TECHNISCHE UNIVERSITÄT WIEN

AIT AUSTRIAN INSTITUTE OF TECHNOLOGY

Definition and Validation of Reference Feeders for Low-Voltage Networks

Supervisor: Prof. Wolfgang Gawlik, TU Vienna

Submitted at the Faculty of Electrical Engineering and Information Technology
Technische Universität Wien in partial fulfillment of the requirements for the degree of
Doktor der technischen Wissenschaften (Dr. techn.)

Matriculation number 0425089
Dipl. Ing. Serdar Kadam, BSc
Töllergasse 13/110
1210 Vienna
Austria

Contents

Abstract

Distribution grids were initially planned to supply small and medium sized customers with electricity. However, in the recent decades a significant amount of distributed energy resources (DER) have been integrated into the electricity grid. Thereby, a high amount of DER, e.g. photovoltaics (PV), are now also connected to distribution grids. Furthermore, there is little knowledge of the real hosting capacity of distribution grids, and in particular, how low-voltage grids (LV-grids) can be classified. Therefore, a better knowledge of LV-grids, to integrate as much DER as possible, without grid reinforcements, is needed. Furthermore, the study of all LV-grids of a distributed system operator (DSO) is a complex and computationally intensive task.

This work presents a validated set of reference feeder topologies and a methodology to parametrize these topologies in LV-networks based on real feeder data. For that, in-depth analysis of the LV-grid data of two Austrian DSOs with a total of 14.000 networks was carried out. Thereby the hosting capability (HC) of feeders is defined. A set of measures were identified, hosting capability studies were performed for a high number of scenarios with varying voltage limits, DER locations and reactive power control strategies ($cos\varphi(P)$, VoltVAr). The developed reference feeder methodology was validated using the computed hosting capability results of the real feeders. Consequently, it was demonstrated to what extent the hosting capability (with and without reactive power control strategies) of real feeders can be estimated, using only a set of specific parameters available in the geographical information system (GIS) systems of DSOs. Furthermore the distribution of classified feeder topologies in real LV-grids is analyzed.

Keywords: Distribution grids, low-voltage networks, network classification, feeder taxonomy, reference networks, representative feeders, Smart Grids, DER, hosting capacity, hosting capability

Kurzfassung

Die ursprüngliche Aufgabe von elektrischen Verteilnetzen war die Versorgung von Kunden mit kleinem und mittleren Bedarf an elektrischer Energie. In den letzten Jahrzehnten wurde eine hohe Anzahl an Anlagen zur Stromerzeugung aus regenerativen Quellen an elektrische Netze angeschlossen. Dabei wurde ein erheblicher Anteil in Verteilnetzen, die ursprünglich für die Verteilung von elektrischer Energie geplant wurden, angeschlossen. Daher ist es eine herausfordernde Aufgabe das Aufnahmevermögen von Verteilnetzen, insbesondere von Niederspannungsnetzen zu bestimmen, sowie Niederspannungsnetze zu klassifizieren. Dies ist jedoch nötig, um einen möglichst hohen Anteil von erneuerbaren Energieerzeugungsanlagen zu integrieren bei gleichzeitiger Vermeidung von erheblichen Netzausbaukosten. Durch die hohe Anzahl von Niederspannungsnetzen im Vergleich zu Mittel- und Hochspannungsnetzen, ist die Analyse der untersten Spannungsebenen mit einem erheblichen Aufwand verbunden.

In dieser Arbeit werden Niederspannungs-Referenzstränge definiert und validiert, sowie eine Methode zur Bestimmung von Referenzsträngen für reale Stränge eingeführt. Als Datenbasis dienten alle Niederspannungsnetzdaten von zwei österreichischen Netzbetreibern mit etwa 14000 Niederspannungsnetzen. Dazu wurde die Hosting Capability (HC - die Aufnahmefähigkeit von Niederspannungssträngen bezüglich dezentraler Erzeugungsanlagen) definiert, die für alle realen Niederspannungsstränge berechnet wurde. Hierbei wurde eine hohe Anzahl von Szenarien berücksichtigt, die sich aus verschiedenen Spannungsgrenzen, Durchdringungsszenarien und Regelungsstrategien (ungeregelt, Q(U) und $cos\varphi(P)$) ergeben. Außerdem wurden für alle untersuchten realen Stränge statistische Parameter bestimmt. Basierend auf diesen Ergebnissen wurden eine Methode zur Bestimmung von Referenzsträngen entwickelt und mit den Ergebnissen der realen Netzdaten validiert. Dies ermöglicht, die HC für die bestimmten Referenzstränge für alle definierten Szenarien zu untersuchen und mit den Ergebnissen der realen Stränge zu vergleichen. Dabei werden nur wenige Eingangsparameter benötigt, die Netzbetreiber mit modernen Geoinformationssystemen bestimmen können. Außerdem wurde die Verteilung der Topologien von Referenzsträngen in den realen Netzen untersucht.

Acknowledgements

Selected publications

Journal papers

F. Andren, B. Bletterie, S. Kadam, P. Kotsampopoulos, C. Bucher: 'On the Stability of Local Voltage Control in Distribution Networks With a High Penetration of Inverter-Based Generation'; IEEE Transactions on Industrial Electronics, Volume 62, Issue 4 (2015), S. 2519 - 2529.

B. Bletterie, S. Kadam, H. Renner: 'On the Classification of Low Voltage Feeders for Network Planning and Hosting Capacity Studies'; Energies, 11 (2018), 3; page 1 - 23.

S. Kadam, B. Bletterie, W. Gawlik: 'A Large Scale Grid Data Analysis Platform for DSOs'; Energies, 10 (2017), 1099; page 1 - 24.

F. Kupzog, R. Schwalbe, W. Prüggler, B. Bletterie, S. Kadam, A. Abart, M Radauer: 'Maximising low voltage grid hosting capacity for PV and electric mobility by distributed voltage control'; e & i Elektrotechnik und Informationstechnik, Volume 131, Issue 6 (2014), page 188 - 192.

Book chapters

'Advanced Applications of DPL: Simulation Automation and Management of Results', Matthias Stifter, Serdar Kadam, Benoît Bletterie in: 'PowerFactory Applications for Power System Analysis', Springer International Publishing, 2014, ISBN: 978-3-319-12957-0, page 323 - 341.

'Smart Network Planning-Pareto Optimal Phase Balancing for LV Networks via Monte-Carlo Simulations', Benoît Bletterie, Serdar Kadam, Roman Bolgaryn in: 'Advanced Smart Grid Functionalities Based on PowerFactory', Springer International Publishing, 2018, ISBN: 978-3-319-50531-2, page 49 - 66.

Presentations and Poster Presentations

B. Bletterie, S. Kadam, A. Abart, R. Priewasser: 'Statistical analysis of the deployment potential of Smart Grids solutions to enhance the hosting capacity of LV networks'; Talk: EnInnov 2016, 14. Symposium Energieinnovation,

Graz, Austria; 2016-02-10 - 2016-02-12; in: "Proceedings EnInnov 2016", (2016).

B. Bletterie, J. Le Baut, R. Bolgaryn, A. Abart, S. Kadam: 'Hosting capacity of LV networks with extended voltage band'; Talk: 2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST 2015), Vienna, Austria; 2015-09-08 - 2015-09-11; in: 'Proceedings of the 2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST)', (2015), ISBN: 978-1-4799-7735-2; Paper ID 0531-vf-004928.

B. Bletterie, A. Zegers, S. Kadam, Z. Miletic et al.: 'On the Effectiveness of Voltage Control with PV Inverters in unballanced Low Voltage Networks'; Poster: CIRED 2015, The 23rd International Conference and Exhibition on Electricity Distribution, Lyon; 2015-06-15 - 2015-06-18; in: 'Proceedings CIRED 2015', (2015), Paper ID 1082.

S. Kadam: 'An overview of different tools for smart grid studies: Experiences from several European projects'; Talk: Tools for the Smart Grid: BNL/ISGAN Workshop on Modeling, Simulation and More, Brookhaven National Laboratory - Upton, New York, USA; 2016-04-14 - 2016-04-15.

S. Kadam, B. Bletterie: 'A comprehensive study on the actual potential of grid support functions to enhance the hosting capacity of distribution networks with a high PV penetration'; Talk: 6th International Workshop in Integration of Solar Power into Power Systems (SIW 2016), Vienna, Austria; 2016-11-14 - 2016-11-15; in: 'Conference Proceedings', SIW16-273 (2016), ISBN: 978-3-9816549-3-6.

S. Kadam, W. Friedl, R. Bründlinger, B. Bletterie: 'Potential der Wirkleistungsbegrenzung zur Erhöhung der Erzeugungsaufnahmefähigkeit in Verteilernetzen'; Talk: EnInnov 2016, 14. Symposium Energieinnovation, Graz; 2016-02-10 - 2016-02-12; in: 'Proceedings EnInnov 2016', (2016), 1 - 10.

S. Kadam, J. Le Baut, P. Zehetbauer, B. Bletterie: 'Methods and tools for large scale network data analysis'; Poster: Smart Grids Week 2016, Linz, Austria; 2016-05-11 - 2016-05-13; in: 'http://www.smartgridsweek.com/', (2016).

Research projects

Relevant work was performed within the project IGREENGrid[1]. In particular:

- Deliverable D5.1 Technical and economic Evaluation of replicability and scalability of solutions to increase the DER

- Deliverable D5.2 IGREENGrid simulation and evaluation framework

---

[1]https://cordis.europa.eu/project/rcn/106399_en.html

Executive Summary

Validated reference feeders for low-voltage networks (LV-networks) are a missing step in scaling and replicating outcomes and results of specific network simulation studies and field tests of a local area. Indeed, several measures have been defined and taxonomy studies performed to classify feeders. However, validation of reference feeders in terms of admissible amount of distributed energy resources (DER) is missing.

In this work, an in-depth analysis of network data provided by two Austrian distributed system operators (DSOs), with 14,000 LV-networks was performed to define reference feeders and validate them with the available grid data.

First of all, reference feeders should be unique from the topological point of view and easy to understand. Moreover, the number of reference feeders should be rather low. As a result, the problem was split into two smaller problems, first the classification by means of the topology of feeders and secondly, the parametrization of the reference topologies.

Secondary transformer types vary from network to network and may be changed over time based on the change of the demand of the supplied area. To begin with, the study was carried out on feeder level. The first and most important reason for this is to avoid misclassification or multiple classes of similar feeders due to different secondary transformer types. Consequently, the secondary transformers were replaced by a slack node in all networks. The hosting capability of a LV-feeder is introduced as the amount of DER that can be installed for a given DER-scenario while meeting voltage and loading prerequisites in the feeder disregarding loads and the secondary transformer. To classify feeders, two DER-scenarios were utilized to introduce a new measure, namely the hosting capability sensitivity (HCS). This parameter is the ratio of the achievable hosting capability for two particular DER-scenarios. To achieve this, the DER-scenario *uniform* and end of feeder (*eof*) are utilized. Within the DER-scenario *uniform*, the in-feed is assumed to be equally distributed across all connection boxes in the feeder (fair-case). The DER is placed only at the 'weakest' node in the feeder (worst case) in case of the DER-scenario end of feeder (eof). The parameter HCS was calculated for several voltage limits (from $1.01p.u.$ to $1.08p.u.$). The analysis of this parameter finally paved the path to define reference feeder topologies. Based on these prerequisites, in total 20 reference feeder topologies were defined.

Secondly, the parametrization of the reference feeders should be as simple as possible and a minimal number of parameters should be used.

Therefore, the correlation between the HC and feeder parameters were investigated. In Figure 1 the Spearman rank correlation coefficient of feeder

parameters and the hosting capability. High correlation coefficients are found for the parameters $dvdP$ and R_k ('Lastbus R') for rather low voltage limits ($1.01 p.u.$ to $1.03 p.u.$) and the DER-scenario *eof*. In general the higher correlation coefficient with a PF=1 control scheme decrease under a PF=0.9-control (ind.) scheme. Alongside the two parameters R_k and $dvdP$, additionally the sum impedance R_Σ, the feeder length and total cable length of the feeder show a higher correlation coefficient compared to other parameters. For the correlation coefficients of the parameter $I_{nom-Min}$ and the DER-scenario *eof*, the correlation coefficient is increasing for higher voltage limits. The dependency of the correlation coefficient on the voltage limit can be explained with the decreasing share of voltage constrained feeders for higher admissible voltage limits.
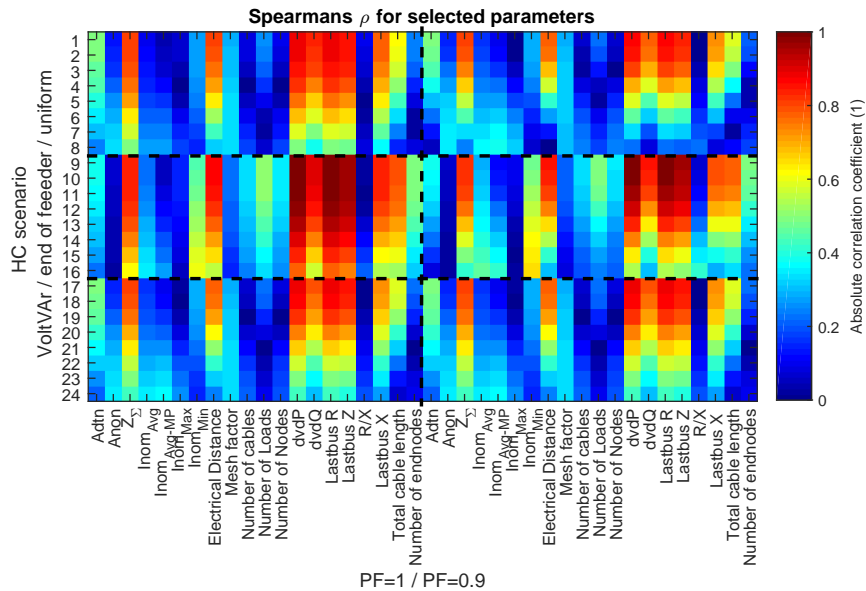

Figure 1: DSO 1 and DSO 2 - Heat map of the Spearman correlation. The plot shows the correlation between statistical parameters and the hosting capability for various scenarios (y-axis). The plot is divided into 6 areas according to the DER-scenario and powerfactor. The bottom left and right part are identical (The VoltVAr-control is overruling PF=1 and PF=0.9)

Two evident parameters were found by the correlation analysis feeder parameters and the calculated hosting capability:

- The short circuit resistance R_k of the end node (the most critical). Since a slack is used instead of the secondary transformer, the short

2

circuit resistance depends solely on the resistance of the cables on the main path from the slack to the end node. Hence, in order to obtain this resistance only the laying of the cables on the main path is required. Thus, R_k is the first evident parameter to parametrize feeders.

- The second evident parameter of a feeder is the cable with the lowest ampacity $I_{nom-Min}$ (weakest cable). This value can be retrieved, if the cable type data of all cables in a feeder is accessible (e.g. from an asset database).

The validation of the methodology and the reference feeders was performed using the LV-network data of two Austrian DSOs with about 14,000 networks. The methodology allows the estimation of the HC of feeders with a relatively low error depending on the admissible voltage limits. Furthermore, the reference feeders were also validated considering a reactive power control scheme (PF=0.9 ind.).

In conclusion, three parameters are needed to classify feeders (HCS, R_k and $I_{nom-Min}$). First, the reference topology is selected according to the parameter HCS. After that the reference feeder is configured with the electrical parameters R and $I_{nom-Min}$.

In Figure 2 the distribution of the HC of all feeders for different voltage levels are depicted for the DER-scenario *uniform* and PF=1 for both DSOs. Up to $1.07p.u.$, a steady increase of the HC can be observed. At $1.07p.u.$ and $1.08p.u.$, the 95th percentile remains stable, but the median and the 5th percentile are still increasing by about 10kW and 5kW, respectively per voltage step. Concerning the allowed voltage rise, allowing a voltage rise of $1.02p.u.$ instead of $1.01p.u.$, leads to a significant HC increase. In that case, the median value can be doubled from about 25kW to more than 50kW. If the admissible voltage rise can be increased to 4%, more than 50% of the feeders have a HC higher than 100kW. For higher admissible voltage limits a higher share of feeders become loading constrained. Hence, a voltage increase has no impact on the achievable HC anymore, since a cable in the feeder may be fully loaded. Therefore, the maximal HC is reached.

The absolute HC mismatch in case of DSO 1 is depicted in Figure 3. In this figure, the absolute mismatch in MW is shown for all investigated voltage limits. For example, the overestimation of the HC by more than 10kW is exceeded for 10% of the feeders at a voltage limit of $1.04p.u.$ and the underestimation by more than 10kW for 7% of the feeders at the same voltage level. Hence, at a voltage limit of $1.04p.u.$, the mismatch is within a $\pm10\%$kW range for about 83% of the feeders. For lower voltage limits, even more feeders are within this range. At a voltage limit of $1.08p.u.$, the
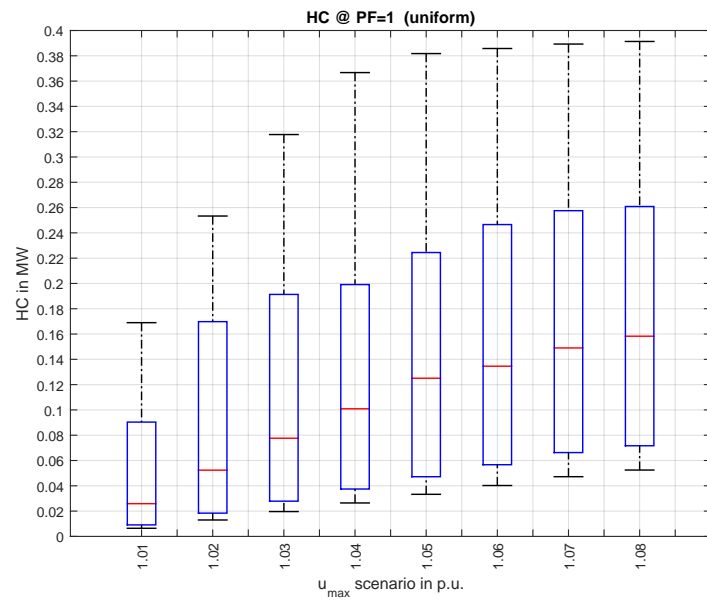
Figure 2: DSO 1 and DSO 2 - Hosting capability for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 1)

mismatch is underestimated by more than 10kW for already almost 14% of the feeders and overestimated by more than 10kW for 33% of the feeders. Hence, only 53% of the feeders remain within a ±10%kW range, indicating that the methodology is not suitable to accurately estimate the HC at higher voltage levels. The parameter describing the topology (HCS) is the ratio of
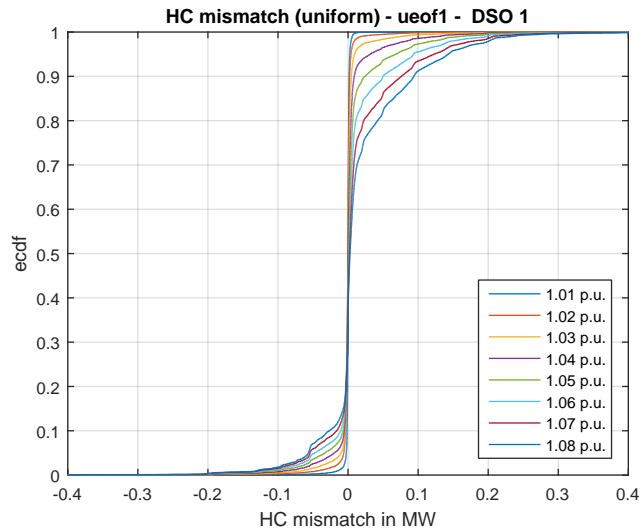
Figure 3: DSO 1 - Empirical cumulative distribution function (ecdf) of the hosting capability mismatch for different admissible voltage rise limits (DER-scenario *uniform*, power factor 1)

the HC for to particular scenarios (uniform and eof) and is rounded to the nearest quarter. In the case, that a feeder contains only a load at the feeder end, both HC values are identical, and the feeder can be modeled by a feeder consisting of two terminals and a line. The distribution of the parameter describing the topology (HCS) for the investigated DSOs is shown in Figure 4. In the case of DSO 1, approximately 11% of the feeders and in the case of DSO 2, only about 17% of the feeders, could be modeled by a two node equivalent. For a value of 1.25 and values above 2.25, similar shares are observed for both DSOs. DSO 1 has a higher share of HCS values between 1.5, 1.75 and 2 when compared to DSO 2.

Additionally, the distribution of RFT in all LV-networks was investigated. Thereby, a high variation of LV-networks was observed, leading to a high number of reference LV-networks. Even though some simplifications were considered, the number of reference LV-networks could not be reduced to a reasonable low number.
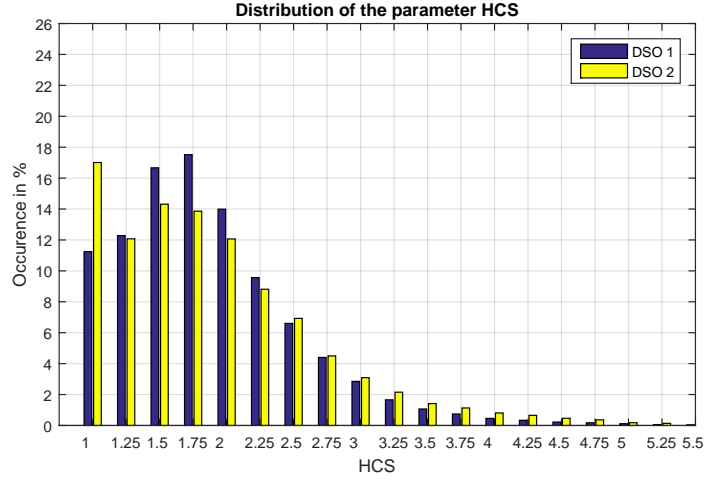
Figure 4: DSO 1 and DSO 2 - Distribution of the hosting capability sensitivity (rounded to quarters). Blue: DSO 1, yellow: DSO 2

However, the proposed methodology allows future comprehensive studies (e.g. hosting capability studies) on the complete supply area of DSO (or even on country level) by reducing the required grid data to a rather small number of reference feeders. Additionally, the proposed methodology can be used to replicate findings of particular field tests to equivalent representative feeders.

A number of reference feeder topologies were defined. In Figure 5 a reference feeder topology is depicted with a 0.4kV slack (External Grid) connected to Terminal 1. Line 12 connects Terminal 1 and Terminal 2. The number below the line name defines the number of parallel lines (in this case 1) of the same line type. This labeling of lines applies to all lines. At Terminal 2, two feeder segments are connected. Segment I consists of the Line 24 (1 cable), Terminal 4 and Load 3 . Segment II contains of Line 23 (2 lines in parallel), Terminal 3 and two loads (Load 1 and Load 2). By convention, all cables have the same line type and length. The impedance of each cable is determined by number of lines in parallel. Eight reference feeder topologies were found that are suitable to describe the topology 90% of the real feeders.
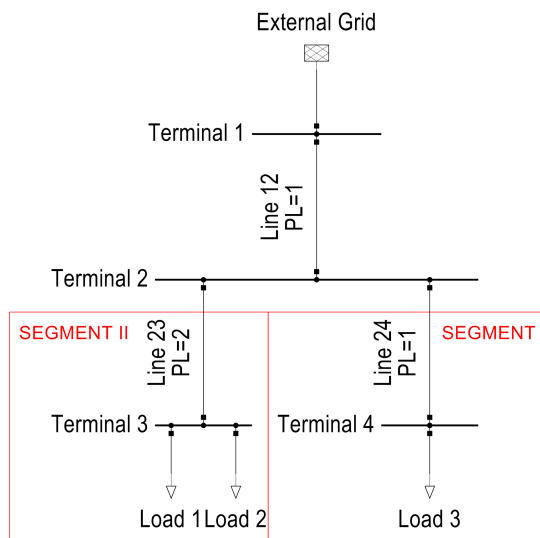
Figure 5: Feeder segments and labeling convention used for the developed reference feeder topologies

Nomenclature

ADTN	(m)	Average distance to neighbor nodes
ANON	(-)	Average number of neighbor nodes
dvdP	(%/kW)	Real power sensitivity
dvdQ	(%/kVAr)	Reactive power sensitivity
HC	(kW)	Hosting capability
I_{nom}	(kA)	Ampacity of a line in a feeder
$I_{nom-Avg}$	(kA)	Weighted average ampacity of all lines in a feeder
$I_{nom-Avg-MP}$	(kA)	$I_{nom-Avg}$ only on the main path of a feeder
$I_{nom-Max}$	(kA)	Ampacity of strongest line in a feeder
$I_{nom-Min}$	(kA)	Ampacity of the weakest line in a feeder
l	(km)	Electrical distance between slack and end node (Feeder length)
L	(km)	Total cable length of all cables in a feeder
R	(Ω)	Resistance
R_k	(Ω)	Short circuit resistance at the end node
u	(p.u.)	Voltage
u_{max}	(p.u.)	Maximal admissible voltage level
X_k	(Ω)	Short circuit reactance at the end node
Z_k	(Ω)	Short circuit impedance at the end node
Z_Σ	(Ω)	Equivalent sum resistance

Abbrevations

CAPEX   Capital Expenditure
DER   Distributed energy resources
DG   Distribution Grids
DNO   Distributed network operator
DPL   DIgSILENT Programming Language
DSO   Distributed system operator
ecdf   Empirical cumulative distribution function
eof   End of feeder, Feeder end
GIS   Geographical Information System
HC   Hosting Capability
HCS   Hosting Capability Sensitivity
HV   High-voltage
IGC   Installed generation capacity
LDF   Load-flow calculation
LV   Low-voltage
LVR   Line voltage regulator
MF   Mesh factor
MV   Medium-voltage
NEN   Number of end nodes
NON   Number of neighbors
OLTC   On-load tap changer
OPEX   Operational Expenditure
PCA   Principal Component Analysis
PF   Power factor $\cos\varphi$
PL   Parallel lines
PV   Photovoltaics
RES   Renewable Energy Resource
RFT   Reference Feeder Topology
SHC   Short circuit calculation
SRA   Scalability and replicability analysis
Term.   Terminal
ueof   HC ratio of DER scenario uniform and eof
ueof1   HCS obtained at $u_{max} = 1.01 p.u.$
ueof2   HCS obtained at $u_{max} = 1.02 p.u.$
ueof3   HCS obtained at $u_{max} = 1.03 p.u.$
ueof4   HCS obtained at $u_{max} = 1.04 p.u.$
ueof5   HCS obtained at $u_{max} = 1.05 p.u.$
ueof6   HCS obtained at $u_{max} = 1.06 p.u.$
ueof7   HCS obtained at $u_{max} = 1.07 p.u.$
ueof8   HCS obtained at $u_{max} = 1.08 p.u.$
uni   uniform
VRDT   Voltage regulated distribution transformer

1 Introduction

In this section, an introduction of challenges that network planners and operators are facing, as a result of the increasing number of DER in the electricity grid is given. First, the development of DER in distribution grids is discussed, showing the need to have a better knowledge of the distribution grids. Secondly, the developed smart grid functionalities, to increase the hosting capacity and to maintain voltage quality requirements, are discussed. Finally, the research question of this work is presented in the context of increasing numbers of DER, particularly for low-voltage networks.

1.1 Development of distributed energy resources in Distribution Grids

In [1] it is reported, that solar PV is already covering more than 7% of the electricity demand in three countries worldwide (Italy, Greece and Germany). Additionally, in 2014 the power generation capacities for Wind and PV in the EU 28 increased by 11,791 MW and 6,574 MW respectively. Globally, about 40 GW PV have been installed in 2014. Moreover, the total amount of installed PV in the world reached 178 GW and covers more than 1% of the world electricity demand. In the next years, the installed capacity is still expected to rise. However, in Germany for example, in certain times of the year, electricity generation from renewables cover more than 50% of the power supply, meaning that these situations have to be accounted for today, so as to smoothly pave the path for the transition to an energy system with a dominant share of fluctuating DER. According to [2], the overall PV deployment potential in Germany (depending on the efficiency) varies between 229 GW and 569 GW. Hence, compared to the generation statistics presented in [3], only 10% of the average PV deployment potential is already utilized.

Authors in [4] report, that the energy transition has a significant impact on the overall power system infrastructure and its operation. As an example, the fact that PV-systems and Wind generators in Germany, combined, are capable of covering more than 50% of the peak load under favorable weather conditions and low demand (e.g. weekend). Thus, accurate forecast and real-time data has to be available, for a large area to maintain grid stability. In the case of a particular German TSO, which uses forecasts from four different sources, the root mean square error reaches 5-7%, revealing a need to research on how to improve the forecasts.

In Table 1 the total installed generation capacity (ICG) for the year 2012 is listed for seven European countries according to [3]. The total installed

capacity varies between 8.6 (Ireland) and 171.6 GW (Germany). The amount of the total installed generation capacity (IGC) in distribution grids is given in the second row. The total IGC grids is at a comparable level for France, United Kingdom, Belgium, while the installed capacity in Germany is higher than the sum of the other countries. In the third row of the table, the share of DER in distribution grids is given. The share of renewables[2] in the distribution grids vary between 48% (Belgium) and 99% (Italy). The remaining countries can be grouped in two groups with a share of renewables in distribution grids (DG) of about 65% or 82%. The table shows that a significant share of renewables are connected to distribution grids. This highlights the importance of understanding how an even higher share of DER can be reached in distribution grids without jeopardizing the quality of supply on the one hand while limiting the costs of network reinforcements on the other hand.

Table 1: Total installed generation capacities (IGC) and share of renewables (2012) according to [3]

	IGC	Installed Capacity in DGs	Share of IGC in DGs
	(GW)	(GW)	(%)
Belgium	20.8	14.3	69
France	128.7	15.5	12
Germany	171.6	98.7	58
Ireland	8.6	1.3	15
Italy	128	22	17
Portugal	18.5	6	32
United Kingdom	77.9	14.7	19

The cable length of transmission and distribution grids varies significantly. In Germany, for example the total cable length in the 110kV-grid is about 66,000km, which is more than one and a half times the length of the equator. The length of all cables in the MV distribution grids (20kV and 10kV), is higher than 500,000km, covering 1.3 times the lunar distance[3]. Finally, the total cable length in all 0.4/0.23kV LV-networks in Germany is above 1.15 million km, which is equivalent to about 80% of the diamter of the sun[4]. The length of the lines in LV-networks is more than 17 times the length of the 110kV-grid. According to [5], about 90% of the lines in LV-networks are underground cables . In Austria, the length of all cables of the 110kV

---

[2]Wind, biomass, biogas, hydro, solar, other renewables
[3]The distance between earth and moon is about 384,400km - www.esa.eu
[4]The diameter of the sun is nearly 1.4 million km - www.esa.eu

grid is 11,365 km. Whereas the total length in the LV-networks is about 172,000km of which about 80% are underground cables. In Austria, the length of the lines in LV-networks is more than 15 times higher compared to the length of the 110kV-grid [6], which is in a similar range as in the case of Germany. This comparison reveals the dimension and the complexity of describing and classifying distribution grids and in particular LV-networks. Distribution grids were not designed to integrate a large number of DER. According to [7], about 70% of the voltage band available in existing LV-grids is typically allocated to the voltage drop at maximum load. Furthermore, authors report that the impact of single-phase components requires consideration of the voltage drop in the neutral wire. Therefore, power, that is six times higher has to be taken into account when compared to balanced components, reducing the hosting capacity significantly. In [8] required grid development and innovations for the energy transition, which are estimated until the year 2030 are discussed. Furthermore it is addressed that regulations have to be adapted for the foreseen ambitious integration of DER plans. More than 60% of the expected grid reinforcement costs are related to the transmission system. The remaining expenses are split with two thirds for MV-grids and one third for LV-grids. Authors report, that the hosting capacity of a network area is approximately as high as the peak consumption. Grid reinforcements are required if the DER penetration is further increased above the peak consumption. The network areas are clustered according to the expected development of Wind and PV installations, the population density and the area of municipalities. According to these properties, eleven cluster centers were found. Another important finding is, that the role of DSOs will change due to a high number of participant in the electricity market. It is foreseen that DSOs will have an important role in the coordiation of ancillary services in their supply area.

The way in which the security of supply and the reliability can be maintained with a targeted share of 80% of DER in the year 2050 for the German electricity grid was analyzed in [9]. It is discussed, how ancillary services such as frequency and voltage control, re-establishment of the power generation and network operation could be provided mostly by fluctuating DER in the future. Moreover, recommendations are given on how the share of ancillary services provided from conventional powerplants could be further provided by DER.

Even though e.g. the reliability of distribution grids is assessed with probabilistic tools, the quasi-stationary voltage stability is evaluated deterministically. This issue is addressed in [10]. It is reported, that the quasi-stationary evaluation is related to a unlikely extreme network condition resulting in an over sized network or rejection of DER connection requests. In this work, a

probabilistic tool was developed to evaluate the quasi-stationary voltage stability. Thereby for each node of a network, realistic voltage ranges and their particular probability of occurrence is calculated while considering planned and unplanned outages. The developed methodology reveals higher transmission capacity reserves as well as a higher hosting capacity of the grid. The issue of active voltage control strategies of DER are not being considered in network planning tools is also addressed in [11] and a study on statistical distribution network planning is presented.

Not only statistical methods are suitable for application to electricity grids. The ware-house-location-problem for example, was utilized to support network planners in selecting the most suitable locations for the secondary substation in [12]. Even though radial LV-feeders are rather facile to describe with graph theory, this work is an important example, demonstrating that tools and methods known from other fields could be applied to distribution grids and particularly integrated in network planning and operation of distribution grids. Thus reducing capital expenditures (CAPEX) and operational expenditures (OPEX) of network operators and additional costs for customers. It is also apparent that, the energy transition and the trend of DER being mostly installed in distribution grids requires better knowledge and more powerful tools to estimate the hosting capacity of grids. Moreover, tools and methods are available to reduce the impact of decentralized active power injection. Additionally, the costs of grid reinforcements should be as low as possible to support the further deployment of DER.

1.2   Smart Grid Functionalities

The increasing number of DER in distribution networks raises several challenges in the field of network planning and operation. In particular, the hosting capacity in some distribution grids may already be exhausted. On the one hand, DER provide major benefits to reduce CO_2 emissions and to cover local demand. On the other hand, if generation and consumption are imbalanced, over-voltages and reverse power flows occur leading to increased network losses. Due to the objective of PV owners, which is to maximize the yield of their installation, the tilt and angle of PV installations are optimized, resulting for nearby installations, in a rather high coincidence factor for PV in-feed. Even though the full load hours of PV installations is rather low (e.g. 1100h per year in Central Europe [13]) the nominal power of the installations has to be considered as a worst case in conventional network planning. The integration of the so called "last kWh" of a DER is cost intensive. Smart grid functionalities aim to increase the hosting capacity while reducing DER-integration costs, maintain voltage quality and avoid

the overloading of elements.

The way in which the hosting capacity can be increased with extended planning and voltage controls strategies was investigated e.g. in [14]. Authors report, that both unbalanced loads and generation reduce the effectiveness of centralized voltage control strategies. Moreover, the hosting capacity of LV grids can be increased by reducing the imbalance by reconfiguring the phase connection of single-phase DER. In [15], it is reported that with reactive power control of PV plants in medium-voltage grids the voltage rise caused by DER in-feed can be reduced to between 16% and 49% for specific line types and a PF of 0.95. From a regulatory point of view, reference networks could also be interesting for benchmarking the network planning principles of DSOs. In [16], as an example, a reference network model is presented that can be used to design a reference grid (LV/MV and HV) for a given area with a green field approach as well as considering the street map of that area. Significant differences between both approaches occur, that are discussed. The proposed methodology is capable of efficiently and economically planning a large-scale distribution area with millions of customers.

Further reactive power control strategies were developed to reduce the voltage rise in distribution networks caused by active power in-feed. In [17] for example, the VoltVAr and VoltWatt control were identified as most promising solutions for increasing the hosting capacity of low-voltage distribution grids from a technical and economical point of view. Reactive control strategies were simulated, tested in the laboratory and evaluated in the field under real grid conditions e.g. in [18] and [19]. In [20], the effectiveness of reactive power control strategies as a function of the voltage of three-phase inverters under unbalanced conditions is discussed. Further, in [21] active and reactive power control concepts were investigated in terms of effectivity, performance, additional reactive power flows, network losses and active power curtailment. In this simulation based work it was shown that with such control concepts the voltage rise caused by DER can be reduced and that the hosting capacity of low-voltage networks can be increased significantly.

Already today, a high number of fluctuating DER dominantly supply electricity demand (e.g. in Germany) for certain periods of the year. Under such conditions, the stable operation of inverters and their voltage support functionalities must be guaranteed. This issue was already tackled in several contributions. The stability of the VoltVAr control was for example investigated in [22] and recommendations to parametrize this control strategy are given. In [23], the open loop stability of the VoltVAr control was investigated and limit curves have been obtained which depend on the X/R ratio at the point of common coupling and the settling time constant. Further, in [24], the impact of reactive power controls strategies and the OLTC were

14

investigated in terms of stability for six sample networks. Moreover, in [25] the results of a stability study conducted on a single-inverter system and on a multi-inverter system are presented. In conclusion, that the requirements for reaching a stable operation can easily be met for integrated systems but could be a significant constraint for systems relying on communication.

In LV-networks with underground cables, the effect of reactive power control strategies to increase the hosting capacity of networks is limited in comparison to that of overhead-line networks. Thus, active power has a higher impact on the voltage than reactive power in networks with a high R/X ratio. Regarding an obliged active power curtailment for PV owners, the information of the percentage of the yield that could not be fed into the grid is important. In [26], a methodology is presented how active power curtailment could be compensated from an economical point of view. Thereby a set of parameters that need to be considered, is developed. Further work on evaluating yield reductions caused by active power curtailment is presented in [27]. Additionally, a decision matrix is given to decide whether the active power of an installation shall be curtailed statically (e.g. to 70% of the nominal power) or dynamically (requiring communication). Alternatively, inverters may also provide ancillary services to reduce losses and voltage unbalance in four-wire distribution grids. Injecting the produced power of a three-phase PV inverter into the phase with the highest power consumption, or to transfer power from highly loaded to less loaded phases, without overloading the PV inverter is investigated in [28]. By balancing the grid, losses as well as the used voltage band can be reduced and the hosting capacity can be increased.

The general impact of smart grid functionalities is important from the DSO point of view as well as for regulatory bodies and stakeholders. Until now, smart grid functionalities are tested in specific demo areas in various countries. Hence, impacts can only be studied in the investigated networks and results cannot be translated into general conclusions easily. To study the expected impact and benefits of a smart grid functionality deployed in the entire supply area of a DSO, or on country level, a scalability and replicability analysis (SRA) is needed. In the project Grid4EU ([29]), innovative concepts were tested in demo areas of six European DSO. The aim of these concepts are to improve network operation in distribution grids, to optimally and smoothly integrate a high number of DER, make use of active demand and demand side management, to balance the supply and demand and assess islanding of supply areas to increase grid reliability. It is also mentioned that for scaling-up and replicating the impact on network hosting capacity, the characteristics of distribution network plays an important role. In particular, the feeder length, R/X ratio and the topology as well as

the location of DG units are pointed out. The scalability and replicability analysis developed in the project GRID4EU is presented in [30]. It is mentioned that scalability may focus on the density (e.g. more DER flexibilities) or in size (large-scale deployment). One of the lessons learned is that the characteristics of distribution networks and operational approaches are relatively homogeneous across European countries. Furthermore, in this work representative networks were defined for the technical SRA. For the considered six demos in six different European countries, seventeen MV and eight LV representative networks were built. The MV networks were classified as urban, sub-urban, rural concentrated and rural scattered. The LV-networks classified were classified as urban, sub-urban, rural, rural overhead lines and rural underground. The IGREENGrid project ([31]) focused on increasing the hosting capacity for Distributed Renewable Energy Sources (DRES) in power distribution grids without compromising the reliability or jeopardizing the quality of supply. During this project, key performance indicators were developed for comparing and objectively evaluating best practices for increasing the hosting capacity of distribution grids. Furthermore, a scalability and replicability study was performed including a cost-benefit-analysis. In [32], the main type of disruptions in distribution grids caused by a large penetration of distributed renewable energy sources such as voltage violations, local balancing of generation and demand are addressed. Further strategies for recovering from faults and blackouts through network reconfiguration or islanding operation are discussed. Moreover, monitoring and automation systems, state estimation, optimal power flow and congestion forecasting solutions are mentioned to prepare and adapt to changing grid conditions and to withstand and recover from disruptions.

Smart grid functionalities are promising solutions for efficiently increasing the hosting capacity of grids at competitive costs compared to grid reinforcements. These functionalities were tested using exhaustive simulations, laboratory tests and tests in dedicated demo areas in real networks. However, the unknown scalability and replicability potential of such functionalities is a barrier in the large-scale deployment. Various regions and federal states have more ambitious targets regarding the integration of renewables in their area, when compared to country level. In [33] for example, a grid study was conducted for the federal state of Rheinland-Pfalz. Rheinland-Pfalz targets to supply the entire federal state by renewables until 2030. The grid requirements and suitable smart grid solutions for such a development are investigated in the grid study to identify potential saving opportunities. Other similar studies are e.g. [8] or [34]. For that, the distribution grids have to be studied and modeled such that findings of network simulations can be replicated and scaled to federal state level, since the individual simulation of

all LV-networks of a DSO is a time-consuming and complex task. Therefore, simplified and validated models of LV-networks are needed to focus on main impacts and to be able to generalize results of a network study based on a low number of networks. With a SRA, these results could be generalized and recommendations given.

1.3   Research Question

To study the effect of DER and voltage control strategies in LV networks on the (MV) distribution and (HV) transmission grid, the entire grid of DSOs has to be modeled in detail with a high number of nodes. However, due to the high number of nodes and voltage dependent control strategies, the computation time for power flows may be significantly higher compared to networks with a few nodes. Hence, only very specific power flow calculations may be performed in a reasonable amount of time for the distribution grid. Simplified LV-feeder models would allow a reduction in the number of nodes and hence the computation time while maintaining the accuracy of power flow calculations and voltage control strategies. Therefore, there is a need for a methodology to replace full models of LV-feeders by average feeder models. Such average models have to be developed and tested for using load flow calculations.

Consequently, for this bottom-up approach a rather high amount of grid data is required to first gather results for a high number of scenarios and then validate the average feeder models against the real feeder results. The aim of this work is to develop an average feeder methodology, to select and parametrize equivalent feeders suitable to replace real feeders for large scale grid analysis. The admissible DER penetration level of real feeders for various scenarios is thereby used to measure the accuracy of the developed average feeder models.

Research questions:

- How can simplified models for low-voltage feeders be defined?

- How can similarities between low-voltage feeders be identified?

- What is the impact on the accuracy when running studies on simplified models compared to the full model?

The aim of this work is to answer these questions by running in-depth analysis on the network data provided by two Austrian DSOs. Furthermore it is investigated, if equivalent feeders can be designed to accurately replace real LV-feeders considering both active and reactive power flows. This would allow the effects of voltage control strategies in LV-networks in a multi-voltage

level grid to be studied. Further, the methodology shall provide a suitable measure to test if particular feeders have similar topology and a similar configuration of the equivalent feeders. If it is the case, results of one feeder will be also valid for the other feeder and vice versa. The developed equivalent feeder models are then suitable for reducing the number of nodes for scalability and replicability studies in multi-voltage level grids.

2 Related Work

Depending on the focus of a study, classification or clustering of electricity grids can be performed at the feeder or network level. In this section related work on feeder and network taxonomy and classification is presented to give an insight into the identified most relevant parameters in literature.

In [35], the rising challenges in distribution grids to host a high number of DER is addressed and a methodology to estimate the hosting capacity of residential neighborhoods, when loads are considered. Thereby, the obtained hosting capacity shall be apportioned equally to all customers. Networks were grouped into nine settlement types characterized by the floor-space index, site occupancy ratio and buildings per hectare. The settlement types together with electrical parameters of the network shape the characteristics of the networks. The discussed electrical parameters are the rating of the secondary transformer, number of feeders, length of feeders, used line types, equivalent load location and number of housing units per house connection. Per housing unit, a hosting capacity of 2.2kW to 3.1kW was obtained. The author reports for a particular settlement type, that the network losses decrease until a penetration rate of 20% of the hosting capacity. At higher penetration levels, the network losses are increasing and are exceeding the initial losses at already 50% of the estimated hosting capacity. Furthermore, deploying the maximum amount of DER leads to more than 3 times higher losses. Another contribution towards the estimation of the hosting capacity of LV-networks is presented in [36]. The impact of PV in LV-networks was investigated considering the rooftop PV-potential. The author concludes that the hosting capacity of networks in villages, rural and sub-urban areas is rather critical. In this work, seven typical LV-networks for villages, rural and sub-urban areas are presented. Furthermore, it is discussed, that due to the rather low number of customers in LV-networks an assumption with regards to a base load is rather risky.

In [37] a new method to create synthetic power grid topologies is proposed and differences between electrical and topological structure of power grids is discussed. For that, a weighted graph is utilized. Another contribution utilizing graph theory can be found in [38]. In this work, a minimum spanning tree algorithm is proposed to design distribution networks with optimal configuration to minimize network losses considering that only radial topologies are allowed.

In [39] it is stated, that feeder data is required for network taxonomy studies. Thereby the five most significant parameters were identified: the total cable length, the length of underground cables and overhead-lines, the number of delivery points as well as the rated voltage and power of the

distribution transformer. However, this work did not lead to a classification of feeders into urban, suburban or rural feeders. Several short, middle and long feeders with a specific number of branches are proposed as reference feeder types.

In [40] new local voltage support strategies for increasing the hosting capacity of LV-networks are presented. Thereby a 2D-pane is utilized to illustrate the hosting capacity constraint reason and the remaining reserve in terms of admissible voltage limit and maximal loading of network elements. Hence, LV-networks can be classified in two groups, namely voltage or loading constrained. Three different critical networks were used to investigate the effect of the support strategies. Furthermore, for each critical network, feeder relevant parameters are listed: length of the feeder, distance between neighbors, length of service lines, number of consumers per connection point, feeder line type and service line type. Five feeders are supplied in the suburban network and three feeders in the rural network. In three networks, only one particular feeder type is supplied.

In [41], the feeder hosting capacity for distributed solar PV was analyzed stochastically. In this work, four base load levels were considered: absolute maximum and minimum, irrespective of time of day and the maximum and minimum load around midday. Further, a high number of PV distribution scenarios were considered. One of the conclusions in this work is that the impact that distributed PV will have on a specific distribution feeder can only be determined based on knowledge of that feeder's characteristics. These characteristics include, but are not limited to load, voltage, regulation, and impedance' [41]. Moreover it is stated, that feeder characteristics are a key factor in the integration of PV. Nevertheless, the size, location and output of the PV system have a high impact on the hosting capacity of feeders. Moreover, a hosting capacity study was performed such that 20% of the nodes were allowed to violate the voltage threshold. Indeed, this addresses an important issue concerning the definition of the hosting capacity of feeders.

Furthermore, the expected costs of grid reinforcements in Germany caused by the expected rise of DER were investigated in [42]. Thereby potential savings by utilizing smart grid functionalities and their requirements are discussed. In this work, ten representative LV-networks were defined, whereas only 35% of the DSOs are assigned to representative network classes, where significant network reinforcement is expected. Moreover, for only 8% of the 500,000 German LV-networks are network reinforcements foreseen. Authors report, that curtailing 1% of the annual yield of DER could reduce the network reinforcement costs by 30%. For each of the representative network classes, 100,000 network models were generated. The assignment of DSOs to network classes was based on three parameters: Average installed PV

power per point of connection, average installed wind power per point of connection and average yearly peak consumption per point of connection of DSOs. Structural parameters, such as the number of feeders, average feeder length, distance to neighbors and the ratio of underground and overhead-line cables were utilized as well. A Monte-Carlo approach is used (based on real parameter distributions) to generate networks of a given class.

In [43], a method to classify distribution feeders into clusters and to select representative feeders from each cluster was developed. Identified parameters of feeders that affect the hosting capacity of feeders amongst others are nominal voltage, feeder length and main conductor type, three-phase vs. single-phase feeder length. Authors state that "*unfortunately, the response of each feeder is unique and the time and effort involved in the detailed study is too great to replicate on thousands of additional feeders. Therefore, the impact results found in the detailed study cannot be mapped directly to other feeders*" [43]. However, a streamlined hosting capacity method was developed, which is presented in [44], that takes into account feeder characteristics by calculating a power flow and a short circuit combined with a characterization of the feeder topology. The streamlined hosting capacity method is presented. "*The streamlined hosting capacity method captures the electrical and consumer characteristics through load-flow and short-circuit analysis, runs through a sequence of impact studies, and then provides the feeders ability to accommodate PV (hosting capacity)*" [44]. Authors report that this methodology is suitable for calculating the distribution system impact of PV, in a fast and effective way. Further, "*optimal*" locations for PV can be identified. Furthermore, authors report that the streamlined method is accurate regarding the lower hosting capacity of feeders and is more conservative at the higher hosting capacity.

According to [45], there are over 300,000 feeders in the LV grid of the Netherlands. Authors clustered 88,000 feeders of the largest Dutch DSO with a fuzzy k-means algorithm. Main network parameters, such as impedances, cable length number of branches, branch depth and the number and type of connected customers are used in combination with the graph theory concepts of degree distribution, sequence and the centrality of the power, impedance and length. Further, a logarithmic transformation was applied to several non-discrete parameters to obtain a more symmetric distribution. The eight most common LV-feeder types are presented in this work. For these common types the length of the feeder, number of branches, impedance, main line type, number of customers and occurrence are given. However, the LV-feeder with highest occurrence represents 6.4% of all feeders, because in total 94 classes were found. The authors conducted load flow studies for the year 2014 with 15-min load profiles and calculated the minimum and maximum voltage as

well as the maximum loading of cables. In this study, the voltage at the LV-side of the secondary transformer was assumed to be at 1p.u., neglecting voltage drops at the secondary transformer and voltage fluctuations in the upstream MV-network. Moreover, risk levels have been defined for voltage and loading. (e.g. voltage drop and rise more than 8%, or smaller than 6% - loading below or higher 90%). This work demonstrates the complexitiy of clustering and classifying LV-feeders.

In [46], power grids were investigated from topological point of view by means of graph theory. In conclusion, it was stated, that no single model can be defined that accurately reflects all power grids and that: "*power grids seem to be poor candidates to treat, both by modeling and analysis, as general complex networks*". it is apparent, that analyzing radial or continuous feeders with graph theory (clustering coefficient, betweenness, node degree etc.) does not generate significant additional benefits in terms of feeder taxonomy.

MV and LV grids were investigated from a topological point of view in [47]. Authors report, that the node degree in distribution grids follows rather a power-law distribution. Moreover, betweenness plots are presented and differences to HV-grids are discussed. While in literature there is a consensus that the node degree for HV-grids is exponential, a power-law distribution is observed for LV-grids.

In [48] a Matlab GUI for the generation of distribution grid models is presented, which generates statistically correct distribution grid models. Thereby the node degree and degree distribution is used. For LV-networks, two statistics have to be given. The normal distribution of the total number of connections in a single secondary substation and the normal distribution of the connections per LV feeder.

In [49], LV-networks were classified as dispersed settlement, hamlet, village, suburban, urban and industry. For each of these classes, a typical network was defined. Moreover, a network for each worst case assumption in terms of voltage, loading and secondary transformer was developed. Thereby, the feeders of a network have different topologies, lengths and line types. In this work, the rating of the secondary transformer as well as the number of house connections and the distance between house connections are the most significant and discriminatory factors. Furthermore, line models are reduced to their impedance and the nominal current.

In [50] a two-bus equivalent model, which may be used to estimate the maximum voltage in an LV area due to PV generation over time is presented. Thereby a network is replaced by a slack with a defined reference voltage, an equivalent network impedance and a node connecting a load and a PV-installation (the defined equivalent impedance is in analogy to [36]). Authors report, that approximately 96% of the areas could be accurately

modeled with this approach, while the absolute error remains below 0.3% under balanced conditions. Further for the investigated area covering 10,213 LV-networks, statistics of the most relevant network parameters are given.

In [49] 1.35% of the grid data of a DSO was analyzed to estimate future network reinforcement costs according to different scenarios. In this work, networks were split into feeders (considering the secondary transformer) and three main feeder topologies were identified, which are suitable for representing the topology of the analyzed network data.

In [51], a methodology is presented to classify distribution feeders (MV) into clusters and to select representative feeders from each cluster. Thereby the cubic clustering criterion is used to determine a reasonable number of clusters to cluster with the K-means clustering methodology. For each of the investigated utilities, a rather small number of five to twelve clusters seems reasonable. Furthermore, correlation maps were used to reduce the number of clustering variables.

Self organized maps (with 3×3 neurons) were utilized to classify distribution feeders into specific groups of representative feeders in [52], with the objective of quickly and accurately obtaining the hosting capacity of feeders (in MV-networks). In this work nine cluster were found based on seven selected variables. The chosen variables are: summer peak, total 3-phase length, transformer count, total kVA, residential kVA, commercial kVA and industrial kVA. However, the validation of the clusters regarding the Hosting Capacity of the feeders is not given. Since the authors performed the analysis on feeder level, it is also revealed that feeder clustering is also suitable in MV-networks.

In [53], 358 networks with 1550 feeders of a German DSO were analyzed and reference networks (LV) based on the typical networks developed in [36]. In total 29 electrical parameters and two geographical parameters were calculated. The target of this work, is to investigate the effectiveness and stability of voltage control strategies in distribution networks, including their validation. In total three network classes are defined: rural, village and suburbs. Furthermore, for the classes rural and village, a dominant overhead line network and a dominant cabled network is presented. All network classes contain a secondary transformer with $u_k = 4\%$, but a different nominal rating. The feeders of all presented networks are radial feeders with a particular length and a number of nodes. The nodes are placed equidistant along the feeder with one customer at each node. Moreover the lines from point of connection to customer premises are modeled with a fixed line type and length.

In [54], the insufficiency of the classification of networks based on the population density of areas is is addressed, since the population density can-

not be matched accurately with LV-networks. Therefore, clustering based on electrical parameters of networks is proposed. From 27 calculated parameters, five most suitable parameters were identified: line length, rating of the secondary transformer, age, specific impedance and number of PV installations. To find the optimal number of clusters, the ward method was used. Finally, eleven clusters were defined and for each cluster a reference network was selected.

In [55], authors report that performing the cluster analysis on feeder level instead of network level, yields in better results. For both cable and overhead-lines five clusters were found. Moreover, the standard deviation of the voltage difference between the detailed models and cluster reference grids is analyzed.

In [56], a techno-economical evaluation of on-load tap changer (OLTC) deployments compared to conventional network reinforcement is performed. An important result of this work is, that network topologies could not be reduced to a few representative networks. Consequently, results of a few networks cannot be scaled up to the entire supply area of a DSO. Instead, the usage of synthetic networks was proposed. Further, three feeder models are introduced and that they are sufficient for covering all topologies in the investigated supply area. Furthermore, typical line types and lengths of feeders are given.

In [57], six sample networks (three typical and three extreme) for rural, sub-urban and villages are presented. Thereby the methodology in [36] is utilized. Three parameters are used to classify networks: specific transformer rating per customer, equivalent sum impedance per feeder and the average distance between customers.

In [58], fractal geometry is used to design and expand distribution networks. With the box-counting algorithm, it was proven that electrical grids can be described as fractals. Additionally, the dielectric breakdown model is utilized to generate synthetic distribution networks. The approach is validated by comparing the maximal voltage drop and the total power losses of the fractal generated networks with the real networks.

In [59] distribution grids of western Australia were classified and prototypical feeders are selected at both MV- and LV-level. On LV-level, 26 parameters were investigated, such as secondary transformer rating, feeder length, energy consumption (for residential, commercial and industrial respectively), number of customers and many more. The seven identified most significant parameters are similar to previously mentioned studies: underground cable length, load capacity of residential customers, number of residential customers, overhead-line length and others. The cluster analysis on LV-feeder level resulted in an optimal cluster number of eight and a descrip-

24

tion of the LV-feeders subsets.

A taxonomy of continental United States radial distribution feeders (MV) was performed in [60]. In total 575 distribution feeders models from 151 substations from 17 utilities located in the five climatic zones, were clustered. Thereby 89 feeder parameters were calculated (including construction and operational characteristics). The number of parameters were then reduced to 35. In conclusion, 24 prototypical feeders were defined and described (three to seven for each climatic zone).

In the Italian project ATLANTIDE ([61]), a representative network was defined for each class of urban, industrial and rural MV networks. Thereby the urban network has eleven feeders, the industrial and rural network seven feeders each.

In [62], 34 low-voltage networks with 247 feeders were clustered on both network and feeder level based on electrical parameters as well as on statistical parameters. Thereby a large set of input parameters were reduced with a principal component analysis to the most relevant ones. In total, nine clusters were defined based on electrical parameters and for each cluster a representative (median) feeder was identified. Furthermore it was shown, that clustering , based on the consideration of only statistical parameters of feeders, leads to a different number of clusters and different grouping of the feeders compared to the clustering results when using electrical parameters.

In the study [63], the losses per voltage level are investigated for Austrian DSOs. Authors report that urban feeders can be clustered from a topological point of view into three groups. In conclusion, three topologies are used to estimate network losses in urban and rural LV-networks.

- radial feeders, 70% of all feeders, 150-300m long

- feeders with two branches, 25% of all feeders, 200m-350m long

- feeders with three branches, 5% of all feeders, 300m-500m long

For rural areas, the same feeder topologies are listed with following particular distributions:

- radial feeders, 30% of all feeders, 200-400m long

- feeders with two branches, 50% of all feeders, 300m-500m long

The deliverable D5.1 (Technical and economical evaluation of the replicability and scalability analysis [31]) of the project IGREENGrid, contains the evaluation of investigated smart grid solutions and recommendations. The most promising solutions were tested in a high number of grids of participating DSOs. One developed tool is the hosting capacity screening methodology,

which is a Monte Carlo based approach to calculate the empirical distribution function of the hosting capacity of feeders in MV- and LV-networks. Further, in-depth studies were performed to rate the performance of smart grid solutions in terms of increased hosting capacity, network losses and reactive power balance. Moreover, a statistical analysis and classification of LV-feeders was performed. Thereby feeders were split into two groups (voltage-constrained and current-constrained feeders). Supervised machine learning techniques were used to distinguish between voltage and current-constrained feeders. Thereby four parameters were identified leading to the highest accuracy among other parameters. These parameters are the equivalent sum impedance, feeder length, short circuit impedance at the end node and average line length per network connection. While the true class agreement of voltage-constrained feeders was reached with an accuracy of 98.5%, the accuracy for current-constrained feeders was unsatisfactory. Only 63.8% of the feeders were correctly classified. Hence, a relatively high share of current-constrained feeders were classified as voltage-constrained feeders. This leads to the issue that for incorrectly classified feeders, reactive power control would be mistakenly suggested to increase the hosting capacity of these feeders. Unfortunately, this decision would lead to a decreased hosting capacity for current-constrained feeders, or an increased overloading of network elements.

3 Theoretical background

In this section, terms and definitions are presented that are used in this work. First, the end node of a feeder is defined. After that, the hosting capability (HC) definition used in this work is presented and followed by the hosting capability sensitivity (HCS). After that, two feeder segments used in the reference feeder methodology, are presented and the labeling of reference feeders is discussed. Next, the closeness centrality is discussed as method to assign consumption data of loads to most suitable loads in equivalent feeders followed by the discussion of selected statistical parameters. Finally, considered reactive power control strategies and the utilized correlation coefficients are briefly discussed.

3.1 Feeder End Node Definition

In a radial feeder without any branches the most distant node is the end node of the feeder. However, the definition of the feeder end node is relevant for feeders with several branches. In such kind of feeders, several nodes have only one connected line and could be considered as being the feeder end node. Therefore a clear definition of the feeder end nodes is needed. Due to varying line types (e.g. overhead lines and underground cables or cable cross sections), number of connected customers and branches in feeders, the most distant end node may not be the most critical node in a feeder. In this work, the end node is defined as follows: Firstly, for all network connection points (e.g. single family homes, buildings, industry etc.) in a feeder, a consumption of 1kW is assumed (regardless of the customer type). Second, a load flow calculation is executed. After that the node with the highest voltage drop is selected as the feeder end node. The advantage of this approach is, that this definition is also suitable for ring topologies. Even though there does not exist a node with only one connected line in such topologies. Nevertheless, in this work only radial feeders are studied. The decision to define a consumption of 1kW per load was selected under the assumption that a effect on the voltage in every feeder can be noticed while network elements are not overloaded.

3.2 Hosting Capability Definition

A general hosting capacity definition can be found e.g. in [64], where a multidimensional performance index is used, which is decreasing with a higher amount of DER. Authors give the EN50160 as main source of incidences to be considered in the performance index. This definition is e.g. utilized in [65] and [66], where also loading requirements of network elements are

considered in the performance index, which was used in [67] to identify the hosting capacity of some representative distribution grids. An overview of common practices for the determination of the DER hosting capacity can be found in [68]. In many works, the term hosting capacity is used to specify the amount of DER that can be integrated without violating requirements, such as voltage and loading limitations, usually considering the secondary substation, as well as, the voltage variation of the upstream network and load demand. For example, also in [69], only the voltage limit violations are considered in regard to the maximal amount of installed DG capacity in the network.

However, e.g. the rating of the transformer could be a limiting factor for the DER integration. Overloading the transformer caused by overestimated admissible DER-penetration levels resulting in reverse power flows to the MV-network may be critical. Moreover, if the HC of one feeder in a network is scaled to reach a limitation (voltage or loading), the voltage in the other feeders of that network is affected as well. Thus, the voltage rise or drop at a secondary transformer is coupling the feeders. Furthermore, for a obtained HC in a feeder, the voltage is influenced by estimating the HC of another feeder. This could be avoided by evaluating each feeder individually considering the transformer. However, this would result in classifying the feeder-secondary transformer pair, which would be obsolete if the secondary transformer is reinforced for any reason or the number of supplied feeders is changed. Furthermore, all feeders supplied by a secondary substation benefit from a reinforced transformer, while the benefits of reinforcements in a feeder are rather limited. For that reason, secondary transformers are neglected and the HC is evaluated on feeder level. Tackling the analysis on feeder level, disregarding the secondary transformer guarantees that the obtained HC of a feeder is not influenced by the HC of other feeders. From this point of view, feeders can be considered to be the smallest 'building block', which are linked together at the secondary substation. Therefore the term hosting capability is defined as follows:

Definition 1. *The hosting capability (HC) of a LV-feeder is the amount of DER that can be installed for a given DER-scenario, while meeting voltage and loading prerequisites in the feeder, while disregarding the loads, the secondary substation (voltage drop and thermal constraints) and the protection scheme.*

The HC of a feeder is defined in three steps:
In the first step, an active power injection at all connection boxes is assumed for the given DER-scenario. Second, the allowed voltage rise is defined. Another requirement is, that none of the network elements in the feeder shall be

28

overloaded at any time. Therefore the loading of network elements for any scenario should not exceed 100%. Thirdly, the defined active power injection is scaled with a Newton-Raphson algorithm to reach either the allowed voltage rise or loading limit at the most critical element in the feeder. The obtained HC defines the maximal peak power that can be installed without violating the defined voltage and loading limit in a feeder. Furthermore, installing a higher amount of DER would require reinforcements in the feeder or smart grid functionalities, if applicable.

Indeed, load assumptions have a high impact on the achievable hosting capacity. In [70] it is reported, that "The worst cases of voltage rise occur when the PV-DG is at or close to its maximum output, and the difference between the load and the generation is at its highest" and further, "Using deterministic models with average values for the loads underestimates the voltage impact of increasing PV-DG because the real loads on a feeder vary". Indeed, the base load can be well monitored at secondary transformers supplying a particular area with a high number of customers. These customers are usually supplied by more than one feeder. However, only a few loads and generation are considered, such as in a feeder, the existence of a certain load level may not be given permanently at all locations in the feeder. In [71], a stochastic analysis was performed on the smart meter data of 1,077 customers (households and PV generators) measured over a period of one year. Authors evaluated the peak and average load depending on the number of households in the respective area that needs to be considered in network planning. E.g. for a range of one to five households, a significantly higher peak load and a significantly lower average load compared to e.g. 100 households was observed.

According to the HC definition, the remaining constraint that could be violated is that the total HC of all feeders may exceed the rating of the secondary transformer supplying these feeders. The base load at the secondary transformer for example, may change with network reconfigurations, e.g if a feeder is planned to be supplied by another secondary substation or an additional feeder has to be supplied by the secondary transformer. Hence, the base load level may be prone to changing in the future. However, this issue arising depends on several circumstances:

- Could the full HC potential be utilized (e.g. available PV roof-top potential)?

- What time frame can be expected for the full deployment of PV?

- What is the actual minimal load at the transformer and the expected growth of consumption?

- Are network topology changes foreseen (additional feeders, feeders supplied by other secondary substations)

Hence, the obtained HC could be deployed without the need to reinforce the feeder for the selected DER-scenario in step one. Accordingly, feeders of the same type but supplied by secondary transformers of a different type have a identical HC. Consequently, for feeder classification purposes the secondary transformer is neglected. Therefore the proposed methodology can also be applied in micro grids without a connection to higher network levels (e.g. MV).

In the case of a radial feeder with one DER at the end node, the following relationships can be given: For a determined hosting capability P and an admissible voltage rise Δu of 1%, the resistance leading to that particular voltage rise can be estimated according to eq. 1.

$$R \approx \frac{\Delta u_{1.01} \times U_N{}^2}{P_{1.01}} \tag{1}$$

According to eq. 2, the resulting current for a given pair of admissible voltage rise and hosting capability (disregaring reactive power) can be estimated to be:

$$I_{1.01} = \frac{P_{1.01}}{\sqrt{3} \times (1 + \Delta u_{1.01}) \times U_N} \tag{2}$$

For a radial feeder with one line and one DER located at the end node, an equivalent feeder can be determined utilizing the presented equations. Thereby the feeder is represented by a slack, one line and a node with a connected DER. According to the calculated resulting current, a line type with a minimal required nominal current can be determined and selected from a standard line type library (e.g. IEC standard line types). Furthermore, the length of the single line is defined by the resistance per length (R') of the identified line type and the calculated resistance according to eq. 3:

$$l = \frac{R}{R'} \tag{3}$$

3.3 Hosting Capability Sensitivity

The presented hosting capability definition depends on 3 assumptions: the admissible voltage rise, loading of elements, and DER distribution. In this work, the voltage rise is investigated in several HC-scenarios, whereas the maximal allowed loading of network elements is limited by their nominal current (e.g. nominal current for lines). For a fixed voltage rise limit and

a specific loading limit, the HC is determined by the DER scenario. Three DER-scenarios (*uniform*, *eof* and *weighted*) were considered, each representing a specific DER penetration level. The DER-scenario *uniform* corresponds to an equal distribution of DER installations in the feeder. Thereby the same amount of DER is connected to all connection boxes in a feeder. The DER-scenario *uniform* represents a rather optimal case where the HC is expected to be rather high compared to random DER distributions. Furthermore, for all variations of this DER-scenario where the DER penetration is moved to more distant nodes, feeder constraints are violated (voltage or loading of elements). Contrarily, if the DER are relocated to a node that is near the slack, the HC of the feeder can be increased. In the case of the DER-scenario *eof* only one DER installation is located in the feeder. The DER-scenario *eof* can be described as the worst case scenario. Thereby active power is only injected at the end node of the feeder (according to the definition in section 3.1). This node is the most critical node in terms of the lower limit of a high DER penetration of feeders. The obtained HC of this scenario can be placed at any node in the feeder without violating any constraints. Last but not least, the DER-scenario *weighted* covers a DER-penetration share driven by self-consumption. Thereby the in-feed is weighted with the yearly consumption at each particular connection box in the feeder. In [41], the Minimum and Maximum Hosting Capacity concept is presented:

1. A Hosting Capacity threshold level until all penetrations are acceptable, regardless of location

2. A Hosting Capacity limit, which is still acceptable, however penetration and site location specific

The defined DER-scenarios *eof* and *uniform* can be compared to these two characteristic Hosting Capacity values.

According to the defined DER-scenarios, the hosting capability sensitivity (HCS) is introduced as a measure of covering topological information of feeders by means of HC results:

Definition 2. *The hosting capability sensitivity for a given voltage limit is the ratio of the hosting capability obtained for the DER-scenarios; uniform and eof.*

The quotient in eq. 4 is defined as the ratio of the HC obtained for the DER-scenarios *uniform* and *eof* (optimal- and worst-case). This ratio can be calculated for any voltage limit. Thereby the value of HCS is a continuous value greater equal 1 and the unit is 1 (kW/kW). To reasonably limit

31

the number of feeder topologies, HCS values are rounded to 0.25 increments, assuming a feeder with one line supplying one load at the end of the feeder. In this case, the active power injection in both DER-scenarios *uniform* and *eof* is identical. Therefore the same HC is reached for both scenarios and the calculated HCS value is 1 for any line type, feeder length and admissible voltage rise. According to eq. 4, ueof1 is defined as the ratio considering an allowed voltage rise of 1% (u_{max}=0.01 p.u.). Consequently, for ueof8 a voltage rise of 8% (u_{max}=1.08 p.u.) is allowed. If not stated otherwise, HCS always refers to ueof1.

HCS is a parameter based on results obtained with two power flow calculations and contains topological information of feeders.

$$HCS = \frac{HC_{uniform}}{HC_{eof}} \tag{4}$$

3.4 Reference Feeder Topologies

In this work, a number of reference feeder topologies are defined. For this purpose a specific labeling scheme of elements is proposed, for the defined two basic feeder segments. The feeder segments are needed to explain the pattern of the reference feeder topologies later in this work. In Figure 6 a reference feeder is depicted with a 0.4kV slack (External Grid) connected to Terminal 1. Line 12 connects Terminal 1 and Terminal 2. The number below the line name defines the number of parallel lines (in this case 1) of the same line type. This labeling of lines applies to all lines. At Terminal 2, two feeder segments are connected. Segment I consists of the Line 24 (1 cable), Terminal 4 and Load 3 . Segment II contains of Line 23 (2 lines in parallel), Terminal 3 and two loads (Load 1 and Load 2). By convention, all cables have the same line type and length. The impedance of each cable is determined by number of lines in parallel. Consequently, according to the end of feeder definition in section 3.1, Terminal 4 is the end node in the depicted feeder for any chosen cable type.

3.5 Closeness Centrality

Networks and topologies in general, can be described by graphs. Hence, LV-feeders can also be described by graphs, where two terminal elements such as lines, transformers, switches, etc., represent edges. There are different types of graph such as undirected and directed graphs which can have unweighted or weighted edges [72]. Weighted undirected graphs are suitable for describing electrical networks. A description of a network by a graphs
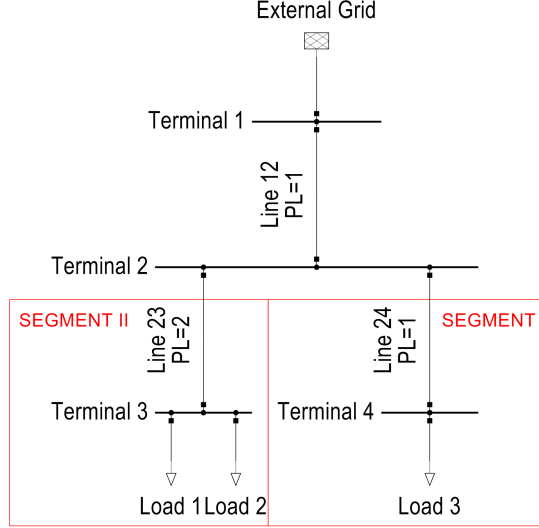
Figure 6: Feeder segments and labeling convention used for the developed reference feeder topologies

allows the evaluation of measures such as the centrality of nodes in electrical networks. The closeness centrality is used to identify the most important vertices (nodes) and was first used in social networks. Basically, the centrality of a node is a fraction of all shortest paths that pass through that node. An introduction to networks and can be found e.g. in [73]. The closeness centrality for node u to all other n-1 nodes is calculated as follows:

$$C(v) = \frac{n-1}{\sum_{v=1}^{n-1} d(u,v)} \tag{5}$$

where v is node v and d(u,v) is the shortest path distance between node u and node v. n is the number of nodes in the graph. Higher values indicate a higher centrality compared to the other nodes. In Figure 7 an example of a feeder is depicted for which the centrality is calculated. The feeder contains six nodes and six loads. The properties of all depicted lines (Line 1 to Line 5) are identical (line type and line length) leading to a resistance of 0.1Ω each. In an unweighted graph, the connection can be represented by a logical one for terminals that are connected by a line and zero for nodes which are not connected directly. Thereby the value on the main diagonal for all terminals is defined as zero. In a weighted graph, instead of the logical one, e.g. the resistance of the line can be given. In Table 2 the weighted graph of the sample feeder is given.
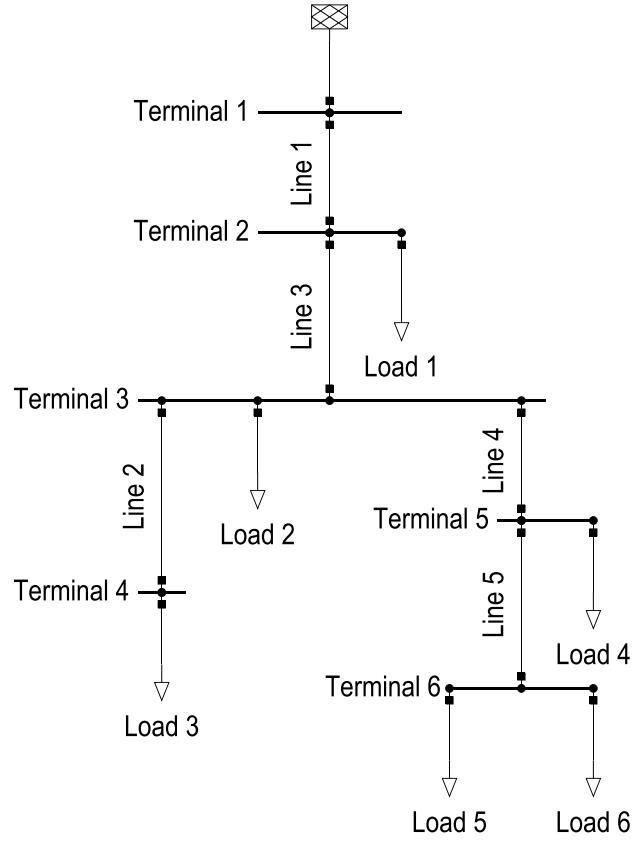
Figure 7: Sample feeder

Table 2: Weighted graph representation of the sample feeder

Distance (Ω)	Term. 1	Term. 2	Term. 3	Term. 4	Term. 5	Term. 6
Terminal 1	0	0.1	0	0	0	0
Terminal 2	0.1	0	0.1	0.1	0.1	0
Terminal 3	0	0.1	0	0	0	0
Terminal 4	0	0.1	0	0	0	0
Terminal 5	0	0.1	0	0	0	0.1
Terminal 6	0	0	0	0	0.1	0

The shortest path distances from all nodes to e.g Terminal 3 is given in Table 3. According to eq. 5, the closeness centrality is 7.143 (5 divided by 0.7). Removing Terminal 3 from the feeder would lead to three clusters (Terminal 4, Terminal 5 + Terminal 6 and Terminal 2). Thereby Load 2 needs to be relocated to another Terminal. Since the distance to the neighboring nodes Terminal 2, Terminal 3 and Terminal 4 is equal, it could be relocated to any of these three nodes. The nodes with the second highest closeness centrality are Terminal 2 and Terminal 5 (5.556).

Table 3: Shortest path to Terminal 3 from all other nodes

Distance	To Terminal 3
From	(Ω)
Terminal 1	0.2
Terminal 2	0.1
Terminal 3	0
Terminal 4	0.1
Terminal 5	0.1
Terminal 6	0.2
sum:	0.7

Assuming, that the RFT depicted in Figure 6 is the equivalent RFT for the given feeder example in Figure 7, the loads in the real feeder would be assigned according to Table 4. Terminal 2 with Load 1 and Load 2 would be assigned to Load 1 in the RFT. Terminal 3 with Load 3 would be assigned to Load 3 and the remaining Loads at terminal 5 and 6 would be assigned to Load 2 of the RFT.

Table 4: Assigment of loads to load locations in the reference feeder topology

Real feeder	RFT
Load 1 at Terminal 2	Load 1 at Terminal 3
Load 2 at Terminal 3	Load 1 at Terminal 3
Load 3 at Terminal 4	Load 3 at Terminal 4
Load 4 at Terminal 5	Load 2 at Terminal 3
Load 5 at Terminal 6	Load 2 at Terminal 3
Load 6 at Terminal 6	Load 2 at Terminal 3

3.6   Selected Statistical Parameters

In this section, the calculated statistical parameters are presented which are invariant to any DER-scenario and voltage limits.

In Table 5 the considered statistical parameters are listed. While many of these parameters can be estimated without calculations in a network simulation environment (such as the number of loads, number of neighbors, etc.), some require a full network model (e.g. dvdP, dvdQ) or even more calculations (e.g. the sum impedance R_Σ). As desired, these parameters shall be sufficient for the estimation of the HC of the previously defined scenarios.

Table 5: Overview of calculated statistical parameters

$ANON$	$ADTN$	Z_Σ	$I_{nom-Avg}$	$I_{nom-Max}$
$I_{nom-Min}$	I_{nom-MP}	$I_{nom-Avg-MP}$	Nr. Cables	Nr. Loads
Nr. Nodes	dvdP	dvdQ	R_k	X_k
Z_k	R/X	L	NEN	l
Mesh factor (MF)	-	-	-	-

The parameters $ANON$ (Average number of neighbor nodes), $ADTN$ (Average distance to neighbor nodes) and R_Σ were introduced in [36] and used in [53] and [62]. These parameters are calculated as follows (eq. 6 to eq. 8):

$$ANON = \frac{\sum_{k=1}^{N} NON_k}{N} \tag{6}$$

$$ANON = \sum_{a=1}^{N} \frac{1}{N} \sum_{i=1}^{NON} distance_{a-i} \tag{7}$$

$$R_\Sigma = \left| \frac{\Delta U}{I_{tot}} \right| \tag{8}$$

Furthermore, the weighted average nominal current of all lines in a feeder $I_{nom-Avg}$ as well as the minimal and maximal nominal current in a feeder are evaluated. Moreover, the minimal current on the main path (I_{nom-MP}) and the weighted average nominal current on the main path ($I_{nom-Avg-MP}$) are gathered for each feeder. The relevant statistical parameters that have been included statistical parameters are also the number of cables, loads, nodes,

end nodes (NEN) as well as the total cable length and the length of the feeder. Additionally, the mesh factor on feeder level was calculated (eq.9):

$$MF = 2\frac{N_{Cables}}{N_{Nodes}} \tag{9}$$

Finally, end node relevant parameters are also calculated, including R_k, X_k, Z_k, $dvdP$, $dvdQ$ and the R/X ratio at the end node.

After obtaining the statistical parameters and the HC results for each feeder, a correlation analysis between statistical parameters and HC results will be performed as well as probability tests. The results of this analysis will indicate, if the statistical parameters are suitable for the estimation of HC results, so as to avoid laborious investigation of detailed network models.

3.7   Reactive Power Control Strategies

As discussed in section 1.2, the voltage rise caused by active power injection in the network can be partly reduced with reactive power control strategies depending on the R to X ratio of lines. In [17] the VoltVAr and VoltWatt control were identified as being the most promising solutions for increasing the hosting capacity of low-voltage distribution grids from a technical and economical point of view. Several control strategies were developed and intensively tested in simulations and in the field (e.g. [15], [18]). Thereby the reactive power can be set as a function of the active power in-feed (PF(P)) or the voltage at the point of coupling (VoltVAr). Another option is to apply a fixed power factor (e.g. PF=0.9 (ind.) independently from active power injection value). The effectiveness of some control strategies was investigated in [20]. However, this introduced additionally reactive power flows in the grid ultimately increasing the loading of network elements and causing additional losses.

The $\cos\varphi$(P)-control is depicted in Figure 8 and is characterized by an active power injection with a power factor of 1 until 50% of the nominal power of the inverter. At higher in-feed levels, the power factor is constantly reduced to 0.9 at nominal power. Hence, with this kind of strategy it is aimed to reduce the voltage rise caused by the active power in-feed. Consequently, the actual grid situation is not considered completely, since the voltage or consumption is not measured. The fixed PF-control (e.g. PF=0.9 (ind.)) is a modification of the PF(P)-control. Considering the steady state at nominal power, these two controls are equivalent. Consequently, the same HC can be reached and the same amount of reactive power flows occcur at nominal power. Therefore, the fixed PF-control was implemented in this work.
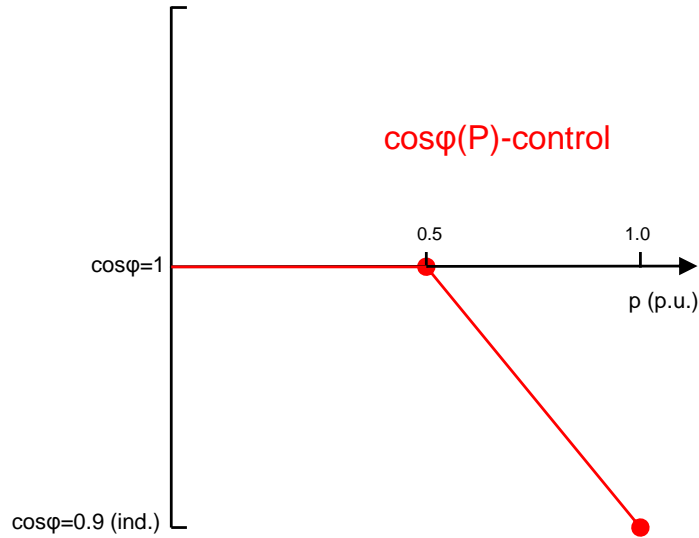
Figure 8: $\cos\varphi(\mathrm{P})$-control

A control strategy as a function of the voltage is capable of considering the grid situations at the point of coupling. Such a control is e.g. the VoltVAr-control. This control is basically characterized by a dead band, drop and saturation areas according to Figure 9. Thus, reactive power flows occur only if the measured voltage drifts outside the dead band. The advantage of this control is that reactive power flows occur only if a deviation of the voltage is observed. For example, in times of high consumption the voltage rise caused by active power injection may be fully compensated by loads. Consequently, the VoltVAr-control would eventually remain in the dead band. However, the VoltVAr-control is a local control strategy and different set points are applied depending on the location of the inverter in the network. By trend, inverters at a more remote location in the network sense a higher voltage sensitivity compared to inverters nearby the secondary substation. The VoltVAr-control is configured by four set points defining the reactive power at certain voltage levels. In this work, the set points are configured as a function of the admissible voltage rise limit. Thereby, 20% of the allowed voltage rise are allocated to the dead band and 50% to the droop area. The remaining share of 30% of the voltage band are reserved for the saturation area.

In this work, the HC is calculated without reactive power control strategies (PF=1) and the fixed PF-control and the VoltVAr-control, as previously discussed.

38
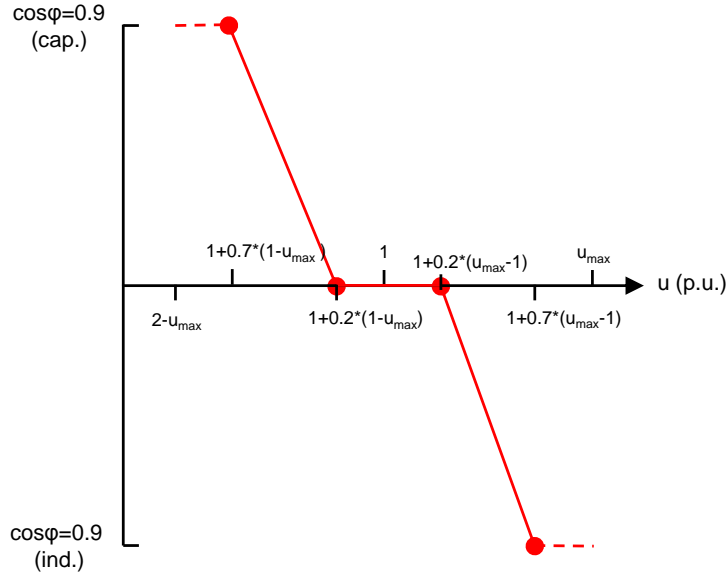
Figure 9: Implemented VoltVAr-control as a function of the admissible voltage rise

3.8   Correlation

The correlation coefficient is a measure used to evaluate the similarity between two parameter vectors. Three correlation coefficients are used in this work. The Pearson, Spearman and Kendall correlation are used to investigate the correlation between hosting capability results and feeder parameters to identify the most promising parameters. The **Pearson correlation coefficient** is a linear correlation measure between -1 and 1. Where a coefficient of -1 indicates full negative correlation, 0 no correlation and 1 full correlation [74]. The **Spearman's rank correlation coefficient** is a nonparametric rank correlation measure. Where the rank of the sorted elements is correlated instead of the values itself. Again, correlation coefficients reach from -1 to 1 [75]. The **Kendall correlation coefficient** is also a rank correlation measure varying from -1 to 1 [76].
In conclusion, a correlation coefficient of +1 means perfect positive correlation and a coefficient of -1 means perfect negative correlation. Since high positive and high negative correlation coefficients is of interest, the absolute correlation coefficient is evaluated in this work.

4 Methodology

In this section the methodology of this work is presented. First the available grid data will be explained, followed by the HC scenario definitions and a list of calculated parameters. Finally the simulation architecture and the reference feeder parametrization will be described.

4.1 Input Data

During the project IGREENGrid (IntegratinG Renewables in the EuropEaN Electricity Grid), two Austrian DSO provided their complete distribution grid data set of about 14,000 LV-networks to perform a scalability analysis. Each DSO provided their grid data in a different format. In Figure 10, the import process is depicted. In case of DSO 1, a GIS/DGS was used to export the grid data from the GIS database to the DigSILENT DGS format which allows importing to PowerFactory. In case of DSO 2, the grid data was stored in Neplan and exported from there (edt, cde and ndt-files), before being imported to PowerFactory.
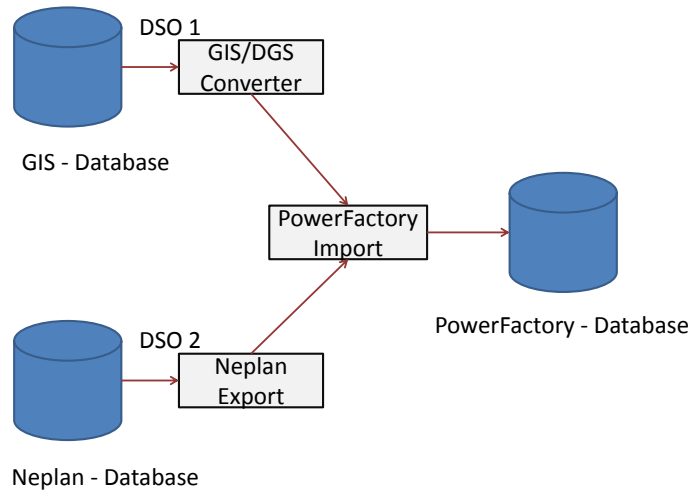
Figure 10: Network data import from different sources to PowerFactory

On LV level, in total 14,400 networks (8,400 and 6,000 networks respectively) were imported to PowerFactory. In total, more than 54,000 feeders were identified. Several plausibility measures were defined to identify feeders with missing or incomplete line type data. These plausibility measures include the short circuit impedance at the end node, the length of the feeder,

Table 6: Plausibility test results per measure

Measure	Lower Limit	Upper Limit	Passed Feeders
Short circuit resistance	> 0	< 1	84.2%
Feeder length	> 0	$< 10 km$	100%
Sum impedance	> 0	< 1	71.4%
Radial feeder	$== 1$	$== 1$	96.6%

the sum impedance and whether feeders have a radial topology. Further information on the input data and basic feeder statistics as well as performed investigations can be found in deliverable D5.1 (Technical and economical evaluation of the replicability and scalability analysis [31]) of the project IGREENGrid.

Based on the selected four plausibility measures, only the feeders passing all four plausibility tests were used in the following sections. The share of feeders passing the individual plausibility test are listed in Table 6. Most of the feeders that failed a plausibility were identified by their short circuit resistance or their sum impedance. The highest number of feeders are skipped due to the circumstance that only one node was found in the feeder. Since the sum impedance can only be calculated for feeders with at least two nodes, these feeders are not considered for the development of the RFT methodology. Hence, a further simplification for feeders with one node is not feasible.

4.2 Hosting Capability Scenarios

In this section, the defined HC scenarios are presented. It is assumed that evaluating the HC per feeder for various scenarios will be an additional benefit for the two DSOs. Furthermore, parts of the HC scenarios results will be utilized to develop and validate the reference feeder methodology. Moreover, it is discussed, why the HC analysis was performed on feeder level disregarding the secondary transformer and the HC scenarios are presented.

The number of HC scenarios are formed by the combination of DER-scenario, control schemes and the admissible voltage limit. The DER-scenarios were already presented in section 3.3. Additionally, the discussed control schemes are used (section 3.7): Uncontrolled (PF=1), PF=0.9 (ind.) and a VoltVAr-control. Furthermore, it is assumed that voltage regulated distribution transformers (VRDT) may be available to decouple low-voltage networks from medium-voltage networks and restraining the voltage on LV-level at nominal voltage, a higher part of the available voltage band could be reserved for voltage rise caused by DER-infeed. Therefore, a deadband

of 2% for an VRDT is assumed, 8% of the upper half of the voltage band ([77]) could be devoted to DER in-feed. Therefore, admissible voltage limits are investigated in 1% steps starting from 1.01 p.u. to 1.08p.u. reserving 0.02p.u. for a possible deadband of a VRDT.

In Figure 11, these assumptions are summarized. HC calculations are performed for each DER-scenario (*uniform*, *eof* and *weighted*), the three control strategies (uncontrolled, PF=0.9 (ind.) and VoltVAr-control), and eight admissible voltage levels (1.01 p.u. to 1.08p.u.). By means of active power curtailment (P(U)-control), the HC could be increased even further. However, the objective is to maximize the HC of feeders according to the DER-distribution. A P(U)-control would curtail the injected power unequally, since the power tends to be more curtailed at the most distant installation where the highest voltage rise is observed. This would cause varying injected powers for the DER-scenario uniform. Hence, considering a P(U)-control in the optimization of the HC would lead to a contradiction that cannot be resolved while maximizing the injected power for all generators in the feeder in a fair way. The amount of curtailed energy however depends on several other factors such as location, connection point, orientation and tilt of the installation, load behavior, etc. In conclusion, the three DER-distribution times three control strategies times eight voltage rise limits results in 72 HC scenarios which are calculated for more than 35,000 LV-feeders.
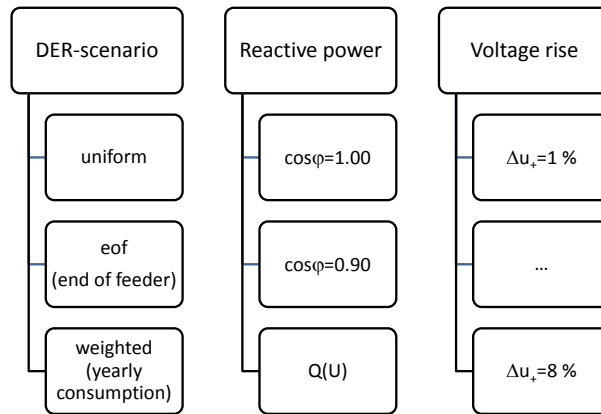
Figure 11: Investigated hosting capability scenarios (combination of DER-scenario, reacitve power control and admissible voltage rise)

The networks provided by the two DSO were imported to DIgSILENT

PowerFactory 15.2.4 for network calculations. In the networks, the DER are modeled as loads and are injecting active power symmetrically into the grid. Balanced load flow calculations were executed to calculate the described scenarios. In this work, the constant power model was used for loads. Further, 3-phase short circuit calculations were used to obtain the short circuit impedance, resistance and reactance at selected nodes in the grid. The impact of the load model was investigated for example in [78].

4.3 Feeder Analysis

In this section the methodology of the feeder analysis is presented. The platform to perform the simulation was published in [79]. In Figure 12, the implemented algorithm is depicted. Each network is assessed for given input parameters (Voltage Limits, control strategies and DER distribution). After a plausibility test, feeders are defined and the secondary substation of the networks is replaced by a LV slack node. Omitting the transformer allows the study of feeder-only relevant parameters and avoids the mischaracterization of identical feeders, caused by different secondary transformer types. First, for each feeder, parameters that are invariant of any assumptions are calculated (static parameters such as the feeder length, Z_k and Z_Σ). Parameters such as the Z_k or I_{nom} are suitable for checking the plausibility of network data (incomplete or corrupt). Next, the HC for each scenario according to Figure 11 is calculated and parameters based on these assumptions are obtained (scenario results). Thereby, i defines the control strategy (uncontrolled, PF=0.9 (ind.), VoltVAr) and j the DER-scenario (uniform, eof, and weighted). For each scenario, the NR-algorithm is utilized to obtain the maximal HC under the given voltage (umax) constraint without overloading any network element.

The problem formulation for the optimization of the HC of a feeder is given in eq. 10, where f is a feeder of Network N, u_i the voltage at bus i, l_j the loading of line j of the feeder. The admissible loading L_{max} of lines is defined as 100% of the respective nominal current of each line. The voltage limit u_{max} is determined by the considered voltage scenarios (1.01 to 1.08 p.u.).

$$\forall f \in N : max(HC_f, u_i \leq u_{max} \wedge l_j \leq L_{max}) \tag{10}$$

Based on the scenario assumptions and the problem formulation in 10, the HC is obtained with an adapted Newton-Raphson algorithm to reach either the voltage constraint or the loading constraint according to the depicted flowchart in Figure 11. For that, two NR-optimizations are performed in each iteration to calculate the highest HC without violating the admissible
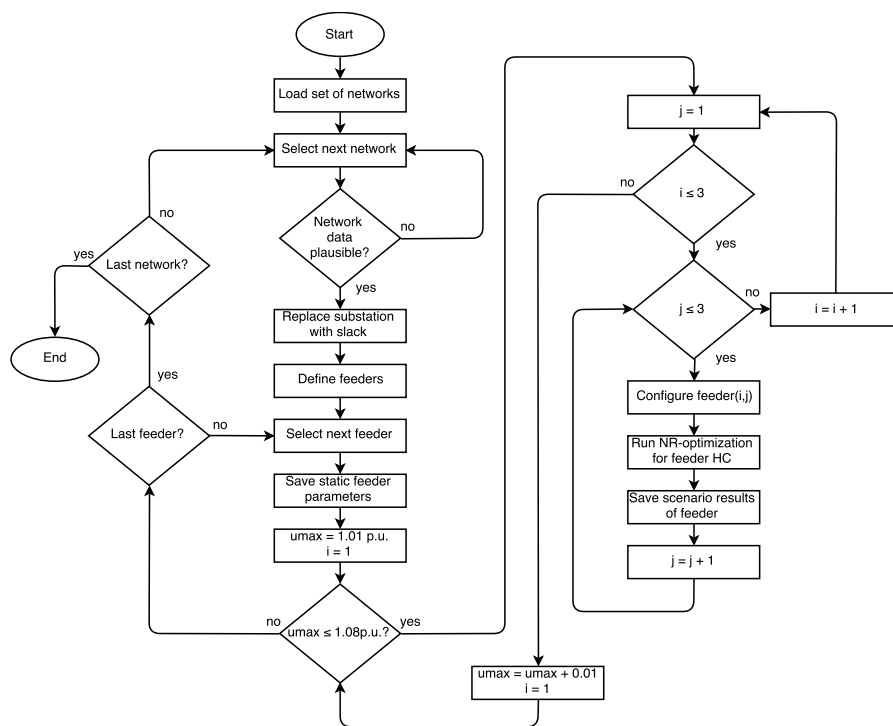
Figure 12: Flow chart for the calculation of the hosting capability of feeders.
i and j are to configure the DER-scenario and the reactive power-control

voltage rise and the loading of network elements (13), respectively. In each step, the lower estimated HC is selected for the next iteration. The detailed algorithms that were developed can be found in appendix G.
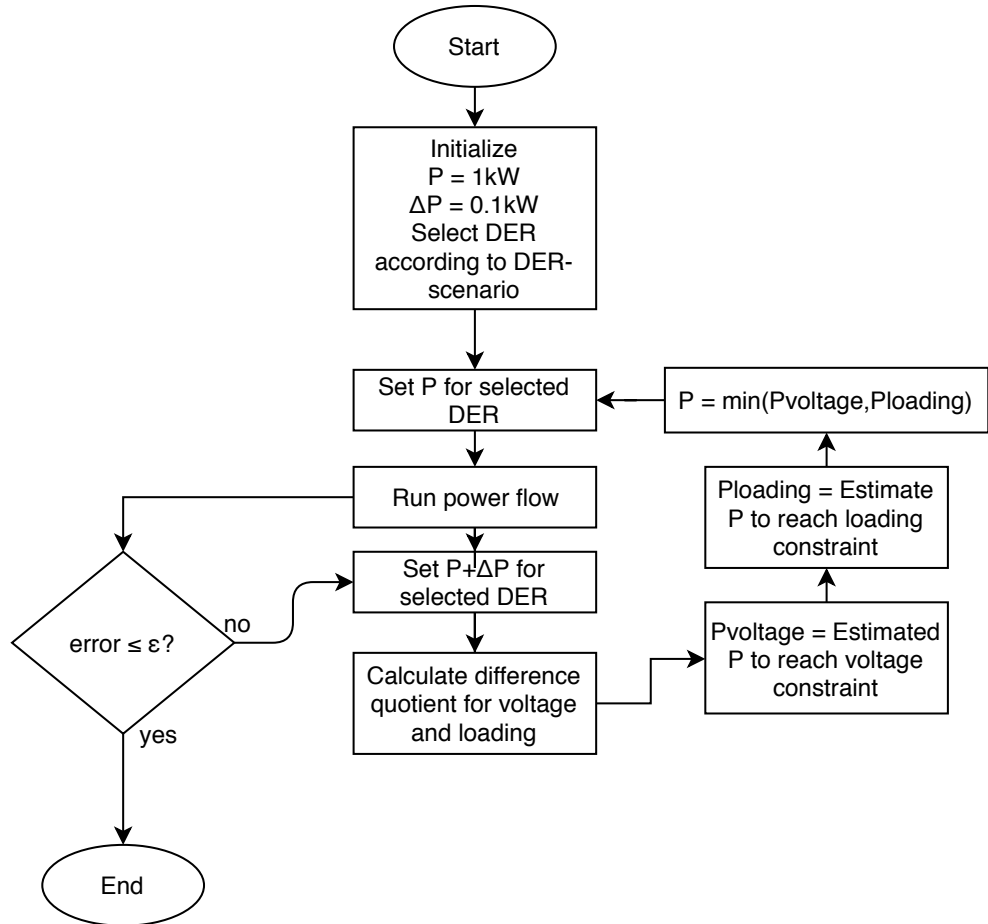
Figure 13: Adapted Newton-Raphson (NR)-optimization to calculate the hosting capability of a feeder

4.4 Definition of Reference Feeders

In this section, the approach to define reference feeder is presented. In general, a feeder is characterized by the used line types and its topology. Roughly, the most relevant properties (in LV-networks) of used line types are the resistance and the nominal current which varies between line types. Although, the topology itself does not contain any electrical parameters, the influence of the topology on HC results cannot be overseen. For example,

the topology contains information on the relative location of cables in the feeder. Further branches are visualized as well as the location of customers. Therefore the task of generating a reference feeder is split into two subtasks. The first subtask is the selection of an appropriate reference feeder topology and the second subtask covers the parametrization of cables of the reference feeder topology.

In Figure 14 the general approach of the design of reference feeders is depicted. In the best case, the required parameters to define a representative feeder are already available e.g. in the GIS-database of the DSO, thus reducing the amount of time needed to gather and prepare needed input parameters. Then, for a given pair of input parameters a representative feeder can be generated. The parameters that are are most suitable for this task will be selected in the next section.
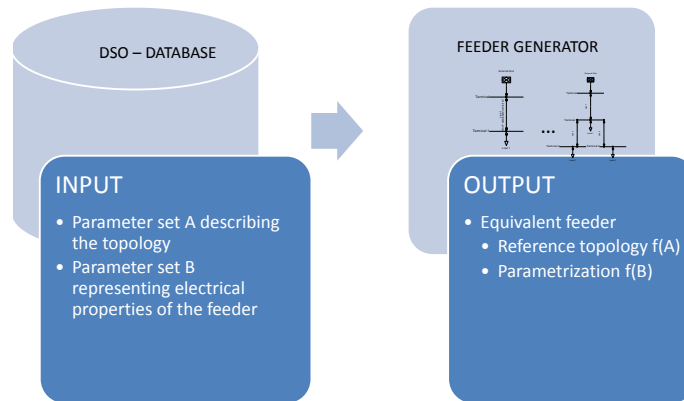
Figure 14: Reference Feeder Generation in case all required parameters are available

Alternatively, if particular parameters are not available, e.g. a neural network could be trained to estimate the required input parameters using other available feeder data. This approach is depicted in Figure 15. The aim of this work is to define and validate a reference feeder methodology. Therefore, the validation of such predictive methods for needed input parameters will be a future work.
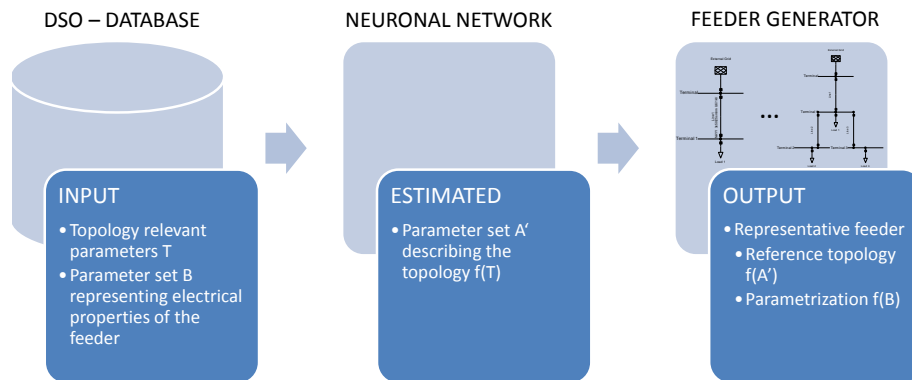
Figure 15: Reference Feeder Generation under uncertainty (e.g. neuronal network)

5 Data Analysis

In this section, the most relevant parameters to define and classify reference feeders are identified. First, the HC study results are discussed. Secondly, the feeder statistics for the grid data of the two DSOs are presented, including a correlation analysis between parameters and HC results as well as probability tests for statistical parameters. Furthermore, the distribution of the parameter HCS for real feeders is discussed. Thirdly, the defined reference feeder topologies are developed and explained.

5.1 Results of the Hosting Capability Study

In this section the results of the HC study based on the defined scenarios are presented. Thereby, boxplots are utilized to present the HC results. The upper adjacent shows the 99th percentile, the lower adjacent shows the 1st percentile and the lower and upper edge of the box show the 5th and 95th percentile. The red line always indicates the median value.

In Figure 16 the distribution of the HC of all feeders for different voltage levels are depicted for the DER-scenario *uniform* and PF=1 for both DSOs. Up to $1.07p.u.$, a steadily increase of the HC can be observed. At $1.07p.u.$ and $1.08p.u.$, the 95th percentile remains stable, but the median and the 5th percentile are still increasing by about 10kW and 5kW, respectively per voltage step. Concerning the allowed voltage rise, allowing a voltage rise of $1.02p.u.$ instead of $1.01p.u.$, leads to a significant HC increase. In that case, the median value can be doubled from about 25kW to more than 50kW. If the admissible voltage rise can be increased to 4%, more than 50% of the feeders have a HC higher than 100kW. For higher admissible voltage limits a higher share of feeders become loading constrained. Hence, a voltage increase has no impact on the achievable HC anymore. Since a cable in the feeder may be fully loaded the maximal HC is reached.

In Figure 17, the results for the DER-scenario *weighted* and PF=1 are depicted. In this scenario it is assumed that consumers with a higher consumption have a higher share of the total installed DER in the feeders (DER-deployment driven by self-consumption). However, up to $1.04p.u.$, there is no significant difference when compared to the DER-scenario *uniform*. Above this threshold, the median HC is approximately 5kW lower compared to the DER-scenario *uniform*. Again, for an admissible voltage limit of $1.04p.u.$, 50% of the feeders have a HC of 100kW.

Next, the HC distribution for the DER-scenario *eof* and PF=1 is depicted in Figure 18. Assuming that only one DER is installed in the feeder and is located at the feeder end node, the HC is significantly lower compared
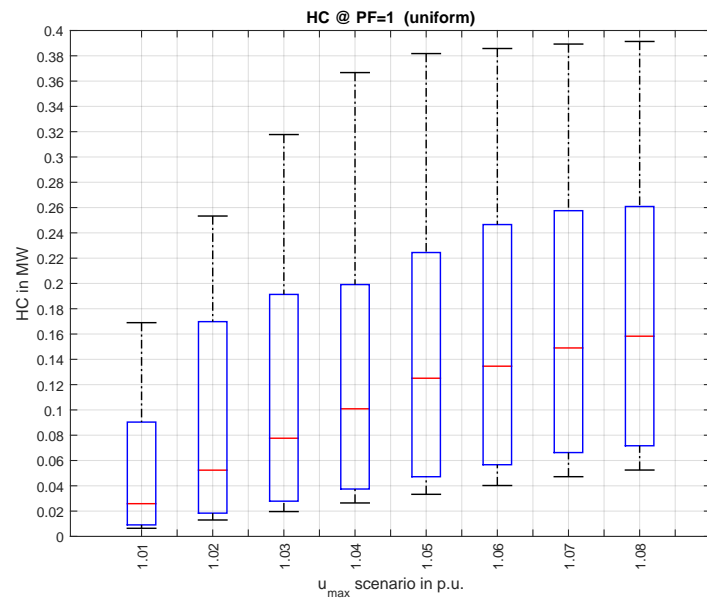
Figure 16: DSO 1 and DSO 2 - Hosting capability for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile(DER-scenario *uniform*, powerfactor 1)
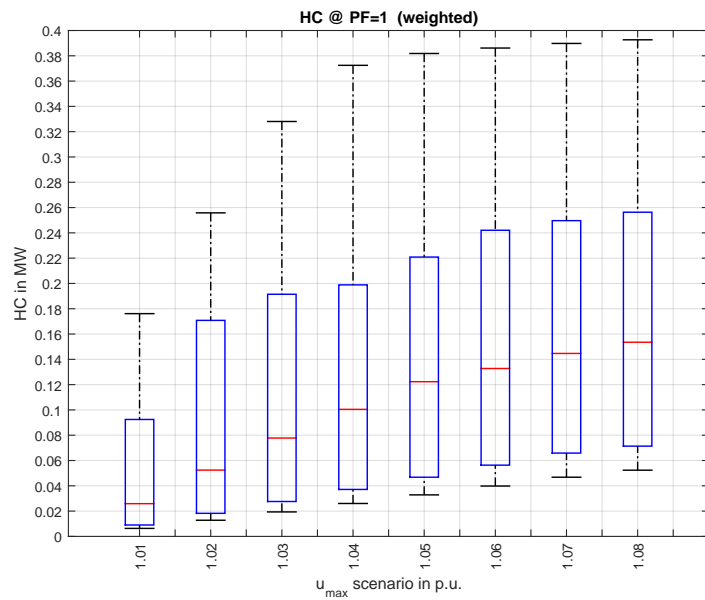
Figure 17: DSO 1 and DSO 2 - Hosting capability for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile(DER-scenario *weighted*, powerfactor 1)
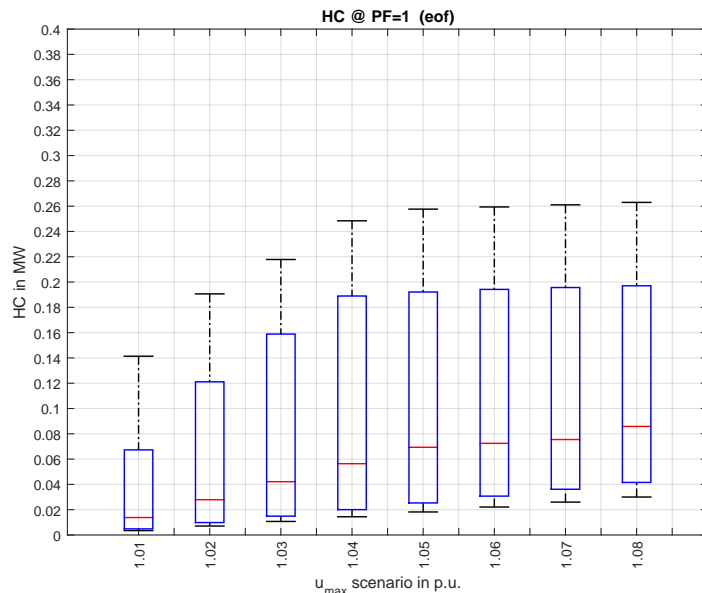
Figure 18: DSO 1 and DSO 2 - Hosting capability for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *end of feeder*, powerfactor 1)

to the DER-scenarios *uniform* and *weighted*. For example, for 1.03*p.u.*, the median value is less than half of the median value for the respective DER-scenario *uniform*. Furthermore, the 95th percentile above voltage limits of 1.04*p.u.* remains almost at the same level. However, the 5th percentile and the median are increasing. Also in this scenario, a voltage limit of 1.01*p.u.* cannot be recommended. In this scenario no median value reaches 100kW. The calculated HC for this DER-scenario is the minimal HC of the feeder that can be deployed at any node without violating voltage or loadings constraints. Therefore this scenario could be used as first benchmark level for the upstream grid, to answer the question of whether the amount of DER could be deployed without further reinforcements or smart grid solutions.

In Figure 19 the results for the DER-scenario *uniform* and VoltVAr-control are shown. Similar to the uncontrolled DER-scenario *uniform*, a steady increase of the HC is monitored, such that the median HC of the VoltVAr scenario is higher for all voltage limits. However, the HC values of the 99th percentiles are not reached (loading constrained). Similar to the DER-scenario uniform, a voltage limit of 1.01*p.u.* cannot be recommended as well. For the voltage limit 1.04*p.u.*, the median value reaches even 120kW.

51
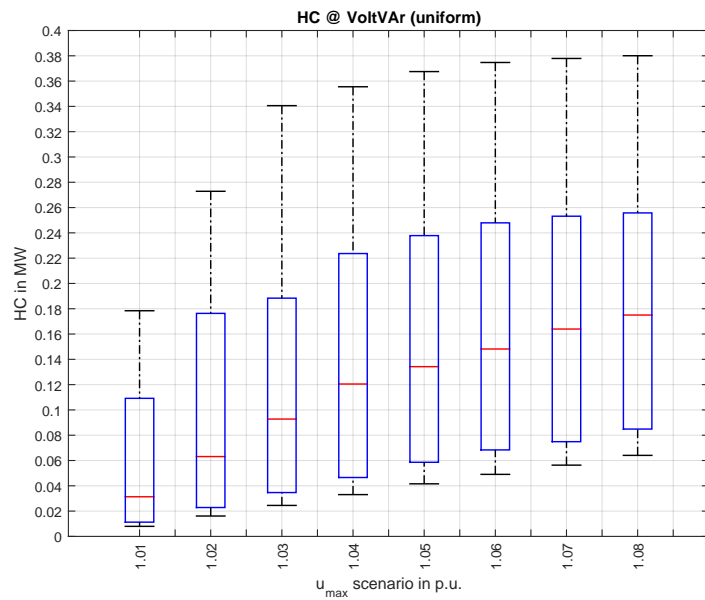
Figure 19: DSO 1 and DSO 2 - Hosting capability for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, VoltVar-control)
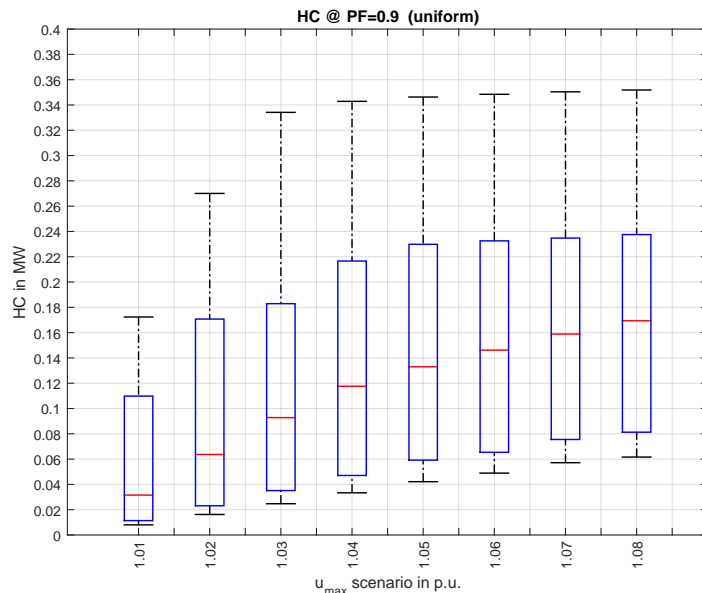
Figure 20: DSO 1 and DSO 2 - Hosting capability for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 0.9 (ind.))

Next, the HC study results for the DER-scenario *uniform* and a PF=0.9-control (ind.) are depicted in Figure 20. In this case, the median is always higher, when compared to the uncontrolled case. The 95th percentile is lower except for $1.01p.u.$, $1.02p.u.$ and $1.04p.u.$, but the 5th percentile is always higher except for the voltage limits of $1.01p.u.$ and $1.02p.u.$. For the voltage limit $1.05p.u.$ and $1.06p.u.$, as well as for $1.07p.u.$ and $1.08p.u.$, similar HC distributions are observed. Compared to the VoltVAr control (DER-scenario *uniform*), a lower 95th percentile is reached, except for $1.01p.u.$. The median HC of both control schemes are comparable for most of the voltage limits. However, regarding the median HC, the VoltVAr control outperforms the PF=0.9 (ind.) for higher voltage limits. At $1.04p.u.$, the median is at a comparable level at 120kW.

The results of the DER-scenario *weighted* and PF=0.9 (ind.) control are presented in Figure 21. Similar results as in the DER-scenario *uniform* PF=0.9 (ind.) were obtained. However, again for higher voltage limits lower HC values are reached. In the case of the two investigated DSO, the distribution of the generation in the DER-scenario *weighted* did not result in significantly different HC results compared to the DER-scenario uniform. At
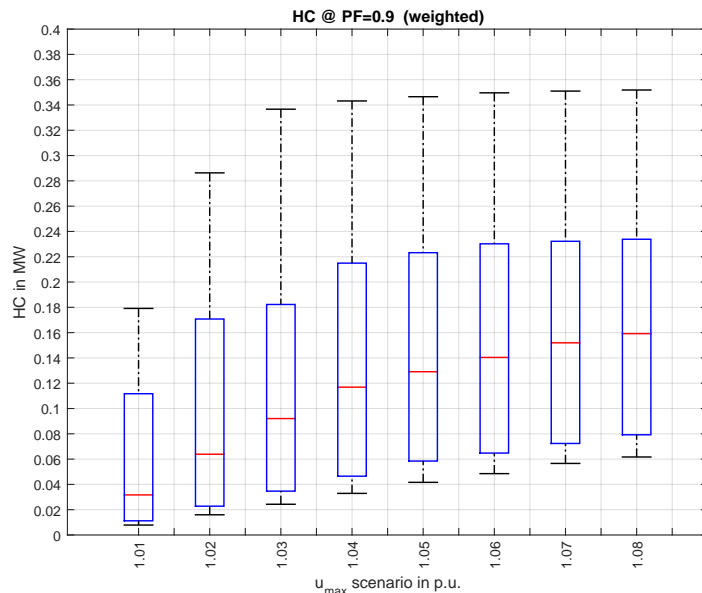
Figure 21: DSO 1 and DSO 2 - Hosting capability for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *weighted*, powerfactor 0.9 (ind.))

1.04*p.u.*, a HC of 120kW can be reached for about 50% of the feeders.

Last but not least, the results for the DER-scenario *eof* and PF=0.9-control (ind.) are given in Figure 22. It is apparent, that above 1.05*p.u.* the median HC remains stable at around 80kW. Thus, a voltage limit of 1.06*p.u.* can be justified for dedicated feeders supplying a single customer. Moreover, for 1.04*p.u.* and 1.05*p.u.* a similar median HC is reached. Also in this case, a voltage limit of 1.01*p.u.* cannot be recommended. This scenario is equivalent to the control scheme VoltVAr and the DER-scenario eof.

In Figure 23, the share of voltage and loading constrained feeders (U/I share) is investigated for the depicted DER-scenarios, voltage limits and PF=1. Additionally, the U/I share is also given for the VoltVAr-control. Obviously, there is no significant difference between the DER-scenarios *uniform* and *weighted*. For $u_{max} = 1.01p.u.$ nearly all feeders are voltage constrained (share below the line). Increasing the available voltage band for in-feed, leads to a higher share of loading constrained feeders. E.g. for $u_{max} = 1.06p.u.$, already 50% of the feeders are loading constrained for the DER-scenarios *uniform* and *weighted*. For the DER-scenario *eof* this share is reached at $u_{max} = 1.07p.u.$ and already at $u_{max} = 1.05p.u.$ for the VoltVAr-control
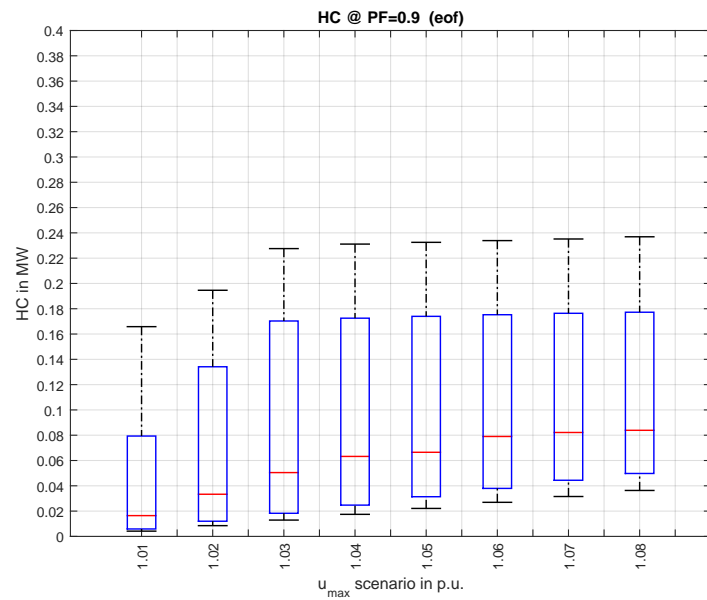
Figure 22: DSO 1 and DSO 2 - Hosting capability for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *end of feeder*, powerfactor 0.9 (ind.))
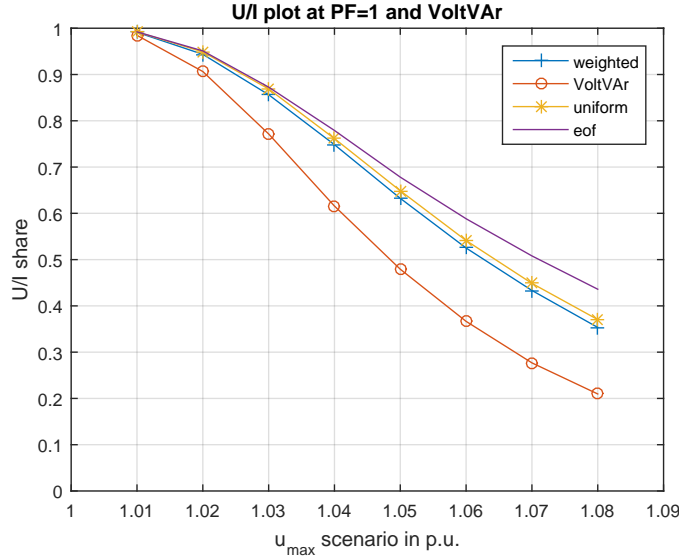
Figure 23: DSO 1 and DSO 2 - Share of U/I constrained feeders as a function of the admissible voltage rise for the investigated DER-scenarios with a power factor of 1 and for the DER-scenario uniform with a VoltVAr-control. One means all feeders are voltage constrained

(DER-scenario *uniform*).

After investigating the HC results for several DER-scenarios, voltage limits and control strategies and the share of voltage and loading constrained feeders, the following recommendations can be given: For the DER-scenario *eof*, a voltage limit could be suggested: $1.04p.u.$ for PF=1 and $1.06p.u.$: for PF=0.9 (ind.). For the DER-scenario *uniform* the voltage limit $1.01p.u.$ cannot be recommended due to the significant lower HC, when compared to the other voltage limits. However, no upper limit can be suggested. In addition, to the discussed boxplots, the ratio between HC values for different voltage limits (according to the definition in section 3.3) may provide further information about feeders. Therefore, the ratio between the DER-scenarios *uniform* and *eof* is calculated for all feeders and selected voltage limits. In Figure 24 for example, the parameter ueof1 and ueof8 are plotted for both DSOs. For this plot, four load flow calculations are necessary: Two HC calculations for each admissible voltage limit of $u_{max} = 1.01p.u.$ and $u_{max} = 1.08p.u.$, for the DER-scenarios *uniform* and *eof* are required. After that, the ratio of the HC obtained for the two DER-scenarios for each voltage limit is calculated. The blue points indicate voltage constrained feeders
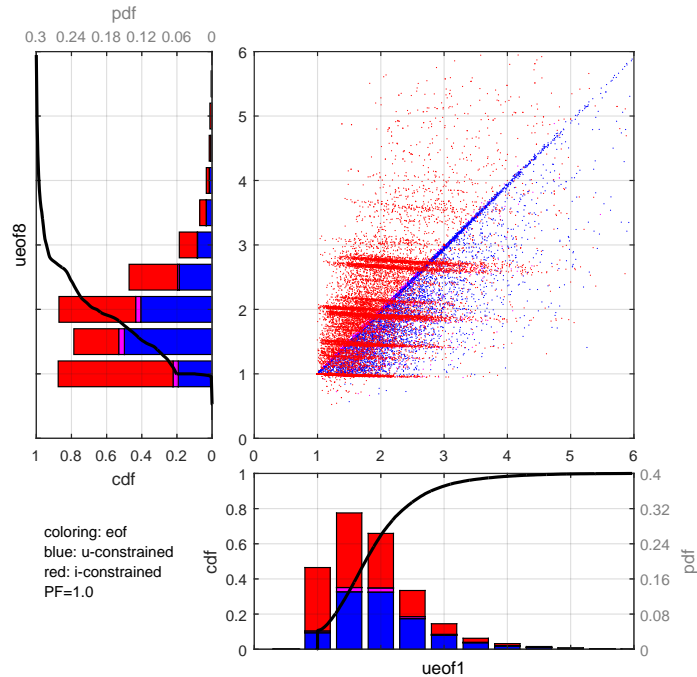
Figure 24: DSO 1 and DSO 2 - hosting capability sensitivity evaluated for an admissible voltage rise of 1% (ueof1) and 8% (ueof8) - Points are colored according to the constraint reason (voltage or loading) for the DER-scenario *end of feeder* at ueof8 (powerfactor 1)

and red points indicate loading constrained feeders at $u_{max} = 1.08 p.u.$ and the DER-scenario *eof*. Logically, feeders on the median (x=y), feeders which have the same ueof value for both voltage limits are mostly invariant to the loading constraint and are in both cases voltage constrained. Remaining blue points are mainly below the median. Points below the median have a higher ueof1 than a ueof8 value. Loading constrained feeders are mostly located along a particular ueof8 value and are horizontally distributed. However, in all areas of the plot, red points can be observed, but about 90% of all feeders have a ueof1 and ueof8 value smaller equal three.

Coloring the same Figure for the DER-scenario *uniform*, gives a different distribution of a significant number of points. Voltage constrained feeders that were below the median now appear above the median and feeders below the median are now loading constrained. It is apparent, that the coloring of the median and the horizontal distributions at particular ueof8 values remain unchanged. For the point ueof1=1 and ueof8=1 the same HC is calculated
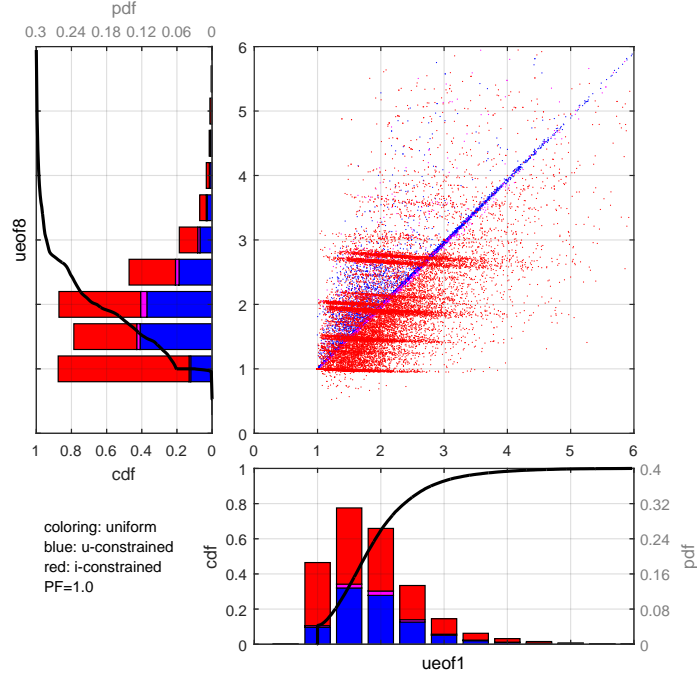
Figure 25: DSO 1 and DSO 2 - hosting capability sensitivity evaluated for an admissible voltage rise of 1% (ueof1) and 8% (ueof8) - Points are colored according to the constraint reason (voltage or loading) for the DER-scenario *uniform* at ueof8 (powerfactor 1)

for all four HC calculations ($u_{max} = 1.01p.u.$,$u_{max} = 1.08p.u.$ and the DER-scenarios *uniform* and *eof*). Hence, a topology could be designed for this particular case. Indeed, if for the DER-scenarios *uniform* and *eof* the same HC value is calculated, the DER-distribution must be identical. If the DER-scenarios *uniform* and *eof* are identical, there must be only one load in the feeder at the end node. Hence the most simple topology containing a slack, one line and a load (two node feeder model) could be applied. However, the overall ueof1 or ueof8 values show, that these variables are continuous.

5.2 Feeder Statistics

In this section, the results of statistical analysis of feeder parameters are presented. Firstly, a justification on why several nominal current definitions are considered is given. Secondly, the correlation analysis between 72 HC scenarios and statistical feeder parameters are presented as well as the results
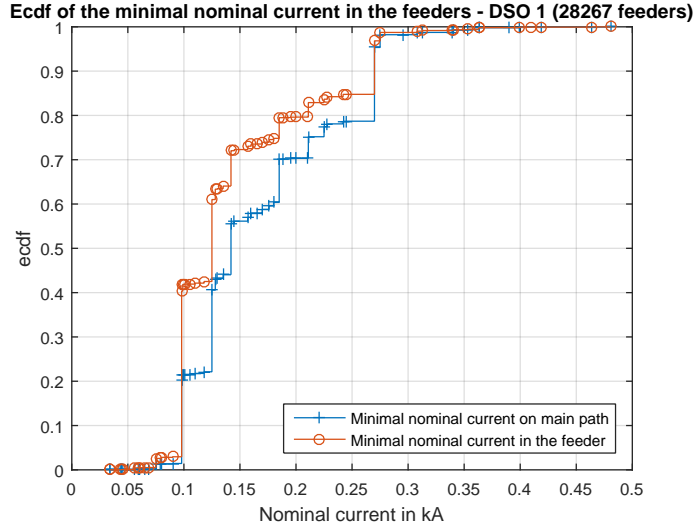
Figure 26: DSO 1 - Empirical cumulative distribution function (ecdf) of the minimal nominal current. Blue: minimal nominal current on the main path - red: minimal nominal current in feeders

of the probability tests for statistical parameters. In addition, the most relevant parameters to classify feeders accurately regarding the HC are identified. Thirdly, the distribution of the parameter HCS is discussed.

To begin with, the justification for considering several ampacity definitions is presented. In Figure 26, for example the ecdf of the minimal ampacity on the main path and in the feeder is depicted for DSO 1. Thereby the data of all feeders with plausible values are considered (including feeders without HC results). The Figure demonstrates that for a high number of feeders the minimal nominal current on the main path is higher compared to the overall feeder. Only for 35% of the feeders are these two parameters are equal. Therefore, a clear distinction between these two parameters is necessary.

Similarly to Figure 26, the difference for DSO 2 is depicted in Figure 27. Once again, the importance of distinguishing between these two definitions of the minimal nominal current is demonstrated. However, in case of DSO 2, for almost 50% of the feeders, the two parameters are equivalent.

The next three figures (Figure 28, 29 and 30) show the correlation coefficient according to Pearson, Spearman and Kendall respectively. Each figure is split into six areas. Thereby, all considered statistical parameters are listed twice on the x-axis and divide the figure in two parts, whereby the left part covers the correlation under the control scheme PF=1 and the right part
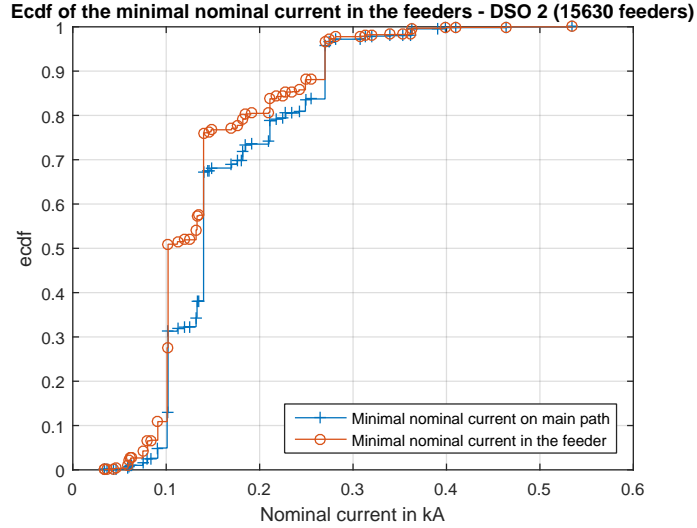
Figure 27: DSO 2 - Empirical cumulative distribution function (ecdf) of the minimal nominal current. Blue: minimal nominal current on the main path - red: minimal nominal current in feeders

PF=0.9 (ind.). The y-axis is split into 3 areas, namely the DER-scenarios *uniform* and *eof* and the control scheme VoltVAr (also DER-scenarios *uniform*). Each block of the y-axis consists of eight lines representing one of the eight considered voltage limits as listed in Table 7 and Table 8. One should note that the combination of the control schemes PF=0.9 (ind.) and VoltVAr are identical for the DER-sceneario *eof* at nominal power. The lower left part of the figure also contains the VoltVAr results, as on the lower right of the figure. Since the approach of this work is to define a DSO-independent reference feeder methodology, the feeder data of both DSOs is considered for the correlation analysis to avoid the identification of parameters that may be suitable for only one DSOs and not for other DSOs.

In Figure 28 the heat map of the absolute Pearson correlation coefficient of the HC and statistical parameters is depicted. Obviously, rather low correlation coefficients are observed for most of the parameters for all simulation cases. The highest correlation between the HC of one simulation case and one parameter is reached for the parameter $I_{nom-Min}$ (Minimal cable rating of a line in the feeder). Another important result of this figure is that the correlation coefficient differs depending on the simulation case. For example, the parameter $I_{nom-Min}$ correlates best (0.7) with the HC of the simulation case: DER-scenario *eof*, $u_{max} = 1.08p.u.$ and PF=0.9 (ind.). Also for PF=1

60

Table 7: Voltage limit of simulation cases (y-axis of figures 28, 29 and 30)

Voltage limit	Sim case	Sim case	Sim case
$u_{max} = 1.01p.u.$	1	9	17
$u_{max} = 1.02p.u.$	2	10	18
$u_{max} = 1.03p.u.$	3	11	19
$u_{max} = 1.04p.u.$	4	12	20
$u_{max} = 1.05p.u.$	5	13	21
$u_{max} = 1.06p.u.$	6	14	22
$u_{max} = 1.07p.u.$	7	15	23
$u_{max} = 1.08p.u.$	8	16	24

Table 8: DER-scenario of simulation cases (y-axis of figures 28, 29 and 30)

Sim case	Scenario	Sim case	Scenario	Sim case	Scenario
1	uniform	9	eof	17	uniform (+VoltVAr)
2	uniform	10	eof	18	uniform (+VoltVAr)
3	uniform	11	eof	19	uniform (+VoltVAr)
4	uniform	12	eof	20	uniform (+VoltVAr)
5	uniform	13	eof	21	uniform (+VoltVAr)
6	uniform	14	eof	22	uniform (+VoltVAr)
7	uniform	15	eof	23	uniform (+VoltVAr)
8	uniform	16	eof	24	uniform (+VoltVAr)
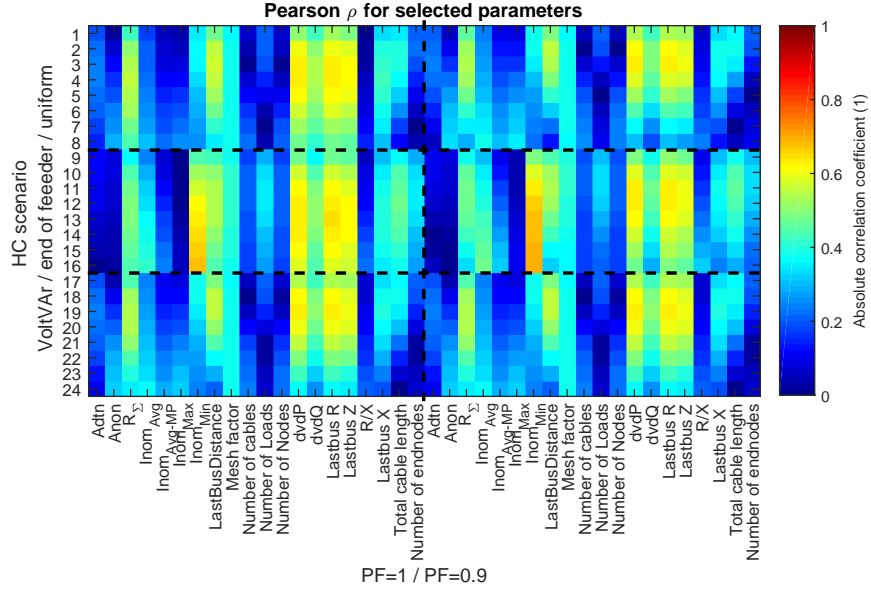
Figure 28: DSO 1 and DSO 2 - Heat map of the Pearson correlation. The plot shows the correlation between statistical parameters and the hosting capability for various scenarios (y-axis). The plot is divided into 6 areas according to the DER-scenario and powerfactor. The bottom left and right part are identical (The VoltVAr-control is overruling PF=1 and PF=0.9)

a relatively high correlation coefficient is reached (0.68). On the contrary, for the parameters $dvdP$ and $dvdQ$ the highest correlation coefficient is reached for lower voltage limits.

In Figure 29 the Spearman rank correlation coefficient is calculated in analogy to the previous plot. It is apparent, that much higher correlation coefficients are observed compared to the Pearson correlation. High correlation coefficients are found for the parameters $dvdP$ and R_k (Resistance at the end node) for rather low voltage limits ($1.01p.u.$ to $1.03p.u.$) and the DER-scenario eof. In general the higher correlation coefficient with a PF=1 control scheme decrease under a PF=0.9-control (ind.) scheme. Alongside the two parameters R_k and $dvdP$, additionally the sum impedance R_Σ, the feeder length and total cable length of the feeder show a higher correlation coefficient compared to other parameters. For the correlation coefficients of the parameter $I_{nom-Min}$ and the DER-scenario eof, a similar trend as in Figure 28 can be observed. The correlation coefficient is increasing for higher voltage limits. The dependency of the correlation coefficient on the voltage limit can be explained with Figure 23, since the share of voltage constrained
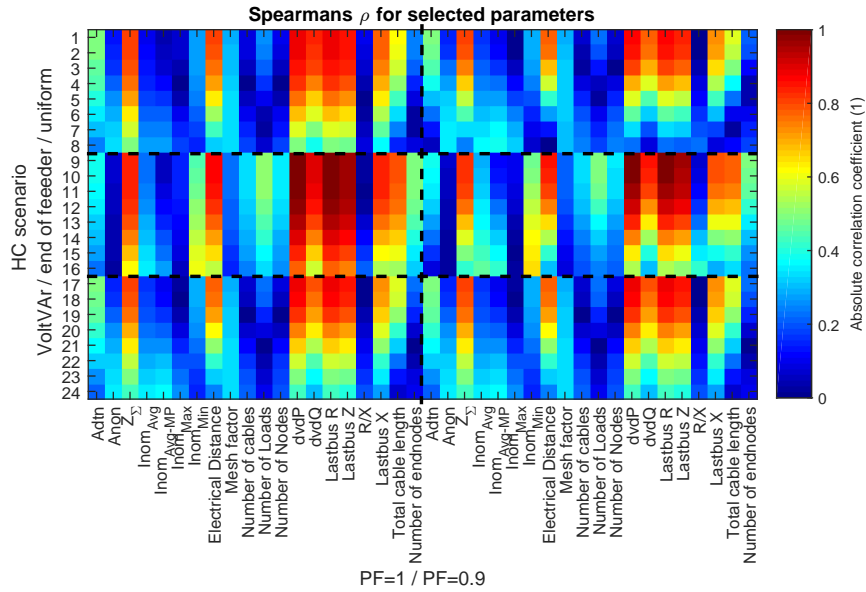
Figure 29: DSO 1 and DSO 2 - Heat map of the Spearman correlation. The plot shows the correlation between statistical parameters and the hosting capability for various scenarios (y-axis). The plot is divided into 6 areas according to the DER-scenario and powerfactor. The bottom left and right part are identical (The VoltVAr-control is overruling PF=1 and PF=0.9)

feeders decreases for higher admissible voltage limits.

The third correlation analysis according to Kendall, is presented in Figure 30. Since the Kendall correlation is also a rank correlation, similar results compared to the Spearman correlation were obtained. However, the correlation coefficients reach a lower level compared to Figure 29. Nevertheless, the same parameters R_k, $dvdP$, R_Σ, the feeder length and total cable length of the feeder show a higher correlation coefficient compared to the other parameters.

The aim of this work is to identify parameters that are suitable for defining representative feeders. Identifying the most relevant parameters to generate representative feeders requires, that the parameters identified also be DSO independent. Therefore, the hypothesis that particular parameters follow a probability distribution was tested for the considered statistical parameters. Tests were conducted, if parameters follow one of the a Weibull, Rayleigh, Lognormal, Normal or Exponential distributions with a significance level of 5%. Even though all feeders with valid parameters could be considered, only feeders which passed the plausibility test (plausible feeder data and valid
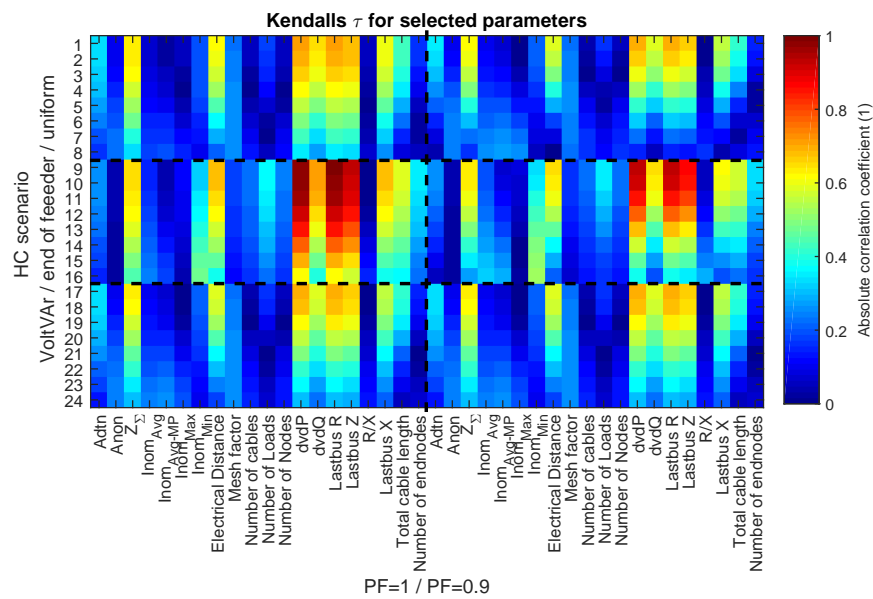
Figure 30: DSO 1 and DSO 2 - Heat map of the Kendall correlation. The plot shows the correlation between statistical parameters and the hosting capability for various scenarios (y-axis). The plot is divided into 6 areas according to the DER-scenario and powerfactor. The bottom left and right part are identical (The VoltVAr-control is overruling PF=1 and PF=0.9)

HC results) were considered. It turned out that only two parameters follow a normal distribution for both DSOs. These parameters are the ADTN (average distance to neighbors) and the R_Σ (equivalent sum impedance). However, the correlation between ADTN and the HC results is rather low. In contrast, the correlation between R_Σ and the HC results is high for the DER-scenario *eof* and a admissible voltage rise of 1%. However, the correlation coefficient for the parameter R_k is always higher.

The correlation analysis showed that several parameters have a rather high correlation coefficient for particular simulation scenarios only. This demonstrates the dependency of the correlation coefficient on the admissible voltage limit and DER-scenario. With the Pearson and Spearman correlation, the minimal nominal current in the feeder (parameter $I_{nom-Min}$) was identified as the only parameter that shows an increasing correlation coefficient for higher voltage limits, particularly for the DER-scenario *eof*. In Table 5 several parameters regarding the nominal current of feeders were defined: the minimal and maximal nominal current of lines as well as the weighted average nominal current of all lines in the feeder and on the main path. However, only one of these parameters (minimal nominal current in the feeder $I_{nom-Min}$ has a high correlation coefficient with the HC. Using the Spearman and Kendall correlations, several parameters related to the end node of the feeder and the total cable length were discovered as highly correlated parameters regarding the HC of conservative voltage limits (1.01$p.u.$ and 1.02$p.u.$). In conclusion, two parameters, namely R and $I_{nom-Min}$ could be selected as most relevant parameters (for covering the voltage constraint and loading constraint, respectively).

Using the correlation analysis, the two most relevant parameters of feeders, with regards to the defined HC scenarios, R and $I_{nom-Min}$, were found to be sufficient to parametrize a radial feeder. For a two node feeder model (slack, line and load), the parameter R_k could be reached with various line types and an appropriate length. However, considering the parameter $I_{nom-Min}$, the line type can be determined. As a consequence, with the parameter pair R_k and $I_{nom-Min}$, the length of the feeder can be defined. For the DER-scenario *eof*, more complicated topologies (e.g. including branches), can still be reduced to a two node feeder model. Hence, the parameters R_k and $I_{nom-Min}$ could still be utilized to parametrize the simplified topology. Nevertheless, for the DER-scenario *uniform* no accurate calculations could be performed with a simplified topology. Therefore, there is still a lack of information available to define reference feeders, in particular a lack of some topological information. The topology of a feeder contains information that is crucial for HC studies. In section 4.4, three points were listed: the relative location of cables in the feeder, branches as well as the location of customers.
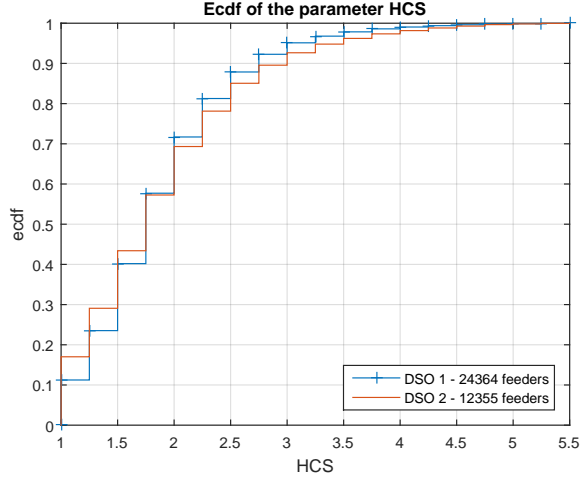
Figure 31: DSO 1 and DSO 2 - Empirical cumulative distribution function of the hosting capability sensitivity. Blue: DSO 1, red: DSO 2

There is a need for a simple measure covering topological information which is influencing HC results. In section 5.1 measures were presented, that contain topological information in form of HC ratios. For example, it was discussed, that for an ueof1 or ueof8 value of one, an accurate two node feeder model can be defined. Since these ratios can be defined for any admissible voltage limit (ueof1, ueof2, ueof3, ..., ueof8), a new measure is introduced: HCS is the ratio of the HC for the two DER-scenarios uniform and eof at $u_{max} = 1.01 p.u.$, thus a measure that implicitly contains the topology of the feeder is derived. Unfortunately, HCS values depend on the voltage limit (e.g. different values of ueof1 and ueof8 for the same feeders as depicted in Figure 24). However, while for ueof1, most of the feeders are voltage constrained, for ueof8 more than 50% of the feeders are loading constrained. Therefore, to consider only one constraint in the defined topology measure, HCS relates to ueof1. Furthermore, HCS is a continuous value and therefore not directly suitable for defining topologies. Therefore, the parameter HCS obtained for real feeders is rounded to the nearest quarter. The ecdf of the measure HCS, rounded to the nearest quarter, for both DSOs is depicted in Figure 31. The ecdf of the plot shows that less than 10% of the feeders have a HCS value greater than three. Hence, the eight most relevant HCS values can be identified (1 to 3 in 0.25 steps). Since for a HCS value of 1, a two node reference feeder model was already defined, however there still remains seven more reference topologies to be developed.

The distribution of the parameter HCS for the investigated DSOs is given
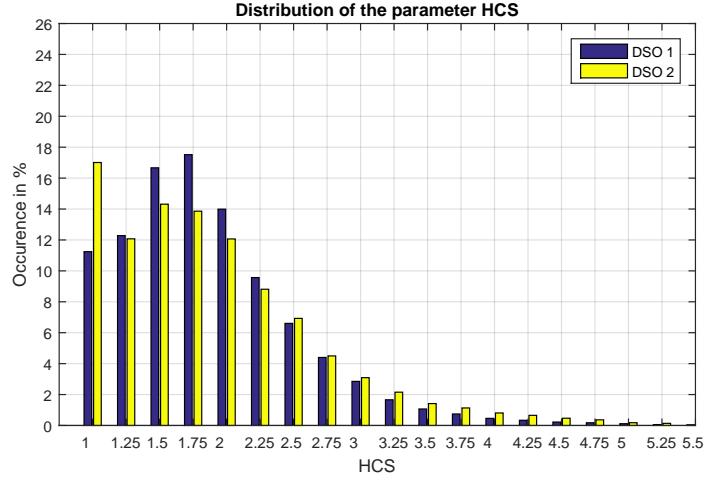
Figure 32: DSO 1 and DSO 2 - Distribution of the hosting capability sensitivity (rounded to quarters). Blue: DSO 1, yellow: DSO 2

in Figure 32. In the case of DSO 1, approximately 11% of the feeders and in the case of DSO 2, about 17% of the feeder, could be modeled simply by a two node feeder model. For a value of 1.25 and values above 2.25, similar shares are observed for both DSOs. DSO 1 has a higher share of HCS values between 1.5, 1.75 and 2 when compared to DSO 2.

In Figure 33, the distribution of the end node resistance R_k for both DSO is shown. For about 90% of the feeders, the resistance is below $300m\Omega$ and for more than 70% of the feeders below $200m\Omega$. The ecdf below $200m\Omega$ is almost linear. The distribution of the parameter R_k may be used to select a number of values and permutate with the other parameters $I_{nom-Min}$ and HCS to define reference feeder input data. Thereby the number of reference feeders has to be chosen according to the desired accurcy of results.

In this section, the parameters R_k and $I_{nom-Min}$ were identified as most relevant parameters according to the correlation analysis. Together with the introduced topology measure HCS, a input parameter triple was defined that will be used to define and validate reference feeders. The parameters R_k and $I_{nom-Min}$ can be relatively easily obtained for any feeder by DSOs. However, the parameter HCS may require more efforts (currently). In conclusion, the problem to define a reference feeder can be split into two smaller problems, namely, the topology selection and the parametrization of a reference feeder as already shown in Figure 14. Ideally these three parameters can be retrieved
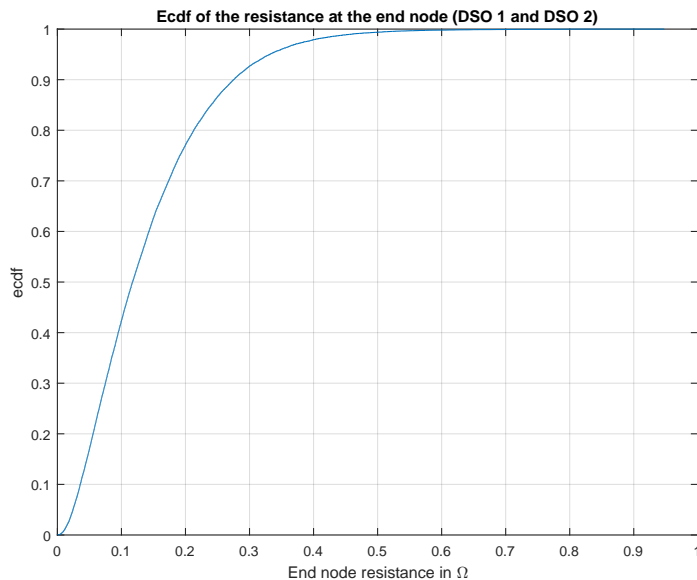
Figure 33: Empirical cumulative distribution function of the end node resistance for the investigated DSOs

from the DSO. Based on this parameter triple, a equivalent reference feeder can then be generated. A total of 24.364 feeders in the case of DSO 1 and 12.355 feeders in the case of DSO 2 were used for the validation of the reference feeder methodology in this work.

5.3   Reference Feeder Topologies

In this section, the reference topologies developed during this work are presented. One should note that the HCS is a continuous parameter. Consequently, there is a priori an infinite number of reference feeder topologies that could be defined. For both DSOs only very few feeders have a HCS value greater than 4. As previously discussed, to reduce the complexity for the utilization of the reference feeder methodology and at the same time limiting the number of reference topologies, HCS values were rounded to the nearest quarter from 1 to 5.75. In total, 20 reference topologies were defined. The reference feeder topologies were developed by designing topologies, where the same HCS value for the investigated voltage limit borders $1.01p.u.$ and $1.08p.u.$ is obtained. The reference feeder topologies are on the median in Figure 25 ($ueof1 = ueof8$). Hence, a line type and length was used to ensure that the parametrized reference topology remains voltage constrained. Nev-

ertheless, during the parametrization of reference topologies with the correct parameters R_k and $I_{nom-Min}$, the loading constraint effect should automatically arise with the considered ampacity value. The reference topologies are grouped by their first digits (e.g. 1.xx = A, 2.xx= B, 3.xx=C, etc.) and the quarters are counted from 1 to 4 (e.g. HCS=1.75 equals reference topology A4). One should note that the presented reference feeder topologies for particular HCS values are not inimitable. Other reference topologies leading to the same HCS value exist. Furthermore, in real feeders, many different line types with a specific length can be expected. Furthermore, even cables in parallel may be found in real grids. Moreover, either underground cables or over-head lines cables could be used to connect customers. Having acknowledged this, the need to consider this variation in the reference feeders is evident. However, since this information is not covered by the identified input parameter triple, another approach is required. Even though in the developed reference feeder topologies the same line type for all cables is used, lines with cables in parallel, have different nominal currents and resistances per length. Thus, the discussed variations are considered in the presented reference topologies in this section. Furthermore, a pattern describing the relation between reference feeder topologies can be described.

Figure 34 shows the first and most trivial reference topology. At Terminal 2, a feeder segment I is connected, supplying a single load. Basically, the feeder topology A1 could be equivalently reduced by bypassing Terminal 2 with a single line between Terminal 1 and Terminal 3 (Two node feeder equivalent). However, to ease understanding of the following reference topologies, Terminal 2 is introduced. The HC for the DER scenario *uniform* and *eof* will be identical, which gives an HCS value of 1. It is apparent, that this particular topology, independently from any line type or voltage limit, leads to a HCS value of 1.

The second reference topology (HCS=1.25) is depicted in Figure 35. This topology is characterized by two feeder segments I at Terminal 2. Thereby, by definition the line type, line length and the number of lines in parallel of Line 12, Line 23 and Line 24 are identical. Consequently, customers located at Terminal 3 and Terminal 4 may not use the full transmission capacity of their dedicated lines (Line 23 and Line 24) simultaneously, otherwise Line 12 would be overloaded by 100%. Therefore, this kind of topology is suitable for supplying customers with a rather low coincidence factor. In regard to DER-integration - with a rather high coincidence factor on LV-level - Line 12 would be limiting in this case the HC. Therefore, for RFT A2, a reinforcement of Line 12 could be a suitable option to increase the HC of the feeder. This would lead to the RFT, which is presented next.
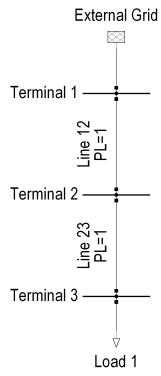
Figure 34: Reference feeder topology A1
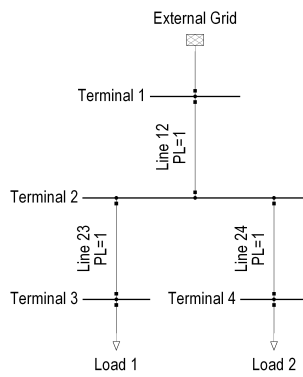
Figure 35: Reference feeder topology A2
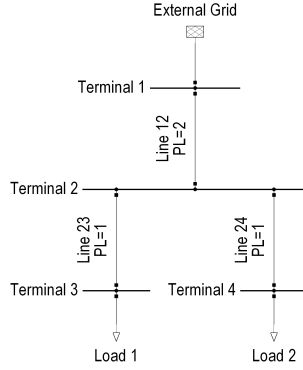
Figure 36: Reference feeder topology A3

In Figure 36, the reference topology for the HCS value of 1.5 is depicted. Compared to Figure 35, the restriction that both customers could together overload Line 12 is lifted and the full capacity of Line 23 and Line 24 can be utilized in terms of HC. Thus, reference topology A3 is a 'reinforced' version of topology A2.

The last RFT of group A is A4 and is depicted in Figure 37. This topology is characterized by a load introduced at Terminal 2 and with a stronger connection to the slack compared to Load 1 and Load 2. However, the three customers have to share the Line 12. Technically, a concentrated placement of DER at Terminal 2 would be favored to reduce network losses caused by the in-feed which would be unfair against other customers. Allocating the same amount of DER on the Terminals 2, 3 and 4, will lead to higher losses in the feeder compared to placing the DER at Terminal 2 only. Also in this topology, all lines have the same length but a different number of lines in parallel. It is apparent, that a 'design rule' can be noticed. Starting from the base topology of a group, e.g. A1 (Figure 34), three extensions are needed to reach A4. A feeder segment I is added at Terminal 2, which leads to the reference topology A2 (Figure 35). In a next step the nominal line rating of Line 12 is doubled - introducing a parallel line of the same line type - and topology A3 is obtained (Figure 36). Last but not least, Load 3 is connected to Terminal 3 and characterizes the reference topology A4 (Figure 37).

71
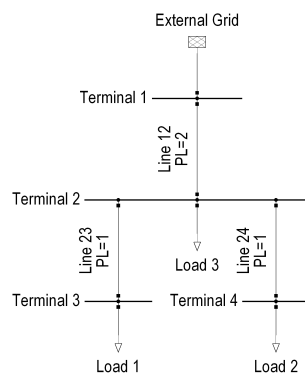
Figure 37: Reference feeder topology A4
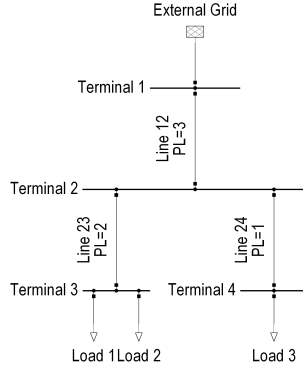
Figure 38: Reference feeder topology B1

Next, the first reference topology of the group B is depicted in Figure 38. This topology contains a feeder segment I and a feeder segment II with a total number of loads of three. Line 12 has three lines in parallel and Line 23 has two lines in parallel. There are two customers located at Terminal 3. Thus, the full transmission capacity of the Line 23 has to be shared by Load 1 and Load 2. Also in this topology the length of the lines is equal. Thus, Terminal 4 is the end node (weakest node), since there is only one line in parallel. For the remaining reference topologies of group B, the same pattern as for group A applies.

Moreover, in Figure 39 the first reference topology of group C is depicted. The feeder contains 4 feeder segments II and one feeder segment I. In analogy to A1 and B1, Line 12 is not a 'bottleneck' concerning simultaneous in-feed. If a fair distribution of DER per customer is chosen (determined by the HC at Terminal 7), no restrictions occur concerning overloading a line. The line type and the length of the lines is equal, therefore Terminal 6 is the node with the highest impedance. Thus, another 'design rule' can be observed between the base topologies (A1, B1, C1, D1, E1). For example B1 (Figure 38) consists of the reference topology A1 with one additional feeder segment II at Terminal 2. Thereby the number of parallel cables of Line 12 is the sum of the number of loads in the feeder (for A1, B1, C1, D1, E1). C1 is based on B1 plus an additional feeder segment II at Terminal 2. This 'design rule' can be extended respectively for D1 and E1. Starting from each base topology, the described pattern for the group A1 to A4 can be applied. The complete set of reference topologies is given in appendix A.

73
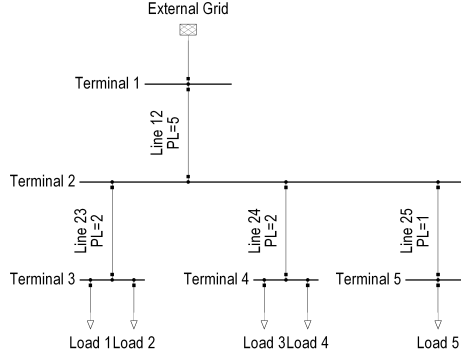
Figure 39: Reference feeder topology C1

In [48], authors report that the decision to branch, and also how often to branch, is handled randomly provided that sufficient connections are assigned to the branch. This is contrary to the presented methodology whereby the topology of the reference feeder is selected according to the input parameter HCS. Even though only the parameter HCS is used to select a topology, some electrical properties are defined as well, such as the number of parallel lines and that all lines have the same length.

The input parameters for the reference feeder methodology were defined and their distributions discussed. Since RFT were developed, findings based on the available input data can be immediately discussed. For all feeders, the input parameter triple R_k, $I_{nom-Min}$ and HCS are available. Hence, a classification can be performed according to the topology measure HCS. In Figure 40, the ecdf of the parameter $I_{nom-Min}$ is depicted for the HCS values between 1 and 1.75 (group A). It is apparent, that significant differences in the ecdf of each group are visible. For dedicated feeders (HCS=1, A1) with one load at the end node, the ampacity of the weakest line is higher compared to the other shown feeder topologies. 40% of these feeders have a rating higher than 0.25kA (cable-cross section higher than NAYY 4x120). Only about 15% of the feeders contain a cable with a rating below 0.1kA (NYY 4x16 which is mainly used for the service line). Contrary to A1, 50% of the feeders related to topology A4 (HCS=1.75) contain a cable with such a low rating. The ecdf of the parameter $I_{nom-Min}$ for topologies A2 and A3 show that four distinct areas are covered. Moreover, for each distribution, the share of the overall feeders is given. About 13% of the investigated feeders can be described by reference topology A1 (HCS=1) and another 13% by A2

74
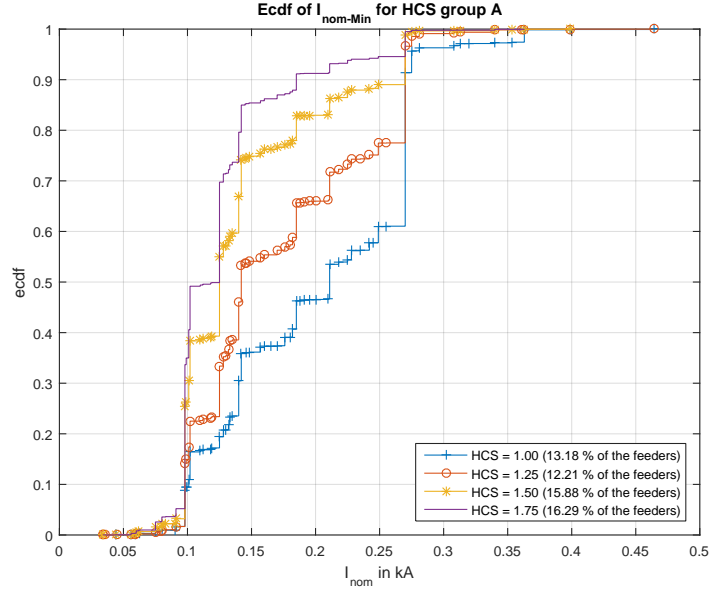
Figure 40: Empirical cumulative distribution function of the minimal nominal current in feeders for the defined hosting capability sensitivity group A

(HCS=1.25). The reference topology A3 (HCS=1.5) or A4 each are suitable to describe even about 16% of the feeders. Hence, 58% of all feeders are linked to HCS group A.

The ecdf of the parameter $I_{nom-Min}$ for HCS group B is depicted in Figure 41. It is apparent, that the ecdf of these subgroups are less distinct compared to Figure 40. For all plots, the share of feeders containing a cable with a rating below 0.1kA is nearly 60%. Furthermore, the overall share of feeders is significantly lower compared to HCS group A. Group B1 (HCS=2), has the highest share of about 13% and the share is continously decreasing. About 9% of the feeder can be described with B2 (HCS=2.25), almost 7% by B3 (HCS=2.5) and only 4% of the feeders by B4 (HCS=2.75). Hence, 33% of the feeders are linked to the HCS group B. Consequently, about 90% of all feeders can be either described with one of the eight reference feeder topologies of HCS group A or B as identified according to Figure 31. The ecdf plots of the remaining HCS groups C and D can be found in appendix B.

The second input parameter R_k, the resistance at the end node, can also be clustered with the developed reference feeder topologies. In case of HCS group A, again more distinct ecdf curves are observed. For example, due to the higher cable-cross section of reference feeder topology A1, the resistance
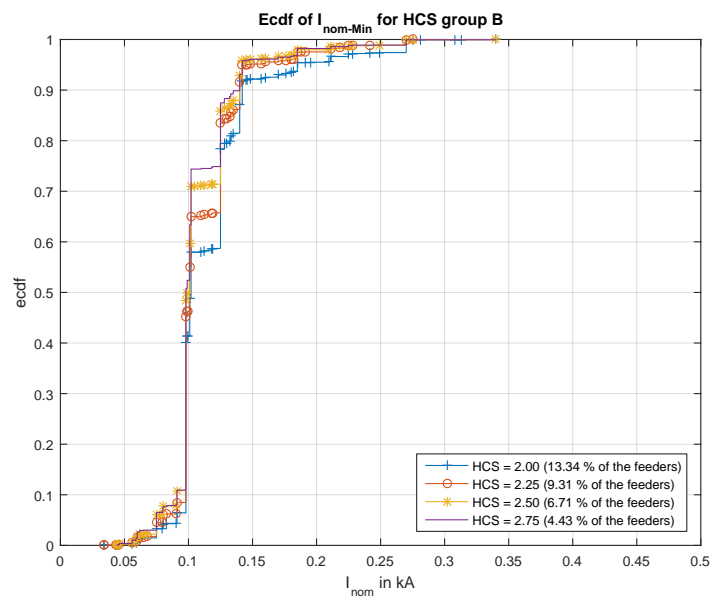
Figure 41: Empirical cumulative distribution function of the minimal nominal current in feeders for the defined hosting capability sensitivity group B
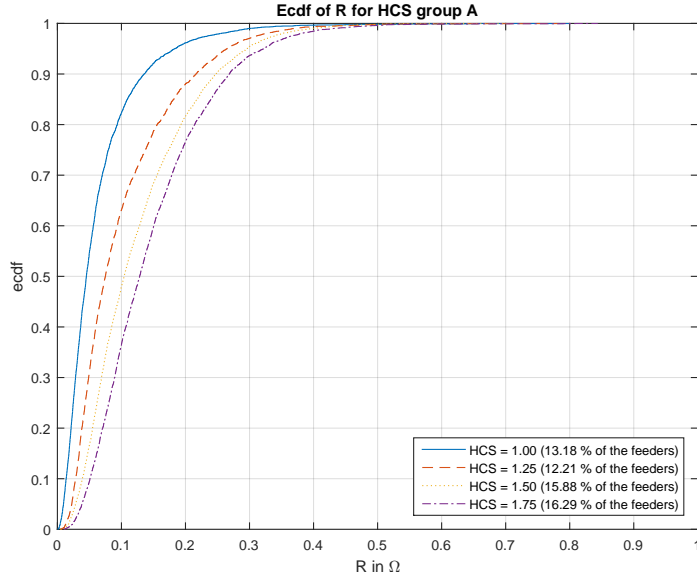
Figure 42: Empirical cumulative distribution function of the end node resistance in feeders for the defined hosting capability sensitivity group A

at the end node is lower compared to the other reference feeder topologies. The 90th percentile for A1 is about $130m\Omega$, for A2 $210m\Omega$, A3 about $240m\Omega$ and for A4 about $260m\Omega$.

The ecdf of the parameter R for HCS group B is depicted in Figure 43. The trend of all group members is similar, hence less variation can be observed compared to HCS group A. However, considering only one distribution, e.g. the ecdf of B2 or B3, for the entire group may lead to inaccurate input parameter R which consequently leads to an error in results regarding the HC. The ecdf plots of the remaining HCS groups C and D can be found in appendix C.

5.4 Reference Feeder Parametrization

In this section, the parametrization of the developed RFT is presented. Therefore the parameters $I_{nom-Min}$ and R_k are required. In the most simple case of Figure 34 (RFT A1), a feeder can be parametrized by one line type and the length of the feeder. The resistance at the end node depends then on the line type and the length of the line. However, the resistance at the end node can be reached by various pairs of line types with an appropriate length or lines in parallel. Therefore, the resistance per length is not a unique measure to select a particular line type. A line type is character-

77
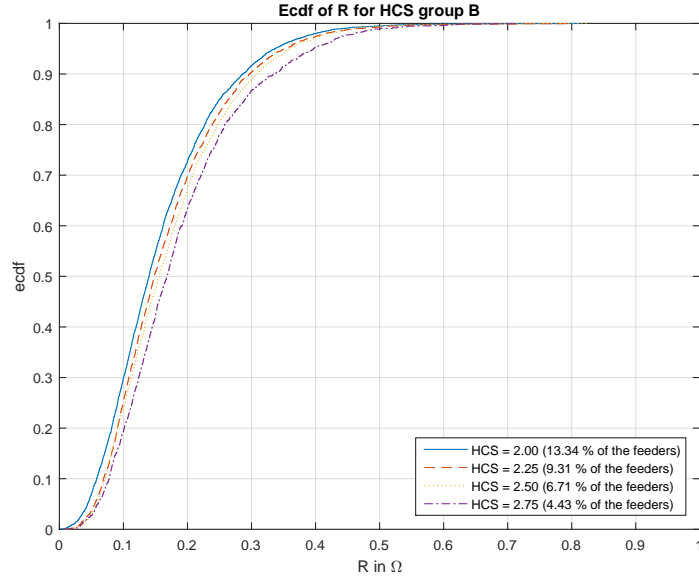
Figure 43: Empirical cumulative distribution function of the end node resistance in feeders for the defined hosting capability sensitivity group B

ized by the ampacity (nominal current) and the resistance and reactance per length. Therefore, the line type to be used is determined by the input parameter $I_{nom-Min}$. Accordingly, the resistance per length of the line type can be retrieved. Finally, the input parameter R_k determines the length of the cables such that the weakest node has the resistance R under the condition, that all lines have the same length. The approach can be summarized as follows:

1. Identification of the parameter HCS of the feeder:

   - With the parameter HCS, the reference topology is selected.

2. Gathering the line with the lowest ampacity $I_{nom-Min}$:

   - With the lowest nominal current of the feeder, the line type is obtained.

3. Calculation of R_k at the end node:

   - With the (short circuit) resistance at the end node, the length of the lines can be calculated.

An example for the parametrization with HCS=2, a particular $I_{nom-Min}$ and R_k is given: According to HCS=2, reference feeder topology B1 is selected (Figure 38). Next the line type is determined by $I_{nom-Min}$ and applied to all lines in the reference feeder topology. Thus, the resistance per length (R') is known. According to eq. 3, the length of a line can be calculated. Since all lines have the same length and Line 23 has two lines in parallel, Terminal 4 is the weakest node in the reference feeder topology. The resistance at the end node (R_{T4} - which is the input parameter R_k) can then be formulated:

$$R_{T4} = R_{L12} + R_{L24} \tag{11}$$

The length of lines is identical:

$$R_{T4} = R_{L12} + R_{L24} = R'_{L12} \times l + R'_{L24} \times l \tag{12}$$

The number of parallel lines has to be considered. Line 12 has three lines in parallel, thus the resistance per length is a third compared to Line 24:

$$R'_{L12} = \frac{R'_{L24}}{3} \tag{13}$$

Applied to eq. 12:

$$R_{T4} = \frac{R'_{L24}}{3} \times l + R'_{L24} \times l \tag{14}$$

Exchanging R_{T4} with the known input parameter R_k and solving equation to l gives:

$$l = \frac{R}{R' \times (1 + \frac{1}{3})} \tag{15}$$

Finally, all three lines in the reference feeder topology are parametrized with the obtained length according to eq. 15. For the validation presented in the next section, a script was written performing the parametrization of feeders, following the calculation approach. In all RFT, a feeder segment I exists. Therefore the calculation can be performed without the need to customize the calculation for any reference feeder topology.

6 Validation

In this section the defined reference topologies are validated based on the HC results. First, the mismatch between the HC of the investigated feeders and the HC of their equivalent reference feeders is presented. After that, the dependency of the parameter HCS on the investigated voltage limits is discussed. Next, the methodology is extended to a fixed PF-control and the resulting HC mismatch is evaluated. Furthermore, the proposed methodology is applied to an external network data set for the uncontrolled and the fixed PF-control scenario. Moreover, Monte-Carlo simulations are performed for selected real feeders and their equivalent feeders. Last but not least, real networks are scanned for the identified RFT and a statistic of most common reference topology combinations in networks is discussed to identify whether a few representative networks exists or not.

6.1 Reference Feeder Validation

In this section, the generated reference feeders are benchmarked against the real feeders. Therefore, the mismatch between the real HC of feeders and the obtained HC for the equivalent reference feeder is compared to the DER-scenario *uniform*. For both DSOs the relative error and absolute error is discussed. Furthermore, the rural and urban feeders of DSO 2 are analyzed separately. Regarding the absolute error, the overestimation of the overall HC of a feeder by more than 10kW for 10% of the feeders is considered as an important threshold level. The replacement of real feeders by their respective reference feeders leads to a loss of information in terms of number of cables, different line types, loads and exact topology. Therefore an exact level of accuracy for all considered voltage scenarios may not be reached. The HC mismatch results for the DER-scenario *eof* can be found in appendix D.

The HC mismatch results evaluated for each investigated voltage limit in case of DSO 1, are depicted in Figure 44. Obviously, the mismatch remains rather stable for the voltage limits of $1.01p.u.$ to $1.03p.u.$ and is then increasing for higher admissible voltage limits. Thereby the HC is overestimated. The mismatch is between $\pm20\%$ up to $1.04p.u.$ for 90% of the feeders (range of the box). At $1.05p.u.$, the range of the box reaches already 60% and increases to about 130% at an admissible voltage rise of 8% of the nominal voltage. For most of the investigated voltage limits, the median remains little higher than 0 mismatch. However, the overestimation error is significantly higher compared to the underestimation error. The figure demonstrates that the HC in case of DSO 1 can be reasonably estimated for 90% of the feeders, up to a voltage rise of 4% and with a mismatch of $\pm20\%$.
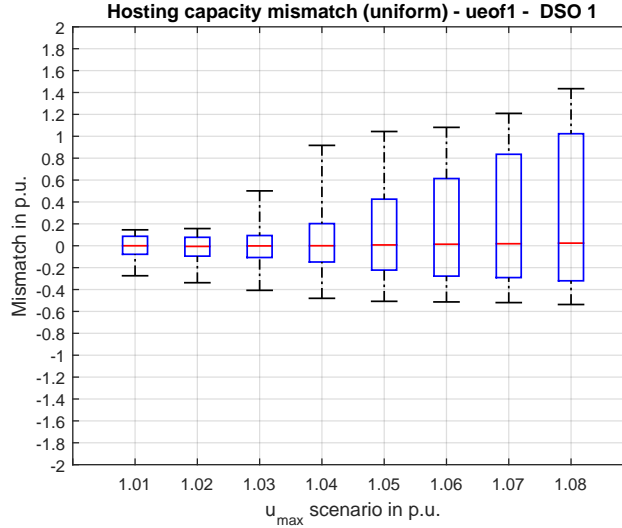
Figure 44: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 1)

The absolute HC mismatch in case of DSO 1 is depicted in Figure 45. In this figure, the absolute mismatch in MW is shown for all investigated voltage limits. For example, the overestimation of the HC by more than 10kW is exceeded for 10% of the feeders at a voltage limit of $1.04p.u.$ and the underestimation by more than 10kW for 7% of the feeders at the same voltage level. Hence, at a voltage limit of $1.04p.u.$, the mismatch is within a $\pm10\%$kW range for about 83% of the feeders. For lower voltage limits, even more feeders are within this range. At a voltage limit of $1.08p.u.$, the mismatch is underestimated by more than 10kW for already almost 14% of the feeders and overestimated by more than 10kW for 33% of the feeders. Hence, only 53% of the feeders remain within a $\pm10\%$kW range, indicating that the methodology is not suitable to accurately estimate the HC at higher voltage levels, since the full feeder model is simplified to a few nodes.

In case of DSO 2, the evaluation of the HC mismatch for the DER-scenario *uniform* is depicted in Figure 46. It is apparent, the mismatch is also increasing proportionally with a higher admissible voltage level. The mismatch is within a $\pm20\%$ range for the lower voltage limits ($1.01p.u.$ and $1.02p.u.$). At already $1.03p.u.$, the $\pm20\%$ range is exceeded and at $1.04p.u.$, already doubled ($\pm40\%$). Furthermore, the size of the box at $1.08p.u.$ is compared to DSO 1 is a little higher and about 140% ($+90\%$ to -50%). In
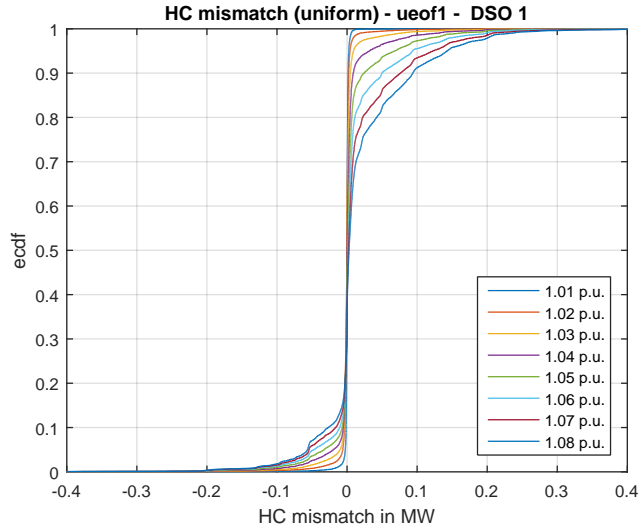
Figure 45: DSO 1 - Empirical cumulative distribution function of the hosting capability mismatch for different admissible voltage rise limits (DER-scenario *uniform*, powerfactor 1)

conclusion, the proposed methodology could only be suggested for a voltage rise limit of 2% of the nominal voltage in case of DSO 2.

The absolute HC mismatch in case of DSO 2 is depicted in Figure 47. The threshold level of overestimating the HC by more than 10kW for 10% of the feeders occurs at a voltage rise limit of 3%. For the same voltage limit, the HC is underestimated by more than 10kW for 11% of the feeders. At $1.03p.u.$, 86% of the feeders remain within a $\pm10\%kW$ range. Hence, compared to the previous figure, the proposed methodology could be extended also to a voltage limit of $1.03p.u.$, if overestimating the HC for 10% of the feeders by more than 10kW would be acceptable.

The HC mismatch for feeders of DSO 2 located in urban areas is depicted in Figure 48. Since the share of urban feeders is rather low compared to all feeders of DSO 2, significant differences can be identified. For example, beside the voltage limits $1.01p.u.$ and $1.08p.u.$ the 5th percentile is always lower compared to Figure 46 while the median shows no significant deviation. However, for the 95th percentile a additional overestimation by 10% is observed for several voltage limits. Hence, the methodology is sensitive to the classification of feeders as being either rural or urban.

The performance of a feeder located in rural areas of DSO 2 is given in 49. One important result is that for a voltage limit of $1.03p.u.$, the mismatch
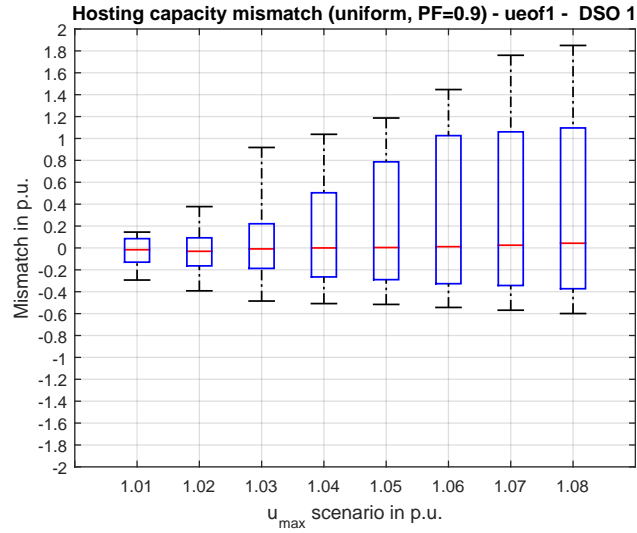
82
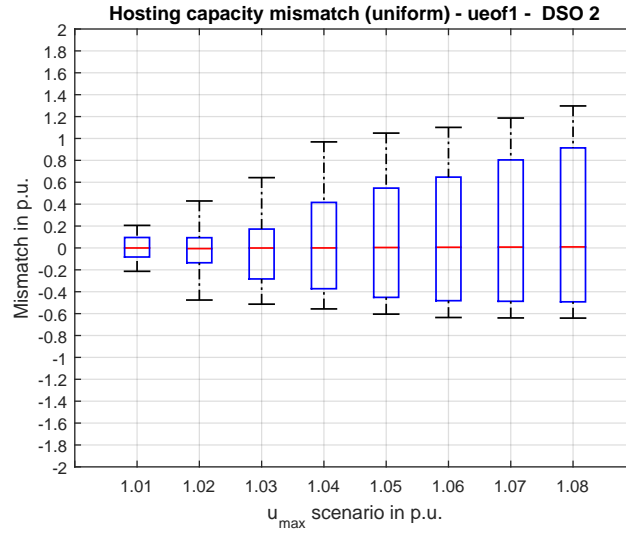
Figure 46: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 1)
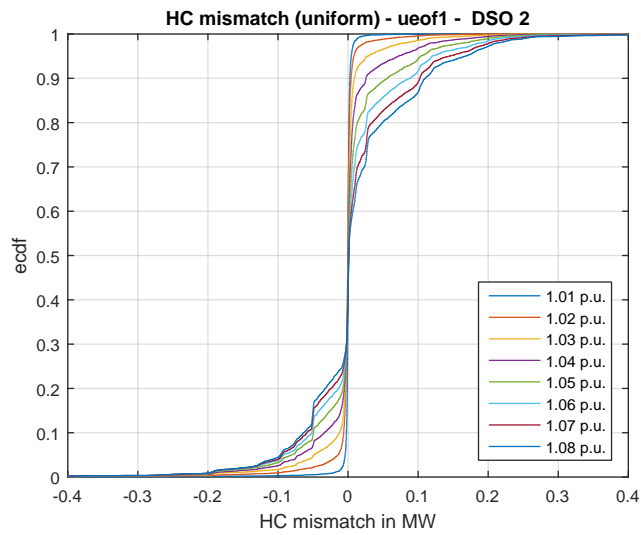
Figure 47: DSO 2, all feeders - Empirical cumulative distribution function of the hosting capability mismatch for different admissible voltage rise limits (DER-scenario *uniform*, powerfactor 1)
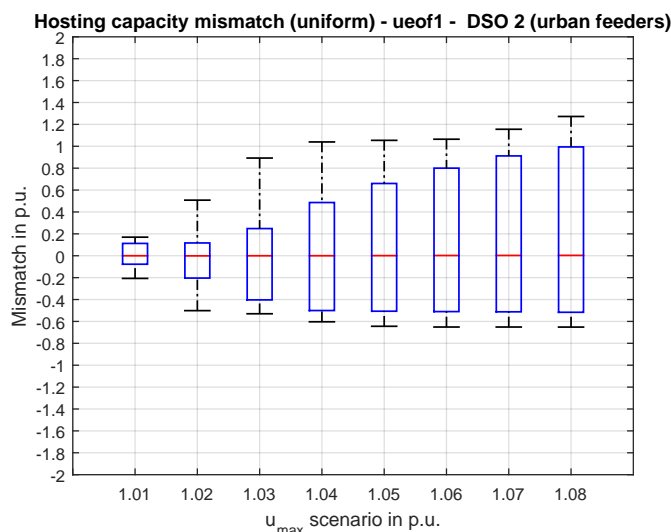
Figure 48: DSO 2 urban feeders - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 1)

for 90% of the feeders remains almost in a $\pm 20\%$ range. Furthermore, the 95th percentile for several voltage limits is lower compared to the overall HC mismatch results (Figure 46). Hence, the size of the box from the 5th percentile to the 95th percentile is smaller compared to the combined evaluation of rural and urban feeders.

The evaluations presented in this section showed, that the HC mismatch is sensitive to the investigated voltage level. Thereby the sensitivity is DSO dependent. While for DSO 1 the proposed methodology could be applied to voltage limits up to $1.04 p.u.$, this is not valid in the case of DSO 2. For DSO 2 a more conservative voltage limit of $1.02 p.u.$ or $1.03 p.u.$ (for rural feeders) may be suitable. In conclusion, the sensitivity may depend on several assumptions. Firstly, the RFT were developed to meet a HCS value for the voltage constraint case only. Hence, for a high share of loading constrained feeders, such as at the upper voltage limits, the developed RFT may not be suitable any more. Secondly, only three parameters are used to describe feeders, which leads to a loss of information of feeders. Thirdly, the parameter HCS which is used to describe topologies is also voltage dependent.
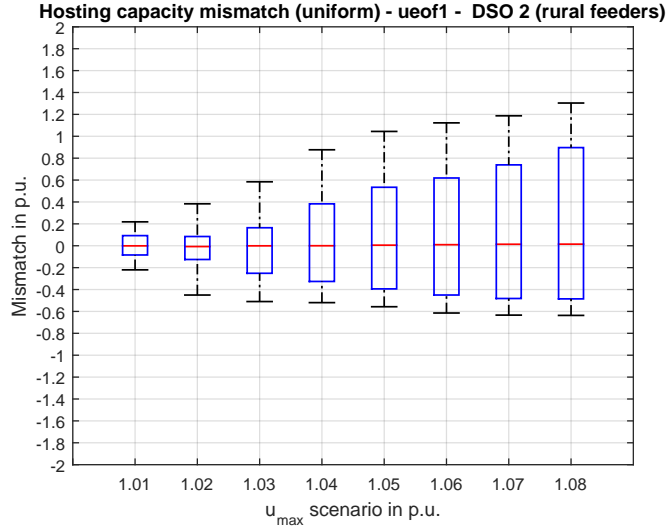
Figure 49: DSO 2 rural feeders - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 1)

6.2 Hosting Capability Sensitivity Dependency Analysis

In this section, the voltage sensitivity of the parameter HCS is investigated. The parameter HCS is obtained for the voltage limit of $1.01p.u.$ where most of the feeders are voltage constrained. In this section, the agreement of HCS values for different voltage limits (ueof2,ueof3, ..., ueof8) is analyzed. Moreover, for ueof4 and ueof8, the HC mismatch between real feeders and reference feeders is evaluated.

The parameter HCS can be calculated for other voltage limits as well. Therefore the HC results for the DER-scenarios *uniform* and *eof* obtained for a particular voltage limits are required. In Figure 50, the ratios obtained at other voltage limits (ueof2 to ueof8) are compared to the parameter HCS (ueof1). At $1.02p.u.$, for more than 90% of the feeders ueof1 and ueof2 are equal. Therefore, HCS is still accurate for a $u_{max} = 1.02p.u.$ However, for higher voltage limits, the parameter HCS does not match the ueof value of the particular voltage limit. In case of $u_{max} = 1.08p.u.$ for example, 50% of the feeders have a different HCS than a ueof8 value. Hence, another RFT would be selected. Referring to Figure 23, more than 50% of the feeders are loading constrained for both the DER-scenarios uniform and eof at $u_{max} = 1.08p.u.$
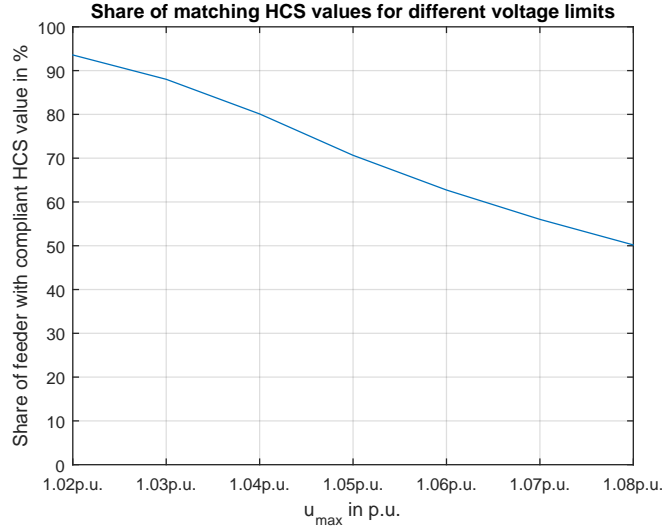
Figure 50: DSO 1 and DSO 2 - Share of calculated hosting capability sensitivity values evaluated at particular admissible voltage rise limits (2% to 8%, ueof2 to ueof8) that are identical to the hosting capability sensitivity at 1% voltage rise (ueof1)

Hence, the ratio ueof1, where nearly all feeders remain voltage constrained is dissimilar to ueof8, where a significant share of feeders is loading constrained. Furthermore, due to the simplification of feeder topologies and feeder parameters there is an impact on the accuracy as well.

The ecdf of the continuous parameter HCS is depicted for all considered voltage limits in Figure 51 for DSO 1 (PF=1). The continuous values, without rounding to quarters, were obtained from the real HC of feeders. Moreover, increasing the voltage limits leads to a higher share of the ueof value of 1 (Topology A1). This means, that A2 topologies (Ratio 1.25) are transformed to the topology A1 at higher voltage limits. The ueof ratio near the value 2 (Topology B1) are rather constant and are not changing. However, above 2.75 (Topology B4), again a small voltage sensitivity is observed. However, due to the rounding to quarters, insignificant changes do not affect the RFT of a feeder.

Figure 52 demonstrates, that in case of DSO 2 the HCS dependency on the voltage limit is higher. First of all, a higher share of the topology A1 (1) is observed (almost 25% of all feeders) if ueof8 would be considered instead of ueof1. For increasing voltage limits, a few reference topologies become more dominant (e.g. between 1.75 and 2, or 2.5 and 2.75). Consequently,
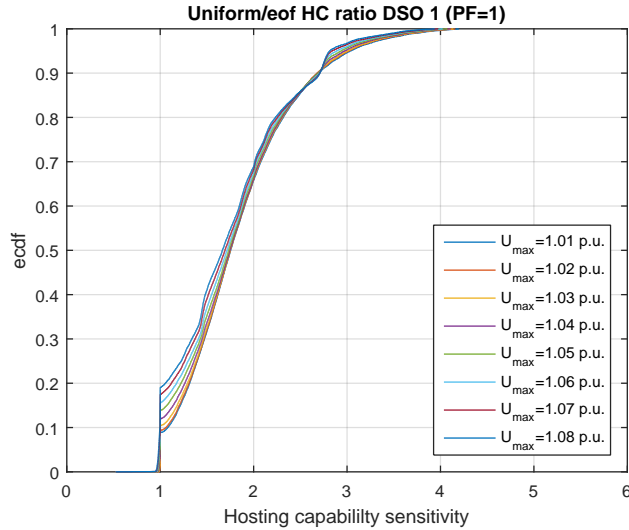
Figure 51: DSO 1 - Dependency of the hosting capability sensitivity on the admissible voltage rise (powerfactor 1)

a higher sensitivity of the parameter HCS is observed. Topologies above 3 occur more often (10%) compared to DSO 1 (5%).

Next, the ratio ueof is presented with a fixed PF-control (PF=0.9 (ind.)) in Figure 53. Compared to the uncontrolled case, more feeders can be described with the reference topology A1 (1) for higher voltage limits (about 23% instead of 19%). Further, for ratios between 2.5 and 2.75, a difference can be noted: the ecdf becomes steeper, meaning that nearby values can be grouped to one RFT. Moreover, the distance between the lines is higher compared to the uncontrolled case. Hence, rounding to the nearest quarter is less suitable for reducing the dependency of the parameter HCS on the powerfactor.

Activating a fixed reactive power control (PF=0.9 (ind.)), in the case of DSO 2 (Figure 54) leads to a even steeper ecdf-curve. The higher the voltage limit, the steeper the ecdf becomes. This means, that most of the feeders can be described by a small number of reference topologies (compared to Figure 52). While the shares at ueof=1 do not change significantly, the difference between ueof2 and ueof3 are significant. In conclusion, a significant dependency of the parameter HCS on the voltage limit under a PF-control scheme is observed.

After revealing the dependency of the parameter HCS on the investigated voltage limit, the question arises if the ueof value and the respective RFT at
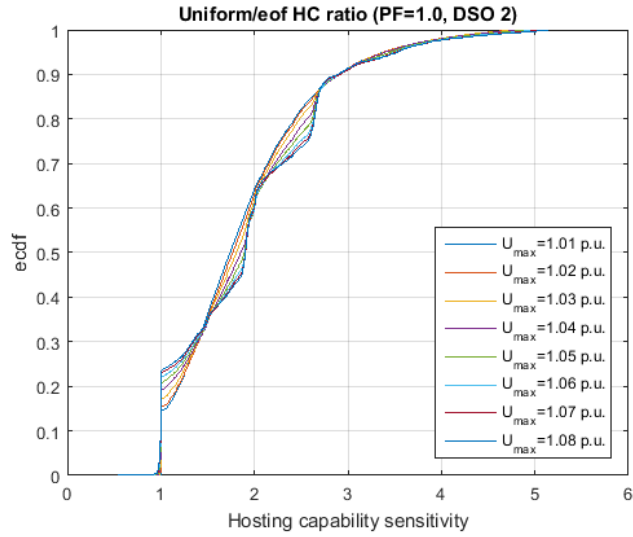
Figure 52: DSO 2 - Dependency of the hosting capability sensitivity on the admissible voltage rise (powerfactor 1)
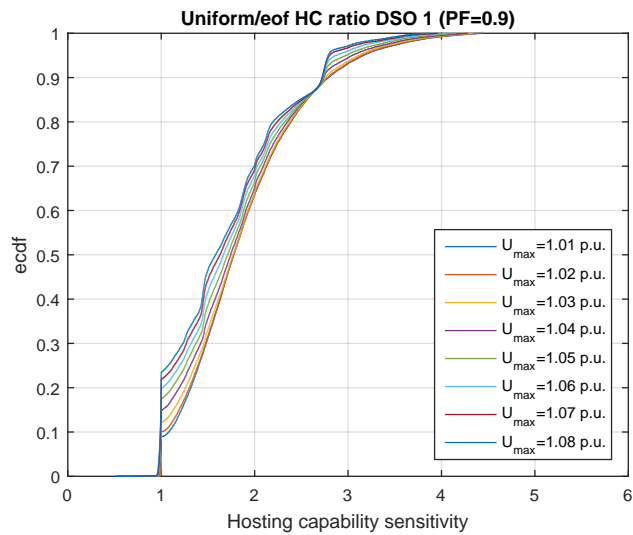
Figure 53: DSO 1 - Dependency of the hosting capability sensitivity on the admissible voltage rise (powerfactor 0.9 (ind.))
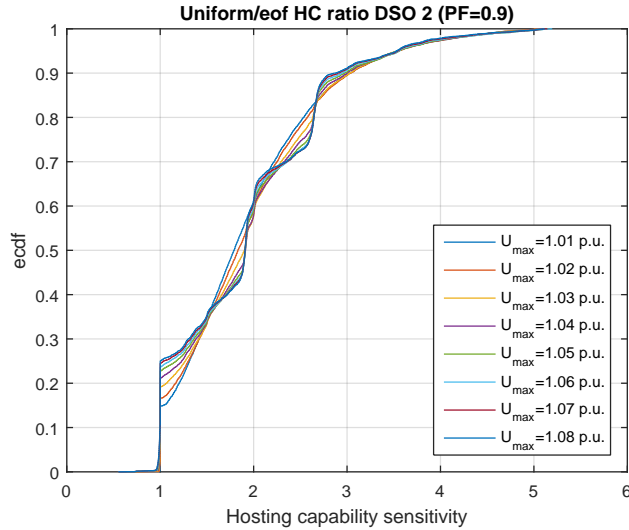
Figure 54: DSO 2 - Dependency of the hosting capability sensitivity on the admissible voltage rise (powerfactor 0.9 (ind.))

a higher voltage limit leads to a generally higher accuracy of the calculated HC of reference feeders. Therefore, a sensitivity study of the HC for different ueof values and voltage limits was performed. Thereby the RFT are not selected according to the parameter HCS values (at ueof1), but other ueof values (ueof2 to ueof8). The parameter R_k and $I_{nom-Min}$ were not changed during this analysis. In Figure 109, for example, the HC mismatch is depicted when ueof4 is utilized instead of HCS (ueof1) to select the RFT. A decrease of the HC mismatch until $1.04p.u.$ is observed, followed by a significantly higher increase at higher voltage limits . The variation of the HC mismatch for 90% of the feeders reaches a minimum at an admissible voltage rise of 4%. Compared to 106 the HC mismatch at lower voltage limits is increased. For the first and 99th percentile the HC mismatch increase is significantly higher. For an admissible voltage rise of 1% or 2%, the range of the box violates the $\pm 20\%$ range criterion for a number of feeders where the HC is underestimated. However, the HC mismatch for overestimated values is reduced at higher voltage limits significantly, while changes for underestimated feeders are negligible. E.g., at $1.08p.u.$, the 99th percentile is reduced from more than 140% to 110% and the 95th percentile from more than 100% to about 70%. The analysis demonstrates that utilizing ueof4 instead of HCS would allow the methodology to be extended to cover another 1% of the voltage band to $1.04p.u.$ in case of DSO 1.

89
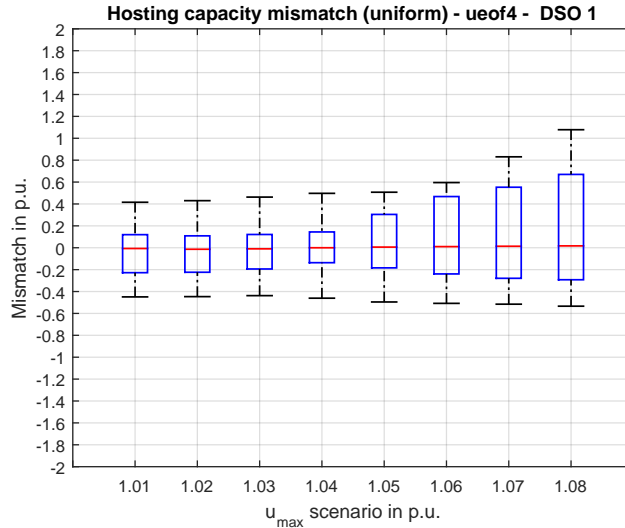
Figure 55: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 4% (ueof4) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 1)

Considering the use of ueof4, and not ueof1, moved the minimum of the HC mismatch to $1.04p.u.$, the question then arises on whether this is also true for the highest voltage limit ($1.08p.u.$). According to Figure 113, where ueof8 is utilized instead of HCS to select the RFT, this presumption cannot be confirmed. Rather a constant error over all voltage limits is observed. For almost all voltage limits the range of the box (90% of the feeders) would fit a ±40% range with an offset of 5% to 10%. The 5th and 95th percentile vary by about 10% of the full range of investigated voltage limits. Hence, the dependency of the parameter HCS as an RFT selection criterion, for the admissible voltage limit, could be significantly reduced if HCS would be determined by ueof8 instead of ueof1. Compared to Figure 106 the HC mismatch is about 4 times higher from $1.01p.u.$ to $1.03p.u.$ for the 5th and 95th percentile. At $1.04p.u.$, the HC mismatch of the 5th and 95th percentile is doubled and at $1.05p.u.$ the HC mismatch is at a comparable level. From $1.06p.u.$ to $1.08p.u.$, the HC mismatch of the 95th percentile is reduced by 10%, 30% and 50%, respectively. Hence, a trade-off between HC accuracy and voltage dependency of the parameter HCS can be achieved.

The evaluation of selecting a different measure as HCS to select the equivalent RFT was also performed for DSO 2. In Figure 117, the HC mismatch
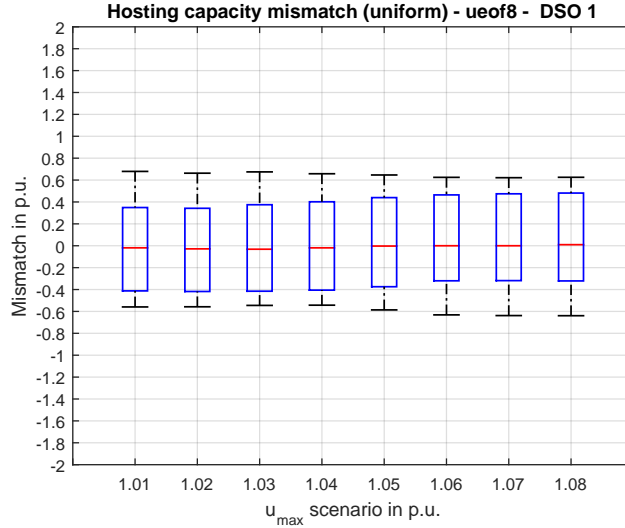
Figure 56: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 8% (ueof8) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 1)

for HCS=ueof4 is depicted for the investigated voltage limits. The trend observed for DSO 1, that the minimum of the HC mismatch is moved to $1.04p.u.$, can only be confirmed for the 99th percentile. While the HC mismatch of the 1th and 5th percentile is rather stable for the first 4 voltage limits. The HC mismatch of the 95th percentile is increasing from about 40% at $1.01p.u.$ to little more than 60% for the highest investigated voltage limit. Compared to Figure 114, the HC mismatch at higher voltage limits can be significantly decreased. For example, at $1.07p.u.$ and $1.08p.u.$, the HC mismatch of the 95th percentile is reduced by 20% and 30% and also the mismatch of the 5th percentile is reduced by about 10%. Unfortunately, the HC mismatch is increased for the first, 5th, 95th and 99th percentile. Furthermore, the HC mismatch is also significantly higher compared to DSO 1, when using the same approach (Figure 109).

Finally, the HC mismatch caused by considering ueof8 instead of HCS to select the RFT is depicted in Figure 121. As already observed for DSO 1, a rather stable HC mismatch for 90% of the feeders is obvserved. An equilibirium of the HC mismatch between utilizing ueof1 and ueof8 is reached at an admisible voltage rise of 5%. Above 5%, the HC mismatch is lower compared to Figure 114. Unfortunately, the HC mismatch is significantly
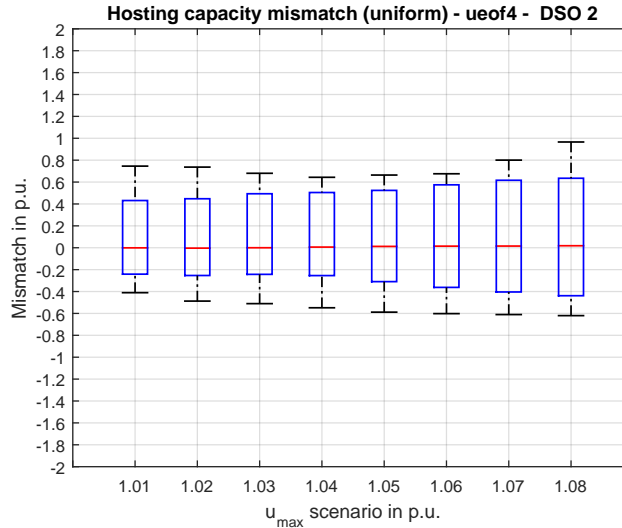
Figure 57: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 4% (ueof4) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 1)

higher at lower voltage limits. The range from the 5th percentile to the 95th percentile reaches 100% with the trend that the HC of feeders is more overestimated than underestimated.

In conclusion, it was shown, that the identified parameter HCS is voltage dependent. This voltage dependency has an effect on the HC mismatch, however the results in this section showed, that a trade-off between accurate results for lower voltage limits and equal mismatch over all limits can be partly performed. Depending on the DSO, another measure than HCS=ueof1 could be suggested for specific investigations. Nevertheless, the favorable HC accuracy for lower voltage limits ($1.01p.u.$ to $1.03p.u.$) justify the selection of ueof1 as a measure for selecting the RFT. The full analysis for different ueof parameters for both DSOs can be found in appendix E and F, respectively.

6.3 Extension of the Methodology for a Fixed Powerfactor-Control

Until now, the validation of the HC accuracy was only performed for active power in-feed and not for reactive power control strategies. Nevertheless, the dependency of the parameter HCS on the reactive power control strategies
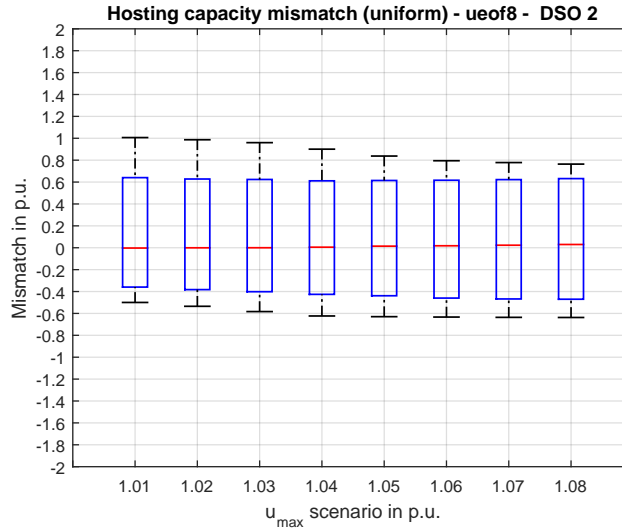
Figure 58: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 8% (ueof8) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 1)

was already demonstrated in the previous section. The developed methodology with the identified parameters to select and parametrize a representative feeder may suggest that the methodology can be easily extended by introducing the reactance as a fourth parameter. In this section the HC results with a fixed PF=0.9-control (ind.) of the real feeders and their respective representative feeders is validated for the investigated voltage levels. Aim of the validation is to clarify if the developed methodology and reference topologies are suitable for the study of reactive power control strategies as well. This would be of course an additional benefit of the proposed methodology. In analogy to the end node resistance R_k, the end node reactance X_k is considered in parameterizing a reference feeder topology. This information could be gathered simultaneously with the resistance at the end node and the used line types within the main path of the feeder. In the previous sections, only one line type was identified for each reference feeder depending on the cable with the lowest ampacity of the feeder. However, the selected line type determines the resistance and reactance per length. Since the reactance was not considered until now, every value of a given resistance at the end node could be reached for standard line types by adapting the length of the lines. Nevertheless, due to varying resistances and reactances of line

93

types, no existing line type can be found fulfilling both the resistance and reactance requirements for a reference topology with a given length at the same time. Achieving both a required resistance and reactance at the end node of the feeder with only one selected line type requires an adaptation of the methodology. The selected solution to solve this issue is for a selected line type and determined feeder length to match a required R_k at the end node the following: The reactance per length of the line type is changed to meet the desired reactance at the end node. Consequently, the line type is not a standard line type any more but a customized line type that meets the required nominal current, resistance and reactance of the reference feeder simultaneously.

Figure 59 shows the mismatch of the HC with a fixed PF=0.9-control (ind.) of the real feeders and their respective representative feeders and the investigated voltage levels for DSO 1. It is apparent, that the mismatch is rather low for 1 and 2% voltage rise. At 3% voltage rise, the mismatch for 90% of the feeders is almost within the range of $\pm 20\%$. For higher voltage limits, the trend that the hosting capacity is overestimated is obvious and for a voltage limit of $1.08p.u.$, the mismatch of the 95th percentile exeeds 100% while still for almost 50% of the feeders the HC is underestimated between 0 and -40%. Thus, the extension of the methodology to cover also reactive power control approaches is reasonably limited to 1-3% of the nominal voltage. Maintaining that the HC mismatch of 90% of the feeders are within a $\pm 20\%$ range, means a reduction of the admissible voltage rise by 1% compared to the uncontrolled case.

The ecdf of the absolute mismatch of all feeders of DSO 1 is depicted in Figure 60. Depending on the voltage limit, the HC is underestimated for about 10-20% of the feeders. For example, the HC is underestimated for a voltage limit of $1.01p.u.$ by more than 3.5kW for 10% and more than 10kW for about 2.1% of the feeders. The overestimated HC in the case of a voltage limit of $1.01p.u.$ for feeders above the 90th percentile is above 1.6kW and for less than 1% of the feeders above 10kW. Increasing the voltage limit, leads to higher over- and underestimation of the HC. The mismatch for a voltage limit of $1.03p.u.$ is overestimated for 10% and underestimated for 10% of the feeders by more than ± 10kW. Consequently, the mismatch is within a ± 10kW range for about 80% of the feeders. For the highest voltage limit of $1.08p.u.$, the HC is overestimated for about 49% and underestimated for 18% of the feeders by more than 10kW. Hence, the mismatch is lower than ± 10kW for only 33% of the feeders. It is apparent, that the referene feeder methodology is only suitable for the estimatation of the HC for a limited voltage rise from 1% to 3%.

For DSO 2, the mismatch of the HC with a fixed PF=0.9-control (ind.)

94

Figure 59: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 0.9 (ind.))
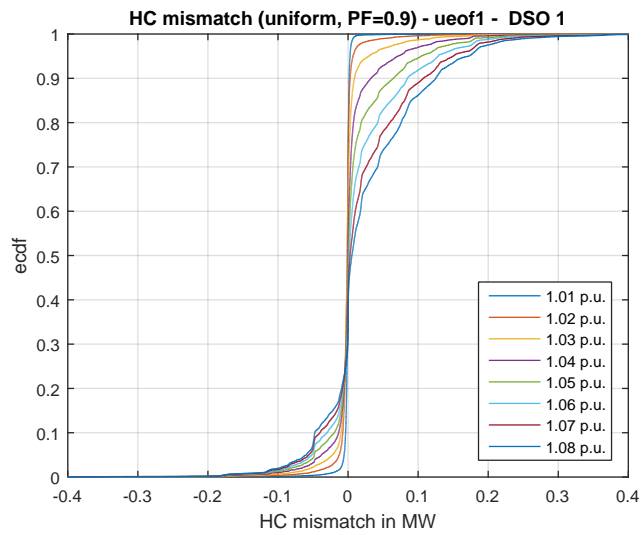
Figure 60: DSO 1 - Empirical cumulative distribution function of the hosting capability mismatch for different admissible voltage rise limits (DER-scenario *uniform*, powerfactor 0.9 (ind.))
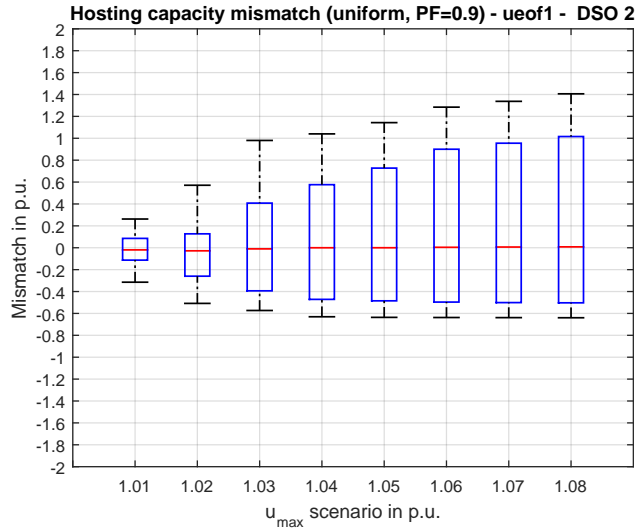
Figure 61: DSO 2, all feeders - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 0.9 (ind.))

of the real feeders and their respective representative feeders and the investigated voltage levels is depicted in Figure 61. It is apparent, that the mismatch is always higher than for DSO 1. The mismatch exceeds the ±20% range, even at an allowed voltage rise of 2% even though the median values are not significantly different compared to DSO 1. The 5th percentile and the 95th percentile are lower for most of the boxplots compared to DSO 1. At an admissible voltage limit of 2% the range of the box is between −26% and +13% and in total below 40% points. Therefore an offset could be introduced to remain within a range of ±20%. In conclusion, the proposed methodology would only be suitable for a allowed voltage rise of 1% considering both rural and urban feeders. The range could also be extended to $1.02p.u.$ due to the performed investigations for this DSO by introducing an offset of about 5%.

The ecdf of the absolute mismatch of all feeders of DSO 2 is depicted in Figure 62. For a voltage limit of $1.01p.u.$, the hosting capability is underestimated by more than 5.3kW and overestimated by more than 2.2kW for about 10% for each of the feeders. The hosting capability is overestimated by more than 10kW for less than 2% of the feeders while for 3.5% of the feeders the HC is underestimated by more than 10kW. For an allowed voltage rise of 3%, the HC is overestimated by more than 10kW for 15% of the feeders and
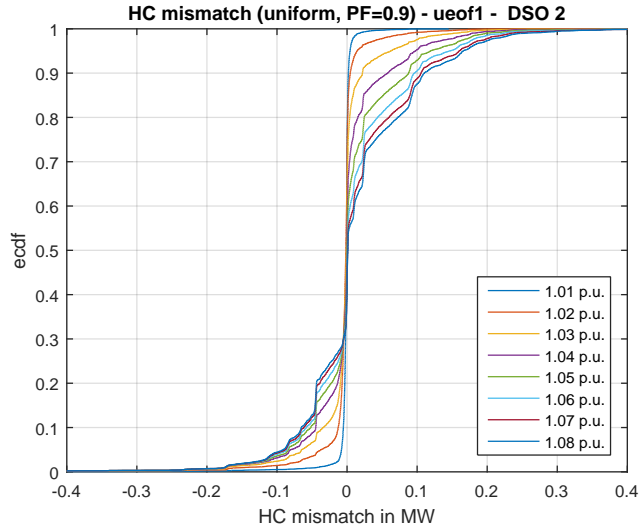
Figure 62: DSO 2, all feeders - Empirical cumulative distribution function (ecdf) of the hosting capability mismatch for different admissible voltage rise limits (DER-scenario *uniform*, powerfactor 0.9 (ind.))

underestimated by more than 10kW for 19% of the feeders. Hence, for 66% of the feeders the mismatch remains within a ±10kW range. The share of feeders within a ±10kW range for a voltage limit of 1.02*p.u.* is 80%. In case of DSO 1, this share was reached at a higher voltage limit (1.03*p.u.*). For the highest voltage limit (1.08*p.u.*), the HC is underestimated by more than 10kW for about 27% of the feeders and overestimated by 10kW for 54% of the feeders. Consequently, only 29% of the feeders remain within a ±10kW range.

When separating the feeder results of rural and urban feeders of DSO 2, differences can be observed at several voltage levels. The HC mismatch for rural feeders is depicted in Figure 63. Since most of the feeders are classified by the DSO as rural, no significant changes can be observed for rural feeders. Only for the 95th percentile at the voltage limits 3% to 5%, a slightly lower value is reached.

In Figure 64, the HC mismatch for urban feeders of DSO 2 is depicted. Compared to the previous Figure 63, the inaccuracy is higher for several voltage limits. For the three highest voltage limits (1.06*p.u.* to 1.08*p.u.*), a similar mismatch can be observed for both rural and urban feeders while the 99th percentile for rural feeders is even higher compared to urban feeders. However, for lower voltage limits, the range of the box (5th percentile to
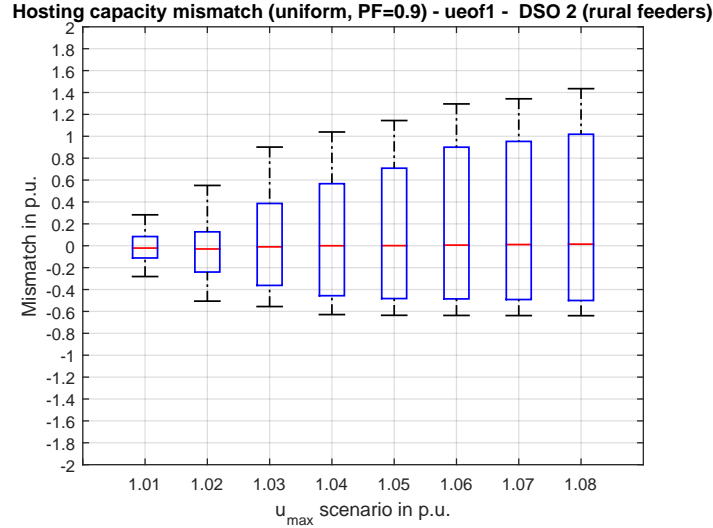
Figure 63: DSO 2 rural feeders - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 0.9 (ind.))

95th percentile) for urban feeders is between 10% (1.05*p.u.*) to 20% (1.03*p.u.*) higher when compared to rural feeders. Hence, the extension of the voltage band for urban feeders cannot be suggested, since the range of the box still exceeds ±20%.

The extension to cover also reactive power control strategies is limited when compared to active power in-feed only. The limited accuracy is caused by several issues. Firstly, with the presented approach, only one line type (with a given R and X) can be considered. Hence, if a feeder contains both overhead-lines and underground cables, the R to X ratio at the end node is utilized to parametrize the feeder. Thus, this has an impact on the effectiveness of reactive power control strategies. In particular the effectiveness of installations near the transformer station may be overestimated if only the last section of feeders consists of overhead-lines. To consider different effective R/X ratio at nodes in the feeder, more input parameters for the reference feeder methodology would be needed. For example, the share of overhead-lines and underground cables in the feeder and their respective lengths. Secondly, only one line type is defined which is applied all over the feeder. This may result in feeders with several branches to an additional HC mismatch. The defined RFT lead to a loss of information of the real feeder topology and also the number of connected loads which also affects the HC
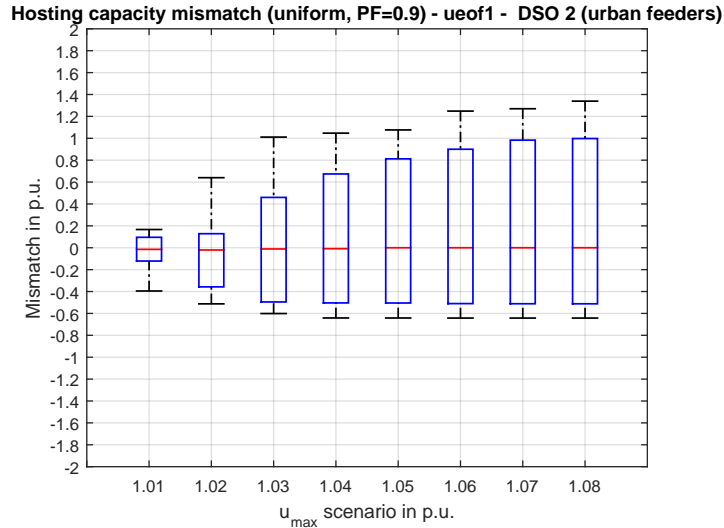
Figure 64: DSO 2 urban feeders - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 0.9 (ind.))

of the representative feeder. The analysis in this subsection shows that the performance is depending on the DSO and applicable to a limited voltage range compared to active power in-feed only. In case of DSO 2, only at 1% voltage rise, the HC mismatch remains within ±20% for 90% of the feeders, while for DSO 1, this criterion is satisfied for an admissible voltage rise up to almost 3%.

6.4 Methodology Performance for External Network Data

In this section, the developed methodology is applied to a number of available LV-networks. In the project U-Control, the defined reference LV-networks in [36] were updated and a new set of representative reference networks were defined. Thereby 358 LV-networks with 1.550 LV-feeders from different regions of Germany were considered. The updated reference networks are presented in [53]. These reference LV-networks contain in total 26 radial feeders without branches and were selected for voltage control studies in a supply area in the south of Germany. The secondary transformers in these networks were replaced by a slack and the methodology presented in this work was applied to each feeder of the LV-networks. For the presented LV-feeders

in [53], only two $I_{nom-Min}$ values occur. This is caused by the consideration of the service line in the network models, with a rather small cable cross section. The corresponding $I_{nom-Min}$ value for the cable cross section of $NAYY - 4x35mm^2$ is *119A* (observed in 19 feeders). For the cable cross section of $NAYY - 4x50mm^2$, the nominal current is *141A* (observed in 7 feeders). The comparison with the ecdf of the minimal nominal current in feeders of DSO 1 and DSO 2 (Figure 26 and 27, respectively) shows that the line type with the nominal current of *119A* is rarely observed. Whereas, the nominal current of *141A* is also observed for the investigated DSOs and significantly in case of DSO 2.

In Figure 65 the share of U/I constrained feeders is depicted for the investigated voltage limits and the DER-scenarios *uniform* (with and without a VoltVAr control) and *eof* (For the investigated feeders, no yearly consumption values were available). Compared to Figure 23, the feeders are more sensitive to a higher admissible voltage rise. Without a VoltVAr-control, at 5% voltage rise, 50% of the feeders are already loading constrained. Due to the given topologies of the feeders, no significant difference can be observed for the uncontrolled considered DER-scenarios *uniform* and *eof* . The share of loading constrained feeders in the case of activated VoltVAr-control is higher. At $1.05p.u.$, already more than 80% of the feeders are loading constrained, while this value was not reached in Figure 23 even for 8% admissible voltage rise.

In Figure 66, the share of the parameter HCS is depicted. More than 50% of the feeders can be represented by the RFT B1. Topology B1 is described by a branched feeder with a feeder segment I and II. Further, 15% of the feeders can be described by RFT A1 which matches the number of dedicated feeders in [53]. Compared to Figure 32 these two topologies are overrepresented and the topology A2 does not even occur.

For the uncontrolled (PF=1) DER-scenario *uniform*, the HC mismatch between real HC and estimated HC with the reference feeders is depicted in Figure 67. The trend of an increased mismatch for higher voltage limits is clearly visible. Thereby the median remains in a band of -5% to +10%. For lower voltage limits, the HC is rather underestimated than overestimated. Compared to Figure 44 and 46, the increase of the mismatch for higher voltage levels is smaller. However, the median of the HC mismatch increases significantly. Nevertheless, at $1.04p.u.$, the mismatch is rather symmetrically distributed within a $\pm20\%$ range for 90% of the feeders. Hence, a suitable mismatch range for half of the investigated voltage range is oberserved.

In Figure 68, the HC mismatch for the DER-scenario uniform with an activated PF=0.9-control (ind.) is depicted. For a voltage limit of $1.01p.u.$, the mismatch is within a $\pm10\%$ range and remains in that range also for a
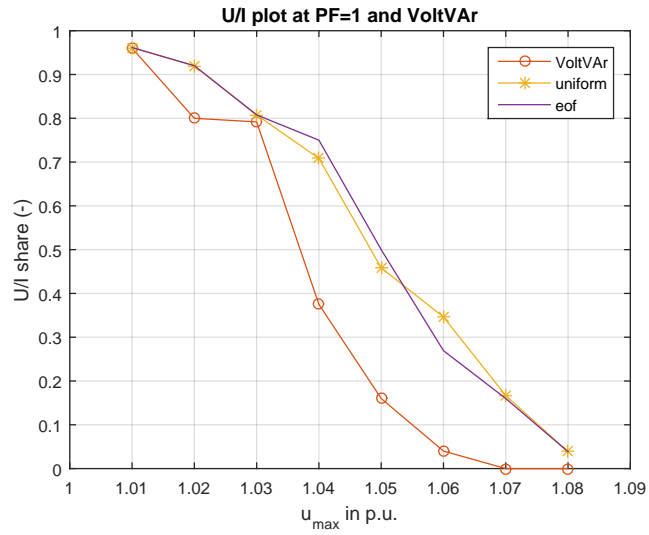
Figure 65: UC feeders - Share of U/I constrained feeders as a function of the admissible voltage rise for the investigated DER-scenarios with a powerfactor of 1 and for the DER-scenario uniform with a VoltVAr-control
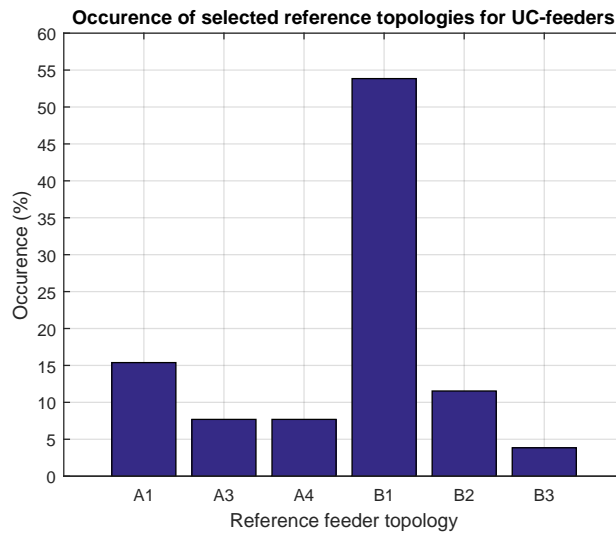
Figure 66: UC feeders - Distribution of the hosting capability sensitivity
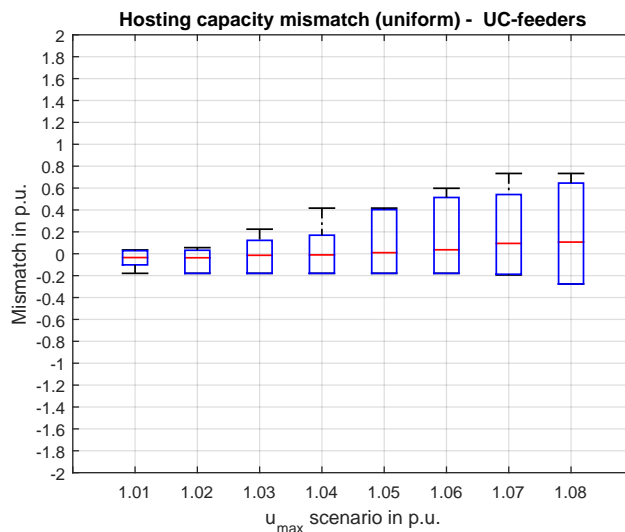
Figure 67: UC-feeders - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 1)

voltage rise of 2%. However, after that, the HC is overestimated by 40% for the 95th percentile with an admissible voltage rise of 3%, whereas at the same time, the error for 50% of the feeders is within 0 and -10%. A similar boxplot is observed for the next higher voltage level. However, after that, the median HC mismatch value increases to 20% and further to almost 40%. It is apparent, that for loading constrained feeders, the mismatch becomes significantly higher according to Figure 65. In conclusion, at higher voltage levels the HC is overestimated for most of the feeders. Thereby the range from the 5th percentile to the 95th percentile ranges between 50% (at $1.03p.u.$) and 100% (at $1.08p.u.$). Hence, even introducing an offset to reduce the maximum error would still result in an error range of $\pm 50\%$. Therefore the proposed methodology is with an activated PF=0.9-control (ind.) only suitable for an admissible voltage rise up to 2%.

In conclusion, the performance for the presented and investigated external data set is comparable to the results of DSO 1. In the uncontrolled case, the mismatch remains within a $\pm 20\%$ range for the first 4 voltage limits ($1.01p.u.$ to $1.04p.u.$). In the case with an activated PF=0.9-control (ind.), the methodology is only suitable for a voltage limit up to 2% of the nominal voltage.

Hosting capacity mismatch (uniform, PF=0.9) - UC feeders
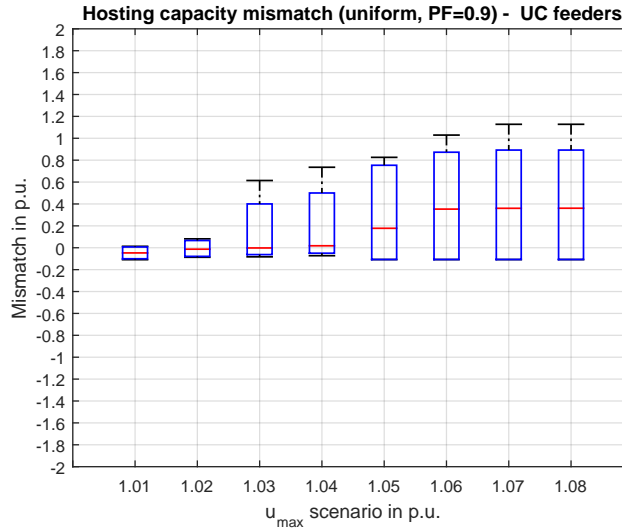
Figure 68: UC-feeders - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, powerfactor 0.9 (ind.))

6.5 Monte-Carlo simulations for real and reference feeders

In [62], about 250 feeders were investigated and classified by electrical and non-electrical parameters on feeder and network level. For each cluster, a median feeder/network were found and discussed. In this section, the median feeders found by clustering the feeders, according to electrical parameters, are considered. For each median feeder, Monte-Carlo simulations were performed and repeated for their individual equivalent feeders. A Monte-Carlo simulation with 1000 iterations was performed where the active power consumption of each load was defined randomly between 0 and 5kW and the reactive power between -2kVAr and 2kVAr. After linking the loads of real feeders with the loads of the respective equivalent feeder, the simulation of the RFT were performed with the consumption data used in the Monte-Carlo simulation. In Table 9 the input data for selecting and parameterizing the RFT feeders for the median feeders presented in [62] is listed. It is apparent, that two RFT topologies (A1 and B1) are selected for more than 2 feeders. However, the remaining input parameters are significantly different, which justifies the investigation of all listed feeders.

During the Monte-Carlo simulations, for each load the active and reactive

103

Table 9: Reference feeder topologies and parameters of investigated feeders

	RFT	R_k	$I_{nom-Min}$	R/X
		(Ω)	(kA)	(1)
Feeder 1	B1	0.1213	0.098	2.1393
Feeder 2	A1	0.0265	0.363	1.0584
Feeder 3	B3	0.1157	0.098	2.631
Feeder 4	B1	0.2303	0.098	1.726
Feeder 5	A2	0.07715	0.098	2.6719
Feeder 6	A4	0.23053	0.098	2.6468
Feeder 7	A1	0.03641	0.128	1.7696
Feeder 8	A3	0.23994	0.098	2.584
Feeder 9	A1	0.12897	0.098	2.0459

power was randomly chosen. Thereby, the range for active power was defined between 0 and 5kW and -2kVAr to 2kVAr for reactive power. Therefore, the range of active and reactive power consumption of feeders was determined by the number of loads per feeder. In Table 10 the resulting range of active and reactive power exchange of feeders with the upstream network is listed. For feeders with the a high number of loads (Feeder 1, 3, 4, 5, 6, 8), a minimal load between 29kW and 75kW is observed. For three feeders with only a few loads (Feeder 2, 7, 9) the lower limit remained between 0 and 1kW. The highest active power consumption was observed for Feeder 6 (130kW).

Table 10: Test range of active and reactive power

	RFT	Pmin (kW)	Pmax (kW)	Qmin (kVAr)	Qmax (kVAr)
Feeder 1	B1	32	90	-18	22
Feeder 2	A1	0	18	-7	7
Feeder 3	B3	29	65	-15	20
Feeder 4	B1	65	115	-20	22
Feeder 5	A2	42	87	-15	18
Feeder 6	A4	75	130	-21	28
Feeder 7	A1	1	15	-5	5
Feeder 8	A3	40	82	-15	17
Feeder 9	A1	0	10	-3.5	4

The accuracy of the results of equivalent feeders for the feeders 1 to 9 is summarized in Table 11. For each feeder three plots are depicted: The results of the active power losses, the maximal loading of lines and the minimal voltage of the feeder. For five of the feeders, the absolute error of the

active power losses is very low and can be compared with the losses of the real feeders. Even though Feeder 1 and Feeder 4 have the same RFT (B1), the accuracy of the losses are significantly different. The loading of lines is only underestimated for feeder 6, meaning that the loading of the highest loaded line in the real feeder is higher compared to the equivalent feeder. In the case of 5 feeders, the mismatch is below 10%. The highest accuracy is reached, of course, for the RFT A1 (A feeder with one line, node and a load). The estimation of the minimal voltage is relatively accurate for most of the feeders. For two feeders (Feeder 4 and 6), the minimal voltage is underestimated, meaning that the voltage in the real feeder is higher compared to the minimal voltage in the equivalent feeder. Related plots of the Monte-Carlo evaluation can be found in appendix H.

Table 11: Accuracy of equivalent feeders

	RFT	Losses (kW)	Max. loading	umin
		Absolute error	Absolute error	Accuracy
Feeder 1	B1	< 0.6	< 10%	good
Feeder 2	A1	< 0.001	< 0.001%	good
Feeder 3	B3	< 0.9	< 15%	good
Feeder 4	B1	overestimated	< 30%	underestimated
Feeder 5	A2	overestimated	overestimated	good
Feeder 6	A4	overestimated	underestimated	underestimated
Feeder 7	A1	< 0.001	< 0.001%	good
Feeder 8	A3	overestimated	overestimated	good
Feeder 9	A1	< 0.001	< 0.001%	good

6.6 Identification of Reference Networks

In this section, the distribution of the identified RFT in the real networks is analyzed. Thereby, the feeder parameters R_k and $I_{nom-Min}$ are neglected. Since, the study was performed on feeder level, networks with a number feeders with a lack of data were not completely discarded, but only the feeder with missing data. Hence, out of the 54,000 feeders, 36,723 feeders were analyzed, meaning that about 30% of the feeders were discarded. Hence, a significant share of feeders was discared without obtaining the parameter triple HCS, R_k and $I_{nom-Min}$. Therefore, the number of feeders per network is underestimated due to the missing data of these feeders.

The first question that arises is, what combination of topologies are present in the networks and how often certain topologies can be found. In
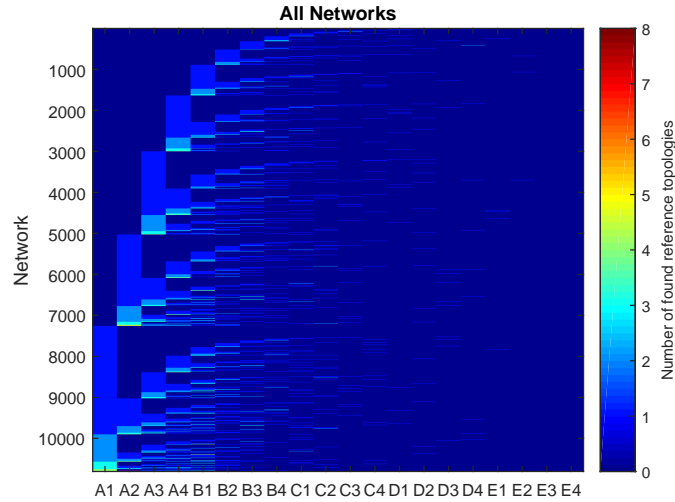
Figure 69: DSO 1 and DSO 2 - Heat map of found reference feeder topologies (RFT) in real networks, colored with the number of occurrences within the same network. The electrical parametrization of the RFTs is disregarded

Figure 69 the identified reference topologies per network are depicted for both DSOs. Each line in the colormap corresponds to a network and the color of each segment indicates the number of feeders of a particular reference feeder topology. In total about 11,000 networks are depicted and the highest number of feeders of one specific reference feeder topology is eight. The networks are sorted by the occurrence of reference topologies starting from A1, A2, to E4. Thereby certain configurations are found several times.

In a next step, the unique configurations shown previously in Figure 69 were obtained. These unique configurations are depicted in Figure 70. For example, about 45% of the networks contain at least one reference topology of type A1. In this way, about 3,100 RFT combinations remain. Thus, each of these 3,100 networks is representative in average for 4 to 5 networks of the investigated 14,000 networks. Hence, 3,100 unique networks are not adequate to be defined as representative networks. Even though the parameters R_k and $I_{nom-Min}$ as well as the secondary transformer was neglected in the analysis, such a high number of unique networks remain. Therefore, a further simplification is necessary to identify reference networks from a topological point of view.
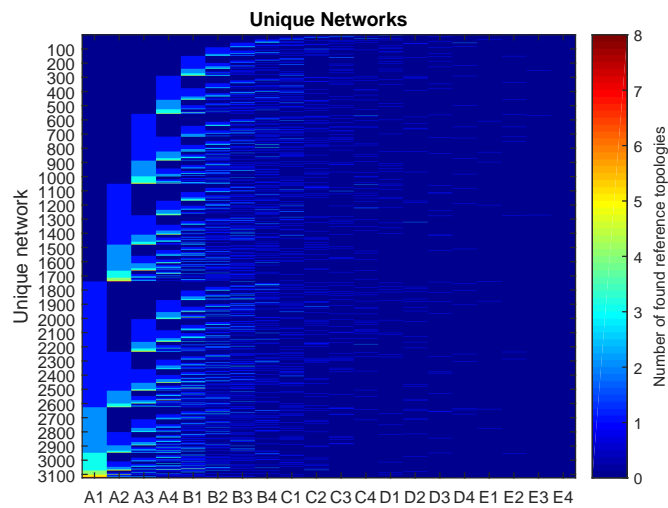
Figure 70: DSO 1 and DSO 2 - Heat map of real networks with an identical reference feeder topology combination, colored with the number of occurrences within the same network. The electrical parametrization of the RFTs is disregarded.
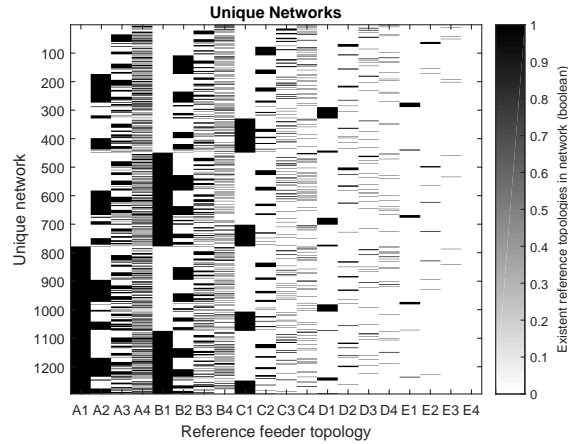
Figure 71: DSO 1 and DSO 2 - Heat map of real networks containing the same reference feeder topology (RFT) combination, where each RFT is at least one time present or not (Boolean). The electrical parametrization of the RFTs is disregarded.

The number of unique networks can be further reduced using the following justification: since secondary transformers are discarded and replaced by a slack, the number of feeders in a network can neither lead to overloading the grid nor to influence results of other connected feeders. Consequently, several feeders with an identical parameter triple HCS, R_k and $I_{nom-Min}$ could be studied by one feeder of that kind. Moreover, if only the RFT is investigated, the reduction of feeders with the same HCS is valid as well. Based on this assumption, the combination of reference topologies in Figure 70 can be further reduced by neglecting how often a certain topology occurs in networks. The analysis whether a RFT combination was identified in a network or not is depicted in Figure 71. Therefore a boolean variable is used to indicate whether a certain RFT is present in a network or not. The networks in this plot are sorted by first A1, B1, C1, D1 to E1. By applying this assumption, the number of reference networks can then be reduced to 1,300 networks. Hence, on average each of the depicted RFT combinations represent about 10 networks. As already discussed in section 5.2, topologies of group D and E are rarely present. Hence, even with the introduced simplifications, a rather low number of representative networks cannot be given for the entire set of networks.

If all RFT combinations are considered, millions of combinations would be possible. Presumably only a limited number would be relevant however. A further reduction of representative networks can be achieved by counting

108

reference topology groups instead of each reference topology. If, for example, reference topologies from A1 to A4 are counted as group A and respectively for the other reference feeder topology groups, 5 reference topology groups remain. Theoretically, 31 combinations would be possible. In Table 12, only occurring combinations of identified reference feeder topology groups (A, B, C, D, E) in the real feeders is listed. The aggregation of the reference feeder topologies (e.g. A1, A2, A3 and A4, etc.) to the corresponding group (A, etc.), leads to 25 combinations. How often these combinations occur is analyzed next.

Table 12: Reference feeder topology group combinations (e.g. A1 to A4 are considered as group A)

Reference Network	A	B	C	D	E
1	0	0	0	0	1
2	0	0	0	1	0
3	0	0	1	0	0
4	0	0	1	0	1
5	0	0	1	1	0
6	0	1	0	0	0
7	0	1	0	0	1
8	0	1	0	1	0
9	0	1	1	0	0
10	0	1	1	0	1
11	0	1	1	1	0
12	1	0	0	0	0
13	1	0	0	0	1
14	1	0	0	1	0
15	1	0	1	0	0
16	1	0	1	0	1
17	1	0	1	1	0
18	1	1	0	0	0
19	1	1	0	0	1
20	1	1	0	1	0
21	1	1	0	1	1
22	1	1	1	0	0
23	1	1	1	0	1
24	1	1	1	1	0
25	1	1	1	1	1

In Figure 72, the share of occurring RFT groups is depicted. It is appar-

109
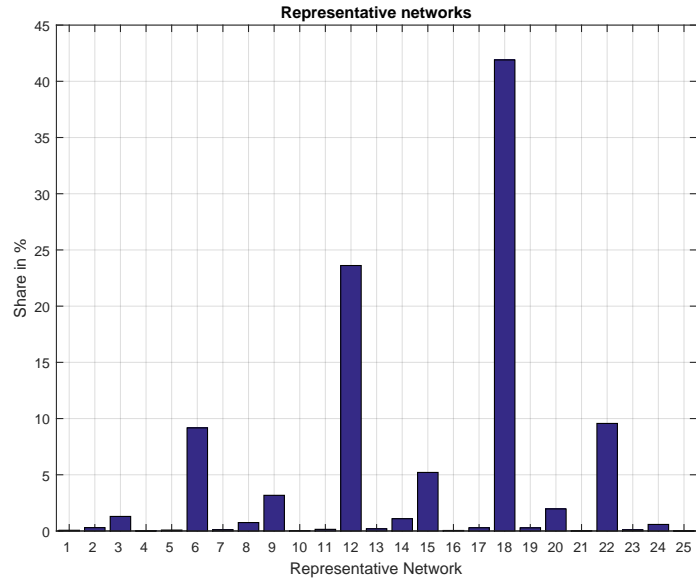
Figure 72: DSO 1 and DSO 2 - Bar plot of identified representative networks

ent, that 5 combinations of RFT with a share higher than 5% are observed. For example, representative network 6 contains only reference topologies of group B and is representative for almost 10% of the networks of both DSOs. The share of representative network 12 reaches almost 25% among all networks and contains only reference topologies of group A. The next network with a share higher than 5% is network 15. In such networks, reference topologies of group A and C are present. The network with the highest share is network 18. In more than 40% of the real networks, reference topologies of group A and group B are present. Last but not least, the representative network 22 with a share of 10% contains reference feeder topologies of group A, B and C.

The number of feeders per network for each DSO is depicted in Figure 73. In this plot, only feeders which passed the plausibility tests and were validated, are considered. Since 30% of the feeders were discarded, findings based on this Figure have to be read carefully. The Figure shows, that for a number of feeders higher than 3, similar shares occur for both DSOs. However, in case of DSO 1, about 11% and in case of DSO 2 nearly 16% of the LV-networks supply a dedicated feeder. Furthermore, about 50% of the LV-networks supply two or three feeders in case of DSO 1 and 44% of the LV-networks supply two or three feeders in case of DSO 2. The share for
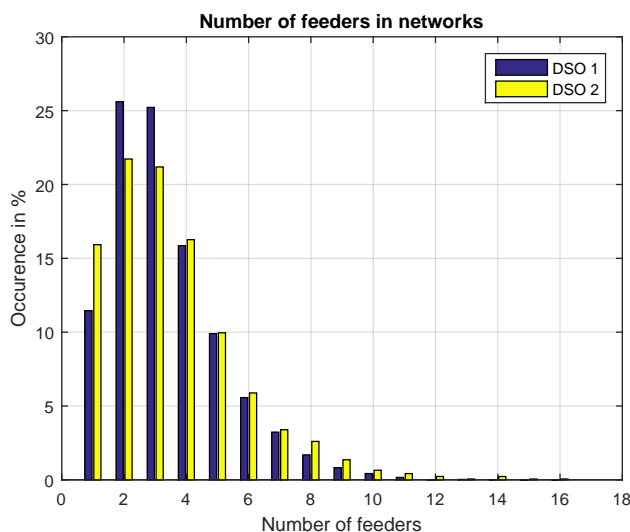
Figure 73: DSO 1 and DSO 2 - Distribution of the number of feeders in networks. Blue: DSO 1, Yellow: DSO 2

both DSOs is thereby almost equal. The share of LV-networks, supplying more than 3 feeders is decrasing significantly. For both DSOs, 16% of the LV-networks supply 4 feeders and about 10% of the feeders supply 5 feeders. Finally, only 6.2% (DSO 1) or 8.4% (DSO 2) of the LV-networks are supplying more than 6 feeders.

The presented data in this section shows that a high variation of reference feeder topologies can be observed within networks. Thereby, the parametrization of the RFT was neglected as well as how often certain RFT occur within the same network. Nevertheless, 5 probable RFT group combinations with a probability of occurrence above 5% were found. To generate representative networks, two tasks have to be performed. First, a methodology to realistically select RFTs of a RFT group for a representative network and second the parametrization of these RFTs. Hence, the definition of a set of fully parametrized reference networks requires further investigations and a validation of course. However, considering all available data in this work a methodology to define such a network generator can be defined and validated in the future.

7 Conclusion

On the basis of the LV-grid data of two DSOs, a methodology was developed and validated to select equivalent representative feeders for particular voltage limits. In total, 14,000 LV-networks with about 54,000 feeders were imported into DigSILENT PowerFactory to perform load flow studies. After plausibility tests (e.g. complete line data, radiality of feeders), about 36,720 feeders were found to be suitable for further investigations.

The hosting capability (HC) was defined as DER-penetration level that leads to either a voltage or loading constraint in a feeder. Furthermore the ratio of two specific DER-scenarios (*uniform* and *eof*) was introduced as a topological measure.

The HC was calculated for the DER-scenarios *uniform* and *eof* for 8 admissible voltage limits ranging from 1 to 8% of the nominal voltage. Moreover, the HC was evaluated with and without reactive power control strategies. In particular, the VoltVAr-control and the fixed PF-control were investigated, in addition to the uncontrolled case for real and their simplified feeder topologies.

Due to the significant HC increase from $1.01 p.u.$ to $1.02 p.u.$, an admissible voltage rise limit of 1% cannot be recommended at all. The HC results showed, that increasing the admissible voltage rise for in-feed is only reasonable if a high share of feeders remain voltage constrained. If an admissible voltage rise of 4% is feasible, for the DER-scenario *uniform* more than 50% of the feeders have a HC higher than 100kW. With a VoltVAr-control, even a median value of 120kW can be reached. In the worst case scenario (DER-scenario *eof*) significantly lower median values are achieved. Moreover, extending the voltage band for loading and utilizing a reactive power control strategy leads to a significant HC reduction for loading constraint feeders.

To find representative feeders that are suitable to estimate the HC of real feeders, a high number of parameters were investigated. It was demonstrated that statistical parameters, which are basically invariant to scenario assumptions, show a dependency on the constraint limiting the HC. For example, the correlation coefficient between the HC and the parameter $I_{nom-Min}$ increases with the admissible voltage rise. Hence, with the increasing share of loading constrained feeders, a higher correlation coefficient is obtained. Furthermore, the correlation coefficient of R_k, the HC for $u_{max} = 1.01 p.u.$ and the DER-scenario *eof* show almost a perfect correlation, due to the fact that all feeders are practically voltage constrained. Therefore the parameters $I_{nom-Min}$ and R_k were identified as the most important parameters describing feeder properties. Moreover, the probability tests showed that for both DSOs only two common parameters follow a normal distribution. These

parameters are ADTN and Z_Σ.

Even though three line types with nominal currents of 100A, 125A and 270A are widely used, the required parameter $I_{nom-Min}$ need to be provided accurately to select the matching cable type. If the exact ampacity is not provided, the error for loading constrained feeders will be too high. However, the parameter $I_{nom-Min}$ is a rather discrete parameter, whereas the resistance at the end node R_k is continuous value. Thus, a limited number of representative feeders cannot be given based on these two electrical parameters.

The parameter HCS was introduced, as the ratio of the HC of the DER-scenarios *uniform* and *eof*. Hence, HCS allows describing the topology of a feeder with an electrical measure (ratio of two HC values). For both investigated DSOs, the 8 most relevant topologies were found (A1 to A4 and B1 to B4) covering about 90% of all feeders. However, it was shown that the parameter HCS is voltage dependent.

A methodology was developed to select RFTs and also parametrize them. So once a RFT is selected, it can be configured with only two electrical parameters (R_k and $I_{nom-Min}$). The RFT are easy to understand and consist of two basic feeders structures. The RFT, determines the number of lines in parallel for each cable. Moreover, by definition only one line type is used in the RFT and all lines have the same length.

It is apparent, that design rules were observed for the RFTs. For each group of RFT (e.g., A, B etc.), a pattern can be given to modify a RFT to obtain the following RFT. Thereby, the pattern can be applied irrespective of the RFT group. Furthermore, a pattern between RFT groups (A1, B1, etc.) was observed. Therefore, well defined inter- and intra-rules for RFT groups were defined.

Thus, a comprehensive kit to select a RFT and configure reference feeders was developed. Moreover, the proposed methodology is scalable to any size in terms of number of feeders and highly adaptive (simple parametrization of the RFT). Furthermore the ecdf of the parameters R_k and $I_{nom-Min}$ were given for defined RFTs which allows to automatically build parametrized RFTs based on distribution functions.

The parametrized reference feeders were validated with the HC results of the real feeders. The validation was performed with and without reactive power provision. The reference feeders were validated first for active power flows. The analysis showed that the HC accuracy is DSO dependent. Moreover, the dependency on the admissible voltage rise was demonstrated. A voltage rise of 4% was found to be the upper limit for the estimation of the HC with a ±20% error for DSO 1. A lower limit was found in the case of DSO 2. The validation of the reference feeders showed that the HC accuracy

113

for all voltage limits cannot be reached, since a loss of information occurs due to the reduction of the full feeder model.

Thus, the proposed methodology allows the clustering of feeders according to the parameter triple HCS, $I_{nom-Min}$ and R_k. A classification based on these 3 parameters allows the parametrization of a RFT, which is suitable for the estimation the HC of real feeders for an admissible voltage rise limit. Consequently, the methodology is suitable for the substitution of real feeders with an equivalent feeder. Thereby the number of nodes is reduced as well, requiring less computation time. This is relevant if an entire grid of DSOs needs to be investigated.

Additionally, the validation was also performed for the PF-control. Therefore the reactance at the end node X_k was considered as fourth feeder parameter to parametrize a custom line type. Thereby the RFT remain unchanged. Unfortunately the accuracy observed for the uncontrolled case could not be met. As a consequence, the extension of the methodology to cover also reactive power flows is only valid for a very conservative voltage range.

The methodology was also applied to some external network data. The obtained results correspond to the results of DSO 1.

Even though some DSOs may have already or will have powerful GIS systems allowing to perform grid studies, the proposed methodology is still suitable for the replication of findings. E.g. for feeders with a similar parameter triple (HCS,$I_{nom-Min}$ and R), corresponding results can be expected.

Due to the reduced number of nodes, the RFT methodology is suitable to provide simplified feeder models for multi-voltage load flow studies. In such studies, where the entire supply area of a DSO is studied simultaneously, loadflows are performed for the LV, MV and HV grids of a DSO. Hence, simplified feeder models for LV-feeders reduce the calculation time and the complexitiy. In the SCADA systems where the medium voltage grids are modeled in detail, the simplified LV-feeders could be integrated as underlined networks, to study the impact of setpoints within the MV-grids on LV-networks and vice verca. Replacing a feeder by the equivalent reference feeder leads to a loss of information. For example, the loads in the real feeder have to be assigned to load location of the respective RFT. The closeness centrality is a suitable method to link loads of real feeders to nodes in equivalent feeders. Thereby, the Monte-Carlo simulation results show that the accuracy of results (voltage levels, loading of lines, losses) is comparable for most investigated feeders.

The development of RFT allowed the classification of feeders from a topological point of view. Unfortunately, a high number of unique network configurations remained, which could not be reduced to a small number of representative networks. The analysis of the occurrence of reference topologies

showed, that more than 40% of the networks contain only RFTs of group A and B. Moreover, almost 25% of the networks contain only RFTs of group A. Networks containing only RFTs of group B have a share of 10%, as well as networks with RFTs of group A, B and C.

In conclusion, a limited set of RFT topologies were defined and validated with the LV-grid data of two DSOs. Furthermore sensitivity studies were performed on the admissible voltage rise and the input parameter HCS. Additionally, the HC accuracy with an activated PF-control was investigated and limitations discussed.

8 Outlook

In this work, a methodology was proposed and validated to select reference feeders that are equivalent in terms of the HC of feeders under defined assumptions. Based on the results of this work, several research questions arise.

In a LV-network with several feeders, the HC of all feeder in a network may exceed the rating of the secondary transformer. However, the HC plausibility of the calculated HC values per feeder need to be checked with e.g. solar resource maps. Further studies could be conducted for the investigated DSOs, to identify such secondary substations that could be theoretically overloaded in the future if a high share of DER is expected. The HC of all feeders supplied by a secondary stations may exceed the rating of the secondary transformer. However, the overloading could be prevented by applying e.g. a VoltWatt-control to all inverters. Since secondary transformers were neglected, the methodology is also suitable for micro grids. The presented statistical analysis of the LV-grid data of the two DSOs, allows identifying the HC of the entire LV-supply area. In combination with solar resource maps, DER integration action plans could be developed while minimizing grid reinforcements costs.

Furthermore, the methodology could be included in GIS-Systems of DSOs to automatically calculate the parameter HCS to identify the RFT of real feeders.

Based on the introduced RFT and their descriptions, a hand book could be developed to support network planners and operators. Such a hand book could for example include guidelines if a voltage band extension or a smart grid solutions could be favored against grid reinforcements for particular RFT and parameters R_k and $I_{nom-Min}$.

The performance of the methodology is distinct for each of the investigated DSOs. Additionally, divergent accuracies were observed for the feeders classified as rural and urban of DSO 2. Additional investigations are required to identify these variances which are not covered by the considered input parameter triple.

The definition of the feeder end node is suitable for radial feeders as well as for meshed feeders. Therefore, another open question is if the proposed methodology performs well for meshed feeders.

In this work, secondary substations were excluded from the running arrangement. However, by considering the resistance, reactance and nominal current of the secondary transformer the methodology could be extended to cover also the secondary transformer.

The methodology is suitable to be extended easily to consider also the end

node reactance X_k. However, the accuracy of representative feeders is lower compared to active power in-feed only. If the performance could be increased, effects of smart grid schemes based on reactive power control could be studied for a higher admissible voltage rise as well.

To run automated power flow studies for a high number of real feeders requires to automate the assigment of loads in real feeders to appropriate locations in equivalent feeders. Hence, an algorithm utilizing the closeness centrality is needed.

Further, the proposed methodology could be improved with investigations to predict HCS precisely. Then, the HC calculation for the DER-scenarios *uniform* and *eof* would be obsolete. However, at the same time it should be considered that the number of DSOs having a powerful GIS-tool to run load flows is increasing.

Indeed, the parameter HCS requires most of the efforts to define a reference feeder. However, if a suitable network model is not available to run load flow calculation, then, a set of parameters could be found to estimate the parameter HCS. With the investigated LV-grid data, neuronal networks or decision trees could be developed to estimate the parameter HCS or even the HC of feeders which would replace power flow calculations. Nevertheless, such an approach would not provide a set of reference feeders that could be utilized for load flow studies.

A remaining open question is the accuracy regarding reactive power flows under a VoltVAr-control scheme of the proposed reference feeders and real feeders. This question is relevant for the investigation of larger grid areas including LV/MV and HV voltage levels.

The defined RFT are suitable for LV-feeders. Another open question is if the methodology could be expanded to MV-feeders.

The proposed RFT could be improved to increase the HC accuracy. For example, radial RFT without any branches could be developed and compared to the proposed RFT.

For training and educational purposes realistic grid data shall be utilized. However, realistic grids may complicate the understanding of concepts and studies for beginners. Hence, the proposed reference feeders are suitable to ease understanding and reduce the complexity of feeders allowing to focus on the study.

The analysis of the RFT distribution in networks showed that a few number of representative network topologies cannot be defined. Further investigations are needed to define such reference networks.

List of Tables

List of Figures

References

[1] Manoël Rekinger and Frauke Thies. Global Market Outlook For Solar Power 2015-2019. Solar power europe, report, June 2015.

[2] Fraunhofer Institut für Windenergie und Energiesystemtechnik. Vorstudie zur Integration großer Anteile Photovoltaik in die elektrische Energieversorgung. Technical report, November 2011.

[3] RWTH Aachen. evolvDSO - Definition of a Limited but Representative Number of Future Scenarios. Deliverable 1.1, March 2014.

[4] Thomas Stetz, Jan von Appen, Fabian Niedermeyer, Gunter Scheibner, Roman Sikora, and Martin Braun. Twilight of the Grids. *Power & Energy*, February 2015.

[5] DIHK Deutscher Industrie und Handelskammertag. Faktenpapier Ausbau der Stromnetze. https://www.dihk.de/ressourcen/downloads/faktenpapier-stromnetze-2015 (accessed 20.12.2017, 2015.

[6] e control. Trassen- und Systemlängen zum 31. Dezember 2015, December 2015.

[7] M. Stifter, B. Bletterie, D. Burnier, H. Brunner, and A. Abart. Analysis environment for low voltage networks. In *2011 IEEE First International Workshop on Smart Grid Modeling and Simulation (SGMS)*, pages 61–66, 2011.

[8] Deutsche Energie-Agentur GmbH. dena-Verteilernetzstudie. Ausbau- und Innovationsbedarf der Stromverteilnetze in Deutschland bis 2030. Technical report, Berlin, December 2012.

[9] Deutsche Energie-Agentur GmbH. dena-Studie Systemdienstleistungen 2030. Technical report, Berlin, November 2014.

[10] Klaus Engels. *Probabilistische Bewertung der Spannungsqualität in Verteilungsnetzen*. PhD thesis, RWTH Aachen, Aachen, December 2000.

[11] A. Kulmala, S. Repo, and P. Järventausta. Using statistical distribution network planning for voltage control method selection. In *Proc. IET Renewable Power Generation Conference*. IET, 2011.

[12] Gerrit Schlömer and Lutz Hofmann. Eine Heuristik zur Umbauplanung von Niederspannungsnetzen ganzer Ortschaften. In *Proc. 14. Symposium Energieinnovation*, Graz, 2016.

[13] EPIA. Connecting the sun - solar photovoltaics on the road to large/scale grid integration. Technical report, September 2012.

[14] R. Schwalbe, H. Brunner, M. Stifter, A. Abart, E. Traxler, M. Radauer, and W. Niederhuemer. DG-demonet smart LV grid - increasing hosting capacity of LV grids by extended planning and voltage control. In *2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST)*, pages 63–69, September 2015.

[15] Denis Mende, Jan Schwarz, Sebastian Schmidt, Daniel Premm, Vitali Sakschewski, Matthias Pfalzgraf, Hannes Homeyer, Johannes Schmiesing, and Johannes Brantl. Reactive power control of PV plants to increase the grid hosting capacity. In *Proc. 28th European Photovoltaic Solar Energy Conference and Exhibition*, pages 4225 – 4230, Paris, September 2013.

[16] Carlos Mateo Domingo, Tomás Gomez San Roman, Álvaro Sanchez-Miralles, Jesús Pascual Peco Gonzalez, and Antonio Candela Martinez. A Reference Network Model for Large-Scale Distribution Planning With Automatic Street Map Generation. *IEEE Transactions on Power Systems*, 26(1):190–197, February 2011.

[17] Thomas Stetz. *Autonomous Voltage Control Strategies in Distribution Grids with Photovoltaic Systems: Technical and Economic Assessment.* PhD thesis, Kassel University press GmbH, 2014.

[18] Martin Heidl and et. al. morePV2grid - More functionalities for increased integration of PV into grid. Final report, December 2013.

[19] Carlos Dierckxsens, Achim Woyte, Benoît Bletterie, Antony Zegers, Wim Deprez, Annick Dexters, Kevin Van Roey, Joris Lemmens, Joris Lowette, Koen Nulens, Yehia Tarek Fawzy, Boštjan Blazič, Blaž Uljanič, and Marko Kolenc. Cost-effective integration of photovoltaics in existing distribution grids: results and recommendations. Technical report, March 2015.

[20] Benoît Bletterie, Serdar Kadam, Antony Zegers, and Zoran Miletic. On the effectiveness of voltage control with PV inverters in unbalanced low voltage networks. In *Proc. Electricity Distribution (CIRED 2015), 23rd International Conference and Exhibition on*, Lyon, 2015.

[21] B. Bletterie, S. Kadam, R. Bolgaryn, and A. Zegers. Voltage Control with PV Inverters in Low Voltage Networks #8212;In Depth Analysis of Different Concepts and Parameterization Criteria. *IEEE Transactions on Power Systems*, 32(1):177–185, January 2017.

[22] Peter Esslinger. Studie Q(U). Technical report, Munich, August 2012.

[23] Adrian Constantin and Radu Dan Lazar. Open Loop Q(U) Stability Investigation in Case of PV Power Plants. In *Proc. 27th European Photovoltaic Solar Energy Conference and Exhibition*, pages 3745 – 3749, Frankfurt, September 2012.

[24] Marco Lindner, Markus Gödde, and Fabian Potratz. Ergebnisse der FNN-Studie zu neuen Verfahren der statischen Spannungshaltung. In *Proc. 3. Konferenz Zukünftige Stromnetze für Erneuerbare Energien*, Berlin, January 2015.

[25] Filip Andren, Benoit Bletterie, Serdar Kadam, Panos Kotsampopoulos, and Christof Bucher. On the Stability of local Voltage Control in Distribution Networks with a High Penetration of Inverter-Based Generation. *IEEE Transactions on Industrial Electronics*, pages 1–1, 2014.

[26] Yves-Marie Saint-Drenan, Stefan Bofinger, and Kurt Rohrig. Bestimmung von Koeffizienten für die Entschädigungszahlungen von PV Anlagen. Technical report, January 2013.

[27] Jan von Appen, Martin Braun, Bastian Zin\s ser, and Dirk Stell-bogen. Leistungsbegrenzung bei PV-Anlagen-Anpassung der Model-lierungsmethoden und Vergleich verschiedener Standorte. In *Proc. 27th Symp. Photovoltaic Solar Energy*, pages 47–52, 2011.

[28] Sam Weckx, C. Gonzalez, and Johan Driesen. Reducing Grid Losses and Voltage Unbalance with PV inverters. In *Proc. IEEE PESGM*, Washington, July 2014.

[29] Rémy Garaude Verdier, et.al. Grid4eu final report, March 2016.

[30] Andrea Rodríguez Calvo, et. al. Grid4eu gD3.5 Scalability and Replica-bility rules. Technical report, January 2016.

[31] IGREENGrid - Homepage. http://www.igreengrid-fp7.eu/, September 2015. [Online; accessed 19-July-2017].

[32] Jesus Varela, Nikos Hatziargyriou, Lisandro J. Puglisi, Gareth Bissel, Andreas Abart, Marco Rossi, and Robert Priewasser. The Best of IGREENGrid Practices: A Distribution Network's Contribution to Re-siliency. *IEEE Power and Energy Magazine*, 13(3):81–89, May 2015.

[33] Thomas Ackermann and et. al. Verteilnetzstudie Rheinland-Pfalz. Tech-nical report, January 2014.

[34] RWE AG, ABB AG, Consentec GmbH, and Technische Universität Dortmund. Abschlussbericht - BMWi-Förderprojekt Netze für die Stromversorgung der Zukunft". Technical report, October 2011.

[35] Jörg Scheffler. *Bestimmung der maximal zulässigen Netzanschlus-sleistung photovoltaischer Energiewandlungsanlagen in Wohnsiedlun-gen*. PhD thesis, TU Chemnitz, June 2002.

[36] Georg Kerber. *Aufnahmefähigkeit von Niederspannungsverteilnetzen für die Einspeisung aus Photovoltaikkleinanlagen*. PhD thesis, München, 2011.

[37] Paul Hines, Seth Blumsack, E. Cotilla Sanchez, and Clayton Barrows. The topological and electrical structure of power grids. In *System Sci-ences (HICSS), 2010 43rd Hawaii International Conference on*, pages 1–10. IEEE, 2010.

[38] D. P. Montoya and J. M. Ramirez. A minimal spanning tree algorithm for distribution networks configuration. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–7, July 2012.

131

[39] Joerg Dickert, M. Domagk, and Peter Schegner. Benchmark low voltage distribution networks based on cluster analysis of actual grid properties. In *PowerTech (POWERTECH), 2013 IEEE Grenoble*, pages 1–6. IEEE, 2013.

[40] Erhan Demirok. *Control of Grid Interactive PV Inverters for High Penetration in Low Voltage Distribution Networks.* PhD thesis, Aalborg University, Aalborg, 2012.

[41] M. Rylander and J. Smith. Stochastic Analysis to Determine Feeder Hosting Capacity for Distributed Solar PV. Technical report, EPRI, Knoxville, December 2012.

[42] Moderne Verteilernetze für Deutschland (Verteilernetzstudie). Technical report, September 2014.

[43] Matthew Rylander, Robert J. Broderick, Matthew J. Reno, Karina Munoz-Ramos, Jimmy E. Quiroz, Jeff Smith, Lindsey Rogers, Roger Dugan, Barry Mather, Michael Coddington, Peter Gotseff, and Fei Ding. Alternatives to the 15% Rule, 2015.

[44] M. Rylander, J. Smith, and W. Sunderman. Streamlined Method for Determining Distribution System Hosting Capacity. *IEEE Transactions on Industry Applications*, 52(1):105–111, January 2016.

[45] Michiel Nijhuis, Madeleine Gibescu, and Sjef Cobben. Clustering of low voltage feeders from a network planning perspective. In *CIRED 23rd International Conference on Electricity Distribution*, Lyon, 2015.

[46] B. Cloteaux. Limits in modeling power grid topology. In *Network Science Workshop (NSW), 2013 IEEE 2nd*, pages 16–22, April 2013.

[47] G. A. Pagani and M. Aiello. Towards Decentralization: A Topological Investigation of the Medium and Low Voltage Grids. *IEEE Transactions on Smart Grid*, 2(3):538–547, September 2011.

[48] Eran Schweitzer, Kanali Togawa, Tim Schlösser, and Antonello Monti. A Matlab GUI for the Generation of Distribution Grid Models. International ETG Congress 2015:79–84, November 2015.

[49] Gunther Gust. Analyse von Niederspannungsnetzen und Entwicklung von Referenznetzen. Diploma thesis, KIT, Karlsruhe, July 2014.

[50] D. Santos-Martin and S. Lemon. Simplified Modeling of Low Voltage Distribution Networks for PV Voltage Impact Studies. *IEEE Transactions on Smart Grid*, 7(4):1924–1931, July 2016.

[51] Robert J. Broderick, Joseph R. Williams, and Karina Munoz-Ramos. Clustering Method and Representative Feeder Selection for the California Solar Initiative. October 2015.

[52] F. Dehghani, H. Nezami, M. Dehghani, and M. Saremi. Distribution feeder classification based on self organized maps (case study: Lorestan province, Iran). In *2015 20th Conference on Electrical Power Distribution Networks Conference (EPDC)*, pages 27–31, April 2015.

[53] Marco Lindner, Christian Aigner, Rolf Witzmann, Frank Wirtz, Ibrahim Berber, Markus Gödde, and Robert Frings. Aktuelle Musternetze zur Untersuchung von Spannungsproblemen in der Niederspannung. In *Proc. 14. Symposium Energieinnovation*, Graz, 2016.

[54] Gerhard Walker, Ann-Kathrin Krauss, Simon Eilenberger, Willi Schweinfort, and Stefan Tenbohlen. Entwicklung eines standardisierten Ansatzes zur Klassifizierung von Verteilnetzen. In *VDE-Kongress 2014*. VDE VERLAG GmbH, 2014.

[55] Gerhard Walker, Haiko Nägele, Fabian Kniehl, Alexander Probst, Marc Brunner, and Stefan Tenbohlen. An application of cluster reference grids for an optimized grid simulation. In *Proc. CIRED 23rd International Conference on Electricity Distribution*, Lyon, 2015.

[56] Andreas Becker, Tobias Lühn, Michael Mohrmann, and Gerrit Schlömer. *Netzausbauvarianten in Niederspannungsverteilnetzen : Regelbare Ortsnetztransformatoren in Konkurrenz zu konventionellen Netzausbaumaßnahmen.* Cuvillier, E, 1. edition, July 2014.

[57] FNN. Schlussbericht Vergleich von technischer Wirksamkeit sowie Wirtschaftlichkeit zeitnah verfügbarer Verfahren zur Sicherung der statischen Spannungshaltung in Niederspannungsnetzen mit starker dezentraler Einspeisung. Technical report, December 2014.

[58] F. Barakou, D. Koukoula, N. Hatziargyriou, and A. Dimeas. Fractal geometry for distribution grid topologies. In *PowerTech, 2015 IEEE Eindhoven*, pages 1–6, June 2015.

[59] Yingliang Li and Peter J. Wolfs. Taxonomic description for western Australian distribution medium-voltage and low-voltage feeders. *IET Generation, Transmission & Distribution*, 8(1):104–113, January 2014.

[60] K. P. Schneider, Y. Chen, D. Engle, and D. Chassin. A Taxonomy of North American Radial Distribution Feeders. In *Proc. Power & Energy Society General Meeting, 2009. PES '09. IEEE*, 2009.

[61] G. Celli, F. Pilo, G. Pisano, and G. G. Soma. Reference scenarios for Active Distribution System according to ATLANTIDE project planning models. In *2014 IEEE International Energy Conference (ENERGY-CON)*, pages 1190–1196, 2014.

[62] Serdar Kadam. Systematical Analysis of Low Voltage-Networks for Smart Grid Studies. Diploma thesis, Technische Universität Wien, Vienna, June 2012.

[63] Consentec. Kurzgutachten zur Ermittlung von Verlustanteilen je Netzebene. Report, Concentec GmbH, Aachen, July 2013.

[64] C. Schwaegerl, M. H. J. Bollen, K. Karoui, and A. Yagmur. Voltage control in distribution systems as a limitation of the hosting capacity for distributed energy resources. In *Proc. Electricity Distribution, 2005. CIRED 2005. 18th International Conference and Exhibition on*, pages 1–5. IET, 2005.

[65] B. Bletterie, A. Gorsek, B. Uljanic, B. Blazic, A. Woyte, T. Vu Van, F. Truyens, and J. Jahn. Enhancement of the network hosting capacity – clearing space for/with PV. pages 4828 – 4834, Valencia, Spain, June 2010.

[66] T. Fawzy, D. Premm, B. Bletterie, and A. Goršek. Active contribution of PV inverters to voltage control – from a smart grid vision to full-scale implementation. *e & i Elektrotechnik und Informationstechnik*, 128(4):110–115, April 2011.

[67] Tobias Walla. Hosting capacity for photovoltaics in Swedish distribution grids. Diploma thesis, Uppsala Universitet, Uppsala, June 2012.

[68] Nikos Hatziargyriou, Evangelos Karfopoulos, Achilleas Tsitsimelis, Despina Koukoula, Marco Rossi, and Viganò Giacomo. On the der hosting capacity of distribution feeders. In *Proc. CIRED 23rd International Conference on Electricity Distribution*, Lyon, 2015.

[69] Marko Kolenc, Igor Papic, and Bostjan Blazic. Distribution network development based on stochastic modelling approach. In *Power Engineering Conference (UPEC), 2014 49th International Universities*, pages 1–6. IEEE, 2014.

[70] C. T. Gaunt, E. Namanya, and R. Herman. Voltage modelling of LV feeders with dispersed generation: Limits of penetration of randomly connected photovoltaic generation. *Electric Power Systems Research*, 143:1–6, February 2017.

[71] Thomas Stetz, Hendrik Wolf, Alexander Probst, Simon Eilenberger, Yves-Marie Saint-Drenan, Erika Kämpf, Martin Braun, Daniel Schöllhorn, and Sebastian Schmidt. Stochastische Analyse von Smart-Meter Messdaten. VDE-Kongress 2012, November 2012.

[72] J. A. Bondy and U. S. R. Murty. *Graph Theory with applications*. North Holland, New York - Amsterdam - Oxford, 1976.

[73] Mark Newman. *Networks: An Introduction*. Oxford University Press, March 2010.

[74] Kelly H. Zou, Kemal Tuncali, and Stuart G. Silverman. Correlation and Simple Linear Regression, 2003.

[75] Wayne W. Daniel. *Applied nonparametric statistics*. PWS-Kent Publ., June 2000.

[76] Abdi Hervé. The Kendall Rank Correlation Coefficient. In *Encyclopedia of Measurement and Statistics*. SAGE Publications, Inc., Thousand Oaks, September 2007.

[77] CENELEC. Voltage characteristics of electricity supplied by public electricity networks, January 2011.

[78] Kenan Hatipoglu, Ismail Fidan, and Ghadir Radman. Investigating effect of voltage changes on static ZIP load model in a microgrid environment. In *North American Power Symposium (NAPS), 2012*, pages 1–5. IEEE, 2012.

[79] Serdar Kadam, Benoît Bletterie, and Wolfgang Gawlik. A Large Scale Grid Data Analysis Platform for DSOs. *Energies*, 10(8):1099, July 2017.

A   Reference Topologies

In this section all defined reference feeder topologies are presented. One should note that the parameters R_k and I_{nom} are needed to configure the reference feeder topologies before being able to run load flow calculations.
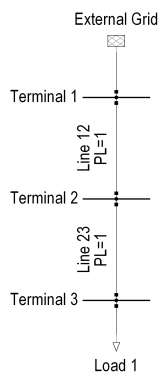
Figure 74: Reference feeder topology A1
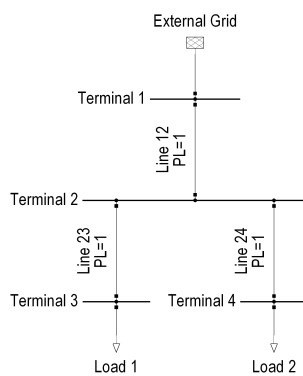
Figure 75: Reference feeder topology A2

136
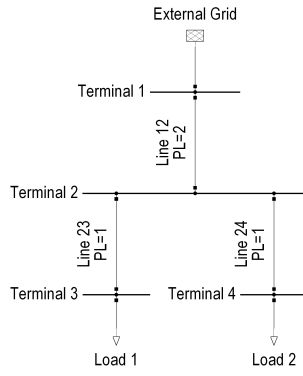
Figure 76: Reference feeder topology A3
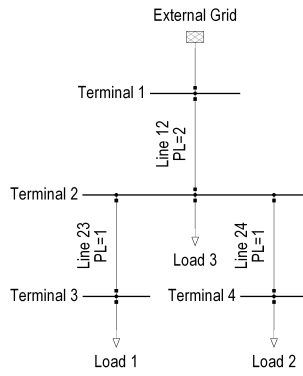
Figure 77: Reference feeder topology A4
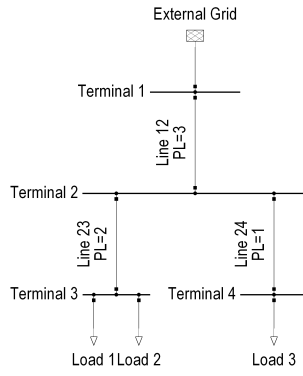
Figure 78: Reference feeder topology B1
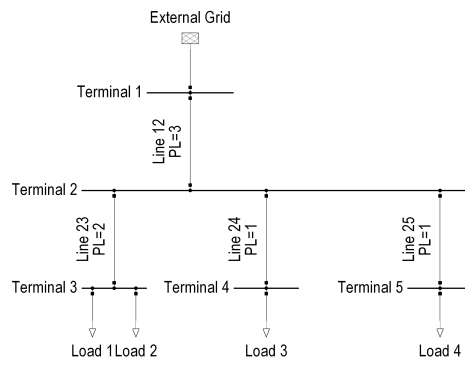
Figure 79: Reference feeder topology B2
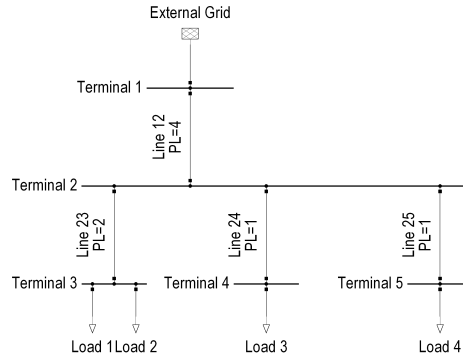
Figure 80: Reference feeder topology B3
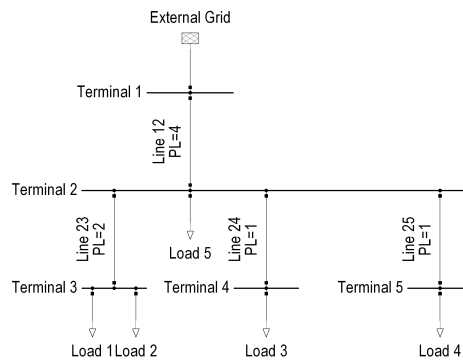
Figure 81: Reference feeder topology B4
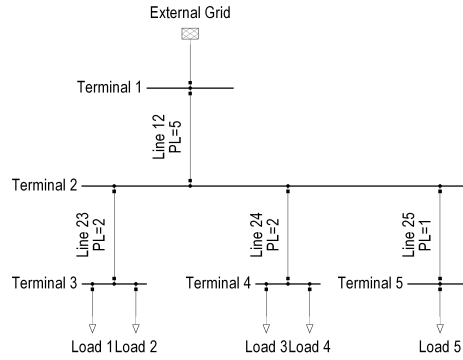
Figure 82: Reference feeder topology C1
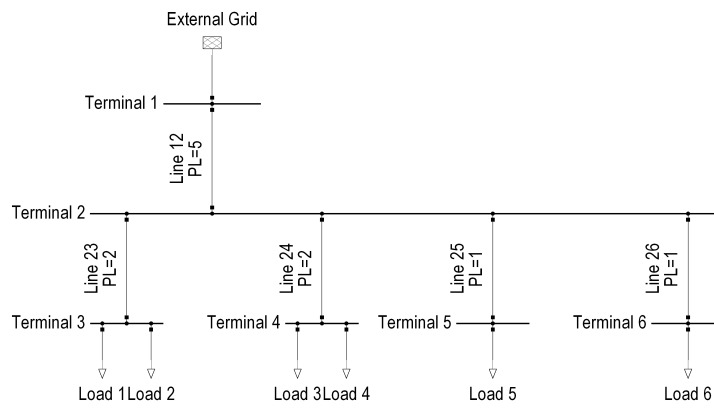
Figure 83: Reference feeder topology C2
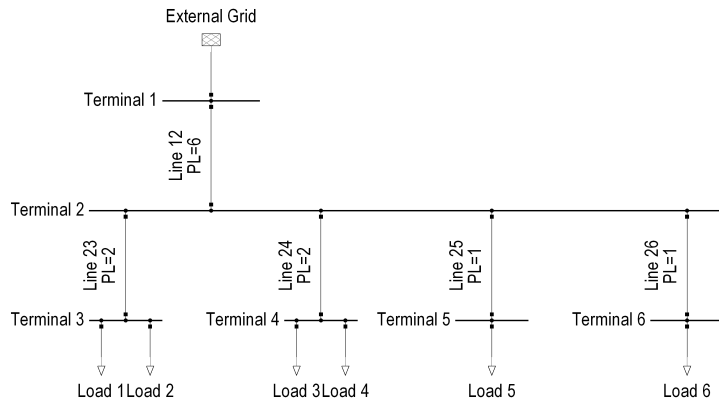
Figure 84: Reference feeder topology C3
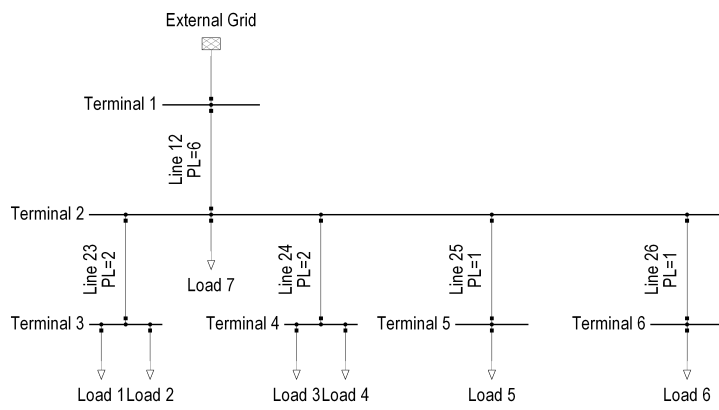
Figure 85: Reference feeder topology C4

141
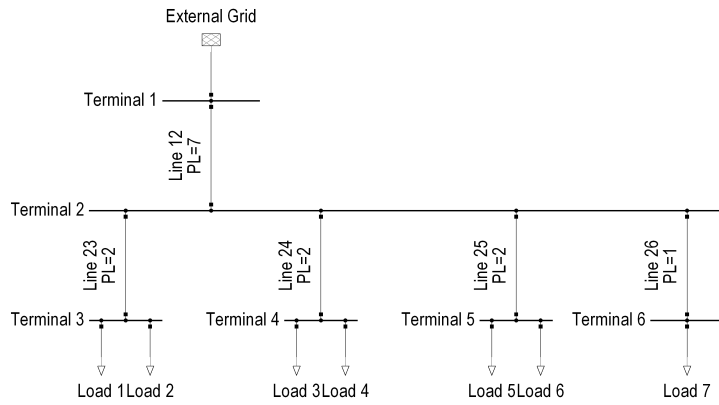
Figure 86: Reference feeder topology D1
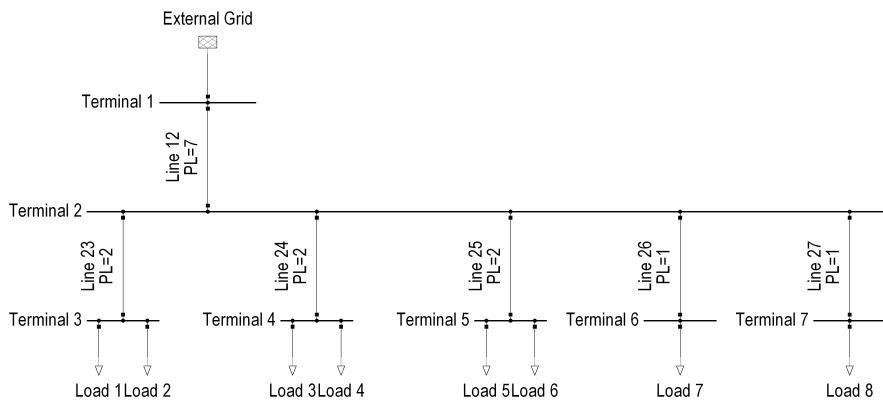
Figure 87: Reference feeder topology D2
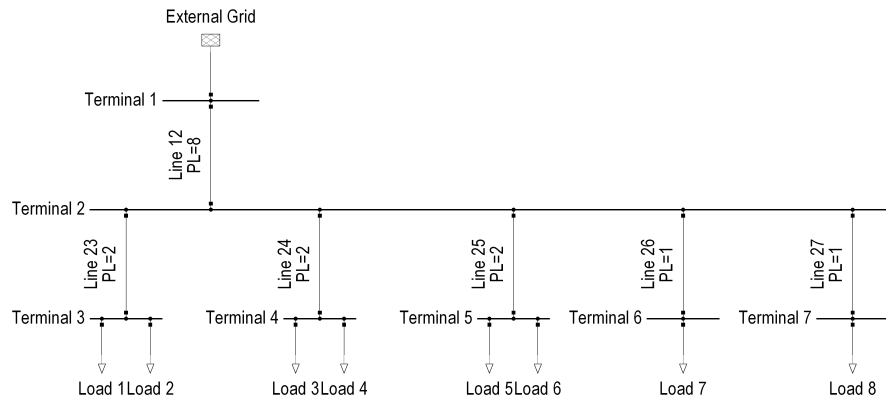
Figure 88: Reference feeder topology D3
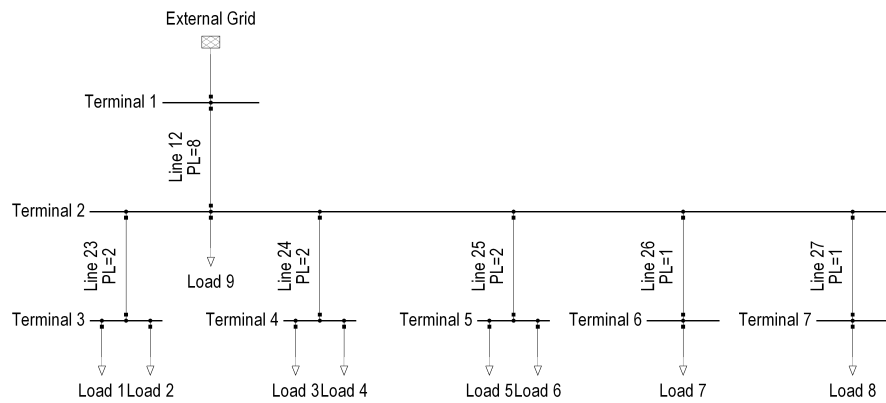
Figure 89: Reference feeder topology D4

143
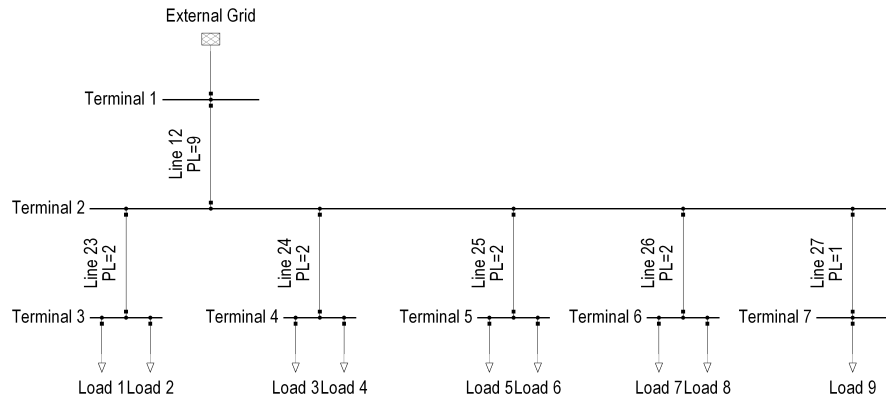
Figure 90: Reference feeder topology E1
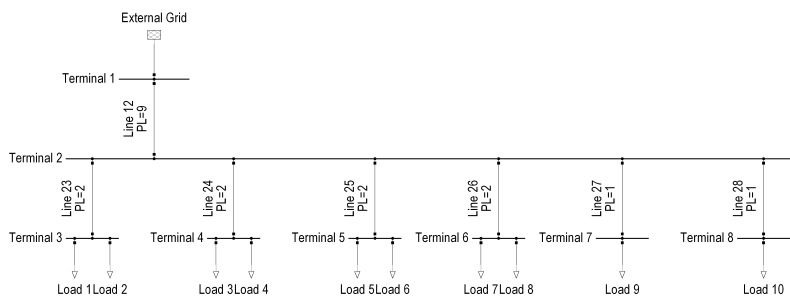
Figure 91: Reference feeder topology E2
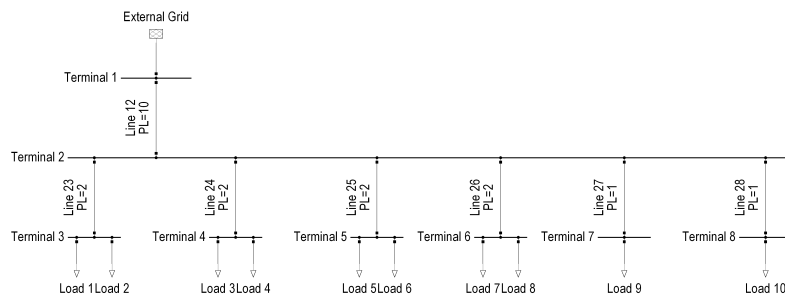
Figure 92: Reference feeder topology E3
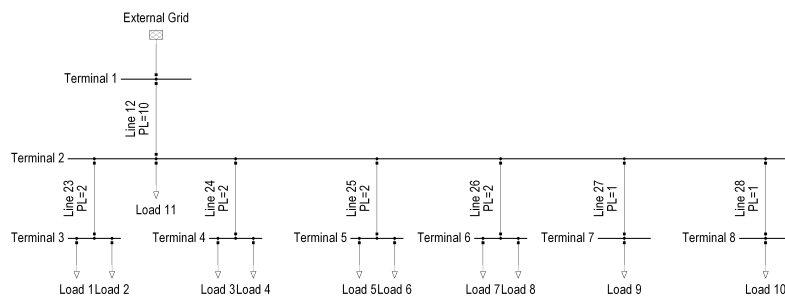
Figure 93: Reference feeder topology E4

B Variation of the Ampacity of the Weakest Line in a Feeder Per Reference Feeder Topology Group

In this section, the ampacity of lines, as one input parameter for the RFT parametrization, is investigated. The ecdf of this parameter is provided for each RFT, to generate a realistically parameterized RFT for synthetic large scale studies.
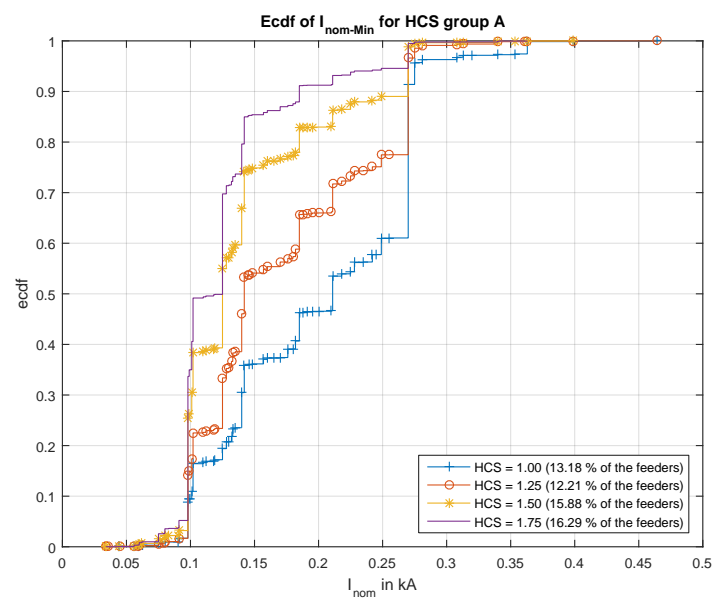
Figure 94: Empirical cumulative distribution function of the minimal nominal current in feeders for the defined hosting capability sensitivity group A
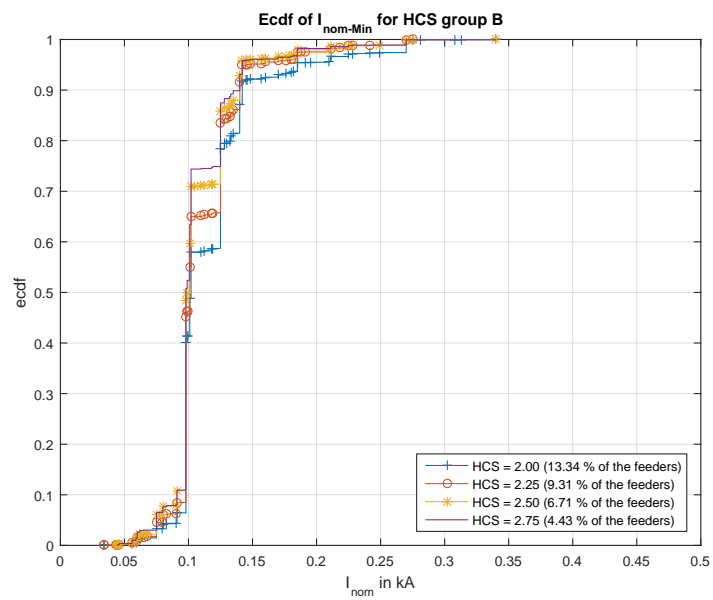
Figure 95: Empirical cumulative distribution function of the minimal nominal current in feeders for the defined hosting capability sensitivity group B
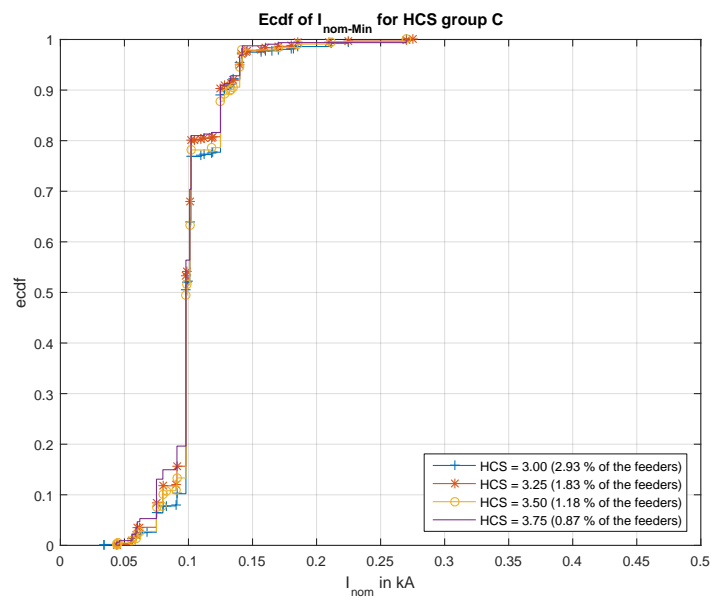
Figure 96: Empirical cumulative distribution function of the minimal nominal current in feeders for the defined hosting capability sensitivity group C
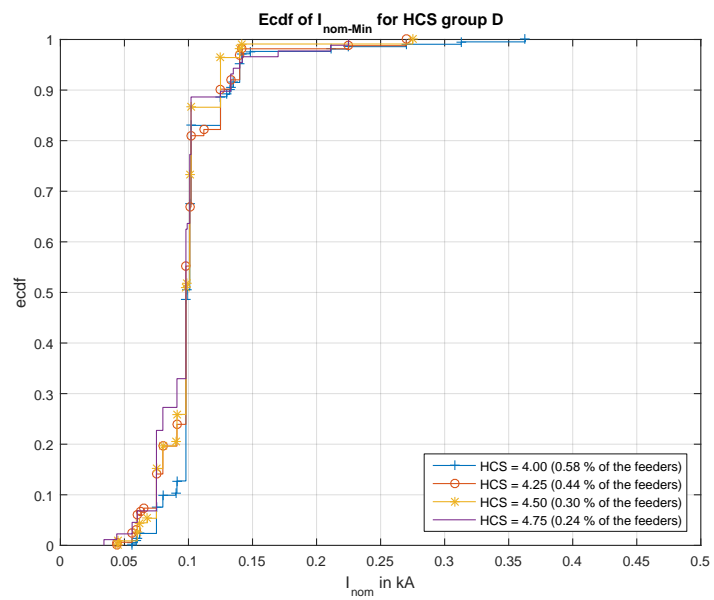
Figure 97: Empirical cumulative distribution function of the minimal nominal current in feeders for the defined hosting capability sensitivity group D
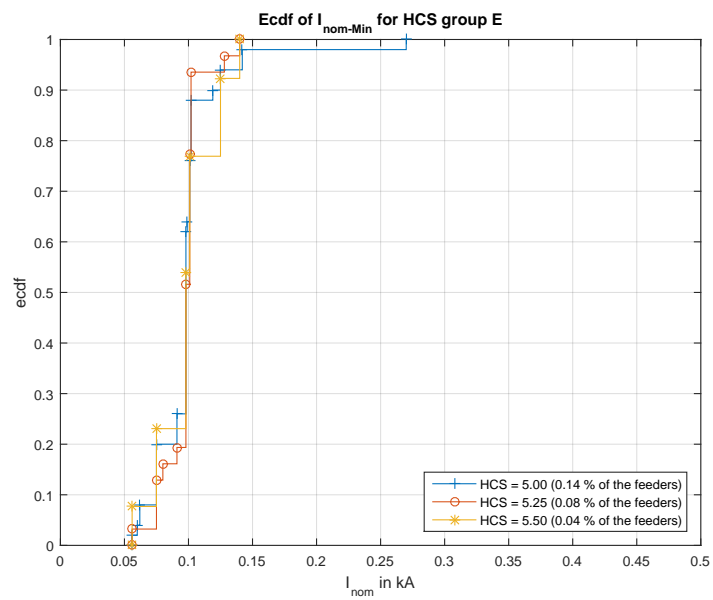
Figure 98: Empirical cumulative distribution function of the minimal nominal current in feeders for the defined hosting capability sensitivity group E

C Variation of the Short Circuit Resistance at the End Node Per Reference Feeder Topology Group

In this section, the short circuit resistance of the end node, as one input parameter for the RFT parametrization, is investigated. The ecdf of this parameter is provided for each RFT, to generate a realistically parameterized RFT for synthetic large scale studies
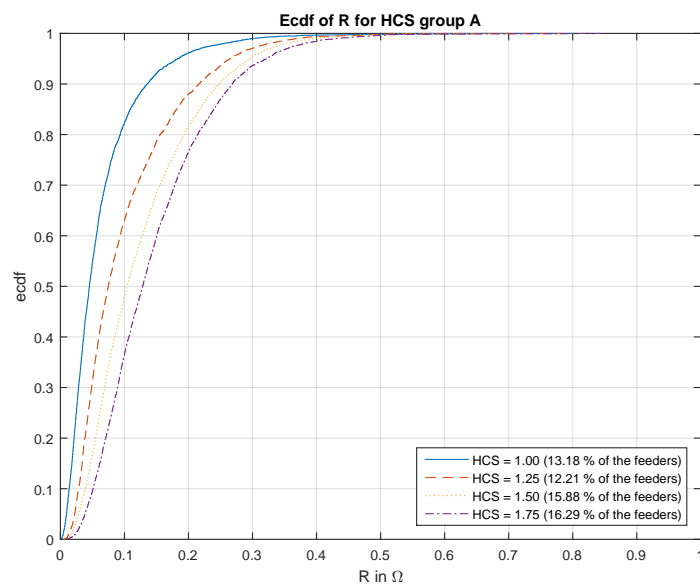
Figure 99: Empirical cumulative distribution function of the end node resistance in feeders for the defined hosting capability sensitivity group A
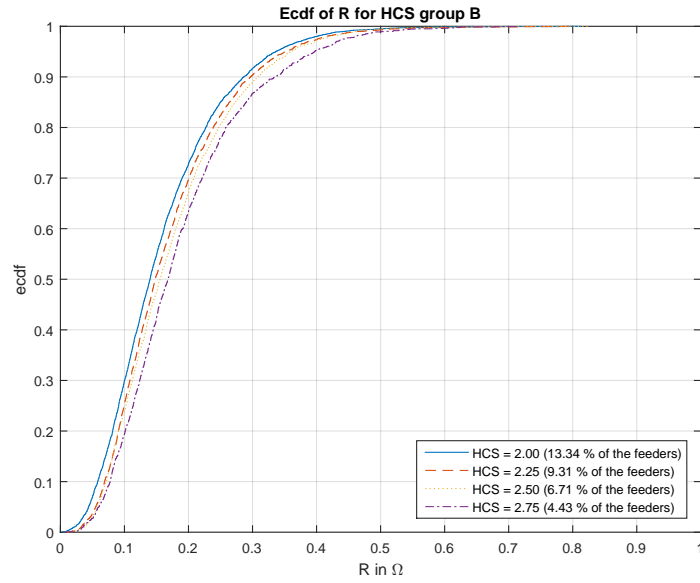
Figure 100: Empirical cumulative distribution function of the end node resistance in feeders for the defined hosting capability sensitivity group B
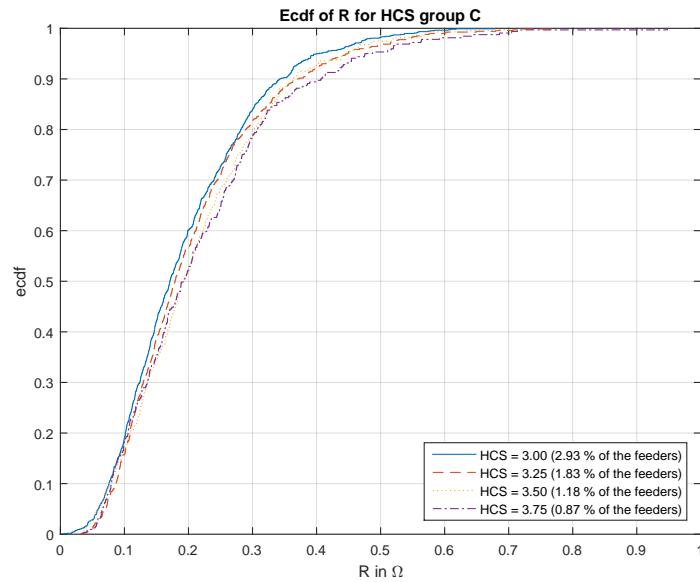
Figure 101: Empirical cumulative distribution function of the end node resistance in feeders for the defined hosting capability sensitivity group C
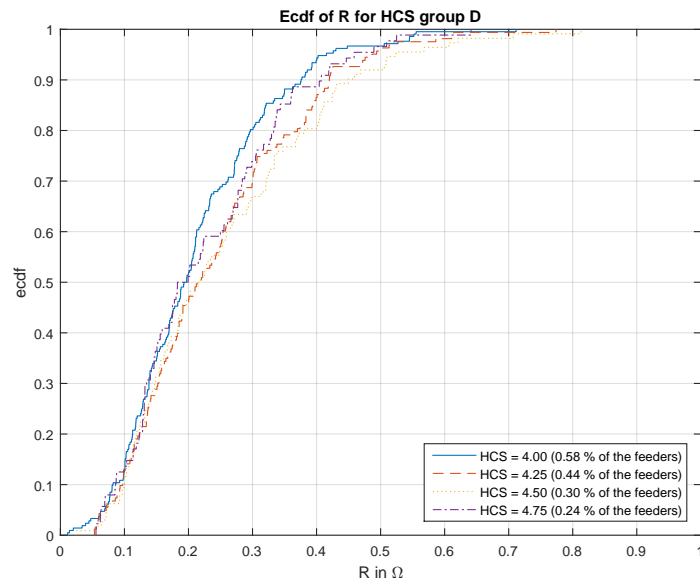
Figure 102: Empirical cumulative distribution function of the end node resistance in feeders for the defined hosting capability sensitivity group D
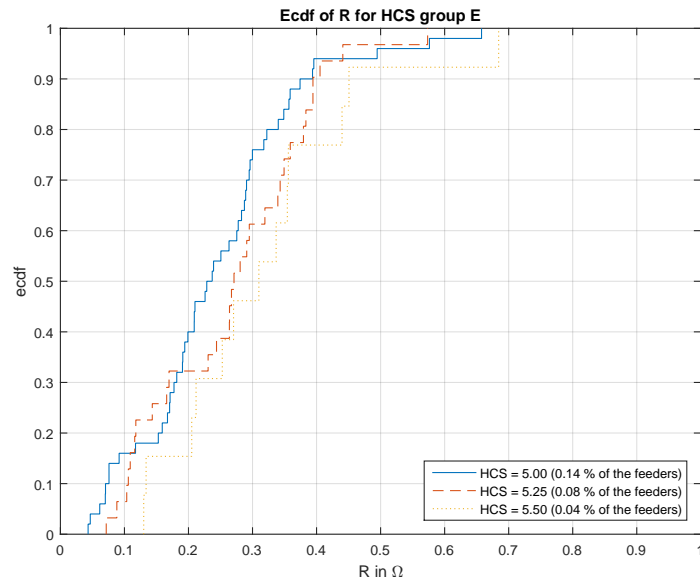
Figure 103: Empirical cumulative distribution function of the end node resistance in feeders for the defined hosting capability sensitivity group E

D Hosting Capability Mismatch for the distributed energy resource-scenario end of feeder

In this section, the HC mismatch results for the DER-scenario *eof* are presented. In the methodology, the lowest ampacity in the feeder is considered. Therefore, the minimal ampacity on the main path of the feeder, which is the relevant ampacity for the DER-scenario *eof* may be higher. Hence, the HC can only be underestimated for loading constrained feeders. Since the end node resistance R_k is a input parameter of the methodology, the mismatch for voltage constrained feeders is neglible. Thus, the mismatch is basically caused by the mismatch between the parameters I_{nom-MP} and $I_{nom-Avg}$.
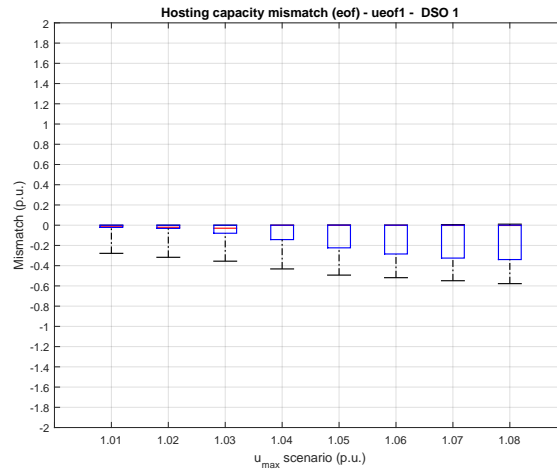
Figure 104: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *end of feeder*, power factor 1)
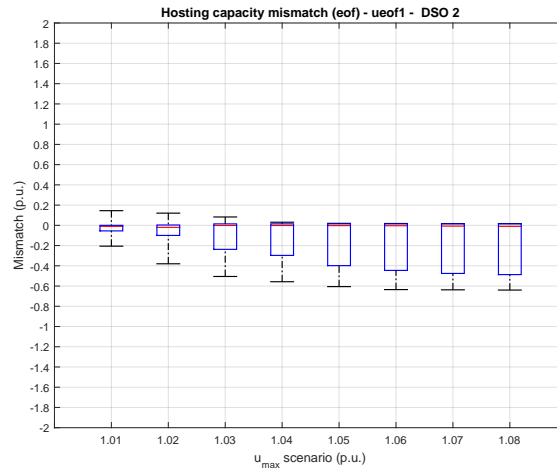
Figure 105: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *end of feeder*, power factor 1)

E Hosting Capacity Sensitivity Dependency Analysis for Distribution System Operator 1

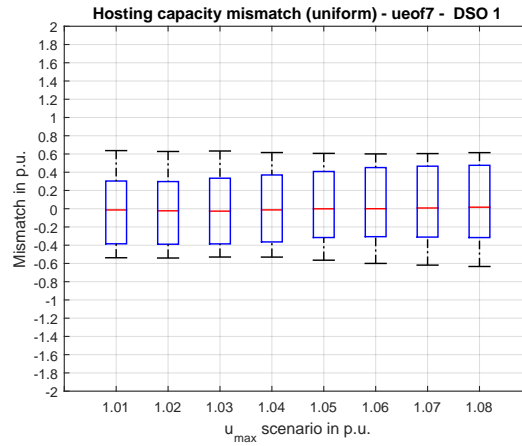In this section, the results of the HCS dependency analysis for DSO 1 are given.

155
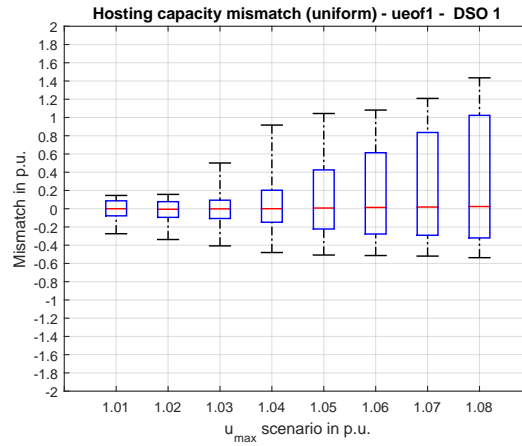
Figure 106: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 1% (ueof1) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)
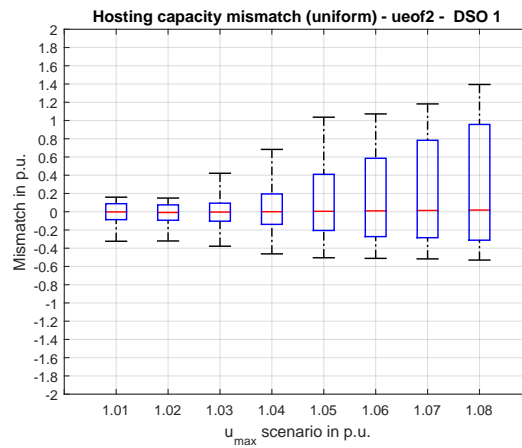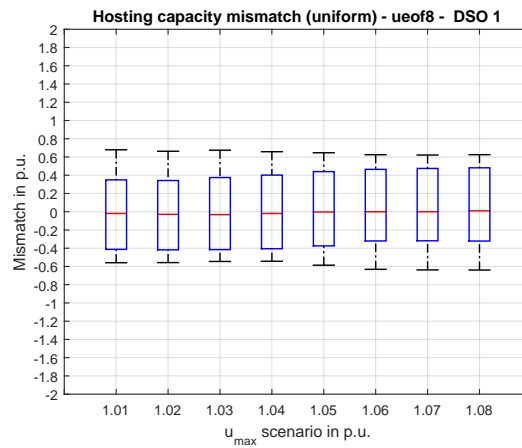

Figure 107: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 2% (ueof2) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)

Figure 108: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 3% (ueof3) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)
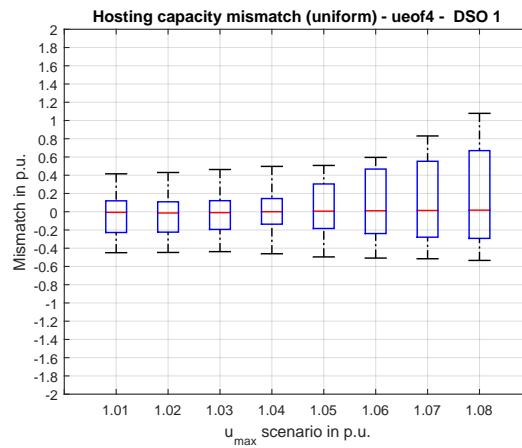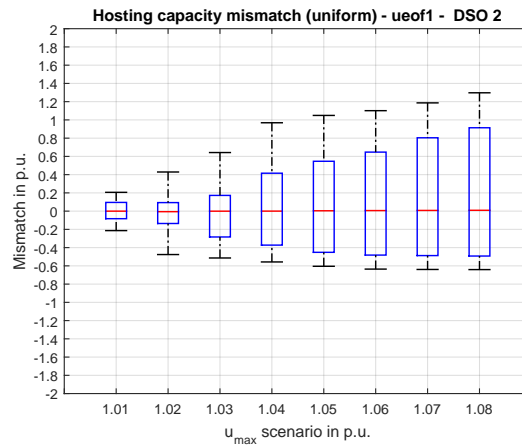

Figure 109: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 4% (ueof4) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)
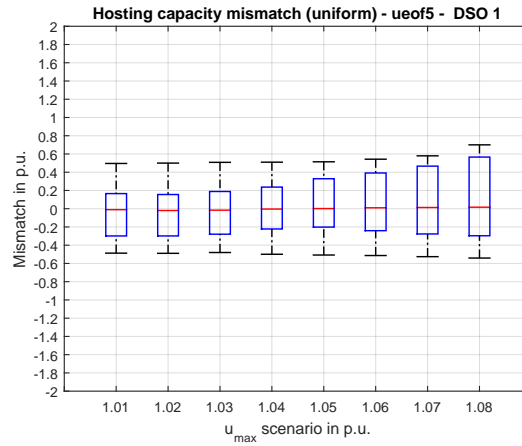
Figure 110: DSO1 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 5% (ueof5) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)
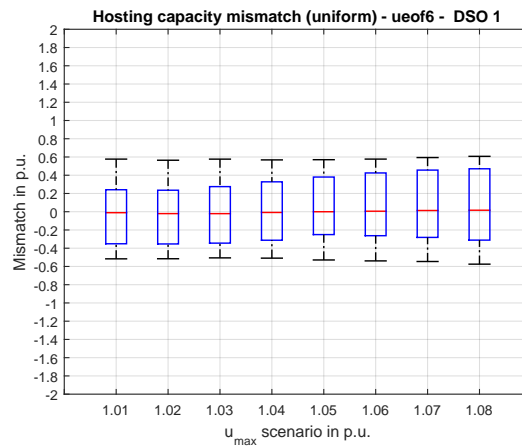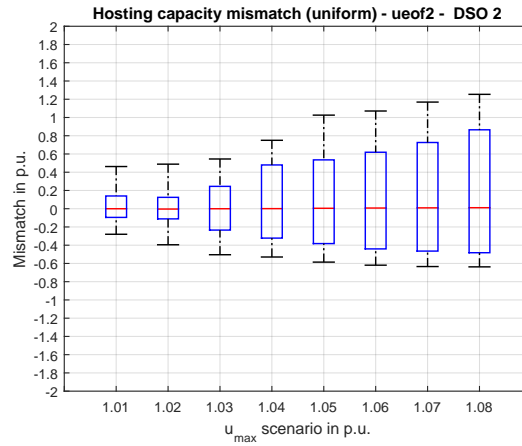

Figure 111: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 6% (ueof6) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)
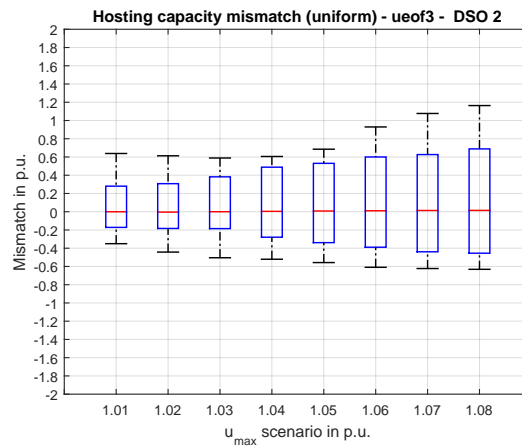
158

Figure 112: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 7% (ueof7) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)
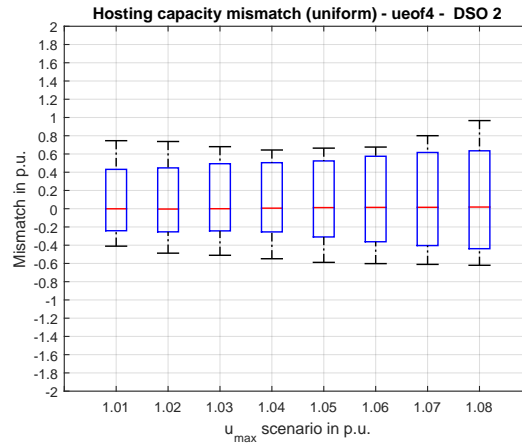

Figure 113: DSO 1 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 8% (ueof8) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)

F   Hosting Capability Sensitivity Dependency Analysis for Distribution System Operator 2

In this section, the results of the HCS dependency analysis for DSO 2 are given.
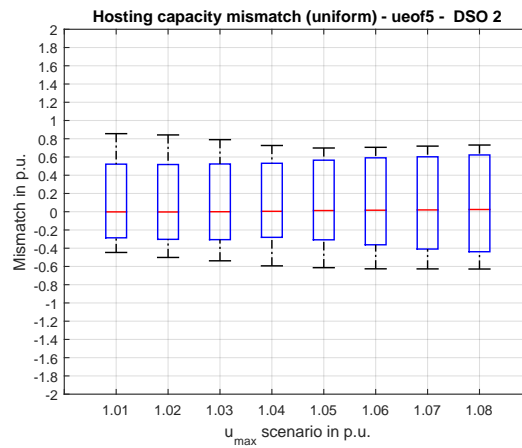

Figure 114: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 1% (ueof1) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)
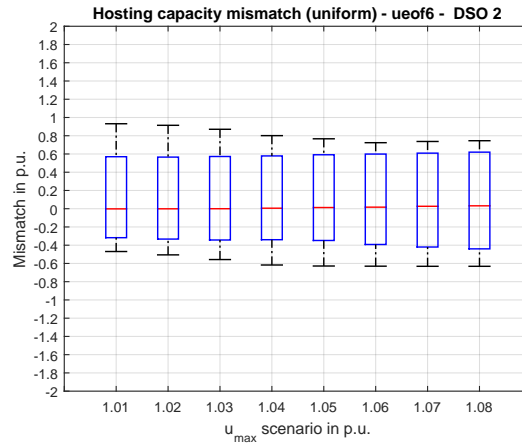
Figure 115: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 2% (ueof2) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)
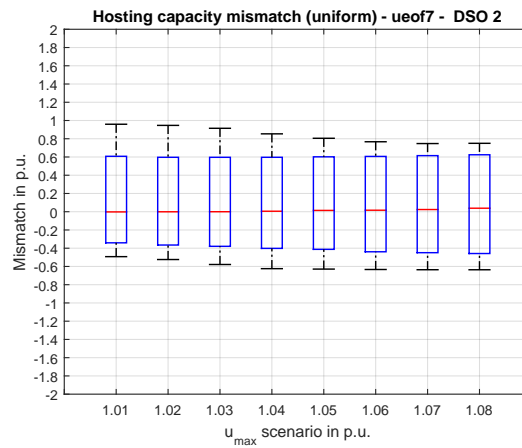

Figure 116: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 3% (ueof3) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)

Figure 117: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 4% (ueof4) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)
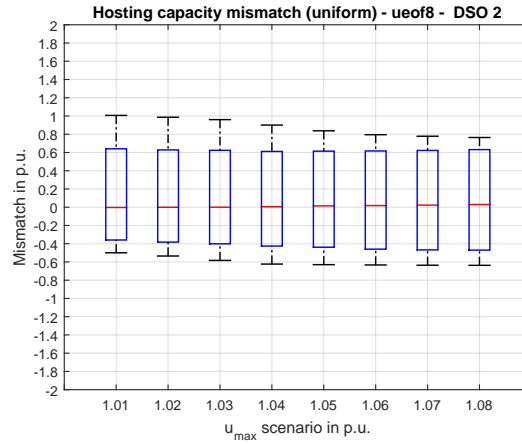

Figure 118: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 5% (ueof5) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)

Figure 119: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 6% (ueof6) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)

Figure 120: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 7% (ueof7) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)

163

Figure 121: DSO 2 - Hosting capability mismatch for different admissible voltage rise limits; Hosting capability sensitivity obtained for an admissible voltage rise of 8% (ueof8) - boxplot showing 1th/5th/median/95th and 99th percentile (DER-scenario *uniform*, power factor 1)

G   Feeder Analysis Algorithms

The described methodology requires performing a high number of scenarios for a large number of feeders. Therefore an adaptable and flexible design of scripts is required to work with this high amount of data and calculate the results in an automated way. An algorithm was developed which was designed to calculate all HC scenarios and statistical parameters (Algorithm 6) in a modular way. First the subscripts are presented, followed by the main algorithm for the feeder analysis.

CheckPlausibility: In this function, the plausibility of the number of nodes, lines and loads is tested. Furthermore, it is tested if a power flow calculation can be calculated without errors. If the plausibility check fails, the network is omitted and no results are written:

Data: Network data
Result: Boolean: Network is suitable for network calculations or not

1 initialization;
2 SetAllLoadsInNetworkToZero();
3 NrLoads = GetNumberOfLoadsInNetwork();
4 NrNodes = GetNumberofNodesInNetwork();
5 iErr = RunLoadFlowCalculation();
6 oTransformer = GetHighestLoadedTransformerObject();
7 transformerLoading = oTransformer.loading;
8 oLine = GetHighestLoadedLineObject();
9 lineLoading = oLine.loading;
10 INVALID = NrLoads < 1 or NrNodes <= 1 or iErr == 1 or
   transformerLoading > 100 or lineLoading > 100;

Algorithm 1: Check network plausibility

DefineFeeders: First, feeders that were already defined in the network are deleted. Next, the feeder starting points are identified. Feeder starting points can be identified since these nodes are drawn both in the 'internal world' of the substation as well as in the 'real world'. Third, the feeders of a network are defined. After the definition of feeders, it is possible to perform calculations for each feeder independently.

Data: Network data
Result: Feeder definitions

1 initialization;
2 DeleteAllFeeders();
3 setOfNodesGrid = FindNodesInTheGrid();
4 setOfNodesInternal = FindNodesInInternalWorld();
5 setFeederNodes = getCommonNodes(setNodesGrid,setNodesInternal);
6 **for** *each node in setFeederNodes* **do**
7 | DefineAFeederStartingAtNode(node);
8 **end**

Algorithm 2: Feeder definition

SetSlackToLV: In some LV-networks, more than one secondary substations may be found. To calculate the hosting capability of feeders, all secondary substations need to be deactivated and replaced by a slack on low-voltage side of the transformers.

Data: Network data

Result: All secondary substations replaced by a slack

1 initialization;

2 ListOfTransformers = FindTransformersInSecondarySubstations();

3 for *each transformer in ListOfTransformers* **do**

4   DeactivateTransformer(transformer);

5   slack = CreateSlack(transformer);

6   ConnectSlackToLVsideOfTransformer(transformer,slack);

7 end

      Algorithm 3: Replace secondary substation by a slack

RunStatisticalAnalysis: With this script, statistical parameters that do not require either a power flow calculation or a short circuit calculation are obtained both on feeder and network level.

Data: Network
Result: Feeder and Grid Statistics

1 initialization;
2 NrLoadsGrid = GetNumberOfLoadsInNetwork();
3 NrNodesGrid = GetNumberOfNodesInNetwork();
4 NrCablesGrid = GetNumberOfCablesinNetwork();
5 TotalLineLengthGrid = 0;
6 ListOfLines = GetAllLinesInNetwork(); **for** *each Line in ListOfLines*
   do
7 | TotalLineLengthGrid += Line.length;
8 end
9 GridADTN = CalculateAverageDistanceToneighbors(Network);
10 GridANON = CalculateAverageNumberOfneighbors(Network);
11 TotalRating = 0;
12 for *each SecondaryTransformer* **do**
13 | TotalRating += SecondaryTransformer.rating;
14 end
15 for *each Feeder in Network* **do**
16 | LoadsFeeder = GetNumberOfLoadsInFeeder(Feeder);
17 | NodesFeeder = GetNumberOfNodesInFeeder(Feeder);
18 | CablesFeeder = GetNumberOfCablesInFeeder(Feeder);
19 | ListOfLines = GetAllLinesInFeeder(Feeder);
20 | Line = GetLineWithHighestRating(ListOfLines);
21 | InomMax = Line.Inom;
22 | cable = GetCableWithLowestRating(ListOfLines);
23 | InomMin = Line.Inom;
24 | TotalLineLengthFeeder = 0;
25 | **for** *each Line in ListOfLines* **do**
26 | | TotalLineLengthFeeder += Cable.length;
27 | **end**
28 | InomAvg = CalculateAverageCableRating(setCables);
29 | FeederADTN = CalculateAverageDistanceToneighbors(Feeder);
30 | FeederANON = CalculateAverageNumberOfneighbors(Feeder);
31 end

Algorithm 4: Gather statistical parameters

RunGeographicalAnalysis: Calculation of the straight line distance and electrical distance from the secondary transformer to the most distant node.

Data: Network

Result: Electrical length and linear distance to end node; BooleanRadialFeeder

1 initialization;

2 RunPowerFlow();

3 setNodes = GetAllNodesInNetwork();

4 refNode = GetReferenceNode(setNodes);

5 **for** *each Feeder in Network* **do**

6   ListOfNodes = GetAllNodesInFeeder(Feeder);

7   endNode = GetEndNode(ListOfNodes);

8   electricalDistance = endNode.distance;

9   linearDistance = CalculateLinearDistance(refNode,endNode);

10   BooleanRadialFeeder = Feeder.isRadial;

11 **end**

<div align="center">Algorithm 5: Run geographical analysis study</div>

RunShortCircuitStudy: A short circuit calculation is performed for each feeder in the network. The short circuit impedance, resistance and reactance for the end node is obtained.

Data: Network

Result: Short circuit parameters of end node

1 initialization;

2 **for** *each Feeder in Network* **do**

3   SetStandardLoadConsumption();

4   ListOfNodes = GetAllNodesInFeeder(Feeder);

5   endNode = GetNodeWithLowestVoltage(ListOfNodes);

6   RunShortCircuitCalculation();

7 R_k = endNode.Rshc;

8 X_k = endNode.Xshc;

9 Z_k = endNode.Zshc;

10 **end**

<div align="center">Algorithm 6: Run short circuit study</div>

RunSumImpedanceStudy: The equivalent sum impedance of the feeder is calculated for the DER-scenarios *uniform* and *eof.*

Data: Network
Result: Specific feeder parameters

1 initialization;
2 for *each Feeder in Network* **do**
3   SetLoadConsumptionWeighted();
4   RunPowerFlow();
5   ListOfNodes= GetAllNodesInNetwork();
6   refNode = GetReferenceNode(ListOfNodes);
7   endNode = GetNodeWithLowestVoltage(ListOfNodes);
8   SumImpedanceWeighted =
      CalculateSumImpedance(refNode,endNode);
9   setPath = GetPath(refNode,endNode);
10   ListOfLines = GetAllLinesInFeeder(Feeder) cable =
      GetCableWithLowestRating(ListOfLines);
11   InomMinMainPathWeighted = cable.Inom;
12   SetStandardLoadConsumption();
13   RunPowerFlow();
14   ListOfNodes = GetAllNodesInNetwork();
15   refNode = GetReferenceNode(ListOfNodes);
16   endNode = GetNodeWithLowestVoltage(ListOfNodes);
17   SumImpedanceStandard =
      CalculateSumImpedance(refNode,endNode);
18   setPath = GetPath(refNode,endNode); cable =
      GetCableWithLowestRating(setCables);
19   InomMinMainPathStandard = cable.Inom;
20   RunSensitivityAnalysis();
21   dvdP = endNode.dvdP;
22   dvdQ = endNode.dvdQ;
23 end

Algorithm 7: Run sum impedance study

RunHCStudy: This algorithm calculates the hosting capability considering all DER-scenarios and voltage levels for a PF of 1 and 0.9 (ind.) for the activated network. The function CalculateHostingCapability utilized the secant method (newton-raphson) for faster convergence. Thereby two scaling factors are calculated: one for the estimated voltage rise and one for the estimated loading. The lower scaling factor then is used for the next interation.

Data: Network
Result: Hosting Capability results for a given scenario

1 initialization;
2 voltageScenario = 1.01:0.01:1.08;
3 ListOfPV = GetAllPVInstallations(Network);
4 setAllPVInstallationsToZero(ListOfPV);
5 **for** *each voltageScenario* **do**
6   **for** *each Feeder in Network* **do**
7   ListOfPV = GetFeederPVInstallations(Feeder);
8   endNodePV = GetPVInstalaltionAtEndNode(ListOfPV);
9   setAllInstallationsToZero(setPV);
10   setInjectedPVPower(ListOfPV,'uniform');
11   setPowerFactor(ListOfPV,1);
12   CalculateHostingCapability(ListOfPV,voltageScenario);
13   setPowerFactor(ListOfPV,0.9);
14   CalculateHostingCapability(ListOfPV,voltageScenario);
15   setAllInstallationsToZero(ListOfPV);
16   setInjection(ListOfPV,'weighted');
17   setPowerFactor(ListOfPV,1);
18   CalculateHostingCapability(ListOfPV,voltageScenario);
19   setPowerFactor(ListOfPV,0.9);
20   CalculateHostingCapability(ListOfPV,voltageScenario);
21   setAllInstallationsToZero(ListOfPV);
22   setInjection(endNodePV,'eof');
23   setPowerFactor(endNodePV,1);
24   CalculateHostingCapability(endNodePV,voltageScenario);
25   setPowerFactor(endNodePV,0.9);
26   CalculateHostingCapability(endNodePV,voltageScenario);
27   setAllInstallationsToZero(endNodePV);
28   **end**
29 **end**

Algorithm 8: Run hosting capability study

RunVoltVArStudy: This algorithm calculates the hosting capability considering the DER-scenario uniform and all voltage levels with an activated VoltVAr control for the activated network. The function Calculate-VoltVArHostingCapability utilized the secant method (newton-raphson) for faster convergence. Thereby to scaling factors are calculated: one for the estimated voltage rise and one for the estimated loading. The lower scaling factor then is used in the next step.

Data: Network
Result: Hosting Capability results for a given scenario

1 initialization;
2 voltageScenario = 1.01:0.01:1.08;
3 ListOfPV = GetAllPVInstallations(Network);
4 setAllInstallationsToZero(ListOfPV);
5 distributionScenarios = 'uniform','weighted','eof';
6 **for** *each voltageScenario* **do**
7   **for** *each Feeder in Network* **do**
8   **for** *each scenario in distributionScenarios* **do**
9   ListOfPV = GetFeederPVInstallations(Feeder);
10   setPVInjection(ListOfPV,scenario);
11   CalculateVoltVArHostingCapability(ListOfPV,voltageScenario);

12   WriteAllFeederResults();
13   setAllPVInstallationsToZero(ListOfPV);
14   **end**
15   **end**
16 **end**

Algorithm 9: Run hosting capability study (VoltVAr)

MainAlgorithm: After the definition of all subscripts, the overall algorithm to calculate all parameters and the hosting capabilities for a particular scenario (DER-scenario, voltage limit, reactive power control strategy) is given in algorithm 10. Starting with an initialization (reading yearly consumption data), for each network a plausibility check is performed. After that the feeders are defined and the secondary transformer is replaced by a LV-slack. Next, for each feeder HC scenario independent analysis are started to gather relevant statistical parameters. Finally, for each voltage, control scheme and DER-scenario, the HC is calculated. The developed algorithm, was implemented in DIgSILENT Programming Language (DPL). The reader will note, that this algorithm contains subscripts that can be run independently from each other. Hence, skipping parts of the main algorithm can be easily implemented by commenting out the appropriate function. Moreover, the analysis can be performed for specific voltage limits only.

Data: Network data
Result: HC and parameters for defined scenarios

1 initialization;
2 **for** *each network* **do**
3   INVALID = CheckPlausibility();
4   **if** *INVALID == 1* **then**
5   │ *continue*;
6   **end**
7   DefineFeeders();
8   SetSlackToLV();
9   RunStatisticalAnalysis();
10   RunGeographicalAnalysis();
11   RunShortCircuitStudy();
12   RunSumImpedanceStudy();
13   RunHCStudy();
14   RunVoltVArStudy();
15   DeleteCustomSlacks();
16 **end**

Algorithm 10: Hosting capability and parameter calculation

The execution of the algorithm is depicted in Figure 122. Thereby the networks of the two DSO are assessed from the PowerFactory Database. Further the annual consumption data of loads needed for the DER-scenario weighted is read. Thanks to the modular scripting of the tasks, each simulation scenario can be performed independently of other scenarios. This is achieved by storing simulation results per network and simulation case. For example, the HC calculation for a particular voltage limit and DER-scenario can be recalculated for a single network, if needed. The results of each simulation case are stored in a specific text file (comma-separated), where the particular results for each feeder are stored in one line. In a second step, the txt-files are aggregated per DSO into a matrix in the correct order of networks and feeders.
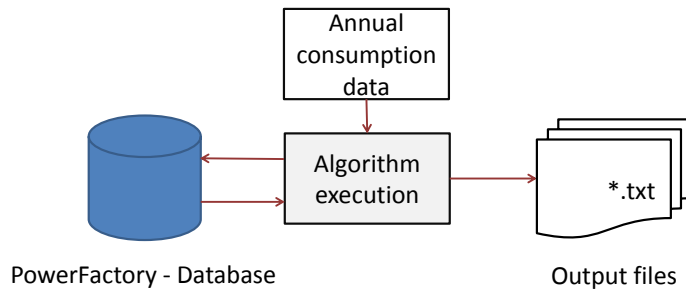
Figure 122: Algorithm execution

H   Methodology Validation with Monte-Carlo Simulations for Selected Feeders

In this section, the detailed results of the investigation of real and equivalent feeders with Monte-Carlo simulations are presented. The closeness centrality descibred in section 3.5 was utilized to assign the consumption data of loads in the real feeder to the most appropriate load location in equivalent feeders. The investigated feeders can be found in [62].

The individual results are depicted from Figure 124 to 149. For that purpose, x-y plots have been utilized to present the results for each power flow for the real and equivalent feeders. Therefore, points on the red line indicate a perfect match between real and equivalent feeders. Points above the median mean that the results of the equivalent feeders are overestimated compared to the real feeders. Respectively, points below the median indicate that results are underestimated compared to the real feeders.
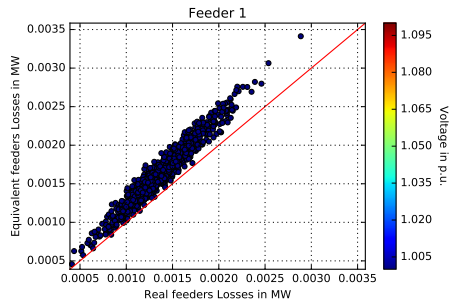
Figure 123: Feeder 1 - Active power losses comparison
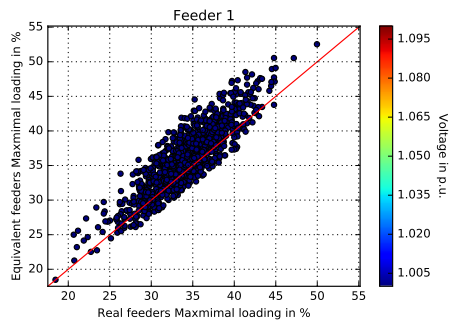
Figure 124: Feeder 1 - Comparison of the maximal loading in the feeder
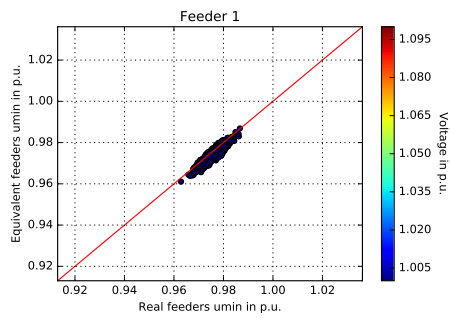
Figure 125: Feeder 1 - Comparison of the voltage drop in the feeder
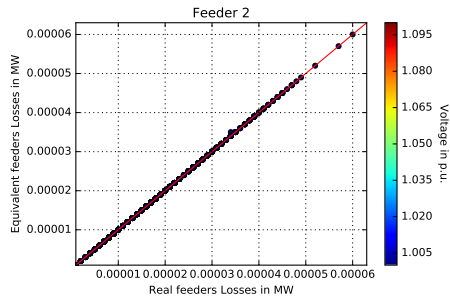
Figure 126: Feeder 2 - Active power losses comparison
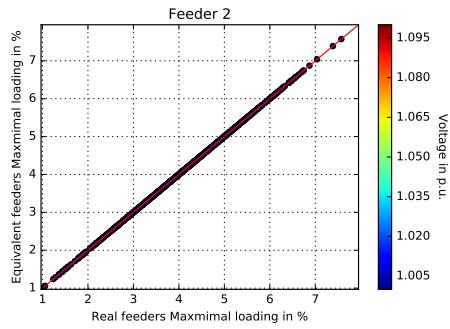
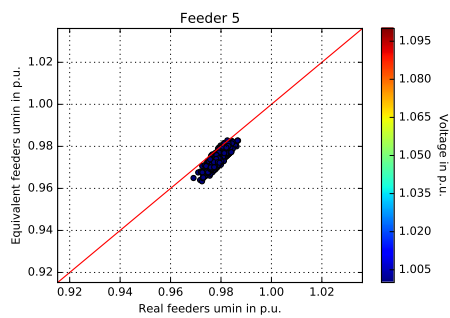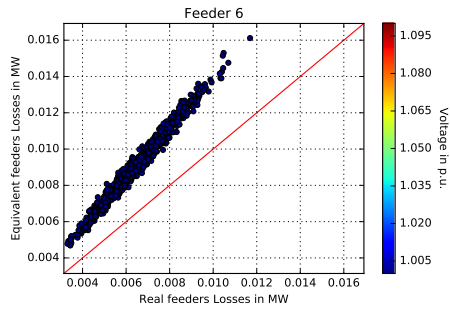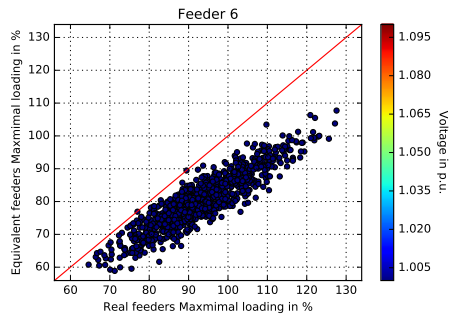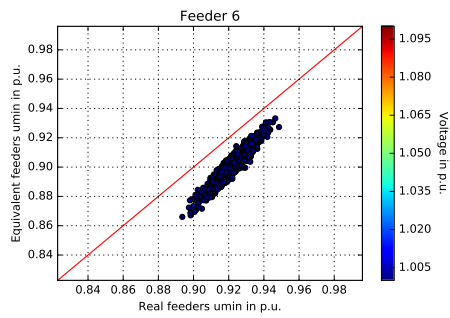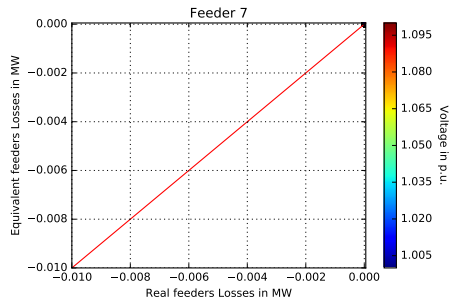Figure 127: Feeder 2 - Comparison of the maximal loading in the feeder
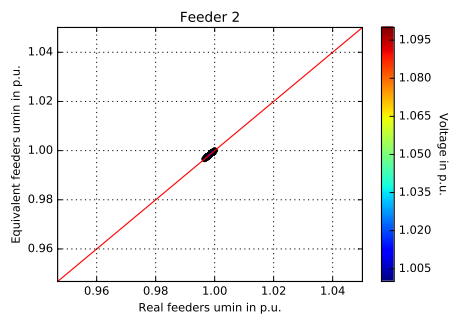
Figure 128: Feeder 2 - Comparison of the voltage drop in the feeder
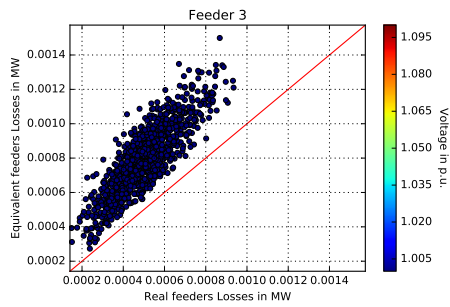
Figure 129: Feeder 3 - Active power losses comparison
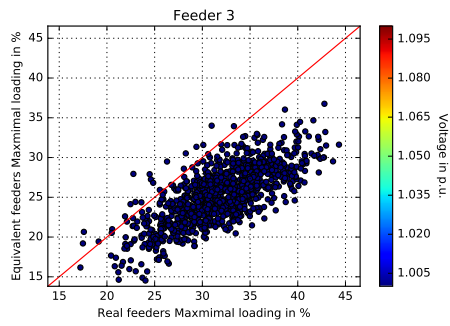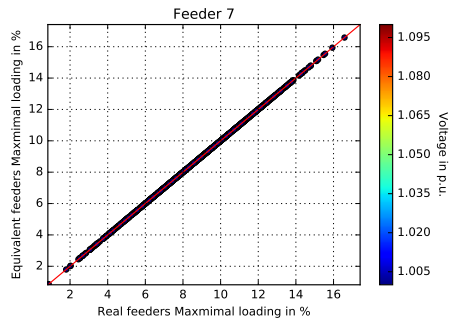

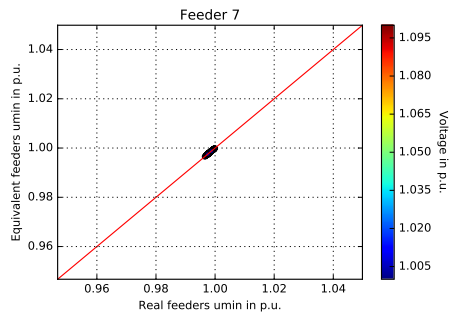Figure 130: Feeder 3 - Comparison of the maximal loading in the feeder
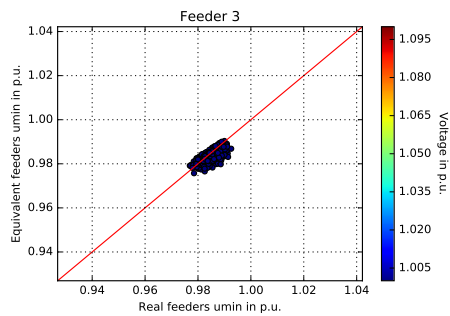
Figure 131: Feeder 3 - Comparison of the voltage drop in the feeder
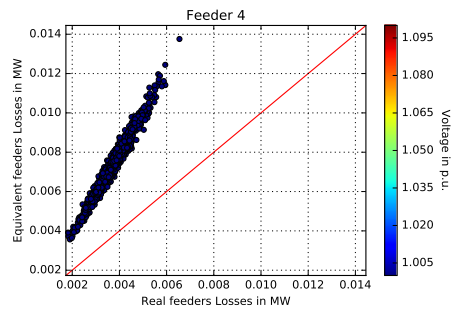
Figure 132: Feeder 4 - Active power losses comparison
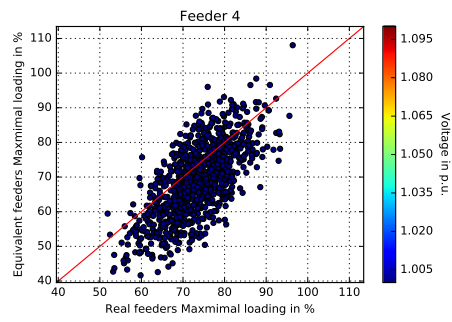
Figure 133: Feeder 4 - Comparison of the maximal loading in the feeder
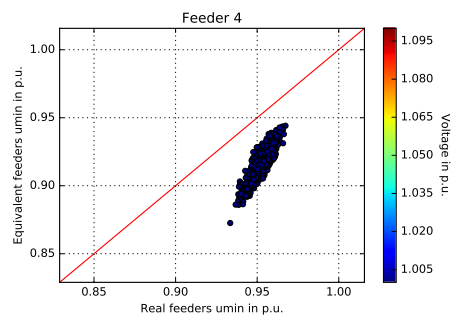
Figure 134: Feeder 4 - Comparison of the voltage drop in the feeder
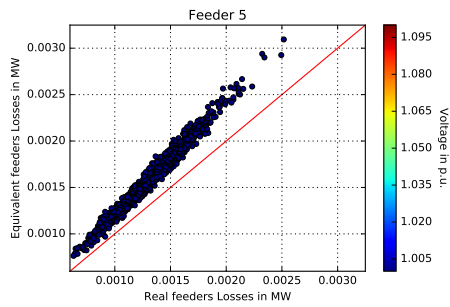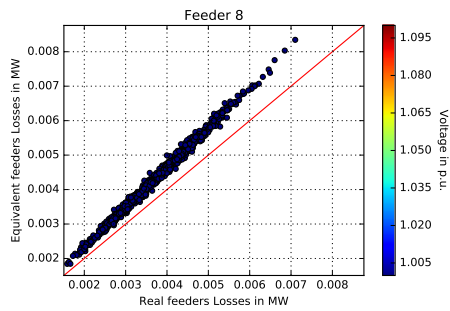
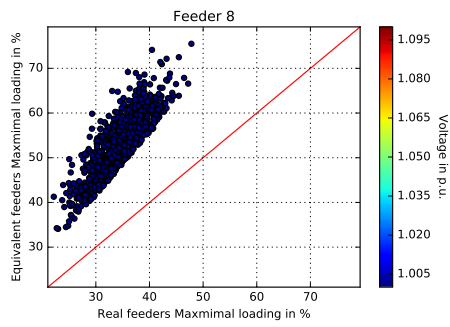Figure 135: Feeder 5 - Active power losses comparison
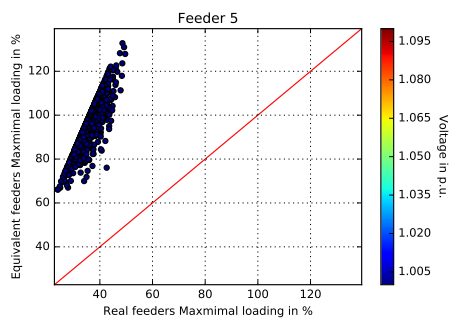
Figure 136: Feeder 5 - Comparison of the maximal loading in the feeder

Figure 137: Feeder 5 - Comparison of the voltage drop in the feeder

Figure 138: Feeder 6 - Active power losses comparison

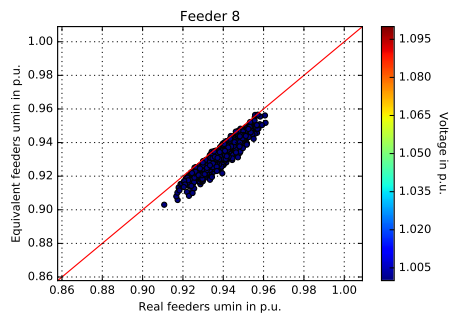Figure 139: Feeder 6 - Comparison of the maximal loading in the feeder

Figure 140: Feeder 6 - Comparison of the voltage drop in the feeder

Figure 141: Feeder 7 - Active power losses comparison

Figure 142: Feeder 7 - Comparison of the maximal loading in the feeder

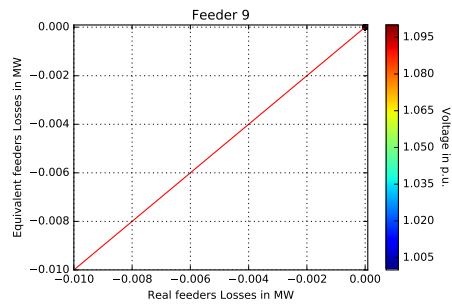Figure 143: Feeder 7 - Comparison of the voltage drop in the feeder

Figure 144: Feeder 8 - Active power losses comparison

Figure 145: Feeder 8 - Comparison of the maximal loading in the feeder

Figure 146: Feeder 8 - Comparison of the voltage drop in the feeder

181

Figure 147: Feeder 9 - Active power losses comparison
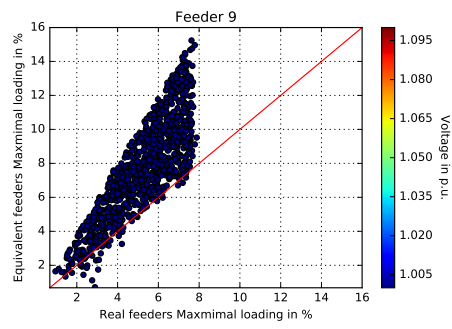

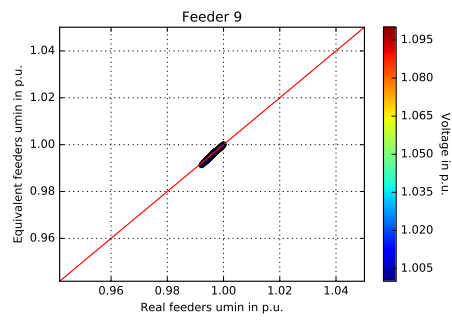Figure 148: Feeder 9 - Comparison of the maximal loading in the feeder

Figure 149: Feeder 9 - Comparison of the voltage drop in the feeder