A unified definition of hosting capacity, applications and review

Amina Benzerga[a,*], Alireza Bahmanyar[a], Guillaume Derval[a], Damien Ernst[a,b]

[a] *Department of Electrical Engineering and Computer Science, Liège, Belgium*
[b] *LTCI, Télécom Paris, Institut Polytechnique de Paris, France*

Abstract

Over the past decade, the landscape of distribution network hosting capacity has been extensively explored through numerous studies. However, a notable gap remains in the absence of a standardised and universally accepted generic definition. Bridging this gap, this paper provides a comprehensive and unified definition of the hosting capacity concept for distributed energy resources (DERs). To illustrate the presented definition, a practical demonstration using a basic illustrative example is provided. Additionally, this paper aligns the most impactful papers in the field, as well as select literature articles, with the proposed definition. This comparative analysis showcases the versatility and applicability of the definition across a spectrum of research contributions. This paper seeks to establish a shared foundation for the determination of hosting capacity that can guide future research.

Keywords: Distributed Energy Resources, Hosting capacity, Renewable Energy Source, Electric Vehicles, Heat Pumps, Photovoltaic, Wind Turbines.

Contents

---

[*] Corresponding author.

Email addresses: `abenzerga@uliege.be` (Amina Benzerga), `abahmanyar@uliege.be` (Alireza Bahmanyar), `gderval@uliege.be` (Guillaume Derval), `dernst@uliege.be` (Damien Ernst)

Preprint submitted to Renewable Energy *November 27, 2023*

1. Introduction

The ongoing energy transition is significantly changing production and consumption dynamics. The number of decentralised energy producers has been steadily increasing for several years. According to the International Energy Agency (IEA), photovoltaic panel (PV) generation increased by a record 270 TWh (up 26%) in 2022 [3]. On the other hand, high consumption technologies like heat pumps (HPs) and electric vehicles (EVs) are also experiencing growth, with the IEA anticipating penetration rates of at least 25% and 40% for 2050, respectively [3].

Technologies influencing the low-voltage distribution network such as photovoltaic panels, electric vehicles and heat pumps are referred to as Distributed Energy Resources, DERs. Figure 1 illustrates the terminology for common technologies. Producers, or distributed generators (DGs), involve for instance PV and wind turbines (WT), while new consumers involve both HPs and EVs. Battery energy storage systems (BESSs) enable energy from renewables to be stored and then released when it is needed most. In some cases, EVs can be considered as BESSs. DERs is the terminology used to address all decentralised technologies, thus all of the aforementioned technologies are DERs. This terminology is used for the remainder of the paper.

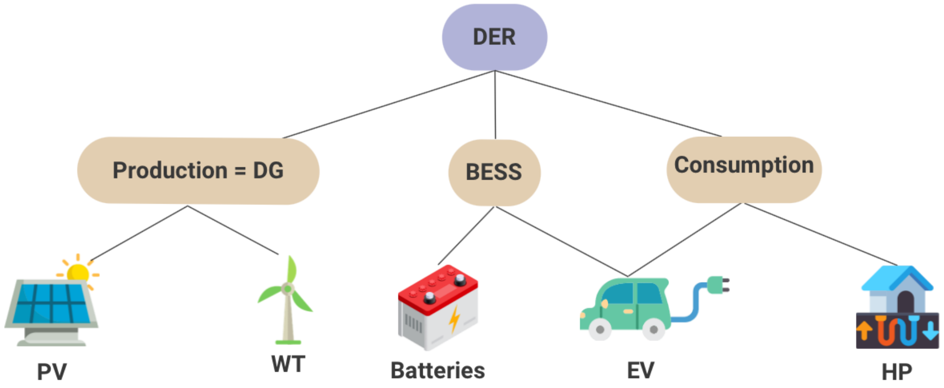


Figure 1: Terminology of the most common emerging technologies at distribution level.

Existing distribution networks, predominantly established in the last century, were not originally designed to accommodate such production and consumption transformations. This situation presents new challenges for distribution system operators (DSOs) in maintaining service levels and ensuring network reliability. Therefore, identifying and minimising the associated costs to alleviate these challenges becomes crucial.

DSOs need to identify the quantity of such technologies their network can host. This is complex as the quantity of installations that can be hosted depends on their location, the installed capacity, the consumption profiles, the potential production and many other factors. In order to provide a summarised view of all these factors, the concept of *hosting capacity* was introduced.

Hosting capacity (HC) is a measure of the capacity of a network to accommodate DER installations before encountering any operational issues. This capacity is generally expressed as one or multiple *penetration*, which can be a number of installations, their power, their consumption, etc. Figure 2 illustrates the concept of the definition. In this illustration, the hosting capacity is a function of the penetration rate, which represents how much of a new technology is added to the network; and a performance index, which determines whether any operational issue happened.

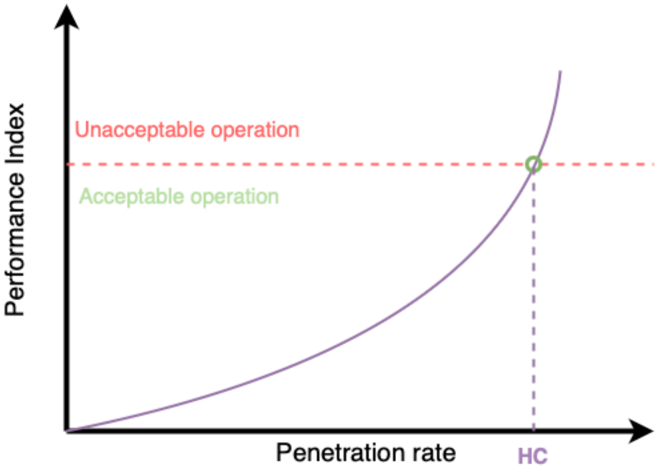


Figure 2: One dimension(performance index) representation of the concept of HC.

HC has become a popular and rapidly evolving subject of interest in the academic community; the large amount of literature available on HC in databases like Scopus and Google Scholar demonstrates its growing importance. Multiple variations of the definition of the penetrations and of the HC computation procedure have been experimented on by the community. There currently is, however, no established common ground for the mathematical definition of the HC problem. Given the extensive interest in HC, it is essential to gather and organise the knowledge into a coherent structure. To address this, this paper aims at providing a unified definition of the hosting capacity problem that encompasses all these different aspects and to show how papers can be re-framed using the mathematical definition. Additionally, a concise review of the predominant literature is presented to highlight the simplicity of organising related work under the lens and terminology of the presented definition. This process will provide guidance for future research in the area of HC.

The main contributions of this paper are:

- Presentation of a formal unified definition of HC. This definition lays out the theoretical HC problem and provides common ground of all 'in practice' HC computations;
- Presentation of how predominant researches on HC falls into the generic HC definition;
- Reviewing similar research according to the proposed definition.

The structure of this paper is as follows: first, Section 2 presents a framework that generalises HC. Then, two papers are reviewed as specific cases of the proposed framework in Section 3 and Section 4 shows how related works can now easily be compared using the framework. Finally, Section 5 concludes this paper with a summary and the main outcomes.

Existing hosting capacity reviews

Various reviews have already been published on hosting capacity such as [4, 5, 6, 7, 8, 9, 10, 11, 12]. While some areas of these reviews naturally overlap, there are notable distinctions from the proposed review.

First, the technologies covered in [4, 5, 7, 8, 9, 10, 11, 12] primarily focus on distributed generation, with [5, 8, 10, 11] specifically concentrating on PV technologies. On the other hand, the recent study in [6] focuses only on EVs. Then, the review in [9] is dedicated to medium voltage (MV). Enhancement techniques aiming at increasing the hosting capacity are discussed in [7, 9, 11] and in addition to that [7] presents an historical overview. Finally, [5] concludes that a general definition is needed.

This paper stands apart from the aforementioned by adopting a more inclusive approach, considering distributed energy resources (DERs), and ultimately presenting a unified definition of hosting capacity.

Paper selection criteria

As aforementioned, a large quantity of papers addressing the hosting capacity issue are available. Therefore, the review in Section 4 intentionally avoids aiming for comprehensiveness. The aim of this review is to show how related works can easily be compared thanks to the presented framework. The reviewed papers were selected on several bases:

- The selection of paper mostly focuses, for conciseness, on low-voltage (LV) distribution networks as most HC studies are done on this voltage level;
- All types of DER were accepted;
- There was no filtering on the input of the methods as this is not the scope of this review.

Furthermore, only papers with a substantial number of citations were selected. The number was set to a hundred, thus papers with more than 100 citations on Scopus with the search terms "Hosting Capacity" were selected. This threshold can be judged as high but it is attributed to the abundance of papers on the topic garnering 100 citations. Some papers were selected outside of this criterion as they add an interesting value for the review and were sometimes heavily mentioned by others. Note that the two papers used in Section 3 were taken from the selection.

2. Definition of hosting capacity (HC)

In this section a unified definition of the hosting capacity is presented. The deterministic definition is first presented in 2.1 as well as an example to illustrate it, and then methods to model usual types of stochasticity and integrate them into the definition are presented in section 2.2. For the sake of readability, in the remainder of this paper, a set is written using a calligraphic uppercase letter (e.g., $\mathcal{X}$) and vectors are written as lowercase bold characters (e.g., $\mathbf{x}$).

2.1. Deterministic definition

The hosting capacity concept was, according to [7], first introduced in [13]. They defined HC as the amount of new resources that can be hosted by a network before facing any issues, i.e., compromising its operational limits or violating safety constraints. The hosting capacity is computed for a given network. Let N denote a modelisation of a network for which one wants to determine the hosting capacity. This network model represents a topology and can contain, for instance, where all the nodes are, the links between them and also a description of the various electrical elements. The definition of N is left fuzzy on purpose to allow flexible representation of the network depending on, for instance, the software used to tackle the problem or the modelisation options such as considering an AC network or a DC one. Note that network and network model are used interchangeably hereafter. In addition to the network model, exogenous data E are defined, which include factors such as sun irradiation or load profiles.

Inside the network are customers that can withdraw or inject power, install new DERs, or more generically have an impact on the network. The set of such customers, also referred as *connection nodes*, is denoted by $\mathcal{C}$ and an element of this set is referenced using the index c. The set of possible types of DERs that customers can add is denoted $\mathcal{H}$. En element h of $\mathcal{H}$ can be for example EV or PV. When deploying a particular technology, like PVs for instance, multiple installation options are available, allowing for decisions to be made on various aspects, such as the installation capacity and phase configuration, among other considerations. The set of possible installation options for a technology $h \in \mathcal{H}$ is defined as $\mathcal{I}_h$. Note that when a technology $h \in \mathcal{H}$ is not installed for customer $c \in \mathcal{C}$, then the corresponding value $\mathcal{I}_h$ is chosen equal to $\emptyset$.

The hosting capacity is determined over a specific time period, denoted by $\mathcal{T}$, during which installations can be modified and their impact on the overall system can be assessed. This time period can be continuous or discrete but is always bounded. Without loss of generality, 1 is always the first time considered and T is the last. Typically, $\mathcal{T}$ is either $[1, T]$ or $\{1, \dots, T\}$. An element of $\mathcal{T}$ is referenced using the index t.

At each time step, a customer has a state for each technology that is its installed option. This state is written as:

$$s_{c,h,t} \in \mathcal{I}_h, \forall t \in \mathcal{T}, c \in \mathcal{C}, h \in \mathcal{H}. \tag{1}$$

Gathering all the states for all customers, time steps and technology types, forms a scenario: a hypothetical evolution of the network over the time period with the given set of technology types. Formally, a scenario is defined as the tuple of states of all customers for all time steps and all technology types:

$$\mathbf{s} = \langle s_{c,h,t} \mid \forall t \in \mathcal{T}, h \in \mathcal{H}, c \in \mathcal{C} \rangle = \langle s_{c_1,h_1,t_1}, s_{c_1,h_1,t_2}, \dots, s_{c_1,h_2,t_1}, s_{c_1,h_2,t_2}, \dots, s_{c_2,h_1,t_1}, \dots \rangle. \tag{2}$$

By design, only one installation type can be installed per technology and per customer. It is not restrictive as $\mathcal{I}_h$ can accommodate for these options. For example, the capacity of a PV can vary as well as the connected phases, in which case $\mathcal{I}_{\text{PV}}$ could be defined as

$$\mathcal{I}_{\text{PV}} = \{\emptyset, (\text{phase 1: 5KW}), (\text{phase 2: 10KW}), (\text{phase 1: 5KW; phase 2: 5kW}), (\text{phase 1: 10KW}), \dots\}.$$

The set of all possible scenarios is $\mathcal{S}$, formally:

$$\mathcal{S} = \bigtimes_{h \in \mathcal{H}, c \in \mathcal{C}, t \in \mathcal{T}} \mathcal{I}_h. \tag{3}$$

Even though only one scenario will realise for a given network, the hosting capacity is independent from this scenario as it aims to evaluate the capacity of the network to host new technologies. Therefore, determining the hosting capacity is subject to several uncertainties. Two categories of uncertainty were defined: aleatory and epistemic [10]. Aleatory uncertainties, alternatively known as certain uncertainties or inherent uncertainties, deal with uncertainties that are known to be stochastic. These includes the new installations' consumption or production, as well as the customer loads. These uncertainties are part of the exogenous data E. Epistemic uncertainties, also referred to as systematic uncertainties or uncertain uncertainties, result from the lack of knowledge or information. These uncertainties are the location of the new installations and their type, the options chosen and the time steps when the installations are added. These uncertainties are taken into account by evaluating the hosting capacity on multiple scenarios.

As the number of possible scenarios could be intractable, a subset of $\mathcal{S}$ is often *considered*:

$$\mathcal{S}^c \subseteq \mathcal{S}, \tag{4}$$

where $\mathcal{S}^c$ is the set of considered scenarios, with the superscript c chosen to emphasise this.

For a given scenario, the penetration is a measure that gauges the amount of resources present on the network. It is represented by a vector $\mathbf{a} \in \mathcal{A}$, the set $\mathcal{A}$ being the one of all representable penetrations. A penetration can be, for instance, the number of new DERs. Let the function $\mathbf{g}(\mathbf{s}) : \mathcal{S}^c \to \mathcal{A}$ compute the penetration for a given scenario $\mathbf{s}$. For instance, $\mathbf{g}(\mathbf{s})$ can return a vector composed of both the number of customers with PVs and the total production of the PVs in scenario $\mathbf{s}$. $\mathbf{g}(\mathbf{s})$ is not injective: multiple scenarios can have the same penetration (for example, if the penetration is defined as the number of PVs installed, two scenarios can have the same number of PVs in different places).

The set of *considered* penetrations can thus be defined from the previously defined set of *considered* scenarios:

$$\mathcal{A}^c = \{\mathbf{g}(\mathbf{s}) \mid \forall \mathbf{s} \in \mathcal{S}^c\} \tag{5}$$

The hosting capacity depends on the previously mentioned penetrations and is primarily governed by the physical constraints of the network; indeed, the addition of new DERs can lead to network issues. Let the set $\mathcal{P}$ be the set of network issues that can be encountered. An example of issue is nodal over-voltage, i.e. a point in the network having a voltage greater than a defined threshold. Let $f_t(\mathbf{s})$ be a binary function which identifies whether, at time t and for the scenario $\mathbf{s}$, any issues from $\mathcal{P}$ occurs in the network N. To evaluate f_t, both the network N data and exogenous data E are required. The function f_t is formally defined as follows:

$$f_t(\mathbf{s}) = \begin{cases} 1, & \text{if at least one issue } p \in \mathcal{P} \text{ occurs at time } t \in \mathcal{T} \text{ in } N \text{ given } E; \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Let f be an aggregation of f_t over the time period based on defined conditions. For instance, f could be defined as returning 1 if at least for one time step there was an issue in the network (f_t is equal to one), i.e.:

$$f(\mathbf{s}) = \begin{cases} 1, & \text{if} \quad \bigvee\limits_{t \in \mathcal{T}} f_t(\mathbf{s}) = 1 \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Where the symbol $\bigvee$ represents the logical *or*. Usually, papers about hosting capacity differ on the way the scenarios are defined as well as how both f and g are defined. This will be addressed in Section 4.

The set of validated scenarios is defined as the set of scenarios where no issues are detected:

$$\mathcal{S}^{\checkmark} = \{\mathbf{s} \in \mathcal{S}^c \mid f(\mathbf{s}) = 0\}. \tag{8}$$

The validated penetration set is derived from this set of scenarios. It is the subset of $\mathcal{A}$ which can be associated with **at least** a valid scenario:

$$\mathcal{A}^{\checkmark} = \{\mathbf{g}(\mathbf{s}) \mid \forall \mathbf{s} \in \mathcal{S}^{\checkmark}\}. \tag{9}$$

Similarly, the set of non-validated scenarios and the set of invalid penetrations are defined as

$$\mathcal{S}^{\times} = \{\mathbf{s} \in \mathcal{S}^c \mid f(\mathbf{s}) \neq 0\}, \qquad \mathcal{A}^{\times} = \{\mathbf{g}(\mathbf{s}) \mid \forall \mathbf{s} \in \mathcal{S}^{\times}\}. \tag{10}$$

The hosting capacity of a network, referred as feasible penetrations $\mathcal{A}^f$, is defined as the valid penetrations that cannot be linked to an invalid scenario:

$$\mathcal{A}^f = \mathcal{A}^{\checkmark} \setminus \mathcal{A}^{\times}. \tag{11}$$

Note that $\mathcal{A}^f, \mathcal{A}^{\checkmark}, \mathcal{A}^{\times} \subseteq \mathcal{A}^c \subseteq \mathcal{A}$, and $\mathcal{A}^f \subseteq \mathcal{A}^{\checkmark}$. Let $\mathcal{S}^f$ be the set of scenarios leading to a penetration in $\mathcal{A}^f$:

$$\mathcal{S}^f = \left\{\mathbf{s} \in \mathcal{S}^{\checkmark} \mid \mathbf{g}(\mathbf{s}) \in \mathcal{A}^f\right\}. \tag{12}$$

The relations between the sets $\mathcal{S}$, $\mathcal{A}$, and their derivatives are schematically explained below.

$$\begin{tikzcd}
\mathcal{S} \arrow[r, "\mathbf{g}(\mathbf{s})"] \arrow[d, "\text{Considered scenarios}"'] & \mathcal{A} & \\
\mathcal{S}^c \arrow[r, "\mathbf{g}(\mathbf{s})"] \arrow[d, bend right, "f(\mathbf{s}) = 0"] \arrow[ddd, bend right=50, "f(\mathbf{s}) \neq 0"'] & \mathcal{A}^c & \\
\mathcal{S}^{\checkmark} \arrow[r, "\mathbf{g}(\mathbf{s})"] & \mathcal{A}^{\checkmark} & \\
 & & \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} \mathcal{A}^{\checkmark} \setminus \mathcal{A}^{\times} = \mathcal{A}^f \arrow[ddll, bend left, "\mathbf{s} \in \mathcal{S}^{\checkmark} \mid \mathbf{g}(\mathbf{s}) \in \mathcal{A}^f"] \\
\mathcal{S}^{\times} \arrow[r, "\mathbf{g}(\mathbf{s})"] & \mathcal{A}^{\times} & \\
\mathcal{S}^f & &
\end{tikzcd}$$

Note that in $\mathcal{A}^f$, having a (valid) penetration $\mathbf{a}$ does not imply that all penetrations $\mathbf{a}' \in \mathcal{A}$ that are dominated by $\mathbf{a}$ ($\mathbf{a}' \leq \mathbf{a}$ for some partial comparison operator $<$) are also valid. This means that for instance one scenario with more installations can have no issues while a scenario with fewer installations can have one. This is mainly due to issues potentially being dependent on the topology and the location of the installations. Therefore, one could define stricter limits for the final hosting capacity.

$\mathcal{A}^s$, the *safe* penetrations, is defined as the subset of $\mathcal{A}^f$ which possesses the above-mentioned property, i.e., all dominated penetrations are feasible (recall that $\leq$ is a partial comparison operator between two penetrations):

$$\mathcal{A}^s = \left\{ \mathbf{a} \in \mathcal{A}^f \mid \nexists \mathbf{a}' \in \mathcal{A}^{\times} : \mathbf{a}' \leq \mathbf{a} \right\}. \tag{13}$$

The corresponding scenario set is denoted $\mathcal{S}^s$ and defined by:

$$\mathcal{S}^s = \left\{ \mathbf{s} \in \mathcal{S}^f \mid \mathbf{g}(\mathbf{s}) \in \mathcal{A}^s \right\}. \tag{14}$$

The defined scenario sets are illustrated in Figure 3.

For penetrations in one dimension, most papers choose to use $\mathcal{A}^s$ rather than $\mathcal{A}^f$ as their definition for the hosting capacity, as it is a single connected space. Such papers generally define the penetrations $\mathcal{A}$ as $\mathbb{R}_+$ or $\mathbb{Z}_+$ (or subsets), thus making $\mathcal{A}^s$ a continuous or discrete range starting at zero and ending at a maximum value that is generally reported as the hosting capacity. Some papers also report these as probability density functions.

In Figure 4, an illustrative example of the different penetration sets defined above is given. All subfigures lie in the plan formed by the set $\mathcal{A}$, here exemplified by $\mathcal{A} = \mathbb{R}_+^2$. The two dimensions of this example can for example represent the number of PVs and EVs in a scenario. The sets $\mathcal{A}^{\checkmark}$ and $\mathcal{A}^{\times}$, representing respectively the penetrations reachable by valid scenarios and invalid ones, are shown in yellow and red respectively, as can be seen in the top subfigures. They can intersect: there can exist two scenarios sharing the same number of PVs and EVs such that one fails and the other is valid, depending, for example, on the location of the PVs/EVs in the network topology.

The feasible penetration set $\mathcal{A}^f$, in green in the lower subfigures, is the the subset of $\mathcal{A}^{\checkmark}$ (in yellow) that is not part of $\mathcal{A}^{\times}$ (in red). A feasible penetration $p = \langle p_1, p_2 \rangle$ that is in $\mathcal{A}^f$ can thus be

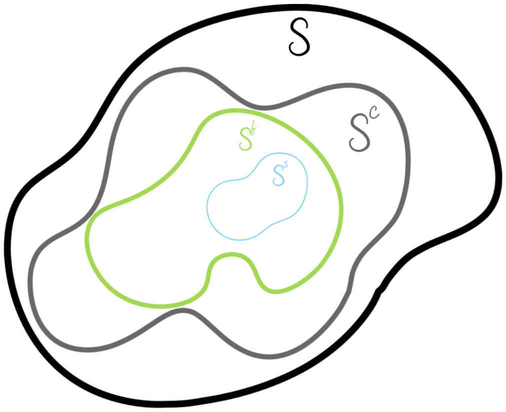


Figure 3: Sets of scenarios, where $\mathcal{S}$ is the set of all possible scenarios, $\mathcal{S}^c$ is the set of considered scenarios, $\mathcal{S}^f$ is the set of feasible scenarios and $\mathcal{S}^s$ is the set of safe scenarios.

reported (for example, to the DSO) as supported by the network ("Having p_1 PVs and p_2 EVs would only lead to valid scenarios").

This may, however, be seen as abusive in certain contexts as shown in the lower right subfigure. For example, the penetration a is in $\mathcal{A}^f$ and, thus, all scenarios having this penetration are valid. However, point b, which lies to the left of a (meaning that it can be, for example, the penetration with the same number of EVs but fewer PVs), is not in $\mathcal{A}^f$; all scenarios having penetration b are not valid. In the notations presented above, a is said to be *dominating* b, that is $b \leq a$, using here the standard element-wise vector comparison operator.

To avoid such contradictions about dominated but invalid penetrations ("having 10 PVs and 100 EVs is supported in the network but not 10 PVs and 50 EVs"), one can instead use the set $\mathcal{A}^s$, defined as the set of points where this situation cannot happen: for any penetration in $\mathcal{A}^s$, all penetrations lower than it are also valid. Thus, the penetration a is not in $\mathcal{A}^s$ due to b being a counterexample. However, c is in the set as the rectangle (in this 2-D example) on its lower left is fully included in $\mathcal{A}^f$. The penetration d is not in the set as the rectangle on its lower left encompasses part of the hole in $\mathcal{A}^f$. The same is true for the penetration e.

Hosting capacity.

Up until here, the hosting capacity was introduced as sets ($\mathcal{A}^f$ and $\mathcal{A}^s$) representing all the penetrations that can be hosted. Alternatively, one can work with the frontiers of the sets, which is easier in one dimension.

In one dimension, most studies evaluate penetrations in $\mathcal{A}^s$ while not dealing explicitly with the set defined above but rather with a scalar: indeed, in one dimension, $\mathcal{A}^s$ is a range between 0 and another scalar, that is sometimes called the hosting capacity as a shortcut. It can be equivalently computed by

$$\max_{\mathbf{a}} : \{\mathbf{s} \mid \mathbf{g}(\mathbf{s}) \leq \mathbf{a}\} \cap \mathcal{S}^{\times} = 0. \tag{15}$$

This is referred to in the literature as *first violation* or *minimal hosting capacity*. Furthermore, some studies also determine the *maximal hosting capacity* (also referred to as *all scenarios with violation*) which is defined as the minimal penetration for which all scenarios encounter at least one issue:

$$\min_{\mathbf{a}} : \{\mathbf{s} \mid \mathbf{g}(\mathbf{s}) > \mathbf{a}\} \subset \mathcal{S}^{\times}. \tag{16}$$

Note that working with penetrations in $\mathcal{A}^f$ would imply several frontiers (in one dimension, multiple ranges).

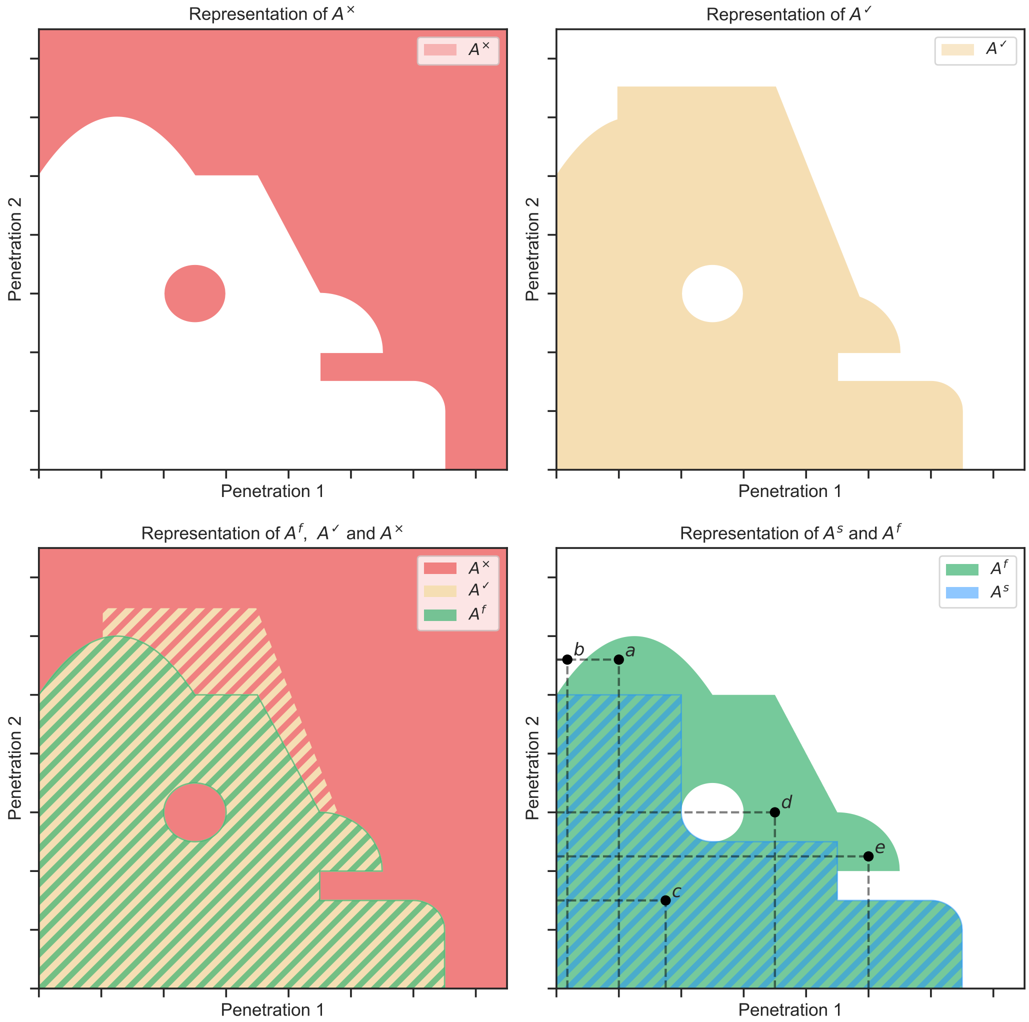


Figure 4: Two-dimensional representations of set $\mathcal{A}^f$, $\mathcal{A}^s$, $\mathcal{A}^{\checkmark}$ and $\mathcal{A}^{\times}$. The axes are a quantification of the penetration of two installations. The stripes mean that sets intersect.

In more dimensions, giving one value as a summary for the hosting capacity is more complex as the frontiers of the sets are not scalars. The concepts presented above can be extended to higher dimension via the use of partial comparison operators as done earlier in this paper.

2.2. *Stochastic definition*

In the following section, two ways to introduce stochasticity in the presented definition are addressed. The first manner to introduce stochasticity is related to the considered scenarios, while the second is on the detection of issues. These are not mutually exclusive: one may consider more than one way to introduce stochasticity in determining the hosting capacity.

2.2.1. *Stochasticity related to scenarios selection*

As previously mentioned, evaluating all scenarios is, most of the time, intractable. Therefore, the subset $\mathcal{S}^c$ is necessary to select a subset of scenarios to consider. Building this set can be done in multiple ways. For example, this set can be constructed by sampling the scenarios in $\mathcal{S}$ using a distribution on $\mathcal{S}$, that may be informed with external information. This distribution, denoted $\mathbb{D}_{\mathcal{S}}$, represents the probability of realisation of each scenario. This allows to built $\mathcal{S}^c \sim \mathbb{D}_{\mathcal{S}}$ with a limited number of scenarios that are nevertheless representative. Note that all scenarios can have the same probability ($\mathbb{D}_{\mathcal{S}}$ being in this case the uniform distribution), which is often the case in the papers reviewed in section 4.

2.2.2. Stochasticity related to issues detection

The functions f_t and f were introduced as binary functions, one identifies if an issue $p \in \mathcal{P}$ occurred in the network N given E at time t and the other is aggregating f_t. A function $h_t \in [0,1]$ and its aggregation $h \in [0,1]$ can be defined as returning the probability of an issue occurring in N given E. The stochastic of function h can come from E that can itself be stochastic, for instance if E contains a distribution of load time series. Given a threshold α for the probability, f_t can now be:

$$f_t(\mathbf{s}) = \begin{cases} 1, & \text{if } h_t(\mathbf{s}) > \alpha \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

This translates into f_t indicating that a scenario that has less than α% probability of having an issue at time t is valid. Or, for all time steps, f can be defined as:

$$f(\mathbf{s}) = \begin{cases} 1, & \text{if } h(\mathbf{s}) > \alpha \\ 0, & \text{otherwise;} \end{cases} \tag{18}$$

i.e. having a scenario that has less than α% probability of having an issue over the whole time range. Multiple variant of this are possible: the framework presented here can accommodate a wide range of settings.

As an example, consider the aggregation f that returns 0 when no issues occur for at most 95% of the time steps of the scenario, i.e.,

$$f(\mathbf{s}) = \begin{cases} 0, & \text{if } \mathop{\mathbb{E}}_{t \in \mathcal{T}} \dfrac{f_t(\mathbf{s})}{|\mathcal{T}|} < 0.95 \\ 1, & \text{otherwise.} \end{cases}$$

2.3. Example

The following subsection illustrates, with a small example, how the formalism of the generic definition can be applied to describe a hosting capacity computation. This example showcases the determination of the hosting capacity, taking into account only over-voltage issues, for a small distribution network where customers can add new photovoltaic capacities.

Following the formalism of the definition, the example is presented as follows:

- The **network model** N, represented in Figure 5, is composed of 8 buses: 7 low-voltage ones and 1 medium voltage one. Let $\mathcal{B}$ denote the set of buses, $\mathcal{B} = \{b0, b1, b2, b3, b4, b5, b6, b7, b8\}$. The network is modelled using PandaPower [14]. The medium-voltage bus is supplied by an external grid with a voltage set to 1 p.u., then a transformer converts medium-voltage to low-voltage. The transformer has a maximum apparent power handling capacity of 0.4 MVA and operates with a primary voltage of 20 kV and a secondary voltage of 0.4 kV. The transformer was chosen from the PandaPower standard library. All lines (L0 to L6) have the same standard type ("NAYY 4x50 SE" from PandaPower). Feeder lines (L0 and L3) have a length of 800m while lateral lines are 200m long except for L2, which is 400m long.
- Five **customers** are modelled by loads added to buses: $\mathcal{C} = \{b3, b4, b6, b7, b8\}$.
- In addition to the loads, customers $b3$ and $b4$ can install PVs as **installation type**, therefore $\mathcal{H} = \{\text{PV}\}$.

- These customers can choose, as **options**, either to install 50 PV panels of $300W_{peak}$ or to install no PVs at all, thus $\mathcal{I}_{\text{PV}} = \{\emptyset, 50 \times 300W_{peak}\}$.
- The **time period** is composed of one time step ($\mathcal{T} = \{1\}$).
- In this example, the **exogenous data** E, are the PV productions and customer loads. PV installations are at their peak production, i.e. a PV installation produces 1.5kW during the time frame, while the customers' loads are 1kW except for $b4$, which is 0.8kW. Note that the model only considers active power.
- All the **considered scenarios** are regrouped in Table 1. Scenario a is the initial topology as shown Fig. 5, scenarios b, c and d are, respectively, presented in Fig. 6a, Fig. 6b, Fig. 6c.
- The **issue** considered is over-voltage ($\mathcal{P} = \{\text{OV}\}$). An over-voltage at time t is flagged when the voltage ($V_{b,t}$) is greater than 1.05 pu for at least one bus b in $\mathcal{B}$, i.e.:

  $$f(s) = f_t(s) = \begin{cases} 1, & \text{if } \exists b \in \mathcal{B} : V_{b,t} > 1.05pu; \\ 0, & \text{otherwise.} \end{cases}$$

  The aggregation f is the same as f_t as there is only one time step.

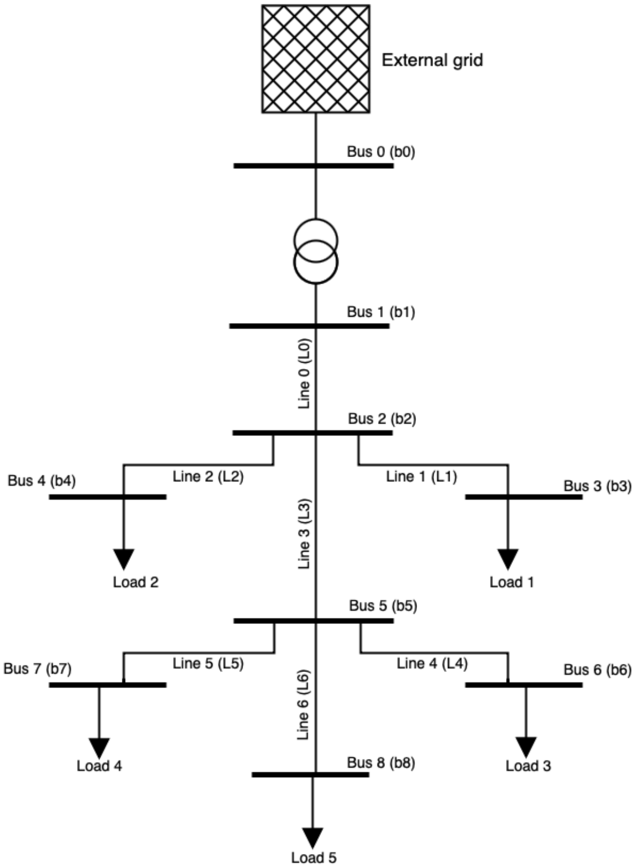
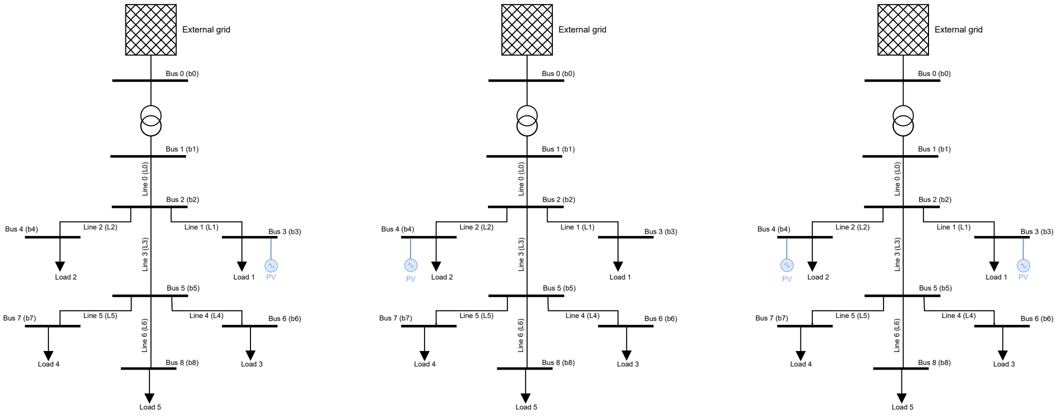


Figure 5: Example network with 5 customers represented by 5 loads.

The power flow is computed using PandaPower with the single-phase model of the network. The outputs are given in Table 2 for each scenario.

Scenarios in $\mathcal{S}^c$	b3	b4
a	$\emptyset$	$\emptyset$
b	$50 \times 300 W_{\text{peak}}$ PVs	$\emptyset$
c	$\emptyset$	$50 \times 300 W_{\text{peak}}$ PVs
d	$50 \times 300 W_{\text{peak}}$ PVs	$50 \times 300 W_{\text{peak}}$ PVs

Table 1: Considered scenarios ($a, b, c, d \in \mathcal{S}^c$) for the example for installing PVs with uncertain location. $\emptyset$ means that no installation was added at that customer.



(a) Scenario b (b) Scenario c (c) Scenario d

Figure 6: Topologies of scenarios b, c and d in $\mathcal{S}^c$. In blue is the added PV installation.

Scenarios c and d have at least an over-voltage at one bus (in V_{b4} for example), while the others have no bus in over-voltage. Therefore,

$$f(a) = f(b) = 0, \text{ and}$$
$$f(c) = f(d) = 1.$$

The function g returns, as a penetration measure, the number of buses with a PV installation:

$$g(s) = |\{b \in \mathcal{B} \mid s_b \neq \emptyset\}|,$$

where s_b designates the installations at bus b in scenario s. Note that $g(s)$ is a scalar and thus not written in bold. The penetration for each scenario is given in Table 3.

The penetration sets and the related scenario sets, are:

- The set $\mathcal{A}^c$ of all **considered** output of $g(s)$, $\forall s \in \mathcal{S}^c$, is thus $\{0, 1, 2\}$;
- The set of **valid scenarios** is $\mathcal{S}^{\checkmark} = \{a, b\}$ and $\mathcal{A}^{\checkmark} = \{0, 1\}$;

$s \in \mathcal{S}^c$	V_{b0}	V_{b1}	V_{b2}	V_{b3}	V_{b4}	V_{b5}	V_{b6}	V_{b7}	V_{b8}
a	1	1	0.98	0.98	0.98	0.97	0.97	0.97	0.97
b	1	1	1.03	1.04	1.03	1.02	1.02	1.02	1.02
c	1	1	1.03	1.03	**1.05**	1.02	1.02	1.02	1.02
d	1	1	**1.08**	**1.09**	**1.10**	**1.07**	**1.07**	**1.07**	**1.07**

Table 2: Powerflow voltage output in p.u. by scenario. Over-voltages are in **bold**.

- The set of **invalid** ones is $\mathcal{S}^{\times} = \{c, d\}$ and $\mathcal{A}^{\times} = \{1, 2\}$;
- By definition of **feasible penetration set** $\mathcal{A}^f$, that is the penetrations that includes only valid scenarios, $\mathcal{A}^f = \mathcal{A}^{\checkmark} \setminus \mathcal{A}^{\times} = \{0\}$.

$s \in \mathcal{S}^c$	$g(s)$
a	0
b	1
c	1
d	2

Table 3: Penetration for each scenario computed using $g(s)$.

The different sets of scenarios and their corresponding penetration sets are summarised in Table 4. The *feasible* $\mathcal{A}^f$ and *safe* $\mathcal{A}^s$ penetration sets are the same, thus the hosting capacity of this example is $\mathcal{A}^f = \mathcal{A}^s = \{0\}$. As this example is one-dimensional, this corresponds to the *first violation* as defined in Section 2. *All violation* HC is 2, as all scenarios with penetration 2 are not valid. Indeed, scenario d, the only scenario with a penetration equal to 2, is not valid.

Scenarios sets	penetration sets
$\mathcal{S}^{\checkmark} = \{a, b\}$	$\mathcal{A}^{\checkmark} = \{0, 1\}$
$\mathcal{S}^{\times} = \{c, d\}$	$\mathcal{A}^{\times} = \{1, 2\}$
$\mathcal{S}^f = \mathcal{S}^s = \{a\}$	$\mathcal{A}^f = \mathcal{A}^s = \{0\}$

Table 4: Scenarios sets and their penetrations.

Deterministic versus Stochastic cases.

This example is deterministic as no uncertainty is taken into account: all possible scenarios are evaluated and exogenous data are unique. A stochastic approach would either sample scenarios or sample exogenous data to account for uncertainty. For instance, adding stochasticity for scenario selection by randomly sampling scenarios: scenarios a and b could be sampled and thus, $\mathcal{S}^c$ is $\{a, b\}$. In such context, the hosting capacity would be $\mathcal{A}^f = \{0, 1\}$.

Notation summary.

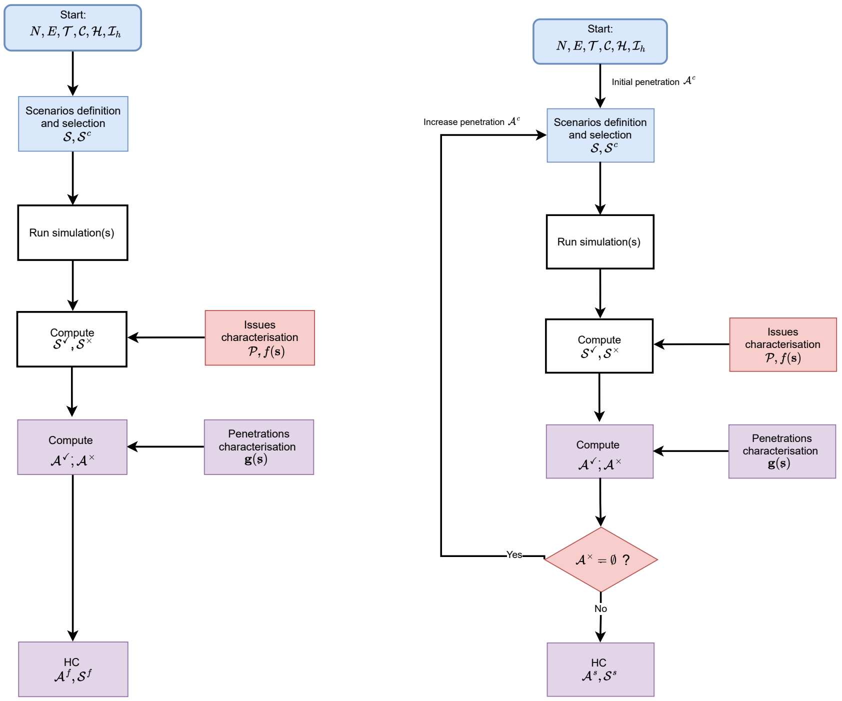
N	Network
E	Exogeneous data
T	Time
Sets	
$\mathcal{P}$	Set of network issues e.g., $\mathcal{P} = \{$ over-voltage, overloading $\}$
$\mathcal{C}$	Set of customer nodes
$\mathcal{H}$	Set of types of technologies { PV, EV, HP, ... }
$\mathcal{I}_h$	Set of installation options of technology $h \in \mathcal{H}$ e.g., $\mathcal{I}_{\text{EV}} = \{$ 0kW, 2.3 kW, 3.7 kW, 7.4 kW, 11 kW, 22 kW $\}$
$\mathcal{T}$	Set of time steps, $\mathcal{T} = \{1, ..., T\}$.
$\mathcal{S}$	Set of all scenarios
$\mathcal{S}^c$	Set of considered scenarios
$\mathcal{S}^{\checkmark}$	Set of valid scenarios
$\mathcal{S}^{\times}$	Set of invalid scenarios

$\mathcal{S}^f$	Set of feasible scenarios
$\mathcal{S}^s$	Set of feasible and safe scenarios
$\mathcal{A}$	Set of all possible penetrations e.g., if the penetration is defined as the number of customers $\mathcal{A} = \mathbb{R}$
$\mathcal{A}^c$	Set of considered penetrations
$\mathcal{A}^{\checkmark}$	Set of penetrations that can be associated to at least one valid scenario
$\mathcal{A}^{\times}$	Set of invalid penetrations
$\mathcal{A}^f$	Set of feasible hosting capacities
$\mathcal{A}^s$	Set of feasible and safe hosting capacity

3. Application of the framework on two papers

In this section, the framework presented in Section 2 is applied to two papers to illustrated how the concepts are applied to the literature.

In this framework, how the HC is actually computed is not explained and there are many ways to do so in the literature. Some of them follows directly the natural steps arising from the framework, while others iteratively refine their scenario definitions until they reach a predefined condition. These two workflows are shown in Figure 7. Note that this paper does not elaborate on how the simulations to determine the physical state of the network for given scenarios are run.



(a) Direct flow

(b) Iterative flow

Figure 7: Flowcharts of a direct and an iterative HC method. In blue are the system assumptions, in red the issue characterisation, and in violet the HC computation, as presented in sections 4.1, 4.2 and 4.3, respectively. The iterative flow 7b computes the *safe* penetrations as the stop condition ensures that all dominated penetrations are feasible. Note that in white are steps not discussed in this paper.

Both workflows follow this general scheme: first the scenario sets $\mathcal{S}$ and $\mathcal{S}^c$ are build for a network N and exogenous data E. These scenarios encompasses the technology types $\mathcal{H}$ and their options $\mathcal{I}_h$.

Then, to construct the sets $\mathcal{S}^{\checkmark}$, $\mathcal{S}^{\times}$ and $\mathcal{S}^{f}$, the function f is introduced given the sets of considered issues $\mathcal{P}$. Finally, to assess the penetration of a scenario and depending on the set in which the scenario belongs, the sets $\mathcal{A}^{\checkmark}$, $\mathcal{A}^{\times}$ and $\mathcal{A}^{f}$ are built using a function $\mathbf{g}$. Each of these three steps is highlighted in a different colour in Figure 7.

For the sake of conciseness and simplicity, two papers were selected to illustrate how the literature fits in the framework.

3.1. Paper 1: "Understanding Photovoltaic Hosting Capacity of Distribution Circuits" by Dubey et al. [1]

This paper aims at determining the HC for PVs and the impact of specific factors on the HC. The second part which deals with the impact of several factors will not be addressed in this review. The study first computes a base case which is the existing configuration and then stochastically assesses the HC by simulating several scenarios with different penetrations levels. The details of this part are given below.

The **Network model** N represents, in OpenDSS, an actual three-phase distribution network with 12.47kV supplied by a 24MVA substation transformer. The base case has 1.196 MW of existing PV. The schema of this network can be found in [1]. No further indications about the network were given.

All **customers** can have new installations and the **installation type** is PVs: $\mathcal{H} = \{\text{PV}\}$. The installation size is the only **option**. The size is randomly picked from two distributions represented by probability density functions (PDFs) depending on the type of customer: residential or commercial.

The study follows the paradigm developed by EPRI [15] for steady-state analysis of HC. This means that one time step was used for the **time period** ($\mathcal{T} = \{1\}$). This time step is derived from **exogenous data**. The study uses two sets of exogenous data: minimum and maximum load both obtained from yearly load demand measured in the substation of the network in 2013. Note that PV production is not exogenous data as the size is used as the production rather than time series.

The **considered scenarios** set contains 5000 scenarios. This set is built by increasing the penetration, here defined as the percentage of the total customers equipped with PV systems. The increase is done by steps of 2% for penetration going from 0% to 100% resulting in a total of $|\mathcal{S}^c| = 5000$. Location of PVs are randomly selected. Note that the considered uncertainties are both epistemic as they are the location and the size of the installations.

The **issues** considered are over-voltage (OV), voltage deviation (VD) and voltage unbalance (VU). Theses issues are explained and formulas are given in Section 4.2.1. In this study, the three issues are considered separately as they are not aggregated and thus leads to three HCs: one for OV, one for VD and one for VU. Also, as the study is steady-state, f is the same as f_t. The thresholds used for each f are 5% more than the nominal voltage for OV, and 3% for both VD and VU. This study used three independent issues to highlight the impact of each for the hosting capacity.

Although penetrations are defined as percentage of the total customers equipped with PV systems for determining the considered scenarios, the penetration output by the function $\mathbf{g}$ and used for the hosting capacity is defined as the total additional PV size in kW.

They consider the *first violation* and *all scenarios with violation* HC as defined in Section2. Additionally, they defined *50% scenarios with violation* HC which is the smallest penetration such that 50% of scenarios encounter an issue, i.e.:

$$\min_{\mathbf{a}} : |\{\mathbf{s} \mid \mathbf{g}(\mathbf{s}) = \mathbf{a}\} \subset \mathcal{S}^{\times}| = |\{\mathbf{s} \mid \mathbf{g}(\mathbf{s}) = \mathbf{a}\} \subset \mathcal{S}^{\checkmark}|. \tag{19}$$

The two sets of exogenous data E, the three limitation functions f, the safe penetration set and the three outputs (*first violation*, *50% scenarios with violation* and *all scenarios with violations*) lead to 18 HC values given in Table III of the paper and reproduced in Table 5.

3.2. Paper 2: "Assessing the Potential of Network Reconfiguration to Improve Distributed Generation Hosting Capacity in Active Distribution Systems" by Capitanescu et al. [2]

This research explores how to increase DG hosting capacity using network reconfiguration. The study is conducted in two parts: the authors first determined the HC and then studied how to use

Cases		Additional PV Size (kW)	
		Max Load	Min Load
Overvoltage	first violation	9,442	5,454
	50% scenarios with violation	9,578	5,536
	All scenarios with violation	9,659	5,722
Voltage deviations	first violation	1,756	2,776
	50% scenarios with violation	1,834	2,970
	All scenarios with violation	1,909	3,088
Voltage unbalance	first violation	0	5,760
	50% scenarios with violation	0	6,101
	All scenarios with violation	0	6,291

Table 5: PV hosting capacity from [1] (copied with authorization).

reconfiguration to improve it. In the context of the present paper, only the first part is of interest and is detailed.

The method proposed is applied on a 34-bus 12.66-kV distribution system **network model**. The network has 37 lines, 34 nodes and one feeder, more information about the network can be found in [16]. In this network, a subset of 8 **customers** (G1,..., G8) are selected to have new installations. The choice of using a subset of customers is justified by (i) the fact that tools can be used to determine most suitable locations and; (ii) these locations might be enforced, for instance, by regulatory rules.

Although the methodology is given for general DG, the test case is made using wind turbines as the **installation type** thus, $\mathcal{H} = \{\text{WT}\}$. The considered **option** for the installation is the size. Two sizes, related to two different wind profiles (WP), are available and, for simplicity, the sizes are named following the profiles: $\mathcal{I}_{WT} = \{WP1, WP2\}$.

This study is time variant and the **time period** is composed of 146 time steps ($|\mathcal{T}| = 146$). The **exogenous data** accounted in these time steps are historic demand and wind data.

Table 6 gathers a summary of the three **considered scenarios** ($|\mathcal{S}^c| = 3$) of the study.

Scenarios in $\mathcal{S}^c$	G1	G2	G3	G4	G5	G6	G7	G8
A	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	WP2	$\emptyset$	$\emptyset$	WP1
B	WP1	WP1	WP2	WP2	$\emptyset$	$\emptyset$	WP1	WP1
C	WP1	WP1	WP2	WP2	WP2	WP2	WP1	WP1

Table 6: Considered scenarios ($A, B, C \in \mathcal{S}^c$) for the [2] for installing WT. The empty set $\emptyset$ means that no installation was added at that customer.

Both voltage and thermal **issues** are considered. The minimum and maximum voltage limits are respectively set to 0.95 p.u. and 1.05 p.u. at all nodes, and the thermal limit of all lines is set to 6.6 MVA. Both limits are further addressed in Section 4.2.

Finally, **penetrations** are defined as the sum of the nominal capacity of the DG and g returns the sum of the maximal nominal capacity. The HC for the three scenarios are 3.622MW, 4.161MW and 7.154MW for scenario A, for scenario B and for scenario C, respectively. This study is considered deterministic as no uncertainty is considered.

4. Review and application

This section aims at showing how related works on hosting capacity problems can be easily compared using the generic framework.

This section is divided into three parts corresponding to the three steps presented in Section 3 :

- First, the **system assumptions** are discussed in Subsection 4.1; the methods used to construct the scenarios sets are explored and a quick overview of the network and data each paper uses is done. How papers address uncertainties is also highlighted.
- Second, the **issues characterisation** (often referred as *limiting factors* in the literature [5, 17]); Subsection 4.2 identifies the common issues used in various research to populate the set $\mathcal{P}$ and how they trigger and aggregate them (i.e., their definition of the functions f and f_t).
- Finally, **the hosting capacity computation**. Subsection 4.3 gathers the most frequent types of penetrations ($\mathcal{A}$) and how they are computed ($\mathbf{g}$).

4.1. System assumptions

This subsection details system assumptions, which are the different hypothesises that are used to build the scenario sets $\mathcal{S}$ and $\mathcal{S}^c$. The construction of the system assumptions can be divided into three phases and the remainder of this section will address these:

- First, *network model* has to be constructed and *exogenous data* has to be gathered;
- Then, the *model scope* has to be defined. The scope determines the accepted values for each previously defined sets ($\mathcal{T}$, $\mathcal{C}$, $\mathcal{H}$ and $\mathcal{I}_{\mathcal{H}}$).
- Finally, *uncertainty* needs to be taken into account. Choices made to manage and model the uncertainty restrict both the scenarios $\mathcal{S}$ and exogenous data E spaces by methodology choices.

A summary of some assumptions can be found in Table 7. The different categories used in the table are explained below.

4.1.1. Network and exogenous data: N and E

The network model N is a representation of a network with a variable degree of sophistication depending on the work and the available data. It can be made from real-world, empirical data or from synthetic data (for example the IEEE 33-bus or 123-bus feeders [42, 24, 20]), or from a mix of both (see, e.g., [41, 42, 18]). This model also encompasses hypotheses made on the network or on its simulation, e.g., the complexity and details of the power flow computation. Exogenous data E are mainly historical data or probabilistic approximations that are used as an estimate for the future.

4.1.2. Model scope: $\mathcal{T}$, $\mathcal{C}$, $\mathcal{H}$ and $\mathcal{I}_{\mathcal{H}}$

The model scope defines the values that populates each of the sets related to the system assumptions, i.e. the set of time steps $\mathcal{T}$, the set of connection nodes for technologies $\mathcal{C}$, the set of types of technologies $\mathcal{H}$ and the set of technologies' options $\mathcal{I}_{\mathcal{H}}$.

Time invariant vs time variant: $\mathcal{T}$.

Electrical systems varies through time. The advantage of using time series, and thus time variant models, is their closeness to reality. Indeed, the model is able to reflect the time related phenomena that occurs in real networks. Nevertheless, time representation is highly dependent on the previous data assumptions and data availability. The model can thus either be time invariant or time variant. Time-invariant modelisations only use one time step. These studies often choose the worst-case time step and are referred to as conservative. Methods using time series varies from one another on:

- The granularity of the time series (i.e., the amount of time between two time steps);
- The considered period. For instance, some papers consider multiple full days, while others consider consecutive hours in a single reference day. Some consider a range of hours in multiple days (e.g., Monday to Friday from 09:00 to 18:00).

Papers	Data (N & E)	Model scope ($\mathcal{T}$, $\mathcal{H}$ and $\mathcal{I}_{\mathcal{H}}$)				Uncertainty ($\mathcal{S}$ & E)
	Empirical Data	Time variant	Technology	Different sizes	Single phase	Considered
[18]	○	●	HP	○	○	○
[2]	○	●	WT	●	●	○
[19]	●	○	EV	●	○	●
[20, 21]	○	○	PV	●	●	●
[22]	○	○	PV	●	?	●
[23]	○	○	PV	●	○	●
[24]	○	●	PV	○	○	●
[25]	○	●	PV	●	○	●
[26]	○	●	PV	●	◑	●
[27]	○	●	PV	●	●	●
[28]	○	●	PV	●	?	●
[29]	◑	●	PV	●	◑	●
[1]	●	○	PV	●	○	●
[30]	●	●	PV	●	◑	●
[31]	●	●	PV	●	?	○
[32]	●	●	PV	○	?	○
[33]	●	○	PV	○	○	●
[34]	●	○	PV	●	○	●
[35]	●	●	PV	●	●	●
[36]	●	●	PV	○	◑	○
[37]	●	●	PV	●	○	●
[38]	●	●	PV	●	○	
[39]	◑	●	PV, EV, HP, μCHP	●	●	●
[40]	◑	●	BESS, [PV]	●	○	○
[17]	●	◑	PV, [EV]	◑	●	●
[41]	◑	●	PV, EV	○	●	●
[42]	○	●	PV, EV	◑	●	●

Table 7: References characteristics. Legend: ● means the characteristic is fully, ◑ partially, ○ not implemented and ? not specified. Note that "Single phase" is ticked if all new installations are connected using only one-phase connection.

Connection nodes: $\mathcal{C}$.

Studies can either consider all customer nodes (e.g., [27, 17, 41, 42, 1]) or only a subset of customers (e.g., [18, 20, 24, 30, 28]) as connection nodes.

Technology type: $\mathcal{H}$.

As the HC definition is not specific to one technology, it is worth mentioning which paper is studying which technology (i.e., set of types of technologies $\mathcal{H}$). The most common technology is DG and more precisely PVs as they represent the first technology to have been used to define the hosting capacity. Note that some studies develop methods to compute HC that are not specific on a technology type and then test the methodology on one or more technology types, often independently (e.g., [39]). In [39], these technologies are referred to as Low Carbon Technologies (LCT). They consider micro combined heat and power units (μCHP) as generator.

Technology options: $\mathcal{I}_{\mathcal{H}}$.

Options varies from one technology type to another. The two options encountered in the reviewed papers are the size of the technology and the connection type. For the size, studies use one size or allow the flexibility of different sizes. Note that some studies use several sizes but only one size is available per run, e.g., they first attempt a run with $\mathcal{I}_{PV} = \{\emptyset,\ 1\ \text{kWc}\}$ and then another $\mathcal{I}_{PV} = \{\emptyset,\ 2\ \text{kWc}\}$. The connection type can be in one, two or three phases, depending on the network, the available data, the authors' choice and the technology itself. Some studies, such as [30], enforce that if the customer is on a single phase the installation has to also be on a single phase.

4.1.3. Uncertainties

Studies may or may not take uncertainties into account. For the one that does, as previously mentioned, the literature identifies two categories of uncertainty: aleatory and epistemic [10]. The aleatory uncertainties deal with exogenous data (in E) such as the technologies' consumption or production and the customers' consumption. These are inherently stochastic. These are, in practice, sampled when computing f and g.

Epistemic uncertainties deal with the lack of knowledge for, for instance, the type $\mathcal{H}$ and options of the technologies $\mathcal{I}_{\mathcal{H}}$ or their locations $\mathcal{C}$. These uncertainties are considered by evaluating several scenarios $\mathcal{S}^c$.

In both cases, these uncertainties can be accounted for by using, for instance, simple Monte Carlo simulations or more complex statistical methods. These uncertainties are not mutually exclusives: studies can use both. Note that on the other hand, studies neglecting all uncertainties, referred to as deterministic, are restricting both $\mathcal{S}$ and E with defined size and values. In recent papers, there is less and less deterministic methods, as the underlying system is, in reality, stochastic.

Table 8 showcase how studies account for uncertainties. Deterministic methods are omitted from the table for conciseness.

Papers	Aleatory		Epistemic		
	Loads	Installation production consumption	Installation location	Installation size	Installation type
[41, 19, 21, 33, 30]	○	○	●	○	○
[1, 22, 38, 23, 34, 25, 35]	○	○	●	●	○
[17]	○	◑	●	○	○
[20]	○	●	●	●	○
[29]	●	○	●	○	○
[37]	●	○	●	●	○
[42, 24]	●	●	○	○	○
[26, 39, 27, 28]	●	●	●	●	○

Table 8: Uncertainties types for reviewed papers that take into account uncertainties. Legend: ● means the characteristic is fully, ◑ partially, ○ not implemented. Partially means that not all simulations account for that uncertainty.

4.2. Issues characterisation

This section discusses the elements inside the set $\mathcal{P}$ and how they are evaluated using the function f, for the papers already analysed above. As a reminder f identifies whether any issues from $\mathcal{P}$ occurs in the network. In some papers, the function f is sometimes not precisely defined or too complex to display here; therefore the focus is put on some readable examples to emphasise the genericity of the formalisation. Motivated readers are encouraged to read the original papers to gather the full definitions of f and $\mathcal{P}$.

There are four main categories of issue [43]: voltage dependant, load dependant, protection and harmonics. These categories are not specific to DER. Papers such as [17, 44, 5, 43, 7] attempt to define all existing issues.

The following subsections are structured using these categories. Each of these subsections showcases how the issues associated with the considered category fit in the definition presented in Section

2. Note that there is no intent to be exhaustive in the list of limiting factors, only the ones encountered in the paper reviewed are mentioned.

4.2.1. Voltage-related issues

Voltage issues are typically defined as voltage unbalance, voltage levels (under- or over-voltage) or deviations:

- **Voltage unbalance** refers to an uneven distribution of voltage magnitudes or phase angles in a three-phase electrical system. It occurs when the three phases of the system have different voltage levels or when the phase angles between the voltages are not equal. This translates Eq (6) into:

  $$f_t(s) = \begin{cases} 1, & \text{if } \exists m \in \mathcal{M} : |V_{m,t,\phi_1} - V_{m,t,\phi_2}| > \beta, \forall \phi_1 \neq \phi_2; \phi_1, \phi_2 \in \Phi; \\ 0, & \text{otherwise.} \end{cases}$$

  where $\mathcal{M}$ is the set of monitored elements, V is the complex voltage of an element of $\mathcal{M}$ for the phase ϕ, Φ is the set of phases, β is a given threshold. Note that the difference can be computed taking into account only magnitudes, or the angles, or both.

  In several cases the hosting capacity could be improved by balancing the network [45, 46].

- **Voltage level** issues happen due to an unusual use of the network. In the case of PVs, or any other DG device, when there is greater production compared to consumption, the remaining production is injected in the network. This process is called **reverse power flow** as the power flows in the opposite direction compared to the direction it was originally intended (i.e., from the centralised generation to the distribution as opposed to from the prosumers DG to other consumers). This can cause an over-voltage. In the case of EVs, or any other bigger electricity consumer than first intended as normal network use, the consumption is bigger than that planned. And the consumers, when peak consumption occurs, withdraw too much electricity leading to an under-voltage as the production and the capacity of the cables were not made for such peak consumption. This translates into:

  $$f_t(s) = \begin{cases} 1, & \text{if } \exists m \in \mathcal{M} : \underline{v} > V_{m,t} \vee V_{m,t} > \overline{v}; \\ 0, & \text{otherwise.} \end{cases}$$

  where $\mathcal{M}$ is the set of monitored elements, V is the voltage of a node and $\underline{v}$ and $\overline{v}$ are given lower and upper bound thresholds. This definition is a base that often varies from one paper to another. For instance, the measured element can vary between studies (e.g., $V_{m,t}$ can be $\max_m V_{m,t}$ as in [33]) and some studies may have different thresholds when they have more than one type of measured element (e.g., [30]).

- **Voltage deviation** is the deviation in voltage from the initial network (i.e., with no installation) to a full-penetration network.

  $$f_t(s) = \begin{cases} 1, & \text{if } \exists m \in \mathcal{M} : |V^b_{m,t} - V_{m,t}| > \mu; \\ 0, & \text{otherwise.} \end{cases}$$

  where $\mathcal{M}$ is the set of monitored elements, V is the voltage of a node and V^b is the voltage of the same node in the initial network, and μ is a given threshold.

Note that the first two issues exist without new installations while the last one is exclusively linked to the addition of new ones. Most researches use standards to define the thresholds (β, ν, μ) for issue detection f. Several standards were used to limit these thresholds, for instance EN-50160 is a European standard, ANSI is an American standard, and VDE-AR-N is a German standard. These standards might have some variations such as BS EN-50160 which is the British version of EN-50160. Table 9 gathers the papers by issues and standard. Some papers do not clearly specify the standard, but the closer fit is assumed, and some do not explicitly use these standards, for instance [19] uses a threshold set by the Brazilian government for distribution.

Papers	Unbalance	Levels	Deviation	Standard
[37, 33, 23, 34, 26, 27, 41, 24]	○	●	○	ANSI
[25, 35, 21, 1]	●	●	●	ANSI
[32]	○	○	●	EN50160
[22, 38]	○	●	●	ANSI
[20, 30]	●	●	○	ANSI
[29, 36]	●	●	○	EN50160
[17]*, [42], [18]*, [39, 2]	○	●	○	EN50160
[31]	○	●	○	VDE-AR-N 4105
[40]	●	●	●	Australian Act 1945
[19]	●	●	○	Other
[28]	○	●	○	Other

Table 9: List of voltage-related issues by papers. ● means the issue is taken into account, while ○ means it is not. The asterisk * means that the standard was not properly mentioned but corresponds to the given one.

4.2.2. *Load-related issues*

Load-related issues occur when transformers or cables experience thermal overload, indicating excessive current flow that generates heat, potentially causing damage or failure. Here is an example on how to define f to detect overload by limiting the current:

$$f_t(s) = \begin{cases} 1, & \text{if } \exists m \in \mathcal{M} : I_{m,t} > \alpha \times I_m^l; \\ 0, & \text{otherwise.} \end{cases}$$

where $\mathcal{M}$ is the set of measured elements, I is the current of an element and I^l is the rated capacity of the measured element, and α is a given threshold.

Some studies (e.g., [24, 41, 23, 18, 19]) limit the power which is equivalent to limiting the current under constant voltage. To detect the overload, most papers used two categories of measured element: the conductor current as in [30, 37, 23, 24, 18, 19] and the transformer load [37, 41, 23, 24, 19]. The threshold varies from one study to another (from 50% [19] to 187% [30]) and several thresholds can be used in the same study for different types of measured element (e.g.[37]). These thresholds are ratios of a rated capacity, but some papers used defined values (e.g., [24, 23]). Also previously defined standards can be used for overload such as EN50160 in [29, 36].

4.2.3. *Protection-related issues*

Protection-related issues are a vast category. They were defined in [43] as issues taking more time to affect the user level in contrast to voltage and load issues even though both also have direct consequences for the end user.

A first protection issue is rapid voltage magnitude variations [17]. These occur because of variations in production or consumption over a period of less than one minute. These rapid changes can cause flicker. An example of a standard that limits these is IEC 61000-4-30. Standards to limit short-term (Pst) and long-term(Plt) flicker severity are defined in IEC 61000-4-15 [17].

Another protection issue is used in [20] which restricts the PV power factor, i.e. the ratio of real power (in watts) to apparent power (in volt-amperes), between 0.8 and 1 to contribute to the stability of the power grid, stability in the sense of the ability to maintain a balanced and reliable supply of electricity.

4.2.4. Harmonics-related issues

Harmonics issues occur when the current injections are not sinusoidal. Several standard indices exist [17] for individual harmonics, interharmonics, and for total harmonic distortions (THD) (e.g., IEEE 519, IEC 61000-4-7, IEC 61000-4-30, IEC/TR 61000-3-6).

In [17], superharmonics are mentioned: these are waveform distortion in the frequency range between 2 kHz and 150 kHz and are injected by an increasing number of devices connected to the grid. Unfortunately, [17] considers that using this issue as a limiting factor for the hosting capacity is not feasible as no limits for distortion in this frequency range are set. Furthermore, the main barrier for taking into account the harmonics as a limiting factor is the lack of appropriate calculation models, especially when considering low-voltage and medium-voltage networks [17]. Thus, harmonics are not often used as limiting factors for the hosting capacity.

4.2.5. Issues aggregation

Some studies may consider several issues. They may consider them independently and thus have several definitions of the f function giving a hosting capacity per issue; or they may aggregate them to have one f function. The aggregation of issues can be a complicated formula but most aggregations are a simple OR operator over several issues. Furthermore, time aggregation of f_t is necessary for time-variant studies such as in Eq (7). Another example of time aggregation is an average during a time period (e.g., [37, 32]). Most studies do not explicitly give how they aggregate the time and the issues.

Papers	Voltage	Load	Protection	Harmonics
[33, 34, 26, 27, 25, 35, 21, 1, 32, 22, 38, 42]	●	○	○	○
[37, 30, 23, 29, 36, 18, 41, 24, 17], [39, 2, 28, 31, 40]	●	●	○	○
[19]	●	●	○	●
[20]	●	●	●	○

Table 10: List of issues by papers. ● means the issue is taken into account, while ○ means it is not.

4.2.6. Issues summary

In Table 10 is a summary of all papers and the issues categories. Voltage deviation issues are the most limiting factor [7]. Therefore, these issues are addressed in the majority of studies as shown in Table 10. Both Protection and Harmonics issues are less used and mostly not as the only limitation.

Most hosting capacity studies are considered as static as they do not take into account the duration of the issues. Taking into account the duration means that the aggregation f evaluates several consecutive time steps f_t. For instance, with a duration of i:

$$f(\mathbf{s}) = \begin{cases} 1, & \text{if} \quad \exists t \in \{1, \dots, |\mathcal{T}| - i\} : f_t(\mathbf{s}) = 1 \land f_{t+1}(\mathbf{s}) = 1 \land \dots \land f_{t+i}(\mathbf{s}) = 1 \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

This means that f reports an issue only if it occurs for i consecutive time steps. Studies that take into account the duration of the issues are referred to as dynamic hosting capacity (DHC) such as in [37, 32, 30, 25, 39].

4.3. Hosting capacity computation

This section focuses on the final part of determining the hosting capacity: penetration calculation. It is how the $\mathbf{g}$ function is defined for different studies and the output/penetration format choice ($\mathcal{A}$).

As the penetration is not a defined concept shared among all the reviewed papers, to compute the hosting capacity one has to first define the wanted output of $\mathbf{g}$, i.e., how to quantify the set of penetrations $\mathcal{A}$. This penetration can either be an absolute quantity that reflects the amount of new technologies that are installed or a ratio between this absolute quantity and a reference. The encountered possibilities for the absolute quantity are, with their abbreviated name given in bold to aid readability:

- The number of customers with a new installation, abbreviated as **NC**;
- The total production or consumption of a new installation, **TP**;
- The maximal capacity of a new installation, **MC**.

The encountered reference quantities are, again with their abbreviated name:

- The total number of customers that can accommodate new installations, **TNC**;
- The total consumption of customers, **TC**;
- The maximal consumption (load) of customers also referred as peak load, **PL**;
- The total capacity that is accepted by the transformer, also referred as transformer rated capacity and thus abbreviated to **TRC**.

Each absolute and reference quantity has is perks depending on the study. For instance, choosing the number of houses equipped with the technology, as an absolute quantity, is intuitive; but from one house to another the loads vary and thus this quantity can lack precision. Also, since peak demand can put significant stress on the grid, evaluating the hosting capacity based on maximum capacity, as a reference quantity, ensures that the grid can handle high-energy demand without issues. Using the maximum capacity allows for unified approach for evaluating hosting capacity, regardless of the specific technology being added. Neither does it require detailed modelling of various DER operational conditions and thus simplifies the planning process. Despite these advantages, it might not fully reflect the actual average or expected output of the DER installation under typical operating conditions.

Table 11 regroups all reviewed papers and organises them following their HC computation choices. This table also shows the different output types for each paper. These can be divided in four types:

- The set of penetration $\mathcal{A}^f$ or $\mathcal{A}^s$ as a *range*, for instance the penetration is a number of customers, HC $= \mathcal{A}^f = [0-5;7]$ meaning that all considered scenarios with penetration in this set are feasible;
- The set of penetration $\mathcal{A}^f$ or $\mathcal{A}^s$ as a range with probability density function (*PDF*). This means having the scenario distribution;
- A penetration of this set corresponding to one scenario, i.e., *one value*. This would happen for instance in a deterministic study having an iterative workflow (Figure 7b) or considering one scenario;
- Several penetrations of this set corresponding to the output of several scenarios, i.e., *several values*.

Note that studies might consider the same scenario with different E and thus still be considered as outputting *one value* (e.g., [18]).

As mentioned in Section 2, papers computing the one-dimensional HC and with output type *one value*, choose stricter limits for the hosting capacity: all dominated penetrations are feasible ($\mathcal{A}^s$) and output the *minimal hosting capacity*. Some of these papers also choose to output *maximal hosting capacity* such as in [27, 1, 35, 25]. On the other hand, in any dimension, papers that compute either the set $\mathcal{A}^f$ or $\mathcal{A}^s$, use PDFs to showcase the distribution of scenarios over the penetrations e.g., in [42] with two dimensions.

Papers	Absolute quantities	Reference quantities	Units	Types
[17]	NC	-	Customers	PDF
[19]	NC	-	Customers	SV
[41, 39]	NC	TNC	- (%)	PDF
[18]	NC	TNC	- (%)	OV
[36]	NC	TNC	- (%)	SV
[26]	TP	-	Power	Range, SV
[1, 27]	TP	-	Power	SV
[42]	TP	TC	- (%)	PDF
[33]	TP	PL	- (%)	Range
[20]	TP	PL	- (%)	SV
[23, 35]	MC	-	Power	Range
[28]	MC	-	Power	PDF
[38, 24]	MC	-	Power	OV
[2, 32, 31]	MC	-	Power	SV
[40]	MC	-	Energy	SV
[29]	MC	PL	- (%)	SV, Range
[21, 22]	MC	PL	- (%)	SV
[30]	MC	TRC	- (%)	OV
[25]	MC	-, PL	Power, - (%)	Range
[37]	NC, MC	TNC,-	Customers, Power	SV
[34]	NC, MC	TNC,-	Customers, Power	OV

Table 11: Summary of papers' definition of hosting capacity computation. Several acronyms are used for the readability of this table: NC for the number of customers, TP means total installation production or consumption, MC for the maximal capacity of a new installation, TNC for the total number of customers that can accommodate new installations, TC for total consumption of customers, PL for peak load, TRC for transformer-rated capacity, OV for one value, SV for several values and PDF for probability density function. One row reads as follows: Paper [17] uses the number of customers with new installations (NC) as the absolute metrics to gauge the penetration. This value is not compared to a reference, hence the "-", thus the hosting capacity is expressed as the number of customers as a unit. They compute the set of penetrations and choose to represent the distribution of scenarios over the penetrations using a PDF.

5. Conclusion

The need to determine network hosting capacity is now widely acknowledged, and numerous researchers have worked on the subject these last years. However, there was, prior to this work, no well-defined formalism.

This paper introduces a general definition of the hosting capacity problem along with definitions of all the elements necessary for its formulation. First, the deterministic definition has been presented, followed by the exploration of various methods for incorporating uncertainties, thus stochasticity.

The definition is then applied to a small fictitious example and then to two prominent papers from the hosting capacity research domain. These latter papers serve to illustrate the use of the definition in concrete and more complex cases while the small example ensure that the different aspect of the definition are commonly understood.

The central focus of the definition is to establish a common ground for concepts; thus, it does not prescribe how to construct different sets or compute hosting capacity, recognising the diversity in approaches. Nevertheless, the related work section provides several examples from the literature on how these tasks are accomplished. This restricted but systematic review of hosting capacity field demonstrates the ease with which related works can be conducted to compare studies using the presented framework.

Future works stand to benefit from this formalism to clarify terminology for subsequent research and identify with more clarity the gaps in the literature that need addressing.

Abbreviations

BESS	Battery Energy Storage Systems
μCHP	Micro Combined Heat and Power unit
DER	Distributed Energy Resource
DG	Distributed Generator
DSOs	Distribution System Operators
EV	Electric Vehicle
HC	Hosting Capacity
HP	Heat Pump
LV	Low Voltage
MV	Medium Voltage
PV	Photovoltaic
pu	Per unit
RES	Renewable Energy Source
THD	Total Harmonic Distortion
THD_v	Total Voltage Harmonic Distortion
WT	Wind Turbine

Acknowledgements

The authors would like to thank Bardhyl Miftari and Jocelyn Mbenoun for the useful conversations about the elaboration of this paper and their constructive feedback.

Author contributions:

Amina Benzerga: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Writing - Original Draft, Visualization. **Alireza Bahmanyar:** Conceptualization, Writing - Review & Editing. **Guillaume Derval:** Methodology, Formal Analysis, Writing - Review & Editing, Supervision. **Damien Ernst:** Conceptualization, Methodology, Writing - Review & Editing, Supervision.

Declaration of competing interest

The authors declare no conflicting financial or personal interests.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used chat.gpt for minor word checks and rephrasing. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

[1] A. Dubey, S. Santoso, A. Maitra, Understanding photovoltaic hosting capacity of distribution circuits, in: 2015 IEEE power & energy society general meeting, IEEE, 2015, pp. 1–5.

[2] F. Capitanescu, L. F. Ochoa, H. Margossian, N. D. Hatziargyriou, Assessing the potential of network reconfiguration to improve distributed generation hosting capacity in active distribution systems, IEEE Transactions on Power Systems 30 (1) (2014) 346–356.

[3] IEA, Renewables 2022 (2022).
URL `https://www.iea.org/reports/renewables-2022`

[4] M. Zain ul Abideen, O. Ellabban, L. Al-Fagih, A review of the tools and methods for distribution networks' hosting capacity calculation, Energies 13 (11) (2020) 2758.

[5] S. Fatima, V. Püvi, M. Lehtonen, Review on the pv hosting capacity in distribution networks, Energies 13 (18) (2020) 4756.

[6] S. Fatima, V. Püvi, M. Lehtonen, M. Pourakbari-Kasmaei, A review of electric vehicle hosting capacity quantification and improvement techniques for distribution networks, IET Generation, Transmission & Distribution (2023).

[7] S. M. Ismael, S. H. A. Aleem, A. Y. Abdelaziz, A. F. Zobaa, State-of-the-art of hosting capacity in modern power systems with distributed generation, Renewable energy 130 (2019) 1002–1020.

[8] A. Kharrazi, V. Sreeram, Y. Mishra, Assessment techniques of the impact of grid-tied rooftop photovoltaic generation on the power quality of low voltage distribution network-a review, Renewable and Sustainable Energy Reviews 120 (2020) 109643.

[9] S. Mina Mirbagheri, D. Falabretti, V. Ilea, M. Merlo, et al., Hosting capacity analysis: A review and a new evaluation method in case of parameters uncertainty and multi-generator, in: Proceedings-2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe, EEEIC/I and CPS Europe 2018, Institute of Electrical and Electronics Engineers Inc., 2018, pp. 1–6.

[10] E. Mulenga, M. H. Bollen, N. Etherden, A review of hosting capacity quantification methods for photovoltaics in low-voltage distribution grids, International Journal of Electrical Power & Energy Systems 115 (2020) 105445.

[11] A. Rajabi, S. Elphick, J. David, A. Pors, D. Robinson, Innovative approaches for assessing and enhancing the hosting capacity of pv-rich distribution networks: An australian perspective, Renewable and Sustainable Energy Reviews 161 (2022) 112365.

[12] U. Singh, The role of hosting capacity study in power system advancements: a review, arXiv preprint arXiv:2301.04765 (2023).

[13] M. Bollen, M. Häger, Power quality: interactions between distributed energy resources, the grid, and other customers, Leonardo Energy (2005).

[14] L. Thurner, A. Scheidler, F. Schafer, J. H. Menke, J. Dollichon, F. Meier, S. Meinecke, M. Braun, pandapower - an open source python tool for convenient modeling, analysis and optimization of electric power systems, IEEE Transactions on Power Systems (2018). doi:10.1109/TPWRS.2018.2829021.
URL `https://arxiv.org/abs/1709.06743`

[15] M. Rylander, J. Smith, Stochastic analysis to determine feeder hosting capacity for distributed solar pv, Tech. Rep. 1026640, EPRI (2012).

[16] M. E. Baran, F. F. Wu, Network reconfiguration in distribution systems for loss reduction and load balancing, IEEE Transactions on Power delivery 4 (2) (1989) 1401–1407.

[17] M. H. Bollen, S. K. Rönnberg, Hosting capacity of the power grid for renewable electricity production and new large consumption equipment, Energies 10 (9) (2017) 1325.

[18] S. Zhan, T. Gu, W. van den Akker, W. Brus, A. van der Molen, J. Morren, Towards congestion management in distribution networks: a dutch case study on increasing heat pump hosting capacity, in: The 12th IET International Conference on Advances in Power System Control, Operation and Management (APSCOM), Institution of Engineering and Technology (IET), 2022, pp. 364–369.

[19] B. E. Carmelito, J. M. d. C. Filho, Hosting capacity of electric vehicles on lv/mv distribution grids—a new methodology assessment, Energies 16 (3) (2023) 1509.

[20] F. Ding, B. Mather, On distributed pv hosting capacity estimation, sensitivity study, and improvement, IEEE Transactions on Sustainable Energy 8 (3) (2016) 1010–1020.

[21] F. Ding, B. Mather, N. Ainsworth, P. Gotseff, K. Baker, Locational sensitivity investigation on pv hosting capacity and fast track pv screening, in: 2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D), IEEE, 2016, pp. 1–5.

[22] F. Ding, B. Mather, P. Gotseff, Technologies to increase pv hosting capacity in distribution feeders, in: 2016 IEEE Power and Energy Society General Meeting (PESGM), IEEE, 2016, pp. 1–5.

[23] S. Jothibasu, S. Santoso, A. Dubey, Optimization methods for evaluating pv hosting capacity of distribution circuits, in: 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC), IEEE, 2019, pp. 0887–0891.

[24] S. Wang, S. Chen, L. Ge, L. Wu, Distributed generation hosting capacity evaluation for distribution systems considering the robust optimal operation of oltc and svc, IEEE Transactions on Sustainable Energy 7 (3) (2016) 1111–1123.

[25] N. Tang, G. Chang, A stochastic approach for determining pv hosting capacity of a distribution feeder considering voltage quality constraints, in: 2018 18th International Conference on Harmonics and Quality of Power (ICHQP), IEEE, 2018, pp. 1–5.

[26] O. Ceylan, S. Paudyal, B. P. Bhattarai, K. S. Myers, Photovoltaic hosting capacity of feeders with reactive power control and tap changers, in: 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), Ieee, 2017, pp. 1–6.

[27] A. Dubey, S. Santoso, On estimation and sensitivity analysis of distribution circuit's photovoltaic hosting capacity, IEEE Transactions on Power Systems 32 (4) (2016) 2779–2789.

[28] X. Chen, W. Wu, B. Zhang, C. Lin, Data-driven dg capacity assessment method for active distribution networks, IEEE Transactions on Power Systems 32 (5) (2016) 3946–3957.

[29] A. Arshad, M. Lindner, M. Lehtonen, An analysis of photo-voltaic hosting capacity in finnish low voltage distribution networks, Energies 10 (11) (2017) 1702.

[30] R. Torquato, D. Salles, C. O. Pereira, P. C. M. Meira, W. Freitas, A comprehensive assessment of pv hosting capacity on low-voltage distribution systems, IEEE Transactions on Power Delivery 33 (2) (2018) 1002–1012.

[31] T. Stetz, F. Marten, M. Braun, Improved low voltage grid-integration of photovoltaic systems in germany, IEEE Transactions on sustainable energy 4 (2) (2012) 534–542.

[32] P. H. Divshali, L. Söder, Improving pv hosting capacity of distribution grids considering dynamic voltage characteristic, in: 2018 Power Systems Computation Conference (PSCC), IEEE, 2018, pp. 1–7.

[33] A. Y. Saber, T. Khandelwal, A. K. Srivastava, Fast feeder pv hosting capacity using swarm based intelligent distribution node selection, in: 2019 IEEE Power & Energy Society General Meeting (PESGM), IEEE, 2019, pp. 1–5.

[34] H. V. Padullaparti, S. Jothibasu, S. Santoso, G. Todeschini, Increasing feeder pv hosting capacity by regulating secondary circuit voltages, in: 2018 IEEE Power & Energy Society General Meeting (PESGM), IEEE, 2018, pp. 1–5.

[35] J. Smith, M. Rylander, Stochastic analysis to determine feeder hosting capacity for distributed solar pv, Electric Power Research Inst., Palo Alto, CA, Tech. Rep 1026640 (2012) 0885–8950.

[36] J. Hu, M. Marinelli, M. Coppo, A. Zecchino, H. W. Bindner, Coordinated voltage control of a decoupled three-phase on-load tap changer transformer and photovoltaic inverters for managing unbalanced networks, Electric Power Systems Research 131 (2016) 264–274.

[37] A. K. Jain, K. Horowitz, F. Ding, N. Gensollen, B. Mather, B. Palmintier, Quasi-static time-series pv hosting capacity methodology and metrics, in: 2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), IEEE, 2019, pp. 1–5.

[38] S. Jothibasu, S. Santoso, A. Dubey, Determining pv hosting capacity without incurring grid integration cost, in: 2016 North American Power Symposium (NAPS), IEEE, 2016, pp. 1–5.

[39] A. Navarro-Espinosa, L. F. Ochoa, Probabilistic impact assessment of low carbon technologies in lv distribution systems, IEEE Transactions on Power Systems 31 (3) (2015) 2192–2203.

[40] N. Jayasekara, M. A. Masoum, P. J. Wolfs, Optimal operation of distributed energy storage systems to improve distribution network load and generation hosting capability, IEEE Transactions on sustainable energy 7 (1) (2015) 250–261.

[41] A. Benzerga, S. Mathieu, A. Bahmanyar, D. Ernst, Probabilistic capacity assessment for three-phase low-voltage distribution networks, in: 2021 IEEE 15th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG), IEEE, 2021, pp. 1–6.

[42] R. Fachrizal, U. H. Ramadhani, J. Munkhammar, 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 26 (2021) 100445.

[43] M. Rylander, J. Smith, W. Sunderman, S. Uppalapati, H. Li, Distributed photovoltaic feeder analysis: Preliminary findings from hosting capacity analysis of 18 distribution feeders., Tech. Rep. 3002001245, EPRI (2013).

[44] O. Lennerhag, G. Pinares, M. H. Bollen, G. Foskolos, T. Gafurov, Performance indicators for quantifying the ability of the grid to host renewable electricity production, CIRED-Open Access Proceedings Journal 2017 (1) (2017) 792–795.

[45] A. Benzerga, A. Bahmanyar, D. Ernst, Optimal connection phase selection of residential distributed energy resources and its impact on aggregated demand, in: 11TH BULK POWER SYSTEMS DYNAMICS AND CONTROL SYMPOSIUM (IREP-2022), IREP, 2022, pp. arXiv–2207.

[46] A. Benzerga, S. Gerard, S. Lachi, Q. Garnier, A. Bahmanyar, D. Ernst, Optimal connection phase selection for single-phase electrical vehicle chargers, in: CIRED Porto Workshop 2022: E-mobility and power distribution systems, Vol. 2022, IET, 2022, pp. 508–512.