FPS Economy, S.M.E.s, Self-employed and Energy

ECOFLEX

With the support of the Energy Transition Fund

D4.3 Digital twin model of public distribution network predicting maximum hosting capacity

Version number and Date: Version 1; 16/09/2024
Author (s): Rémy Cleenwerck (UGENT)

Abstract for dissemination (PU)

This document provides an overview of the various modelling approaches applied in the field of digital twin modelling of low-voltage distribution systems, in the scope of hosting capacity studies for future scenarios. These include the impact of integrating low-carbon technologies, or the implications that flexibility services can create on the distribution network at local level. To address this, a state-of-the-art review is presented that addresses both established and innovative methodologies (e.g. the use of artificial intelligence) to perform power flow analysis. In addition, a use case demonstrating the applied methodologies is presented.

Contents

List of abbreviations

ADN	Active Distribution Network
AMI	Advanced Metering Infrastructure
BFM	Branch Flow Model
BIM	Bus Injection Model
DER	Distributed Energy Resources
DHC	Dynamic Hosting Capacity
DOE	Dynamic Operating Envelopes
DR	Demand Response
DSO	Distribution System Operator
EV	Electric Vehicles
FSP	Flexibility Service Provider
HC	Hosting Capacity
LCT	Low Carbon Technologies
LV	Low-voltage
LVDS	Low-voltage Distribution System
MLR	Multivariable Linear Regression
PCC	Point of Common Coupling
PV	Photovoltaic
QSTS	Quasi-Static Time Series
RES	Renewable Energy Sources
SDGE	Stochastic Distribution Grid Exchange
VUF	Voltage Unbalance Factor

1. Introduction

The widespread adoption of renewable energy sources (RES) together with the electrification of the heat and mobility sectors poses numerous challenges to network operators in terms of resilience and grid stability. Integrating these assets at residential level often results in a grid reinforcement. Large amounts of RES connected to the distribution network could potentially lead to reverse power flows and consequently voltage rises beyond the regulatory limits, while the uptake of electric vehicles or heat pump with higher loads will result in voltage drops that could violate the lower threshold of the operational limits (cf. IEC EN50160 defines the lower voltage limit (v^-) at -10% and the upper limit (v^+) at $+10\%$). However, the stochastic distribution grid exchanges (SDGEs) imposed by these low carbon technologies (LCTs) are not necessarily a burden for the distribution grid if properly managed (e.g. demand response). As noted in [1], LCTs do also provide a huge potential for aggregators to unlock the flexibility needs at low-voltage level, i.e. demand-side flexibility, thus contributing to reach the sustainability goals.

This deliverable provides an extensive overview of both, (i) conventional and (ii) novel methodologies for evaluating the operating conditions of a low-voltage distribution system (LVDS) under the influence of the LCTs. To this end, the concept of grid modelling is elucidated including the terminology employed, the calculation methods, and a test use case to demonstrate the diverse approaches. These aspects are presented in the sections to follow, with Section 2 describing the terminology, state-of-the-art modelling techniques and challenges associated with it, Section 3 presents the concept of power flow from two distinct perspectives, the (i) physics-based and (ii) data-driven approaches. Subsequently, Section 0 provides the conclusions for this deliverable.

2. Grid modelling: Definitions, state-of-the-art and challenges

2.1. General terms employed

(Dynamic) Hosting capacity

Sometimes referred to as the network's headroom, the term hosting capacity (HC) refers to the amount of consumption or generation that could be connected to the (low-voltage) network without violating the distribution network's boundary conditions, i.e. the performance of the grid [2]. The grid performance deterioration of the grid can be described in terms of various phenomena such as: voltage magnitude, voltage unbalance, losses, harmonics, transformers and line loading. Depending on the scope of the study, a different approach must be handled: (i) deterministic, (ii) stochastic (or probabilistic) and (iii) time series [3].

If the location of the additional load or distributed energy resources (DER) source is known, a deterministic approach can be performed since the hosting capacity is unique. However – *if the location is unknown* – depending on the location of the network where an asset would be allocated, the resulting hosting capacity would be different. This stochastic approach requires a probabilistic method where the outcome results in min. and max. HC limits respectively, as shown in Figure 1 (source: [4]).

Taking the example in the figure below, the region on the left hand side of the min. HC limit (green dotted line) indicates that all the penetration of PV systems are acceptable and will not result in any voltage violation (i.e. overvoltage in this scenario), regardless of the location. On the one hand, the region delimited by both limits implies some penetrations are acceptable but remain location-specific while on the other hand all PV penetration level greater than the max. HC limit (red) will result in overvoltage.

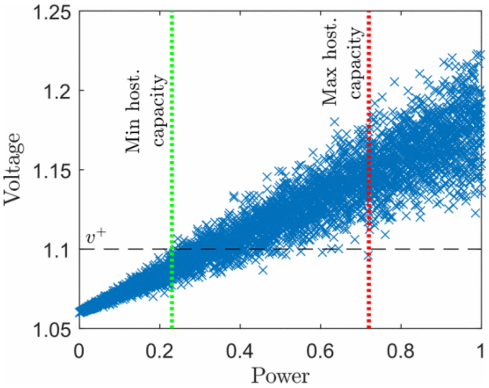


Figure 1. Hosting capacity limits in the case of random allocation of PV systems with green: the lower limit (min.) and red: the upper limit (max.), where v^+ denotes the voltage limit in [p.u] according to the IEC EN50160 standard.

However, the proliferation of DERs as well as the electrification of the mobility is creating significant challenges for distribution system operators. Due to their stochastic nature e.g. variable production or unpredictable charging behaviour of electric vehicles (EVs), the deterministic or probabilistic HC approaches may no longer be valid. Instead time-varying approaches are required to define the HC. In other words, it becomes a temporal-spatial limit that reflects the additional load or injection a grid can accommodate at a specific time, we denote it as the "*dynamic hosting capacity*" or DHC [5, 6]. Interested readers are referred to Table 1, that offers a sample overview of the available literature on the subject.

Table 1. The following overview present a review of the available literature on the topic of hosting capacity for the different types of approaches, with **D**: deterministic, **P**: probabilistic and **T**: time series method.

Ref.	**Type**	**Simulation**	**LCT**	**Voltage**	**Studied impacts**
[7]	D	Power flow	PV	LV	Line loading and voltage magnitude
[8]	D	Power flow	PV	MV/LV	Voltage magnitude
[9]	P	Power flow	PV	MV	TRAFO loading and voltage magnitude
[10]	P	Power flow	PV, EV	LV	Voltage magnitude
[11]	D	Power flow	PV	MV/LV	OLTC operation and voltage magnitude
[12]	P	Monte Carlo	PV	LV	Voltage magnitude and unbalance
[13]	T	Power flow	EV	LV	Voltage unbalance
[14]	T	Power flow	EV, HP	LV	Cable and TRAFO thermal violations, voltage magnitude
[15]	T	Power flow	PV, EV, HP	LV	Cable and TRAFO thermal violations, voltage magnitude
[16]	P, T	Monte Carlo	PV, EV, HP	LV	TRAFO loading and voltage magnitude
[17]	T	Quasi-static power flow	PV, HP	LV	Cable and TRAFO thermal violations, voltage magnitude and unbalance

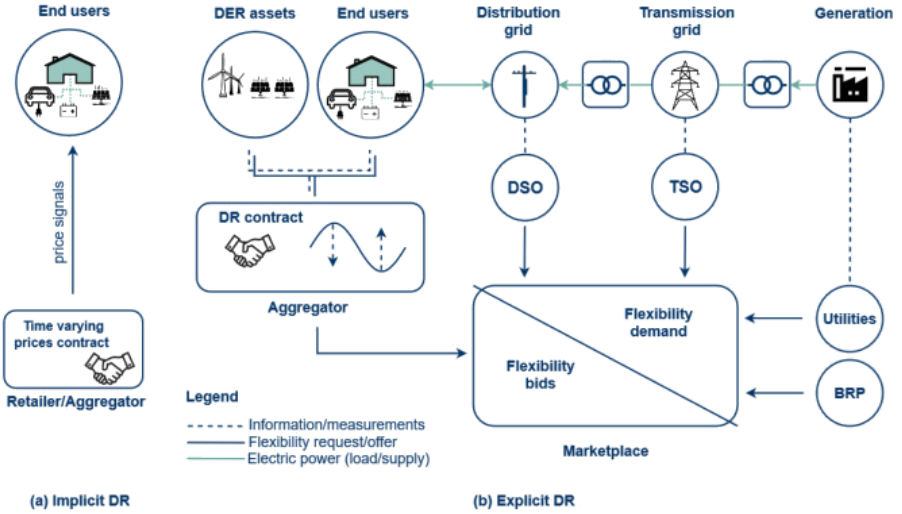
Flexibility: Implicit and Explicit

In the latest CIRED flexibility working group's publication entitled "Network planning and system design with flexibility" [18], the concept of low-voltage flexibility is defined as:

> '*[…] a power modulation of any flexible resources in voluntary response to a need (i.e. a signal). This response allows the power system operator or other third party to optimize their operation condition (e.g. costs, voltage profile) without affection their security and reliability.*'

Depending on the point of view taken, this flexibility can be seen from the user's perspective as the ability to modulate, shift or reduce the power demand or curtail the injected power into the grid for a certain time duration. While from a distribution system operator's (DSO) perspective, this suggests that the available (hosting) capacity of the network is optimally utilised or even that extra capacity is made available.

In this regard, the flexibility provider (e.g. DER asset owner) can either (i) respond to external signals such as prices (e.g. dynamic tariff, imbalance price, etc.) which is referred to as the implicit flexibility. Another option is (ii) to have a bilateral contract with a third party (e.g. aggregator, or flexibility service provider (FSP)) to offer explicitly flexibility.



(a) Implicit DR

(b) Explicit DR

Figure 2. Overview of the demand-response (DR) processes, with (a) the implicit DR and (b) the explicit DR, source: [1].

Dynamic operating envelopes

The proliferation of LCTs including: heat pumps, EVs and DERs has prompted DSOs to consider transitioning from the contemporary fixed network HC limits governing the withdrawal or injection of power towards more flexible limits. True, traditional HC studies typically involve the study of two worst-case snapshots analysing on the one hand the case of max. demand, and on the other hand max. generation within a network. In contrast, the flexible limits, designated as dynamic operating envelopes (DOEs), represent a convex set of active and reactive power that can be consumed or generated while preserving the network within its safe operational range [19, 20, 21]. Consequently, DOEs determine feasible time-varying import and export limits in real-time or in advance (e.g. 24-hour forecast) at the customer's connection point or for an aggregation pool of customers in a designated area, considering the local network's capacity. Therefore, the utilisation of DOEs facilitates the optimal integration of LCTs in a grid, while avoiding any network violations (e.g. over- and undervoltage, thermal stress of lines and transformers). Finally, the calculated envelopes are then disseminated to aggregators who must adhere to them when managing their DER portfolios.

It is important to recognise that DOEs are calculated from the point of interest of the network's safety, which may result in limitations when considering the market participation of controllable assets. Additionally, the extensive adoption of LCTs may have led in certain grid segments experiencing violations of the network constraints. One should note that DOEs are unable to address these pre-existing network issues [22].

Active distribution networks

An active distribution network (ADN) is defined as a network encompassing SDGEs with a control system where DSOs are capable of proactively operating the network, guaranteeing a safe and economical operation. The concept of ADNs relies on advanced data management methods (among other advanced metering infrastructure (AMI)) and information and communication technologies (incorporating two-way communication) to coordinate the distribution-side resources on a centralised platform. Some of the methods applied include demand response (DR) programs, voltage regulators, on-load tap changers and network reconstructions. In DR programs, the enrolled customers are compensated for the capacity that they provide [23].

Grid stability metrics

In an effort to evaluate the technical impacts caused by LCTs, multiple grid stability metrics that reflect both voltage and thermal implications are defined, we refer to the limitations imposed by the IEC EN50160 [24]. Among these indicators are: (i) voltage deviation (i.e. under- and overvoltage), (ii) voltage unbalance, (iii) line overloading, (iv) transformer overloading and (v) network losses:

(i) In accordance with the standard, the 10min. rms voltages are required to remain within the specified bounds of 0.9 and 1.1 p.u. for 95% of the time during each period of one week, i.e. $v^- \leq V_i \leq v^+$. Also, the 10min. rms-values shall not exceed 0.85 and 1.1 p.u.

(ii) Voltage unbalance factor (VUF): under normal operating conditions, the 10min. rms-values of the negative phase sequence component of the voltage shall be within the range of $0-2\%$ of the positive phase sequence for each period of one week. Considering branch i with phase terminals a, b, c and n, the positive (V_i^+), negative (V_i^-) and zero (V_i^0) sequence components are defined,

$$V_i^0 = \frac{1}{3}V_{i_{an}} + \frac{1}{3}V_{i_{bn}} + \frac{1}{3}V_{i_{cn}} \tag{1.1}$$

$$V_i^+ = \frac{1}{3}V_{i_{an}} + \frac{a}{3}V_{i_{bn}} + \frac{a^2}{3}V_{i_{cn}} \tag{1.2}$$

$$V_i^- = \frac{1}{3}V_{i_{an}} + \frac{a^2}{3}V_{i_{bn}} + \frac{a}{3}V_{i_{cn}} \tag{1.3}$$

where $\boldsymbol{a} = \angle 120° = -0.5 + \boldsymbol{j}\frac{\sqrt{3}}{2}$ and $\boldsymbol{a^3} = 1$. VUF is obtained using terms (1.2)–(1.3):

$$VUF = \frac{V_i^-}{V_i^+} \tag{2}$$

(iii) Line overloading refers to exceeding the cable's nominal current capacity, this can be evaluated by considering the hourly average current flowing through the conductors.

(iv) Transformer overloading (T^{OL}) is defined as the hourly average apparent power $\left(S_T^{avg_{1h}}\right)$ divided by the transformer rating, i.e. nominal capacity $\left(S_T^{rated}\right)$ [17], the hourly values are obtained by applying a moving average filter based on the granularity of the dataset used (e.g. 15min. resolution). In case T^{OL} is greater than 0, the transformer is overloaded.

$$T^{OL} = max\left(\frac{\left|S_T^{avg_{1h}}\right|}{S_T^{rated}} - 1, \quad 0\right) \tag{3}$$

To reflect the severity associated with an overload, we introduce σ as the severity factor,

$$\sigma(T^{OL}) = \begin{cases} \text{Permissible} \quad \text{if} \quad T^{OL} \leq 0.10 \\ \text{Critical} \quad \text{if} \quad 0.10 < T^{OL} \leq 0.20 \\ \text{Intolerable} \quad \text{if} \quad T^{OL} > 0.20 \end{cases} \tag{4}$$

(v) Finally, we define the network (or power) losses[1] as variable since they are function of the power flow and are created by the heating effect of current passing through the lines. These losses vary proportionally to the square of the current [25],

$$P_{loss} = \sum_{l \in \mathcal{L}} \Re(V_l \cdot I_l^*) \tag{5}$$

Within ECOFLEX, we consider limitations related to transformer's overloading, power losses, under- and overvoltage violations and voltage unbalance where a distinction is made between the frequency of occurrences and the severity of a violation.

2.2. State-of-the-art modelling approaches

Low-voltage distribution system modelling

The majority of the LVDS models employed for power flow studies assume that the neutral conductor is multi-grounded and that the network is balanced. However, the latter is seldom valid due to the single phase domestic appliances connected to the system. Moreover, multi-grounded neutral conductors are not applicable for most of the European countries, including Belgium. This assumption gives rise to a common practice of reducing the three-phase four-wire model into a three-wire model using Kron's reduction, whereby the resulting impedance matrix is of the shape 3 x 3. This reduces the computational complexity of the power flow by eliminating the neutral voltage and current variables. Although this can be done for balanced multi-grounded networks, this approach is not valid in a three-phase four-wire model with isolated neutral, Figure 3.

Fundamentally, Kron's reduction presupposes that the current flowing through the neutral is zero, thus underestimating the neutral voltage rise and ignoring the risk of zero-point shifting [26]. The topic is well addressed in the literature, for instance Ciric et al. [27] propose a power flow algorithm for three-phase four-wire radial networks considering the neutral wire. While Claeys et al. [28] conclude that Kron's approximations diverge if the assumptions do not hold (cf. four-wire model) and therefore the resulting errors can be cumbersome in decision-making contexts. Furthermore, the authors highlight that ignoring the unbalance can lead to violations of the voltage limits and an underestimation of the network losses. In [29], the authors developed a phase-to-neutral transformation that reduces the lines impedance matrices from 4 x 4 to 3 x 3, while providing the possibility to recover the neutral voltage variables.

---

[1] Network losses are affected by a variety of sources, i.e. technical, non-technical and other factors. Technical losses also include the losses induced by energizing transformers. Non-technical losses on the other hand refer to theft or unmetered supplies. Other factors affecting network losses include phase imbalance and harmonics.

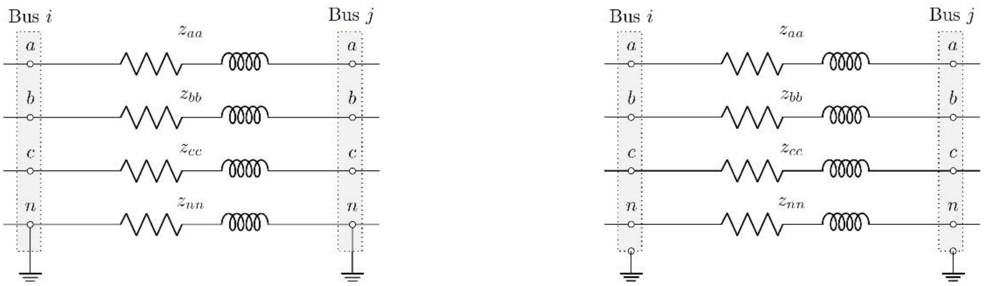


Figure 3: Representation of line l_{ij}, with multi-grounded (left) and isolated neutral (right).

Physics-based or data-driven approaches

It can thus be concluded that conventional physics-based strategies require in-dept knowledge of the line parameters (i.e. length and cable cross-sections) [30], grid topology and phase arrangements in order to perform accurate hosting capacity studies. In the event that any of the aforementioned parameters are not available, it will not be possible to perform an accurate calculation and assumptions must therefore be made. In response to the emergence of digitalisation, promising alternatives are being developed, including so-called model-free approaches [31]. In the context of distribution networks, these data-driven approaches rely on smart meter and/or AMI[2] data. Historical smart meter measurements (at customer level) incorporate instantaneous or average values for the voltage (V) and current (I) magnitudes, on top of the active and reactive power, (P and Q respectively). This data is extracted at varying time intervals, ranging from seconds to minutes in accordance with the specific local necessities and the country in question, but it can be reasonably presumed that this occurs on a 15-minute basis. One possible objective of the data-driven approach is to calculate the voltage among the nodes by exploiting the historical data (i.e. V, P and Q) through regression models [32]. The gathered data may also be employed in the context of other use cases that extend beyond forecasting, with the prospective application of fraud detection.

Hosting capacity expressions

As discussed in Section 2.1, depending on the scope of the hosting capacity study, a different approach can be addressed. In an effort to develop generic mathematical formulations for the HC of PV systems, Koirala et al. [33] present various concepts that can be used for both deterministic and stochastic HC. The expression for a feeder's directional hosting capacity (HC_f) is defined as the sum of additional PV capacity (p) of a subset of customers ($\mathcal{D}_f$) that can be accommodated by a feeder without exceeding the admissible bounds:

$$HC_f = \sum_{p \in \mathcal{P}} p \qquad \text{s.t.,} \quad |\mathcal{P}_f| \leq |\mathcal{D}_f| \tag{6}$$

where, $\mathcal{P}_f$ denotes the set of new PV capacity in feeder f and with the reasonable assumption that the set of new PV capacity is mapped one-to-one to the set of LV customers $\mathcal{D}_f$, thus the cardinality of $\mathcal{P}_f$ is less or equal to 1. Nonetheless, the HC is often expressed relatively to the transformer capacity (S_T^{rated}). In that particular case, the additionally installable PV capacity is normalised to the LVDS transformer ratings within the studied LV network $\mathcal{M}$:

$$HC_t = \frac{\sum_{f \in \mathcal{M}} HC_f}{\sum_{t \in \mathcal{M}} S_T^{rated}} \tag{7}$$

---

[2] Advanced Metering Infrastructure (AMI) is an integrated system that enables two-way communication and has the ability to collect energy usage data.

Deterministic hosting capacity formulations are valid for the worst-case scenarios, resulting in the typical fit-and-forget approach DSOs apply. Within these expressions, the considered variable is often the penetration ratio (γ) of the studied technology and is defined as the proportion of LCTs that are connected to a load location. To illustrate this in the context of PV installations, let N_p represent the number of PV installations [4]:

$$\gamma = \frac{N_p}{|\mathcal{D}_f|} \tag{8}$$

Furthermore, it is assumed that every PV installations within the subset has the same size (in [kWp]). In contrast, probabilistic hosting capacity studies utilise uncertainties and therefore stochasticity is added to the assumptions such as the location of the PV installations and the PV capacity vary as well as the production profile.

2.3. Challenges and uncertainties

As stipulated earlier, the choice of a modelling approach is pivotal to achieving an accurate representation of the system under consideration. Consequently (referring to previous section, 2.2) the implementation of Kron's reduction can significantly impact the outcomes. Additionally, the choice of simulation software needs to be envisaged, as some tools are inherently designed to apply Kron's reduction when executing power flows. In most cases, this will result in a trade-off between accuracy of the model and the computational requirements. Another challenge linked to the modelling is the observability of the LVDS. True, unlike transmission systems, distribution networks are the black boxes of the grid. Due to their size and complexity, existing models or information gathered of the LVDS are often incomplete or erroneous. Examples related to the topology are the incorrect assignment of phases at the consumers or service cables to the correct distribution lines. Further, the number of metering instruments is small compared to the size of the network although this tends to change with the smart meter deployment. Finally, despite the fact that the equations used to determine hosting capacity are well established, there is no consensus on the definition of this term. Consequently, the scientific literature does not provide a concise definition of what is meant by HC. This brings us to the first challenge, which is a streamlined view on the modelling techniques due to the gap in a unified interpretation of HC.

3. Methodology

3.1. Physics-based approach

Multiple approaches are used when doing load flow analysis, here we will briefly introduce two most commonly used ones, i.e. the *Bus Injection model* (BIM) and the *Branch Flow model* (BFM). To illustrate the equations, let us consider the following segment of a network that consists of a set of nodes/buses $\mathcal{N}$ (where buses are indexed $i = 0,1,\dots,n$) and a set of distribution lines $\mathcal{L}$ connecting these buses. We denote a line l by the pair of buses it connects, e.g. (i,j) becomes l_{ij}.

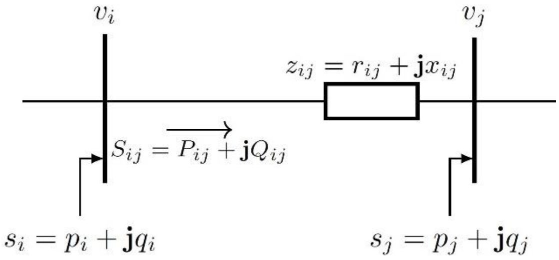


Figure 4: Illustrative example of a line segment l_{ij} feeding bus j from its parent bus i.

For convenience, let bus 0 represent the slack node or point of common coupling (PCC), which has a fixed voltage and variable active and reactive power injection. Further, it is assumed that the complex voltage on the substation bus (V_0) is known. We refer to equations (9)–(11) as the BFM [34]. Let z_{ij} be the impedance on line l_{ij}, S_{ij} the complex power flowing through the line and s_j be the complex net load minus generation on bus j. At steady state, the power flow satisfies the power balance at each bus $j \in \mathcal{N}$ with $(i,j) \in \mathcal{L}$, s.t.:

$$s_j = \left(S_{ij} - z_{ij}^2 \left|I_{ij}\right|^2\right) - \sum_{n:(j:n)\in\mathcal{L}} S_{jn} \tag{9}$$

and according to Ohm's law, the voltage drop at each line can be written as:

$$V_i - V_j = z_{ij} I_{ij} \tag{10}$$

where V_i denotes the complex voltage at bus i. Finally, the complex power on each line $l_{ij} \rightarrow (i,j) \in \mathcal{L}$ is given by (11) with I_{ij}^* the complex conjugate of the current from bus i to j:

$$S_{ij} = V_i I_{ij}^* \tag{11}$$

For notational simplicity we define the squared voltage and current magnitude, $v_i \coloneqq |V_i|^2$ and $\ell_{ij} \coloneqq \left|I_{ij}\right|^2$ respectively. By introducing these simplifications, we can reduce the expressions to the so-called *relaxed branch flow model* (Eq. (12)–(14.1)) if the studied network is radial:

$$p_j = P_{ij} - r_{ij}\ell_{ij} - \sum_{n\epsilon\mathcal{N}(j)} P_{jn} \tag{12}$$

similarly, we obtain for the reactive power flow:

$$q_j = Q_{ij} - x_{ij}\ell_{ij} - \sum_{n\epsilon\mathcal{N}(j)} Q_{jn} \tag{13}$$

Lastly, the voltage at a given bus $j \in \mathcal{N}$ can be retrieved:

$$V_j^2 = V_i^2 - 2\left(r_{ij}P_{ij} + x_{ij}Q_{ij}\right) + \left(r_{ij}^2 + x_{ij}^2\right)I_{ij}^2 \tag{14.1}$$

Replacing the voltage and current magnitude terms with their squared equivalents:

$$v_j = v_i - 2\left(r_{ij}P_{ij} + x_{ij}Q_{ij}\right) + \left(r_{ij}^2 + x_{ij}^2\right)\ell_{ij} \tag{14.2}$$

As it can be noted from above equations, the BFM are accurate models for both radial and meshed networks accounting for the active and reactive power, voltage magnitudes and phase angels. Therefore, it allows the expressions to be formulated as a set of nonlinear equations at each bus. On the other hand, the relaxed BFM define a system of equations which are a subset of the complex variables without the phase angels. Due to the nonlinearity of the power flow equations, an iterative solution method such as Newton-Raphson or Gauss-Seidel is required.

In the event that the system under study consists of a single source (or substation), an approximation of the BFM is frequently employed. These simplified equations, designated as the *DistFlow equations*, are applicable to LVDS wherein typically the shunt elements are disregarded, thereby reducing the computational burden of the calculations (Eq. (14.2)). Although they are approximations, the application of DistFlow within smart grid studies, e.g. for DER implications are frequent. Their goal is to calculate the voltage, current and power flow between buses using active and reactive power along with the line impedances.

Further simplifications by neglecting losses on both series and shunt admittance, and assuming the shunt admittance of each line to be zero, results in Linear Distflow equations [35]:

$$v_j = v_i - 2\left(r_{ij}P_{ij} + x_{ij}Q_{ij}\right) \tag{15}$$

In contrast to the BFM, the Bus Injection model focuses on the bus-level quantities (i.e. active and reactive power injections at each bus) to perform power flow studies [36]. Since the main variables of interest in the BIM are the power injections, voltage magnitudes and angles, this method is typically applied for power flow studies at transmission level. Let G_{ij} and B_{ij} be the real and imaginary parts of the admittance matrix; θ_{ij} is the angle difference of the branch (i,j). The general power flow expressions of the BIM model are expressed in terms of:

$$P_i = V_i + \sum_{j\epsilon\mathcal{N}(i)} V_j\left(G_{ij}cos\theta_{ij} + B_{ij}sin\theta_{ij}\right) \tag{16}$$

and

$$Q_i = V_i + \sum_{j\epsilon\mathcal{N}(i)} V_j\left(G_{ij}cos\theta_{ij} - B_{ij}sin\theta_{ij}\right) \tag{17}$$

While the voltage angle difference is derived as follows:

$$\theta_{ij} = \theta_i - \theta_j \tag{18}$$

Let r_{ij} and x_{ij} denote the resistance and reactance of branch (i,j), following the definition of the admittance matrix G_{ij} and B_{ij} are obtained by:

$$G_{ij} = \frac{r_{ij}}{r_{ij}^2 + x_{ij}^2} \tag{19}$$

and

$$B_{ij} = \frac{x_{ij}}{r_{ij}^2 + x_{ij}^2} \tag{20}$$

3.2. Data-driven approach

In recent years, the increased interest in and development of artificial intelligence has coincided with a proliferation of smart meters, which has generated considerable interest within the energy sector. This, along with the substantial amounts of data generated by the smart meters has paved the way for data-driven models in the field of power flow analysis. A sample overview of the possible methods is given in Table 2. In this work, we only discuss multivariable linear regression (MLR).

In their original form, the MLR can be written as (21), let $\hat{y}_i$ be the dependent variable, β_0 the y-intercept term, m_{ik} the independent (explanatory) variables, β_k the slope coefficients for each explanatory variable and ϵ the error term (i.e. the residuals),

$$\hat{y}_i = \beta_0 + \beta_1 m_{i1} + \beta_2 m_{i2} + \cdots + \beta_k m_{ik} + \epsilon \tag{21}$$

Since the voltage at a branch is influenced by the active and reactive power of the parent nodes as well as the studied node, the BFM can be rewritten in terms of a MLR. In that case the non-linearities of the model such as a variable slack bus voltage are omitted and the resulting expressions are a valid approximation of the true voltages. Let $\hat{V}_{i,t}$ be the predicted voltage of the i-th component of the dependent variable $V_{i,t}$ at timestep t and $let\ x_{i,t}$ denote the active $\left(p_{i,t}\right)$ and reactive power $\left(q_{i,t}\right)$ of the i-th component at timestep t [37]:

$$\sum_{i=1}^{n} \hat{V}_{i,t} = \beta_0 + \sum_{i=1}^{n} \beta_i x_{i,t} + \epsilon_{i,t}, \qquad \text{s.t.} \quad x_{i,t} = \left[p_{i,t}, q_{i,t}\right] \tag{22}$$

The present models utilise this methodology, showcasing a LVDS can be approximated by linear expressions if the slack bus voltage is esteemed to be constant as well as a complete smart meter coverage within the studied grid. Incomplete data will result in a loss of accuracy.

Table 2: State-of-the-art forecasting of voltage and congestion through data-driven approaches with, **ANN**: artificial neural network, **BRT**: bagging regression trees, **MLR**: multivariable linear regression, **NN**: neural network, **RF**: random forests and **SVR**: support vector machines.

Ref.	Type	Simulation	LCT	Studied impacts
[31]	NN	Power flow	PV	Forecasting of voltage profiles for PV penetration
[32]	MLR	Power flow	EV, PV	Forecasting of congestion with EV, PV integration
[38]	SVR	Power flow	PV	Evaluation of the SVR model to predict power flows
[39]	RF, BRT	Power flow	PV	Voltage estimation with DERs
[40]	ANN	Power flow	EV	Voltage prediction in the context of V2G studies

Authors in [41] demonstrate that the use of a two-step regressor will significantly reduce the prediction errors. Therefore, the authors compared a multitude of machine learning-based algorithms including AdaBoost, bagging, Bayesian ridge, gradient boosting, and others in lieu of a multivariable linear regression and evaluated the obtained results in terms of a comparison with an ensemble regressor and a two-step regressor. To conclude, future work will consist of expanding the current framework in order increase the robustness of the used models based on the findings from aforementioned studies.

3.3. Sample use case

A traditional low-voltage distribution network comprises a three-phase four-wire system operated radially where the line-to-line voltages are 400V. In Belgium, the neutral conductor is grounded at the star point of the LV distribution transformer. Since most customers connected to the low-voltage are single phase, it can be assumed that LVDS are unbalanced. This makes available models such as the IEEE European Low Voltage Test Feeder (LVTestFeeder), where the neutral conductor is sparsely grounded at a few nodes, unsuitable for this study in its actual configuration.

The proposed methodology with MLR has proven to be a valuable alternative for the conventional physics-based power flow analysis. In terms of scalability, the model is being evaluated against a larger network encompassing 5 feeders modelled in detail and a total of 91 dwellings shown in Figure 5. To demonstrate its performance, various levels of SDGE penetrations are predicted via the MLR approach and compared to the results obtained by the conventional model-driven approach. The training data for the MLR is obtained through a quasi-static time series (QSTS) power flow simulation in OpenDSS without SDGEs, later defined as the benchmark. A QSTS simulation refers to a versatile study that solves a series of sequential steady-state power flow solutions, wherein the converged state of each iteration serves as the initial condition for the subsequent one. In our case, the iteration step is set to 15min. This approach allows to capture time-varying aspects such as the state-of-charge of the EVs.

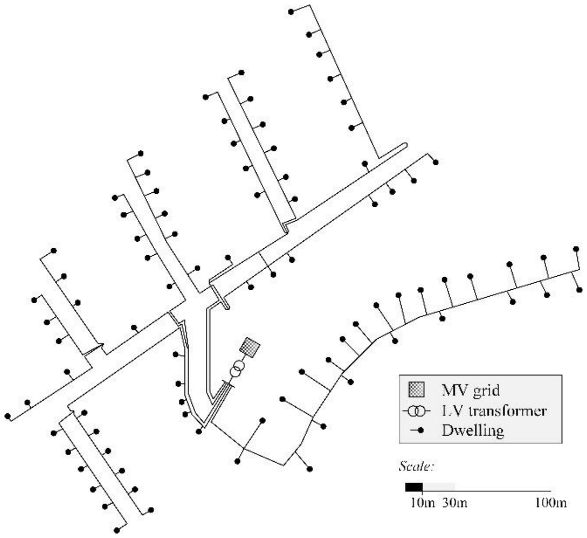


Figure 5: Illustration of the four-wire radial low-voltage distribution network

We define the SDGE levels as continuous uniform distribution {25, 50, 100}%. Here, the coverage i.e. the penetration ratio (γ) is always assumed to be 100 percent, see Eq. (8).Thus a SDGE level of 25 percent reflects a ratio of 25/100 to the max. power consumed or generated by the SDGE, compared to its nominal capacity.

For instance, if the EVs charge at 7,4 kW, a 25% penetration rate denotes that the max. charging power of the EVs is (7,4 · 25/100) kW while the coverage of dwellings having an EV is 100%. This, to avoid stochasticity in the model since the goal is to evaluate the data-driven approach rather than pure HC analysis. The outcome of the simulations is presented in Figure 6 where both, a PV and EV assessment is done.

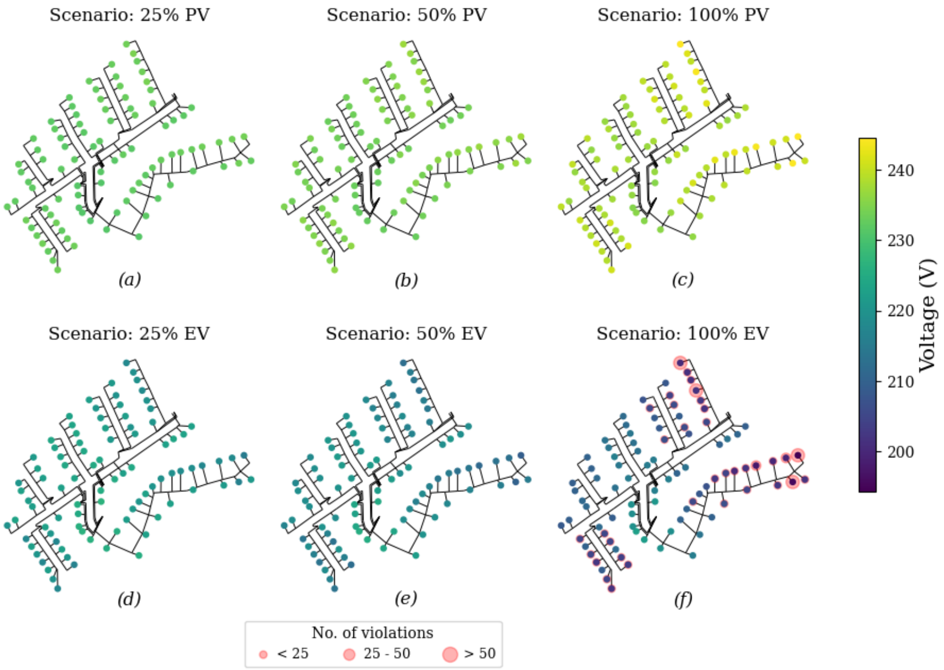


Figure 6: Forecasted voltages via the data-driven method for various SDGE levels {25, 50, 100}% with the upper subplots showcasing the voltage evolution variation originating from PV systems, while the bottom subplots highlight the voltage violations due to EV integration.

As expected, an increased PV generation will result in a greater voltage rise. Due to Ohm's law the voltage rise will manifest itself at the feeder ends. This is particularly the case for longer feeder such as the bottom right of the figure where the feeder does not consist of sub-feeders leading towards a linear voltage rise due to the reverse power flows.

Analogously, due to the high load of EVs, the voltage will drop significantly at feeder ends as a function of the penetration rate. For the particular case under study, it can be noticed that an increase in EVs will have a greater impact on the voltage deterioration, mostly caused by the high charging rates and the simultaneity of the charging. This is also visible in Figure 6 (*f*) where voltage violations are noticeable, i.e. voltage drops below the v^- limit are monitored.

Obtained results are subsequently compared to the outcome of the physics-based model, executed in simulation software OpenDSS through a python interface. Although, it was formerly expected that the outcome of the data-driven model would only reflect accurately the benchmark case, the model excelled in forecasting the voltage for unseen scenarios (e.g. various SDGE penetration rates). It should be noted that this is partially due the aforementioned assumptions: constant slack bus voltage, a complete smart meter coverage and fixed topology, that high accuracy is obtained. Moreover, the model has been benchmarked against other data-driven approaches such as SVM, NN, HistGradientBoosting and unexpectedly outperformed the other models, as shown in Figure 7:

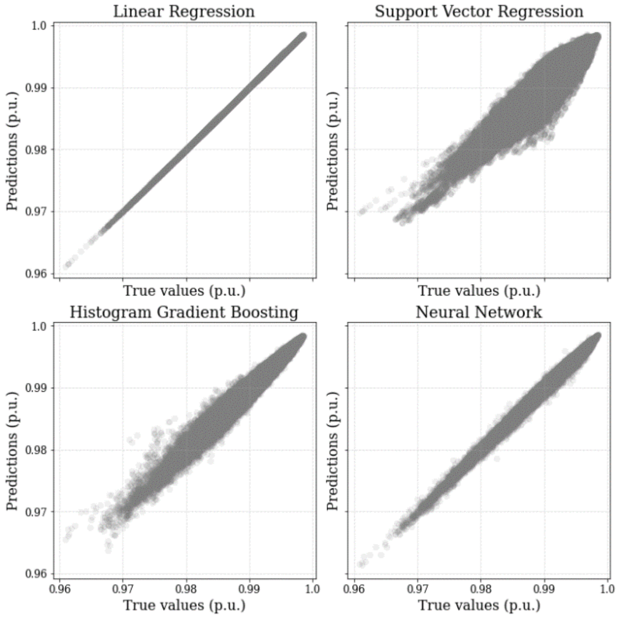


Figure 7: Review of the accuracy for different data-driven models.

4. Conclusions

This deliverable delineates the fundamental concepts pertinent to the modelling of low-voltage distribution networks. In order to achieve this, the most commonly used terminology was addressed. This included terms such as dynamic hosting capacity, implicit and explicit flexibility, dynamic operating envelopes and active distribution networks. These were described based on the existing literature and the author's own interpretations, in the absence of a consensus on a generic definition. To this end, a comprehensive review of the state-of-the-art definition of hosting capacity is given. Moreover, the diverse grid stability metrics employed within the ECOFLEX project are examined, and their mathematical interpretation is provided.

In addition to the aforementioned concepts and nomenclature, Section 2.2 provides a rigorous overview of the state-of-the-art modelling approaches for low-voltage distribution systems. These range from the Kron reduction to physics-based and data-driven approaches, and include a detailed examination of the challenges associated with modelling. Hence, depending on the data availability (e.g. complete or partial coverage) and the required accuracy of the digital-twin model, either physics-based or data-driven approaches can be applied.

Finally, Section 3 outlines the conventional methodology of calculating power flows through physics-based approaches, i.e. the branch flow model and the bus injection model. It also provides a brief overview of potential data-driven techniques, with a literature review and an explanation of the adopted multivariable linear regression model. To conclude the chapter, an illustrative use case is presented comprising a low-voltage distribution network with five feeders and 91 dwellings. This example demonstrates the application of a data-driven approach to evaluate and monitor voltage congestion if stochastic distribution grid exchange are introduced. As highlighted in the results of the study, one can conclude that the data-driven approaches provide a promising alternative to a fast evolving sector where sensors and metering devices are being installed the clock round. Making space for novel algorithms to evaluate the future shortcomings and needs of a network's HC.

Nonetheless, future prospects for simulations should include a variable slack bus voltage and an assessment of the implications that a partial smart meter coverage has. This is of major importance since the current version of the developed models rely on a full coverage of smart metering, while in reality a lot of DSO do not reach these numbers. Furthermore, due to technical failures the data often is incomplete or erroneous while for this study we omitted data limitations. Besides the influence of the input data and training length (e.g. 1 month, 3 month or full year), future work shall include a more probabilistic approach to evaluate the hosting capacity and propose viable solutions as alternatives to grid traditional grid reinforcements in case of HC limitations.

Unleashing the potential of the data-driven approaches to perform hosting capacities studies can be beneficial for both the distribution network operator in terms of grid stability and monitoring, but also for the solution providers who aim to develop novel energy management systems incorporating intelligent services to provide flexibility and thus optimize the use of the available HC.

5. References

[1] F. D'Ettorre, M. Banaei, R. Ebrahimy, S. Ali Pourmousavi, E. Blomgren, J. Kowalski, Z. Bohdanowicz, B. Łopaciuk-Gonczaryk, C. Biele and H. Madsen, "Exploiting demand-side flexibility: State-of-the-art, open issues and social perspective," *Renewable and Sustainable Energy Reviews,* vol. 165, p. 112605, 2022.

[2] P. M. S. Carvalho, L. A. F. M. Ferreira and J. J. E. Santana, "Single-Phase Generation Headroom in Low-Voltage Distribution Networks Under Reduced Circuit Characterization," *IEEE Transactions on Power Systems,* vol. 30, no. 2, pp. 1006-1011, 2015.

[3] E. Mulenga, M. H. J. Bollen and N. Etherden, "A review of hosting capacity quantification methods for photovoltaics in low-voltage distribution grids.," *International Journal of Electrical Power & Energy Systems,* vol. 115, p. 105445, 2020.

[4] M. Deakin, C. Crozier, D. Apostolopoulou and T. Morst, "Stochastic Hosting Capacity in LV Distribution Networks," in *IEEE Power & Energy Society General Meeting (PESGM),* Atlanta, GA, USA, 2019.

[5] T. E. Castelo de Oliveira, M. Bollen, P. F. Ribeiro, P. M. S. de Carvalho, A. C. Zambroni and B. D. Bonatto, "The concept of dynamic hosting capacity for distributed energy resources: Analytics and practical considerations.," *Energies,* vol. 12, no. 13, p. 2576, 2019.

[6] I. Sherif M., S. H. E. Adbel Aleem, A. Y. Abdelaziz and A. F. Zobaa, "State-of-the-art of hosting capacity in modern power systems with distributed generation.," *Renewable energy,* vol. 130, pp. 1002-1020, 2019.

[7] R. A. Shayani and M. A. G. de Oliveira, "Photovoltaic generation penetration limits in radial distribution systems," *IEEE Transactions on Power Systems,* vol. 26, no. 3, pp. 1625-1631, 2010.

[8] R. Tonkoski, D. Turcotte and T. H. M. El-Fouly, "Impact of High PV Penetration on Voltage Profiles in Residential Neighborhoods," *IEEE Transactions on Sustainable Energy,* vol. 3, no. 3, pp. 518-527, 2012.

[9] M. Mirbagheri, M. Moncecchi, D. Falabretti and M. Merlo, "Hosting capacity evaluation in networks with parameter uncertainties," in *2018 18th International Conference on Harmonics and Quality of Power (ICHQP),* Ljubljana, Slovenia, 2018.

[10] R. Fachrizal, U. Hanig Ramadhani, J. Munkhammar and J. Widén, "Combined PV–EV hosting capacity assessment for a residential LV distribution grid with smart EV charging and PV curtailment," *Sustainable Energy, Grids and Networks,* vol. 26, p. 100445, 2021.

[11] S. Hashemi, J. Østergaard, T. Degner, R. Brandl and W. Heckmann, "Efficient Control of Active Transformers for Increasing the PV Hosting Capacity of LV Grids," *IEEE Transactions on Industrial Informatics,* vol. 13, no. 1, pp. 270-277, 2017.

[12] R. Torquato, D. Salles, C. Oriente Pereira, P. Cesar Magalhaes Meira and W. Freitas, "A Comprehensive Assessment of PV Hosting Capacity on Low-Voltage Distribution Systems," *IEEE Transactions on Power Delivery,* vol. 33, no. 2, pp. 1002-1012, 2018.

[13] N. Jabalameli, X. Su and A. Ghosh, "Online Centralized Charging Coordination of PEVs With Centralized Var Discharging for Mitigation of Voltage Unbalance," *IEEE Power and Energy Technology Systems Journal,* vol. 6, no. 3, pp. 152-161, 2019.

[14] C. Edmunds, S. Galloway, J. Dixon, W. Bukhsh and I. Elders, "Hosting capacity assessment of heat pumps and optimised electric vehicle charging on low voltage networks," *Applied Energy,* vol. 298, p. 117093, 2021.

[15] N. Damianakis, G. Ram Chandra Mouli, P. Bauer and Y. Yu, "Assessing the grid impact of Electric Vehicles, Heat Pumps & PV generation in Dutch LV distribution grids," *Applied Energy,* vol. 352, p. 121878, 2023.

[16] A. Navarro-Espinosa and L. F. Ochao, "Probabilistic Impact Assessment of Low Carbon Technologies in LV DIstribution Networks," *IEEE Transactions on Power Systems,* vol. 31, no. 3, pp. 2192-2203, 2016.

[17] S. Meunier, C. Protopapadaki, R. Baetens and D. Saelens, "Impact of residential low-carbon technologies on low-voltage grid reinforcements," *Applied Energy,* vol. 297, no. 1, 2021.

[18] CIRED WG 2021-2, "Network Planning and System Design With Flexibility," CIRED, 2024.

[19] A. Attarha, M. Noori, P. Scott and S. Thiébaux, "Network-Secure Envelopes Enabling Reliable DER Bidding in Energy and Reserve Markets," *IEEE Transactions on Smart Grid,* vol. 13, no. 3, pp. 2050-2062, 2022.

[20] K. Petrou, M. Z. Liu, A. T. procopiou, L. F. Ochoa, J. Theunissen and J. Harding, "Operating Envelopes for Prosumers in LV Networks: A Weighted Proportional Fairness Approach.," in *IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe),* The Hague, Netherlands, 2020.

[21] G. Lankeshwara, R. Sharma, R. Yan, T. K. Saha and J. V. Milanović, “Time-Varying Operating Regions of End-Users and Feeders in Low-Voltage Distribution Networks,” *IEEE Transactions on Power Systems,* vol. 39, no. 2, pp. 4600-4611, 2024.

[22] A. Attarha, M. Noori, M. Mahmoodi, J. Iria and P. Scott, “Shaped operating envelopes: Distribution network capacity allocation for market services.,” *Electric Power Systems Research,* vol. 234, p. 110639, 2024.

[23] V. F. Martins and C. L. T. Borges, “Active Distribution Network Integrated Planning Incorporating Distributed Generation and Load Response Uncertainties,” *IEEE Transactions on Power Systems,* vol. 26, no. 4, pp. 2164-2172, 2011.

[24] International Electrotechnical Commission, *EN 50160:2011 - Voltage characteristics of electricity supplied by public distribution networks,* Geneva, CH: IEC, 2011.

[25] K. Girigoudar and L. A. Roald, “On the impact of different voltage unbalance metrics in distribution system optimization,” *Electric Power Systems Research,* vol. 189, p. 106656, 2020.

[26] F. Olivier, R. Fonteneau and D. Ernst, “Modelling of three-phase four-wire low-voltage cables taking into account the neutral connection to the earth,” in *CIRED Ljubljana Workshop,* Ljubljana, Slovenia, 2018.

[27] R. Ciric, A. Feltrin and L. Ochoa, “Power flow in four-wire distribution networks - General approach,” *IEEE Transactions on Power Systems,* vol. 18, no. 4, pp. 1283-1290, 2003.

[28] S. Claeys, M. Vanin, F. Geth and G. Deconinck, “Applications of optimization models for electricity distribution networks.,” *Wiley Interdisciplinary Reviews: Energy and Environment,* vol. 10, no. 5, 2021.

[29] F. Geth, R. Heidari and A. Koirala, “Computational analysis of impedance transformations for four-wire power networks with sparse neutral grounding.,” in *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems,* 2022.

[30] R. Cleenwerck, H. Azaioud, R. Claeys, T. Coosemans, J. Knockaert and J. Desmet, “An approach to the impedance modelling of low-voltage cables in digital twins,” *Electric Power Systems Reasearch,* vol. 210, p. 108075, 2022.

[31] V. Bassi, L. F. Ochoa, T. Alpcan and C. Leckie, “Electrical Model-Free Voltage Calculations Using Neural Networks and Smart Meter Data,” *IEEE Transactions on Smart Grid,* vol. 14, no. 4, pp. 3271-3282, 2023.

[32] R. Cleenwerck, W. Parys, J. Desmet and T. Coosemans, “Voltage congestion monitoring through machine learning,” in *27th International Conference on Electricity Distribution (CIRED),* Rome, Italy, 2023.

[33] A. Koirala, T. Van Acker, R. D'hulst and D. Van Hertem, “Hosting capacity of photovoltaic systems in low voltage distribution systems: A benchmark of deterministic and stochastic approaches,” *Renewable and Sustainable Energy Reviews,* vol. 155, p. 111899, 2022.

[34] L. Gan, N. Li, U. Topcu and S. Low, “Branch Flow Model for Radial Networks: Convex Relaxation,” in *51st IEEE Conference on Decision and Control (CDC),* Maui, HI, USA, 2012.

[35] H. Lin, X. Shen, Y. Guo, T. Ding and H. Sun, “A linear Distflow model considering line shunts for fast calculation and voltage control of power distribution systems,” *Applied Energy,* vol. 357, no. 122467, 2024.

[36] T. Ding, R. Lu, Y. Yang and F. Blaabjerg, “A Condition of Equivalence Between Bus Injection and Branch Flow Models in Radial Networks,” *IEEE Transactions on Circuits and Systems II: Express Briefs,* vol. 67, no. 3, pp. 536-540, 2020.

[37] R. Cleenwerck, M. A. Putratama, J. Desmet and T. Coosemans, “Smart Meter-Based Re-Phasing for Voltage Imbalance Enhancement Through Topology Reconstruction,” in *IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe),* Grenoble, France, 2023.

[38] X. Liang, H. Zhang, Q. Lui, Z. Liu and H. Liu, “A support vector regression-based interval power flow prediction method for distribution networks with DGs integration,” *Frontiers in Energy Research,* vol. 12, 2024.

[39] M. Markovíc, A. Sajadi, R. Cruickshank and B.-M. Hodge, “Voltage Estimation in Low-Voltage Distribution Grids With Distributed Energy Resources,” *IEEE Transactions on Sustainable Energy,* vol. 12, no. 3, pp. 1640-1650, 2021.

[40] B. Sah, P. Kumar and S. K. Bose, “A Fuzzy Logic and Artificial Neural Network-Based Intelligent Controller for a Vehicle-to-Grid System,” *IEEE Systems Journal,* vol. 15, no. 3, pp. 3301-3311, 2021.

[41] A. F. Bastos, S. Santoso, V. Krishnan and Y. Zhang, “Machine Learning-Based Prediction of Distribution Network Voltage and Sensor,” in *IEEE Power & Energy Society General Meeting (PESGM),* Montreal, QC, Canada, 2020.