid cells were used in the dataset for calibration and validation of the model.

Fig. 4 shows the real-world transformer capacity per grid cell and the population density for the corresponding grid cell. It is clear that there is a general increase in transformer capacity with population, although a large scatter can be observed (the scale is logarithmic). This means that

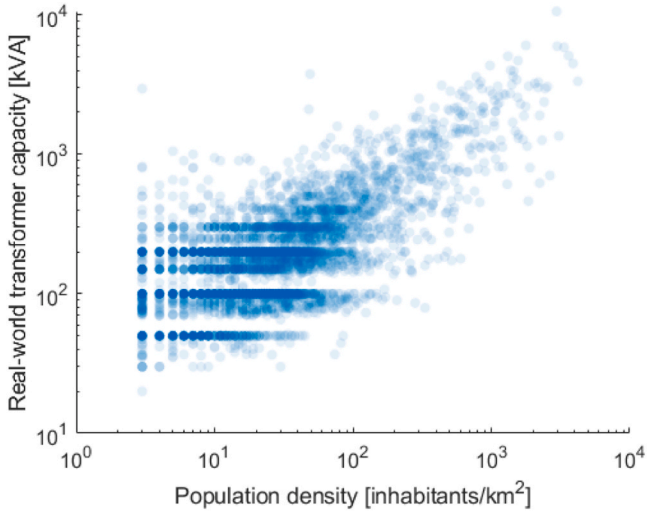


Fig. 4. Transformer capacity allocated to residential users in the real-world data in kVA (*y*-axis) and population density expressed as the number of inhabitants per square km for each grid cell in the calibration dataset (*x*-axis). Scales are logarithmic with a base of 10.

although the population density can be used as a predictor of transformer capacity in a grid cell, it will never predict perfectly the transformer capacity in all grid cells. This is due to several reasons, one being that transformers come in discrete sizes and another being that the locations at which inhabitants reside according to the population statistics may not be the same locations as where they consume electricity. It is also possible that the population in an area has changed significantly since the grid was dimensioned. As the current population statistics form the basis for the model, a deviation from real-world data is expected and will occur.

The transformer margin (α_p), used for dimensioning the transformers in Eq. (2), was varied to study its impact on model accuracy. In previous versions of the model [18], a fixed α_p-value was used based on a dataset derived from fewer cells and only a single grid operator. In this study, the α_p-values were varied to identify those values that deviated the least when comparing modeled and real-world transformer sizes. The model was calibrated to reduce the difference between the model values and the values from the real-world data.

The error and the deviation describe how different the transformer capacity in the model for a grid cell is in comparison to the real-world data. The deviation in transformer capacity between the REGAL model and the real-world values was calculated according to Eq. (10). A positive value means an over-dimensioning of the transformer in the model and a negative value represents an under-dimensioning. The error between the modeled transformer capacity and the real-world data in a grid cell was calculated using Eq. (11). In Eqs. (10) and (11), μ is the residential share of the maximum power, i.e., the residential power demand divided by the total power demand in the grid cell. All errors will be seen as positive values, regardless of whether the errors are due to over- or under-dimensioning of the transformer. When examining how the model performs for a larger area, both the error and deviation are averaged over the grid cells in the area. Thus, the deviation describes the bias for the group when averaging it over a studied area, while the error describes the average error in a grid cell. The notations used are defined in Table 4. When the validation of the initial electricity demand estimation was carried out, deviations from the real-world values for the annual electricity demand were calculated in the same way as for the transformer capacity.

$$d = \frac{T_{\text{mod}} - T_{real} * \mu}{T_{real} * \mu} \tag{10}$$

$$e = \frac{|T_{\text{mod}} - T_{real} * \mu|}{T_{real} * \mu} \tag{11}$$

Validation of the model accuracy for different geographic scopes was made by comparing the average errors and the average deviation on both the grid cell level and for larger geographic areas. Table 5 shows the grouping of grid cells into three categories based on population density (city, urban, and rural) and the population density limits used for the groupings. Grouping was also performed for two other geographic resolutions: municipality, and region. The key features of regions are described in Table 6. The numbers of grid cells in Tables 5 and 6 represent the remaining grid cells after applying the two exclusion criteria previously described.

2.2. Geographic input data

The REGAL model requires input data on the grid cell level for several demographic parameters, including: (i) population density; (ii) the number of apartments; (iii) the number of single-family dwellings; and (iv) the number of vehicles. Of these, only the population density is available for the grid cell resolution. In Sweden, statistics regarding the numbers of apartments, single-family dwellings, and vehicles are available for demographic statistical areas (DeSA, or DeSO in Swedish). The DeSAs divide Sweden into 5984 areas based on population density, where each DeSA contains between 700 and 2700 inhabitants [28]. Table 7 shows the original geographic resolution of the geographic data used in the model.

The geographic areas of the DeSA, for which most of the input data are available, differ significantly in size between rural and urban areas, given that DeSAs by design have roughly similar populations. Each grid cell was allocated to a DeSA by identifying the specific DeSA within which the center-point of each grid cell was located. To downscale the DeSA statistics to grid cells, nonlinear functions describing how the numbers of vehicles, single-family dwellings, and apartments vary with population density were estimated from the data on the DeSA level. These population density relationships were then applied to cells within each DeSA but weighted so that the total numbers of apartments, single-family dwellings, and vehicles within each DeSA matched the original statistics when summing up the grid cells. Although this might not be a perfect representation of reality, it is assumed to be more-accurate than a constant or a linear, population-based distribution, which would not capture the differences between urban and rural areas as accurately.

3. Results

Section 3.1 presents a comparison of the model results and the real-world data, and Section 3.2 provides the results from runs that used the

Table 4
Parameters for Eqs. (10) and (11).

Notation	Description
d	Deviation from real-world data
e	Error compared to real-world data
T_{mod}	Total transformer capacity in grid cell from the model
T_{real}	Real-world data for transformer capacity
μ	The residential share of the maximum power

Table 5
Groupings based on population density. The population is reported as number of inhabitants per km^2.

Group name	Population density, p	Number of grid cells
Rural	$p < 200$	7062
Urban	$200 \leq p < 1000$	261
City	$p \geq 1000$	59
$\sum$		7382

Table 6
Characteristics of the regions.

Area name	Average population per km^2	Number of grid cells
R1	406.2	190
R2	159.7	308
R3	39.7	374
R4	22.5	426
R5	31.9	1401
R6	51.46	1854
R7	30.7	2829

Table 7
The initial geographic resolutions of some input parameters.

Parameter	Initial geographic resolution	Reference
Inhabitants	Grid cell	[29]
Average number of inhabitants in single-family dwellings and apartments	DeSA	[28]
Number of people in single-family dwellings and apartments	DeSA	[28]
Number of cars in traffic	DeSA	[28]
Low-voltage grid regulations	National	[24]

REGAL model to add EV charging to residential loads.

3.1. Comparison of the model with real-world data

Fig. 5 shows a comparison of the annual residential electricity demand calculated in the REGAL model using the annual electricity demand measurements from grid operators. As shown in Fig. 5, there are grid cells in which the annual electricity demand in the REGAL model is overestimated and ones in which it is underestimated. Overall, the estimations in the model follow the measured electricity demand, whereas for low demands the relative scatter of the data-points is large. As the demand estimation from the REGAL model for a given population density will have a similar value for cells with the same population density, only varying depending on the distinction made between apartments and single-family dwellings, most of the variation in deviation arises from differences in the measured demand. Therefore, there will always be scatter in the modeled results as more factors than population density and housing type influence the electricity demand. Fig. 6 shows the average deviation, d, between the measured electricity demand and the estimations of annual electricity demand in the REGAL model [calculated as in Eq. (10)]. The density of data-points in Fig. 6 indicates the occurrence of different values of d for the given population density. A small bias towards overestimation is evident in the electricity demand estimations made by the model. Typically, larger deviations are seen for grid cells with low population densities, and the large span in load estimations seen for areas with low population densities reveals how the same population can have very different electricity consumption levels in different areas. This indicates that there are individual differences between customers' electricity demands that cannot be captured by a general assumption. A smaller deviation with a higher population density is expected because with a higher number of households it is more likely that an average value will be representative, given that individual electricity use variations will cancel each other out.

During the calibration of the model, various α-values were investigated, i.e., different factors for over-dimensioning transformers to represent real-world grid dimensioning principles. As can be seen in

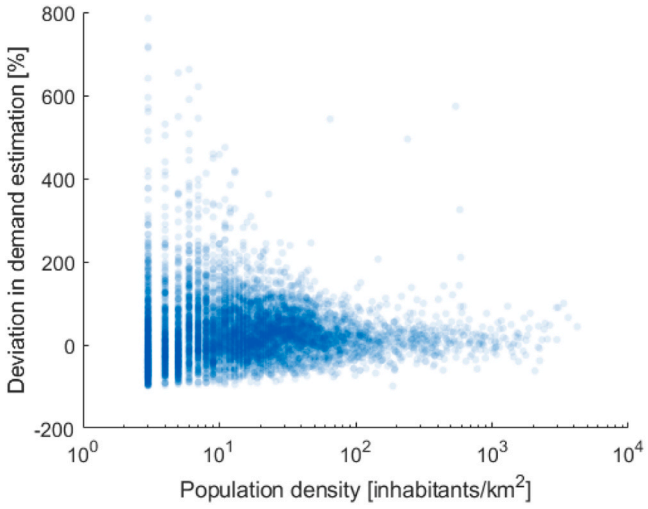


Fig. 6. Percent deviation, d, between the annual residential electricity demand assumed in the REGAL model and the measured residential annual electricity demand from grid operators, for grid cells with different population densities [Eq. (10)]. The x-axis is logarithmic with a base of 10. One data-point corresponds to one grid cell.

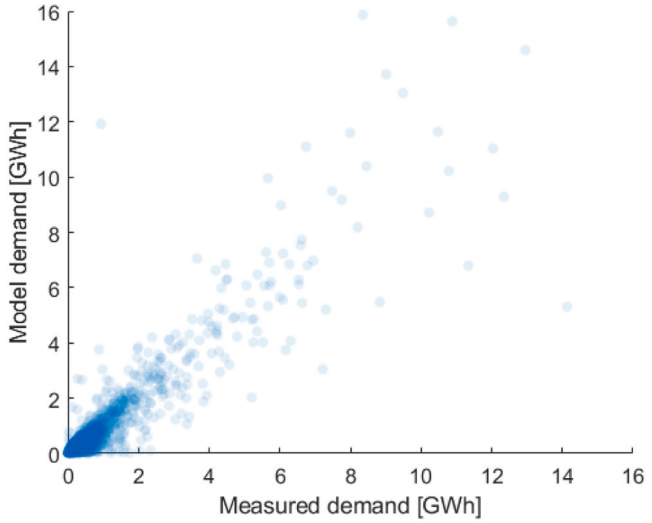


Fig. 5. The annual electricity demands assumed in the REGAL model and the measured residential annual electricity demands from grid operators per grid cell with linear axes. One data-point corresponds to one grid cell.

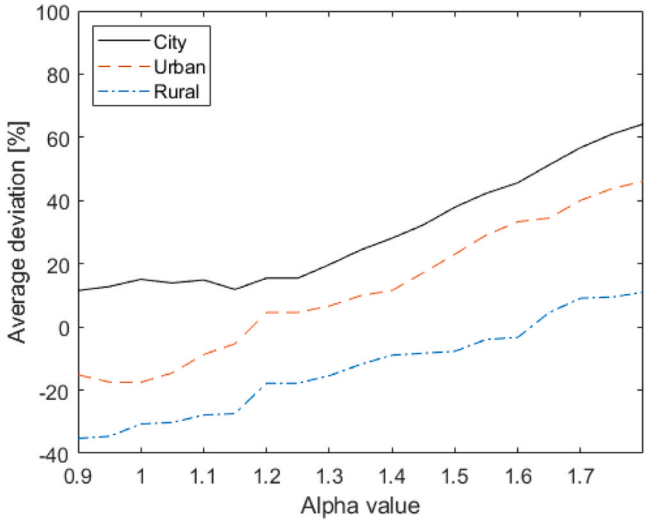


Fig. 7. Percent deviation, d, from the real-world data with respect to transformer capacity, averaged for all the grid cells in rural, city, and urban areas, calculated according to Eq. (10).

Fig. 4, the real-world data are highly variable, so a large error is expected using this methodology. Fig. 7 shows the deviation in transformer capacities between the REGAL model and real-world values averaged over all grid cells for different α-values [see Eq. (10) for calculation of d]. As the deviation differs depending on population density, the results are shown separately for rural, urban, and city areas. As the α-value increases, the transformer capacity increases in the model. A positive value in Fig. 7 means that the transformer capacity in the model is larger than that in the calibration dataset, i.e., there is an overestimation in REGAL, and a negative value means that the transformer capacity is underestimated. As can be seen in Fig. 7, the deviation for a given α-value is the highest in city areas and the lowest in rural areas. This means that transformers typically have a smaller margin in areas with higher population densities.

Fig. 8 shows the error, e, averaged across the grid cells for the rural, city, and urban areas, respectively [with e calculated according to Eq. (11)]. From Fig. 8, it can be concluded that for all α-values, the average error per grid cell is larger in grid cells with higher population densities (urban and city), as compared with grid cells that have lower population densities (rural). For rural areas, increasing the α-value has a weak impact on the error per grid cell, meaning that there are few cases in which the selected transformer size is affected by an increase in the α-value. This indicates that there is typically an over-capacity in transformers in those areas with a low population density, as transformers come in discrete sizes. The average error for all grid cells follows the trend of the rural grid cells, as there are many more rural grid cells in the calibration dataset. For the city grid cells, a low α-value gives a small error, indicating that there is a low level of over-capacity in transformers in grid cells with a high population density. The trend observed for the urban grid cells shares similarities with that for the city grid cells, although the errors are smaller, and a higher α-value is needed to achieve an average deviation of 0 in Fig. 7. The average errors per grid cell for urban areas are similar for α-values of between 1.1 and 1.4, as can be seen in Fig. 8. As the goal of the calibration was to achieve an average deviation of 0 (Fig. 7) while minimizing the average error per grid cell (i.e., as small as possible, according to Fig. 8), the application of different α-values to different population densities yielded the best fit to the real-world data. An α-value of 1.15 for city cells and an α-value of 1.65 for rural grid cells gave a small error and weak deviation. For populations of between 50 and 500 people per km^2, a linear decrease in the α-value from the rural to the urban setting was found to yield both a small error and low levels of deviation. This combination of different α-values depending on the population density resulted in an average deviation over all grid cells of −0.83 % and an average error over all grid cells of 44.5 %.

The percent deviation per grid cell when using the α-values that generate the smallest average error and deviation (α-values of 1.1 and 1.65, respectively) is shown as a histogram in Fig. 9. One value for each of the 7382 grid cells in the validation dataset is included in Fig. 9. A span of the deviation from real-world values of between −100 % and +350 % is evident in Fig. 9 when analyzing the grid cells individually. However, since the real-world data do not strictly follow the current population density (as shown in Fig. 4), some deviation is to be expected. This indicates that the REGAL model is not suitable for analyzing results at the single grid cell level, but instead works for a set of grid cells taken together.

3.1.1. Regionally aggregated results

The average deviation and the average absolute error over all grid cells when aggregating grid cells over seven regions in Sweden are shown in Fig. 10. The regions are sorted in Fig. 10 such that the number of grid cells included in the region increases with the region number. This means that regions R1 and R2 (see Table 6), which have the largest average deviations (above ±10 %) also have the lowest numbers of grid cells. These regions also have the highest population densities. This is as expected since grid cells that have a higher population density typically show the largest errors, as shown in Fig. 8. Fig. 11 shows the deviation values averaged over the grid cells in the different municipalities and regions represented in the validation and calibration dataset. The deviation values when grid cells are grouped into municipalities are shown in blue, and when they are grouped into regions are shown in orange. A dashed line is shown for nine grid cells, as the smallest municipality in Sweden is approximately 9 km^2. This means that all groupings with fewer than nine cells are not complete datasets for the municipalities. Furthermore, grid cells are likely missing for municipalities with only a few more than nine grid cells, as most municipalities are much larger. Fig. 11 shows that the average deviation value decreases when the number of included grid cells increases.

3.2. The modeling results when adding EV charging to residential loads

Fig. 12 shows the share of grid cells that have an average number of voltage violations over different numbers of time-steps in the rural, urban, and city areas, as well as for all cells [calculated according to Eq. (9)], when adding EV charging of a fully electrified fleet of passenger vehicles to the residential loads in the model, and assuming that these

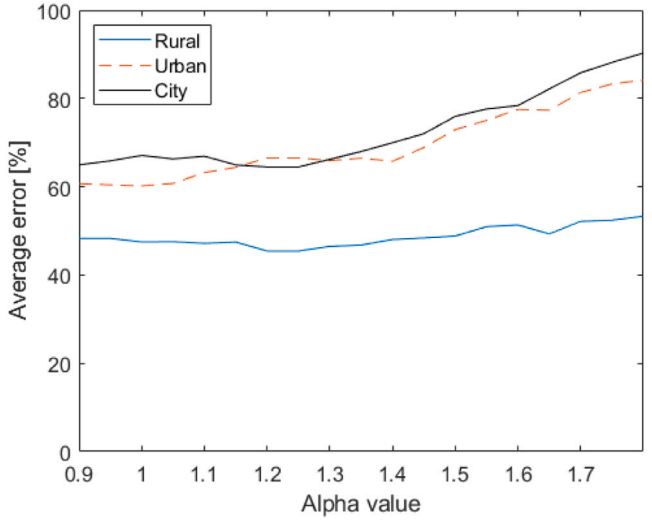


Fig. 8. Error values, e, for the transformer capacity averaged across all grid cells within rural, urban, and city cells, as a function of the α-value, calculated according to Eq. (11).

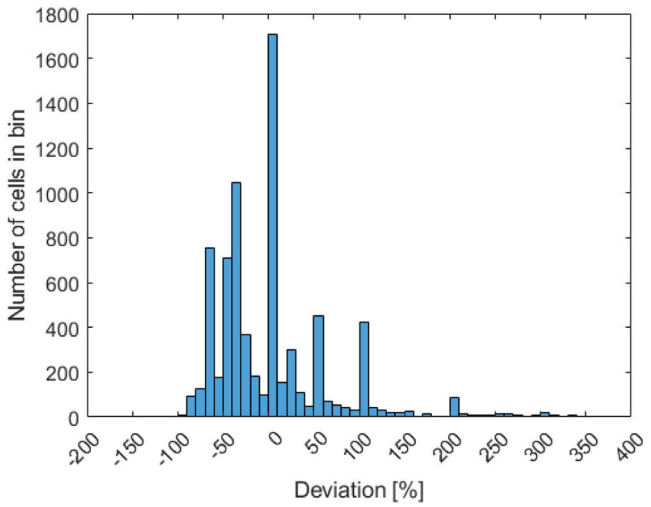


Fig. 9. Percent deviations, d, of the transformer capacity for 7382 grid cells, calculated according to Eq. (10).

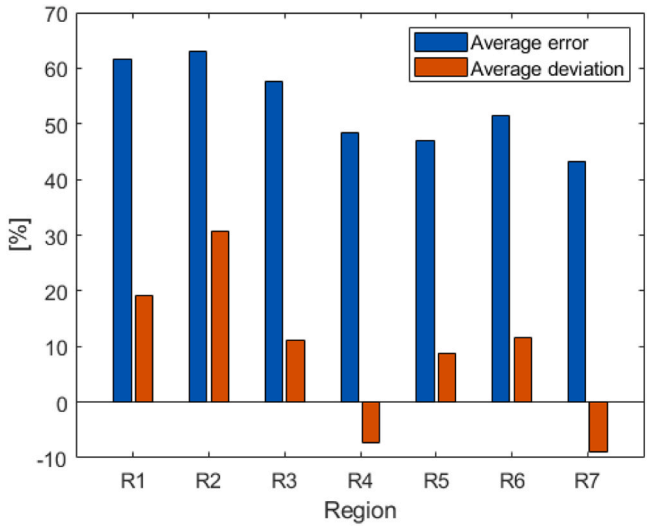


Fig. 10. Error values, e (blue), and deviation values, d (red), of the transformer capacity averaged over all grid cells in the geographic grouping R1 to R7 (see Table 6). The value of d is calculated according to Eq. (10) and that of e is calculated according to Eq. (11). Region names have been anonymized to prevent the identification of sensitive data received from grid operators.

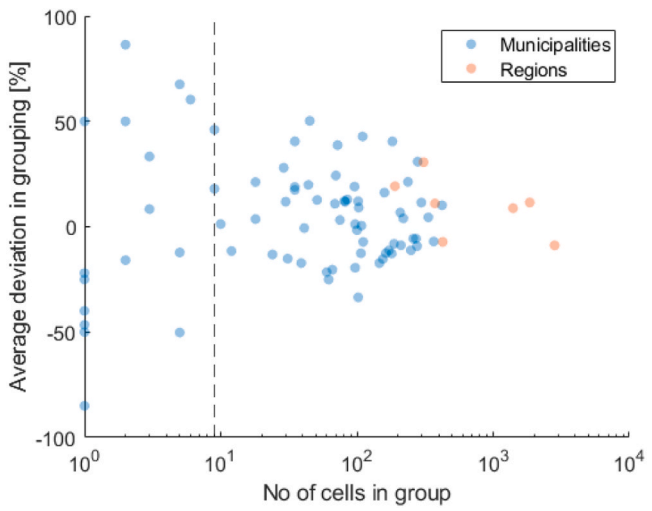


Fig. 11. Deviation values, d, of the transformer capacity calculated according to Eq. (10), averaged over the grid cells in the municipalities and regions (y-axis) and the number of grid cells in the municipality or region (x-axis). The x-axis is logarithmic with a base of 10. A dashed line is shown for nine grid cells, as the smallest municipality in Sweden has an area of approximately 9 km^2.

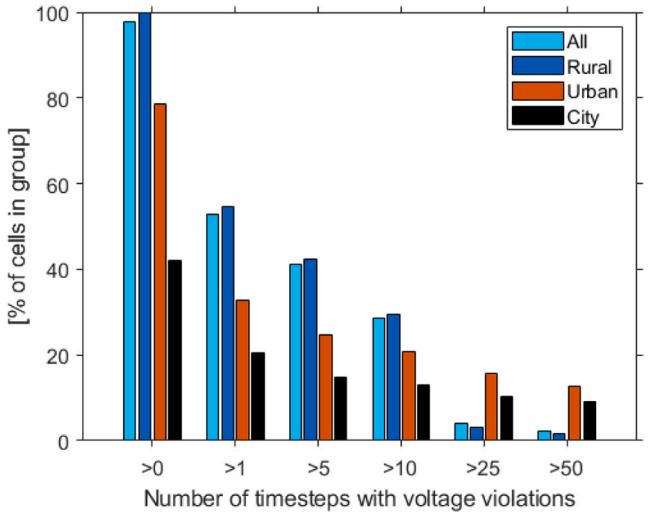


Fig. 12. Share of grid cells with voltage violations above a certain value when adding EV charging. The share is given for different average numbers of 10-minute time-steps during a year with a voltage drop below the threshold for voltage violations. The results are shown for rural, urban, and city areas, as well as for all cells. The results shown are the averages across all iterations.

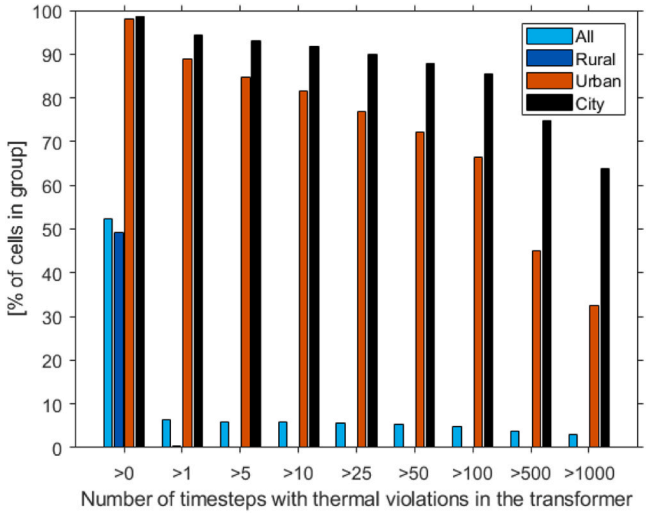


Fig. 13. Share of grid cells with thermal violations in the transformer above a certain value when adding EV charging. The share is given for different average numbers of 10-minute time-steps during a year with thermal violations. Data shown are for rural, urban, and city areas, as well as for all cells. The results shown are the averages across all iterations.

vehicles initiate charging when arriving at home (see Section 2.1.2). The value for all cells is similar to that for the group of rural cells, as rural cells make up by far the largest grouping. The number on the x-axis in Fig. 12 corresponds to the number of 10-minute time-steps during a year with a voltage drop below the threshold for voltage violations averaged over all iterations, when including EV charging. Fig. 13 is similar to Fig. 12, except that it is calculated for thermal violations in the transformer instead of voltage violations. From Fig. 12, it can be concluded that voltage violations are more common in rural areas than in urban and city areas, although some violations are also seen in the groups of grid cells that have higher population densities than in the rural areas (i.e., city and urban grid cells). Typically, the grid cells with voltage violations have very few average number of time-steps with violations in a year. Fig. 13 shows that thermal violations in the transformer are common in the groups with high population densities (i.e., city and urban grid cells). The majority of the city and urban cells have an average number of time-steps with violations that exceed 100 violations (corresponding to a total of 16 hours and 40 minutes), and more than one-third of the urban grid cells and a majority of the city grid cells have an average number of time-steps with violations that exceed 100 violations (corresponding to a total of 6 days, 22 hours and 40 minutes). However, Figs. 12 and 13 do not show by how much the voltage and thermal limits are exceeded, when in time the violations are likely to occur, and whether the time-steps during which violations occur are sequential. Other EV charging strategies, and flexible ones, have not been considered in this study. Inclusion of flexible EV charging could potentially increase the hosting capacity as it could shift EV loads to times when household loads are low.

4. Discussion

Using a synthetic grid representation allows for estimations of hosting capacities in those cases where detailed information on the state of the grid is not publicly available. However, when possible, more accurate results will naturally be achieved with real grid capacities. This means that a model such as REGAL can provide initial indications as to where and when problems will occur in the grid when adding, for example, EV charging as is done in this work, although the model will seldom make predictions that are correct for a small geographic area (such as 1 km^2). As can be seen in the validation of the model, it can make accurate predictions of transformer capacities when assessing the average deviation from real-world data for a larger area, such as a region. For a given population density, some grid cells will have a larger transformer capacity in the REGAL model than is evident from the data provide by the grid operators, and some grid cells will have a smaller capacity. When there is a smaller transformer capacity in the model than in reality, the average number of time-steps with violations will be overestimated, whereas in a grid cell with a larger transformer capacity, the average number of time-steps with violations will be underestimated. As the dimensioning of transformers has been calibrated to have close-to-zero deviation, while maintaining as small errors as possible on the grid cell level, this issue will to some degree cancel itself out, giving correct average estimates for larger groups of grid cells.

The methodologic choices for generating a synthetic grid will depend largely on the purpose of the model, as well as on the types of data that are available for the region studied. The methodologic choices will also determine the accuracy of the model. As was observed during the validation of the REGAL model in the present study, accuracy can be improved by differentiating the assumptions according to population density. The calibration of the REGAL model against real-world data showed a clear difference in the sizing of the transformer margin depending on the population density. This underlines the importance of differentiating the margin used when creating the synthetic grid, so as not to overestimate or underestimate consistently the transformer capacity for certain grid cells. When differentiating the values used depending on population density an increase in the overall accuracy of the model is achieved.

Based on the level of accuracy obtained from the model validation, we conclude that the model is useful for a multitude of purposes related to making general estimations of the capacity of residential LV grids to handle changes to loads and electricity production. This could include further analysis of how different EV charging strategies affect the LV grid, how much solar PV different areas are likely to be able to host, or how the use of stationary batteries in households for ancillary services influences the grid. In addition, the model can be used for estimations of the cost efficiency of alleviating issues on the residential LV grid.

Grid cells that carry a high commercial load are most likely to be less accurate because the REGAL model includes only residential loads that are estimated from population statistics, thereby excluding any commercial or industrial activities. In those areas, the numbers of violations due to EV charging might not be accurate, although the modeling could still indicate what types of violations might be introduced, as well as where and when they are more likely to occur. To generate a more-accurate model for these grid cells, commercial loads need to be added to the model.

The current model is based on the national grid regulations of Sweden. As grid regulations and standards vary between different countries, the grid regulations applied in this model might not be transferrable to other countries. Furthermore, since the α-values were calibrated against data from only one country, it is likely that these values are not representative of other countries. The method is generally applicable to other geographic regions for which population data is openly available, however, more accurate results will be found if national regulations are taken into account and calibration parameters are updated to correspond to the studied region.

The number of residential load profiles used in the model is low, and they are all gathered from one region within Sweden. A clear improvement would be to increase the number of residential load profiles from which the model can choose. In addition, the allocation of household load profiles to different households could be improved by employing a smarter allocation scheme. On way to do this is through weighting the likelihood of a profile appearing in an area using statistics on, for example, heating types or sizes of buildings. This would, however, require a larger dataset of household load profiles with high temporal resolution, generated either from measurements or from deploying an algorithm to increase the temporal resolution of load profiles with a 1-hour time resolution, which are more readily available.

In the model, residential loads are allocated to the places where people are officially registered to live. This means that the loads could be overestimated or underestimated in some areas, especially areas with large quantities of seasonal housing. For such areas, there might be times of the year with intensive private charging of vehicles, which cannot be captured in the model. When predicting future demands for households and EV charging, changes in the inhabitants have not been considered. This is a drawback, as it is likely that some areas will see an increase and some areas will see a decrease in population during the coming years.

Random combinations of household profiles are considered together. However, it is likely that there is a correlation between electricity demands in households that are located next to each other, e.g., through similar heating technologies or house sizes. An example of an important factor is the availability of district heating, which has not been included in the model. As heating has a strong effect on the residential electricity demand and the studied region has numerous cities with district heating grids, introducing a way to take this into account might improve the accuracy of the model. Taking this into account could both increase and decrease the number of violations achieved, depending on how the distribution grid operators have applied their dimensioning criteria in cities with district heating systems. This aspect is, however, not considered in this article.

The current version of the model does not consider distributed electricity supply on the LV grid, for example through PVs. Furthermore, it does not include stationary batteries within households. Including these technologies might alleviate the issues that appear as consequences of large-scale EV charging, as they have the potential to supply some of the electricity needed, thereby decreasing load peaks in the LV grid. Additional work is needed to understand how changes to charging patterns could decrease the number of issues seen in LV grids.

5. Conclusions

In this work, the REGAL model is presented and validated against a large dataset of real-world transformer capacities and electricity demands, such that open data describing the population and distribution of dwellings in Sweden can be used to estimate the electricity grid capacities for the LV grid. Different calibration values are explored and the levels of accuracy of estimations of grid capacities are calibrated using proprietary real-world data from DSOs. Thus, for a region that includes many grid cells, an average deviation from the real-world data of ±10 % is achieved. For an average grid cell, an error of 44.5 % is seen, which means that the model is not suitable for analyses of an area of 1 km^2. However, the level of accuracy is deemed sufficient for initial estimations of hosting capacity for larger geographic areas, such as a region or a country.

CRediT authorship contribution statement

Therese Lundblad: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization. **Maria Taljegard:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Data curation, Conceptualization. **Niclas Mattsson:** Writing – review & editing, Validation, Software,

Methodology, Data curation. **Elias Hartvigsson:** Software, Methodology. **Filip Johnsson:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Therese Lundblad reports financial support was provided by Norwegian Public Roads Administration. Therese Lundblad reports financial support was provided by Swedish Electromobility Center. Maria Taljegard reports financial support was provided by Swedish Energy Agency. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Data availability

The model and the data used to run it are openly available, but it has been validated against confidential information. Results of the validation is shared openly.

Acknowledgments

We gratefully acknowledge the financial support provided by the Norwegian Public Road Administration, Swedish Electromobility Center, and Swedish Energy Agency.

Appendix A1. Additional technical input data

Input data for transformers are presented in Table A 2 [30] For the dimensioning of cable capacities, a maximum power demand per household is estimated using the approximated fuse sizes of households, i.e., 10 A for an apartment and 20 A for a single-family dwelling.

Table A1
Input data for transformers based on *Energimarknadsinspektionen* [30]. The values in the shaded boxes are given in the reference, the values in the boxes without shading are extrapolated from the given data.

Transformer capacity [kVA]	Cost [SEK]	Earthing impedance [mΩ]
1600	234,551	4.85
1500	223,450	5.15
1250	195,272	6.5
1125	179,381	6.7
1000	163,835	7.5
900	151,090	8.3
800	134,751	10
700	124,778	10.5
600	111,211	12.1
500	101,565	13
400	83,255	17.6
315	70,501	20
200	53,509	32
150	46,769	43.9
100	38,446	65
70	34,731	89
50	32,140	130
30	28,647	195

As there can be different numbers of vehicles or households within a grid cell, the numbers of household or EV profiles used to create the aggregated profiles increase approximately logarithmically to cover all possible options. The numbers of households and EVs pre-aggregated to the aggregated load profiles are listed in Table A 2. Each aggregated load profile contains the sum of randomly selected household or EV charging profiles, and a multitude of pre-aggregated load profiles was created. Single-family dwellings, apartments, and EVs were treated separately when aggregating the profiles, to allow for different configurations within the different grid cells. The number of households of a specific type (single-family dwellings and apartments) and the number of vehicles decide which sets of pre-aggregated profiles should be used for all the iterations. Thus, for one model run and one iteration, one of the pre-aggregated profiles, containing the appropriate numbers of households or vehicles summed together, was randomly selected for each point at which loads were added. For a grid cell in which there is a branch or transformer with a number of households and/or vehicles for which there is no set of pre-aggregated profiles, a pre-aggregated profile from the collection nearest above in the number of summed profiles is selected, and it is scaled down to the lower number of households or vehicles.

As an example, assume a case with a total of 64 customers distributed so that there are 4 customers per branch connected to each LV feeder, as in Fig. 3a, and that the grid cell has only single-family dwellings and one EV per household. In this example, two aggregated profiles are selected randomly for each branch, one representing the total load for four single-family dwellings and one representing four EVs. For the customers not on the feeder, aggregated profiles for the remaining households and their vehicles (in this case, 48) are randomly selected. As there is no pre-aggregation of the 48 profiles, random profiles for 50 houses and 50 vehicles are selected, and scaled down linearly to represent 48 profiles. Thus, the pre-aggregation allows for the combination of randomly selected profiles for all 64 households and 64 randomly selected EVs by adding up only five pre-aggregated profiles of each type, as in this example.

Table A2
Numbers of households and vehicles for which the loads are pre-aggregated.

Numbers of households/vehicles									
1	2	3	4	5	6	7	8	9	10
12	14	16	18	20	25	30	35	40	45
50	60	70	80	90	100	120	140	160	180
200	250	300	350	400	450	500	600	700	800
900	1000	1200	1400	1600	1800	2000	2500	3000	3500
4000	4500	5000							

References

[1] M. Nazari-Heris, M. Abapour, B. Mohammadi-Ivatloo, An updated review and outlook on electric vehicle aggregators in electric energy networks, Sustainability vol. 14 (23) (Nov. 2022) 15747, https://doi.org/10.3390/su142315747.

[2] M. Kumar, K.P. Panda, R.T. Naayagi, R. Thakur, G. Panda, Comprehensive review of electric vehicle technology and its impacts: detailed investigation of charging infrastructure, power management, and control techniques, Appl. Sci. vol. 13 (15) (Aug. 2023) 8919, https://doi.org/10.3390/app13158919.

[3] M. Nour, J.P. Chaves-Ávila, G. Magdy, Á. Sánchez-Miralles, Review of positive and negative impacts of electric vehicles charging on electric power systems, Energ. (Basel) vol. 13 (18) (Sep. 2020) 4675, https://doi.org/10.3390/en13184675.

[4] K. Hedegaard, H. Ravn, N. Juul, P. Meibom, Effects of electric vehicles on power systems in Northern Europe, Energy vol. 48 (1) (Dec. 2012) 356–368, https://doi.org/10.1016/j.energy.2012.06.012.

[5] M. Taljegard, V. Walter, L. Göransson, M. Odenberger, F. Johnsson, Impact of electric vehicles on the cost-competitiveness of generation and storage technologies in the electricity system, Environ. Res. Lett. vol. 14 (12) (Dec. 2019) 124087, https://doi.org/10.1088/1748-9326/ab5e6b.

[6] V. Umoh, I. Davidson, A. Adebiyi, U. Ekpe, Methods and tools for PV and EV hosting capacity determination in low voltage distribution networks—a review, Energ. (Basel) vol. 16 (8) (Apr. 2023) 3609, https://doi.org/10.3390/en16083609.

[7] B.E. Carmelito, J.M. de C. Filho, Hosting capacity of electric vehicles on LV/MV Distribution Grids—a new methodology assessment, Energ. (Basel) vol. 16 (3) (Feb. 2023) 1509, https://doi.org/10.3390/en16031509.

[8] M. Zain ul Abideen, O. Ellabban, L. Al-Fagih, A review of the tools and methods for distribution networks' hosting capacity calculation, Energ. (Basel) vol. 13 (11) (Jun. 2020) 2758, https://doi.org/10.3390/en13112758.

[9] E. Mulenga, M.H.J. Bollen, N. Etherden, A review of hosting capacity quantification methods for photovoltaics in low-voltage distribution grids, Int. J. Electr. Power Energy Syst. vol. 115 (Feb. 2020) 105445, https://doi.org/10.1016/J.IJEPES.2019.105445.

[10] R. Luthander, "Self-Consumption of Photovoltaic Electricity in Residential Buildings," Uppsala University, Uppsala, 2018.

[11] R. Leou, et al., Stochastic analysis of electric transportation charging impacts on power quality of distribution systems, IET Gener., Transm. Distrib. vol. 12 (11) (Jun. 2018) 2725–2734, https://doi.org/10.1049/iet-gtd.2018.0112.

[12] J. Zhu, W.J. Nacmanson, L.F. Ochoa, B. Hellyer, Assessing the EV hosting capacity of australian urban and rural MV-LV networks, Electr. Power Syst. Res. vol. 212 (Nov. 2022) 108399, https://doi.org/10.1016/J.EPSR.2022.108399.

[13] C.D. Zuluaga-Ríos, A. Villa-Jaramillo, S.D. Saldarriaga-Zuluaga, Evaluation of distributed generation and electric vehicles hosting capacity in islanded DC grids considering EV uncertainty, Energies vol. 15 (20) (Oct. 2022) 7646, https://doi.org/10.3390/en15207646.

[14] S. Fan, C. Li, Z. Wei, T. Pu, X. Liu, Method to determine the maximum generation capacity of distribution generation in low-voltage distribution feeders, Jan. 2017, J. Eng. vol (13) (2017) 944–948, https://doi.org/10.1049/joe.2017.0470.

[15] E. Veldman, R.A. Verzijlbergh, Distribution grid impacts of smart electric vehicle charging from different perspectives, IEEE Trans. Smart Grid vol. 6 (1) (Jan. 2015) 333–342, https://doi.org/10.1109/TSG.2014.2355494.

[16] R. Luthander, M. Shepero, J. Munkhammar, J. Widén, Photovoltaics and opportunistic electric vehicle charging in the power system – a case study on a Swedish distribution grid, IET Renew. Power Gener. vol. 13 (5) (Apr. 2019) 710–716, https://doi.org/10.1049/iet-rpg.2018.5082.

[17] J. Amme, G. Pleßmann, J. Bühler, L. Hülk, E. Kötter, P. Schwaegerl, The eGo grid model: an open-source and open-data based synthetic medium-voltage grid model for distribution power supply systems, J. Phys. Conf. Ser. vol. 977 (Feb. 2018) 012007, https://doi.org/10.1088/1742-6596/977/1/012007.

[18] E. Hartvigsson, M. Odenberger, P. Chen, E. Nyholm, Generating low-voltage grid proxies in order to estimate grid capacity for residential end-use technologies: the case of residential solar PV, MethodsX vol. 8 (Jan. 2021) 101431, https://doi.org/10.1016/J.MEX.2021.101431.

[19] W.J. Nacmanson, J. Zhu, and L. Ochoa, "Milestone 6: Network Modelling and EV Impact Assessment," Melbourne, 2021. Accessed: Dec. 21, 2023. [Online]. Available: ⟨https://www.energynetworks.com.au/miscellaneous/milestone-6-ev-integration-final/⟩.

[20] T. Lundblad, N. Mattsson, and E. Hartvigsson, "REGAL-public." Accessed: May 13, 2024. [Online]. Available: https://github.com/Atlun/REGAL-public.

[21] J.P. Zimmermann, "End-use metering campaign in 400 households In Sweden - Assessment of the Potential Electricity Savings," 2009.

[22] I.Y. Choi, S.H. Cho, J.T. Kim, Energy consumption characteristics of high-rise apartment buildings according to building shape and mixed-use development, Energy Build. vol. 46 (Mar. 2012) 123–131, https://doi.org/10.1016/J.ENBUILD.2011.10.038.

[23] M. Hyvärinen, "Electrical networks and economies of load density," Aalto University, 2008.

[24] Svenska Elektriska Kommissionen, "SS 424 14 06 (Methods of calculation to safeguard correct disconnection – single cable in a directly earthed system protected by fuses (simplified method))," 2005.

[25] S. Karlsson and L.-H. Kullingsjö, "GPS measurement of Swedish car movements for assessment of possible electrification," *EVS27*, 2013.

[26] M. Taljegard, L. Göransson, M. Odenberger, F. Johnsson, To represent electric vehicles in electricity systems modelling—aggregated vehicle representation vs. individual driving profiles, Energ. (Basel) vol. 14 (3) (Jan. 2021) 539, https://doi.org/10.3390/en14030539.

[27] Statistics Sweden, "Kartor över indelningar, Län, Kommuner," Geografiska indelningar. Accessed: Jan. 16, 2024. [Online]. Available: ⟨https://www.scb.se/hitta-statistik/regional-statistik-och-kartor/regionala-indelningar/kartor-over-indelningar/⟩.

[28] Statistics Sweden, "DeSO – Demografiska statistikområden." Accessed: Jan. 27, 2023. [Online]. Available: ⟨https://www.scb.se/hitta-statistik/regional-statistik-och-kartor/regionala-indelningar/deso—demografiska-statistikomraden/⟩.

[29] Statistics Sweden, "Population density," 2018.

[30] Energimarknadsinspektionen, "Normvärdeslista elnät 2020-2023," 2019.