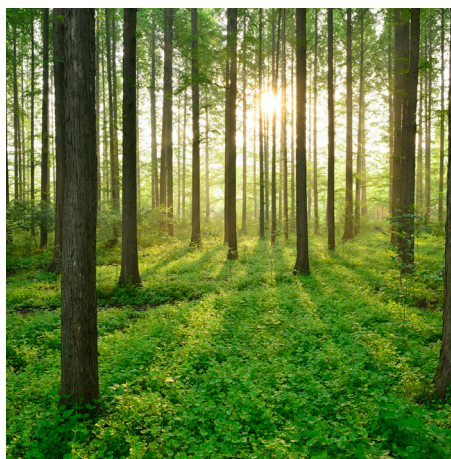


THE BENEFITS OF NEW SMART METERS AND UNBALANCED GRID MODELS FOR INCREASED DIGITALIZATION

RAPPORT 2026:1224



ELNÄTENS HÅLLBARA TEKNIK-
UTVECKLING OCH DIGITALISERING



THE BENEFITS OF NEW SMART METERS AND UNBALANCED GRID MODELS FOR INCREASED DIGITALIZATION

MARKUS ERIKSSON, MAHANTHA AMPAVATINA KAMBAGIRI, STEFAN STANKOVIĆ

ISBN 978-91-89917-67-5 | © Energiforsk October 2026

Energiforsk AB | Telefon: 08-677 25 30 | E-post: kontakt@energiforsk.se | www.energiforsk.se

Förord

Projektet *Nya smarta mätare och obalanserade nätmodeller* ingår i programmet *Elnätens hållbara teknikutveckling och digitalisering* med syftet att jämföra en traditionell symmetrisk modell med en avancerad som inkluderar fasobalans, för att visa hur snedbelastning påverkar förluster och effektivitet. Nätägare kan optimera resurser och förbättra elnätsdriften för ökad hållbarhet och flexibilitet.

Projektet tillgängliggör en modell som inkluderar fasobalans som kan användas i licensfria verktyg. Erik Weihs och senare Stefan Stancović från RISE har varit PL.

Stort tack också till programstyrelsen för deras initiativ och stöd till projektet:

- Olle Bergström, Jämtkraft Elnät
- Arne Berlin, Vattenfall Eldistribution
- Staffan Bjurulf, Sveriges Ingenjörer (MF)
- Magnus Brodin, Skellefteå Kraft Elnät (ordf)
- Emil Frostfeldt, Jönköping Energi Nät
- Josefin Grundius/Christer Flood, Ellevio
- Stefan Ivarsson, RISE
- Magnus Lindström, Grid Diagnose
- Per-Olov Lundqvist/Fredrik Andersson, Elinorr
- Karl-Johan Mannerback, Jönköping Energi Nät
- Giuseppe Martinelli, Svenska kraftnät
- Tilda Nordin, Mälarenergi Elnät
- Dennis Ossman, Göteborg Energi Elnät
- Johan Ribrant, Nacka Energi
- Göran Sandström, Umeå Energi Elnät
- Peter Silverhjärta, Energiföretagen Sverige
- Magnus Sjunnesson, Öresundskraft Elnät
- Claes Wedén, Hitachi Energy Sweden

Följande bolag har deltagit som intressenter till projektet. Energiforsk framför ett stort tack till samtliga för värdefulla insatser.

Ellevio	Karlstads El- och Stadsnät	Blåsjön Nät
Vattenfall Eldistribution	Borås Elnät	Dala Energi Elnät
Svenska kraftnät	Falu Energi & Vatten	Elektra Nät
Göteborg Energi	Borlänge Energi	Gävle Energi
Statkraft Sverige	Nacka Energi	Hamra Besparingsskog
Fortum	C4 Energi	Hofors Elverk
Mälarenergi Elnät	PiteEnergi	Härjeåns Nät
Öresundskraft Elnät	Trollhättan Energi Elnät	Härnösand Elnät
Tekniska Verken i	Skövde Energi	Ljusdal Elnät
Linköping	Hitachi Energy Sweden	Malungs Elnät
Skellefteå Kraft Elnät	Energiföretagen Sverige	Sandviken Energi Nät
Umeå Energi Elnät	Sveriges Ingenjörer, MF	Sundsvall Elnät
Jämtkraft Elnät	GridDiagnose	Söderhamn Elnät
Jönköping Energi Nät	RISE	Åsele Elnät
Eskilstuna Strängnäs	Elinorr ekonomisk förening;	Årsunda Kraft & Belysningsf
Energi & Miljö	Bergs Tingslags Elektriska	Övik Energi Nät

Här redovisas resultat och slutsatser från ett projekt inom ett forskningsprogram som drivs av Energiforsk. Det är rapportförfattaren/-författarna som ansvarar för innehållet.

Stockholm, mars 2026

Susanne Stjernfeldt, Energiforsk

Sammanfattning

Nya smarta mätare och den pågående digitaliseringen av distributionssystemen ger operatörerna möjlighet att få en mer detaljerad och realtidsnära bild av systemdriften. Övergången till grön energi och hållbara transporter medför fler solceller och elbilar i distributionssystemen och med dem nya utmaningar. Några av dessa är obalanserad drift och behovet av att mer exakt uppskatta systemets nuvarande och framtida kapacitet. Avreglering och öppna marknader har fört med sig många olika leverantörer och distributionssystemoperatörer. Detta medför utmaningar med att dela systemmodeller, information om systemdrift och planering samt bred tillgänglighet av viktiga verktyg för analys av distributionssystem. Analys- och modelleringsverktyg med öppen källkod samt standardiserade format för informationsutbyte kan erbjuda lösningen.

Detta projekt syftar till att utnyttja de nya möjligheter som digitaliseringen och nya smarta mätare medför genom att även använda verktyg med öppen källkod för att lösa några av de utmaningar som är relaterade till drift och planering av distributionssystem. I ett specifikt fall av ett svenskt distributionssystem har två studier med smarta mätardata och modelleringsverktyg med öppen källkod genomförts. Den första studien undersökte, identifierade och karakteriserade de icke-tekniska systemförlusterna. Den andra studien uppskattade systemets PV kapacitet under olika scenarier för PV-integration och obalanserad systemdrift. Båda studierna illustrerade potentialen att lösa de väsentliga systemutmaningarna med hjälp av licensfria verktyg med öppen källkod och tillgängliga 15-minutersmätningar från smarta mätare. Projektresultaten banar väg för vidare diskussioner om fördelarna med digitalisering av distributionssystem och användning av verktyg med öppen källkod och filformat för att modellera, analysera och utbyta information om systemets planering och drift.

Nyckelord

digitalisering, distributionsnäts drift och planering, distributionsnäts kapacitet, icke-tekniska förluster, lokala elnät, obalanserade modeller, smart mätare, öppet källkodverktyg

Summary

New smart meters and the ongoing digitalization of the distribution systems are giving the operators opportunity to get more detailed and close-to real time picture of the system operation. Transition toward green energy and sustainable transport brings more PVs and EVs in the distribution systems and with them new challenges. Some of these are unbalanced operation, and the need to more precisely estimate the system current and future hosting capacity. Deregulation and open markets have brought many different vendors and distribution system operators. This brings challenges with sharing the system models, information about system operation and planning as well as wide availability of the essential tools for distribution system analysis. Open-source analysis and modelling tools as well as standardized formats for exchanging information may offer the solution.

This project aims to leverage on the new capabilities brought up by digitalization and new smart meters by also using open-source tools to solve some of the challenges related to distribution system operation and planning. On a specific case of a Swedish distribution system two studies using smart-meter data and open-source modelling tools have been done. The first study investigated, identified and characterized the non-technical system losses. The second study estimated the system's PV hosting capacity under different PV integration scenarios and unbalanced system operation. Both studies illustrated potential to solve the essential system challenges using license-free open-source tools and available 15-minutes measurements from smart-meters. The project results pave the way for further discussions on benefits of distribution systems digitalization and using of open-source tools and file formats to model, analyze and exchange information about the system's planning and operation.

Table of contents

1	Introduction	7
1.1	Project scope	8
1.2	Project relevance	8
1.3	Report's strUcture	9
2	Background and "state of the art"	10
2.1	Open_DSS	10
2.1.1	Solution Capabilities	10
2.1.2	User Interface and Integration	11
2.1.3	Applications	11
2.2	Sharing grid data using CIM	11
2.2.1	CIM to OpenDSS export	12
2.3	Smart meters and improved system observability	13
2.4	Unbalanced operation of distribution systems	13
3	Detection of non-technical losses using smart-meter data	15
3.1	Methodology	15
3.1.1	Categorization of NTL	15
3.1.2	Localization of NTL	16
3.2	Case study	18
3.2.1	Setup and description	18
3.2.2	Results24	
3.3	Discussion	34
4	Impact of the unbalanced operation on the hosting capacity of the system	35
4.1	Methodology	35
4.2	Case Study	38
4.2.1	Setup and description	38
4.2.2	Results39	
4.3	Discussion	42
5	Conclusions	44
6	References	45

1 Introduction

Distribution systems are undergoing fast digitalization. New smart meters with 15 minutes readings are giving the operators opportunity to get more detailed and close-to real time picture of the system operation. More and more PV generation and EV chargers are being hosted in the distribution systems, increasing the need for a proper estimation of the distribution systems' hosting capacity. One of the factors affecting the hosting capacity is unbalanced distribution system operation that requires more detailed analysis and modelling of all three phases and their couplings. With many vendors and distribution system operators, the challenges of sharing information about system modelling, analyses and operation arise. Open-source analysis and modelling tools as well as standardized formats for exchanging information may offer the solution. The mentioned challenges and opportunities come along with distribution systems digitalization, higher penetration of renewables and EV chargers. This project aims to address some of these challenges and leverage on the possible opportunities.

A major part of distribution systems' digitalization was and is introduction of smart meters. Smart meters in Sweden have been introduced in 2003, before which electricity metering was done on a yearly basis [1]. From 2003 to 2023, smart meters went from monthly, hourly to finally 15 minutes metering requirements. While major reasons for their initial introduction in Sweden were to increase consumer awareness and reduce energy consumption, it also led to a better observability of the distribution system operation. This allowed for better understanding of the load profiles, better forecasting of the load growth and better planning of current and future distribution systems' operation.

Smart meter data can be used to better assess hosting capacity of the distribution system under different future scenarios but also for the system monitoring applications such as system's parameters estimation or identifying deviations from predicted system operation. An example of such deviations is appearance of non-technical losses in the system. With smart meters that can nowadays record 15 minutes (and in the future maybe even more frequent) samples of currents and voltages on each of three phases, possibilities to monitor the distribution system operation as well as identify deviations from the predicted system behaviour are substantial as also shown in this project.

Distribution systems, especially local, low voltage ones, operate usually with unbalance over their three phases. To get an accurate picture of the system, these unbalances need to be considered in mathematical/computer models/representations of the real system as well as in the methods used to analyze these models. As some previous works showed [2] [3] [4], inappropriate modelling can lead to the errors in estimation of the system's hosting capacity. While three phase unbalanced operation brings complexities in system modelling and analysis, it also opens new opportunities in detecting and identifying the

undesirable system behaviour such as appearance of non-technical losses. With smart meters' measurements of currents and voltages per phases, it is possible not just to get better awareness of the unbalanced distribution system operation but also to capture footprints of different end-users connected to the distribution grid and their unique unbalanced behaviour.

To address the challenges mentioned and exploit possible opportunities of the unbalanced system operation analysis, tools that have ability to make three-phase unbalanced representations of the distribution systems need to be used. With many smaller and bigger distribution system operators in Sweden, it is important that all of them have appropriate and transparent tools for the unbalanced system analysis and planning. One of the solutions may be to use free and open-source tools as investigated in this project. Furthermore, cooperation and coordination between many different vendors and distribution system operators raises the need for standardized and transparent data and information exchange. To address this need, IEC has introduced Common Information Model (CIM) [5].

1.1 PROJECT SCOPE

The project aims to support distribution system operators in developing and making the distribution system operation more efficient by taking advantage of the opportunities brought by digitalization. Concretely, the project aims to showcase the use of open-source tools together with CIM on analysis of unbalanced operation of a distribution system. Furthermore, the project shows how smart meter data can be used to improve efficiency of the distribution system operation by detecting the non-technical losses.

The project aims to on a selected case study of a Swedish distribution system address the following objectives:

- Showcase three-phase modelling of a Swedish distribution system using open-source tools and data formats;
- Use open-source tools to analyse the unbalanced system operation, compare it to the real smart meter measurements to identify and minimize non-technical losses in the system;
- Use open-source tools to investigate how unbalanced operation can affect the hosting capacity of the system in the case study of PV integration in a Swedish distribution system

1.2 PROJECT RELEVANCE

The project offers several concrete benefits for electricity grid companies and their customers by creating solutions that improve the operation and planning of the power grid. By using open-source (license-free) network models, transparent and flexible modelling is ensured, allowing for adaptation and further development without being locked into specific vendors or tools. Furthermore, the project also aims to develop and use solutions that can be integrated with the network

information systems of the electricity grid companies, enabling process automation.

The project contributes to:

- Fairer electricity distribution: Reduces imbalances and uneven load.
- Reduced non-technical losses: Provides better control and minimizes losses that negatively affect both network owners and customers.
- More efficient analysis tools: License-free and open tools that support decision-making and future development.
- Knowledge sharing: The results can be used by other network owners, promoting collaboration and innovation within the industry.

1.3 REPORT'S STRUCTURE

This report is divided into three separate parts. The first part gives a general background to the work in this project and explains the tools and state-of-the art methods used in the analyses in this project. The second part focuses on how smart-meter data and modelling of the power system in the open-source tools can be used to detect, locate and characterize the non-technical losses in the system. The third part of the report deals with use of open-source tools to analyze the hosting capacity of the distribution system. The hosting capacity is assessed with regards to different factors including maximum allowed power of PV connections, order of locations in which PV systems are connected to the system as well as unbalanced vs balanced operation of the system. The report finishes with the conclusions and the main take-aways from the project results.

2 Background and “state of the art”

2.1 OPEN_DSS

OpenDSS (Open Distribution System Simulator) is a freely available simulation tool for electric power distribution systems made as open source by EPRI in 2008 with regular updates [6]. It was originally developed to support distributed generation interconnection studies and has since evolved into a comprehensive analysis platform for distribution network modelling and planning. It is a script-driven circuit solution engine capable of performing a wide range of steady-state and quasi-dynamic analyses. It is well suited for modern distribution systems with high penetration of distributed energy resources.

OpenDSS is primarily intended for analysis of unbalanced, multi-phase distribution systems [7]. It was designed to handle the North American distribution system networks, which are often highly unbalanced requiring more than a positive-sequence model. It can also model European-style systems, which typically have simpler structures. OpenDSS uses actual voltages and currents (not in per-unit domain) and three-phase detail modelling of the system accounting for unbalanced parameters between the phases and their couplings. It does not internally use symmetrical components or per-unit representation, but can provide this as an output.

2.1.1 Solution Capabilities

Solving the power flow is the most common application of the OpenDSS software. However, OpenDSS is designed as a power system harmonics analysis tool. Being able to primarily do harmonics analysis, it can also relatively simply find the power flow solution. Comparably the tools designed to solve classic power flow problems might not be able to solve a harmonics problem [8].

OpenDSS supports a wide range of RMS steady-state analyses commonly required for distribution system planning, in addition to advanced simulation modes. Its solution capabilities include:

- Static power flow for unbalanced multi-phase power flow
- Quasi-dynamic time-series for daily and yearly time-series simulations
- Duty cycle simulations (1–5 s resolution)
- Fault analysis
- Harmonic analysis
- Dynamics analysis (electromechanical transients)

It is important to note that OpenDSS is not an electromagnetic transients (EMT) program, but it can simulate electromechanical dynamics [9].

2.1.2 User Interface and Integration

OpenDSS is distributed as a standalone executable (OpenDSS.exe) with a text-based, script-driven interface [7]. It provides no graphical input and a limited graphical output. However, this script-based structure makes it well suited for automation and batch simulations. The tool can be integrated with external platforms for extended functionality. For example, the py-dss-interface [10] (which can be augmented with the py-dss-toolkit [11]) provides a python API for interfacing with OpenDSS. This is particularly useful for enabling automated workflows and analysing results. There is also a MATLAB integration available through the GridPV Toolbox [12], which is designed to model and simulate the integration of distributed generation into the electric power system and to assess the impacts on distribution networks from highly variable generation.

2.1.3 Applications

OpenDSS is suited for studies related to Distributed Energy Resource (DER) integration and advanced distribution system planning [7]. Originally developed for distributed generation interconnection analysis, it remains well suited for evaluating the impact of renewable and distributed energy resources on existing distribution networks. The tool can model virtually any distribution circuit configuration and is designed to allow for easy expansion.

A major application area is DER integration studies. A common use case for OpenDSS is therefore to analyse photovoltaic (PV) systems both on a utility and residential rooftop scale, wind turbines, electric vehicles and battery energy storage systems [13]. An example is that through quasi-dynamic time-series simulations, the impact of variable generation and variable load behaviour on feeder voltages, currents, and power flows can be evaluated.

Typical analyses for OpenDSS include studying voltage rise behaviour, thermal overloading of equipment, harmonic distortion, and dynamic or islanding behaviour [13]. Hosting capacity analysis, defined as the maximum amount of DER that can be accommodated in a distribution system without violating operational constraints or requiring infrastructure upgrades, is a usual use case. By incrementally increasing DER penetration and monitoring system performance, acceptable integration limits can be determined.

Beyond DER integration, OpenDSS supports a wide range of planning and advanced control studies [13]. It can be used to evaluate time-varying feeder behaviour, perform load allocation and load growth studies and conduct system reliability assessments. Furthermore, it is able to provide advanced distribution system controls studies such as distributed energy resource management systems studies, Volt-VAR optimization, conservation voltage reduction, and fault location, isolation, and service restoration studies.

2.2 SHARING GRID DATA USING CIM

The Common Information Model (CIM), defined by the IEC 61970, 61968 and 62325 standards, provides a standardized way to represent power system

components and their relationships. An in-depth explanation of CIM is outside the scope of this report and readers seeking a more comprehensive description are referred to [5] and [14] published by PNNL. However, in brief, CIM is designed to facilitate interoperability and data exchange by defining a standardized semantic model for power systems.

CIM describes the structure of a power system in terms of network topology, equipment characteristics, and electrical connectivity, i.e. CIM captures what exists in the network and how components are related. The CIM model is formally defined using Unified Modeling Language (UML). UML class diagrams are used to specify the CIM classes, their attributes, inheritance relationships, and associations with other classes. For data exchange, CIM models are commonly serialized using XML-based formats. These formats provide a standardized way to encode CIM instances in files that can be exchanged between tools. The XML documents contain instances of CIM classes and their relationships as defined in the UML model.

2.2.1 CIM to OpenDSS export

By sharing grid data through CIM, it is possible to automate the modelling process, reducing manual modelling effort and improving model consistency as well as accuracy. In this project, a proof-of-concept conversion script for OpenDSS export was developed. The conversion script acts as a translation layer between CIM XML files and OpenDSS syntax. It is responsible for parsing the CIM XML files and extracting relevant equipment and topology information. The script should then generate structured OpenDSS files based on the parsed information which can be executed with minimal additional interventions. The resulting OpenDSS model reproduces the network topology and core electrical characteristics as defined in the original CIM dataset. This has the possibility to enable a faster model setup for distribution system studies as well as improving repeatability and reduces manual modelling errors.

It is important to note that the overlying grid information is not included in the CIM model. Therefore, additional data must be provided outside the CIM XML files to define the Thevenin equivalent of the overlying network in OpenDSS. Specifically, one of the following is required to define the Thevenin equivalent in OpenDSS:

1. Three phase and single-phase short circuit power (MVA_{sc3} and MVA_{sc1})
2. Three phase and single-phase short circuit current (I_{sc3} and I_{sc1})
3. Positive and zero-sequence impedance ($Z_1 = [R_1, X_1]$ and $Z_0 = [R_0, X_0]$)

In addition, performing time-series simulations (for example, quasi-dynamic simulations) requires operational data for the studied period. Even for static simulations, additional information about load and generator operating points may be necessary. This information is possibly not contained in the CIM model and an additional source of external data, specifically operational data, may therefore need to be incorporated. The main thing to keep in mind with the

operational data is that the naming lines up with the naming in the CIM model in order to correlate the operational data with the correct bus.

A feature that has been considered, but not tested, is the automated creation of a graphical representation of the model in a single-line diagram. As the OpenDSS model only contains the connection information of the network elements along with their electrical parameters, there are not any schematic overview of the network. To create a single-line diagram, relative bus coordinates are required. Since the CIM model contains geographical coordinates for the buses, these can be parsed by the file conversion script and used to automatically generate the graphical network representation.

2.3 SMART METERS AND IMPROVED SYSTEM OBSERVABILITY

Smart meters are digital measurement devices that record electrical operational data with high temporal resolution. Smart meters measure voltage, current, active and reactive power, and sometimes frequency and power factor. While conventional (non-smart) meters only capture energy consumption with no information of when the energy was consumed, smart meters provide a way of measuring electricity consumption in near real-time.

The adoption of smart meters improves system observability in distribution networks by providing insights into demand patterns, voltage levels, and system loading conditions across the network. This helps with enabling more accurate billing but also supports more efficient network operation. For example, by better understanding grid demand through the smart meters, demand-side-management becomes viable to incentivise shifting consumption away from peak periods and thereby using the grid more efficiently.

Smart meter data can be used in several advanced applications. One example is the detection of non-technical losses, such as improper metering or electricity theft. By comparing measured consumption and voltage profiles from smart meters with simulation results from network models, any present non-technical losses can be identified. A detailed example of this is presented in Section 3.

In addition, smart meter data supports the integration of renewable energy sources more effectively. For instance, it can be used in hosting capacity studies for which Section 4 will provide an example. Enhanced system observability through smart meters also improves system state estimation, where a best estimate of network conditions is derived from the available measurements, reducing uncertainties caused by missing data or large measurement errors.

2.4 UNBALANCED OPERATION OF DISTRIBUTION SYSTEMS

In traditional three phase power system analysis, the underlying assumption for using the power flow methods is that the system is balanced. Such an assumption simplifies the computation by allowing us to use a single-phase equivalent representation of the network. However, unbalanced systems are often encountered in real life, distribution systems in specific. To analyze unbalanced

power systems, all three phases need to be modelled separately in some cases or special decomposition methods need to be used.

Firstly, we define what balanced and unbalanced power systems are. A balanced power system is a system where the three phase voltages and currents are perfectly balanced, which means that their magnitudes are equal and are shifted from each other in phase by exactly 120 degrees. Such a system can be represented by an equivalent single-phase system. On the other hand, an unbalanced power system is a system where there is an unbalance in the magnitudes and (or) phases of the voltages and currents. Unbalanced conditions may arise due to several factors such as faulty voltage sources, unequal system parameters, or unequal loads. Depending on the type of unbalance, different methods and approaches are used for simulation and analyses of unbalanced systems.

In cases where system parameters per phase are symmetrical but power flows per phases are unbalanced (for example due to unbalanced loads per phases), decomposition methods such as Fortescue method [15] can be used. Also known as symmetrical components method, it decomposes unbalanced system to be analyzed into three balanced sets: positive, negative and zero sequence [16]. Each sequence is represented as one-phase equivalent model of the system and solved independently from the other sequences. When all sequences are solved, the results can then be recombined giving the solution to the original unbalanced system. For a balanced power system, only the positive sequence component exists. In case of an unbalance in the system, all three components may appear.

If the system parameters are not symmetrical per phases, Fortescue method cannot be used. In this case, the system needs to be modelled in more detail using its three-phase representation. Each phase is described by specific, possibly different (phase-asymmetrical) parameters as well as coupling parameters between the phases. The system model, containing specifics to each phase and their coupling, is therefore expanded compared to the balanced single-phased representation by $3 \times 3 = 9$ times. As a consequence, solving the three-phase model of the system will require around 9 times more computation resources to solve than the balanced one-phase model of the system.

The advantage of using the Fortescue method to analyze unbalanced systems compared to full three-phase modelling of the system is that it reduces the computational complexity of analysis from factor of 9 to factor of 3 (compared to the balanced system analysis). As Fortescue method decomposes the system into three sequences that can all be represented as single phase equivalents, the computational complexity of analyzing unbalanced system increases 3 times compared to analyzing the balanced system (where only one positive sequence exists instead of possibly all three).

3 Detection of non-technical losses using smart-meter data

Non-Technical Losses (NTL) are any losses related to improper metering, meter tampering or power theft. By increasing the observability of the network through smart meters, methods become available for identifying and locating any NTL in a network. This section provides a detailed description of how smart meters can be used in combination with a network simulation model, here employing the open-source simulation tool OpenDSS, for the detection of NTL.

3.1 METHODOLOGY

To detect the presence of NTL in a network, smart meter measurements are compared with results from a network simulation model. Significant deviations between measured losses and simulated losses indicate that there may be NTL in the network unaccounted for in the simulation environment. By then further analysing how the deviations, if present, behaves, insights can be obtained regarding the characteristics of the NTL. Assuming the NTL is a load that is not captured by a meter, the applied methodology proceeds with a categorization of the NTL to gain insight into the type of connected load. The presented methodology also describes how NTL may be located to determine where in the network it originates.

3.1.1 Categorization of NTL

When categorizing the NTL, the network loss was analysed in three ways. These analyses are referred to as the NTL-load profile the NTL-transitions and the NTL-temperature correlation.

NTL-load Profile

To calculate the NTL-load profile, the regression of the expected losses was computed based on the simulations. The difference between actual losses and the expected losses gives the NTL-loading, i.e. the loading of the unmetered load connected to the network. Since the network losses are tied to specific timestamps, this allows for the derivation of a time series of the NTL-load.

To investigate whether a daily pattern exists in the loading, the average NTL load was calculated for each time step of the day. For example, when using a 15-minute resolution time series, all NTL load values for the datapoint 00:00 - 00:15 across the available days were averaged to obtain the first data point in the daily load profile. This procedure was repeated for each time step of the day, resulting in a complete average daily profile.

NTL-transitions

Transitions refers to datapoints where the NTL appears to be increasing or decreasing and are therefore transitioning to a higher or lower value. By studying the transition points, further insights into the NTL pattern can be derived. By grouping transition points by time of day, day of the week, and the month, as well

as keeping track of transitions leading to an increase or to a decrease, distribution can be derived showing when the NTL has a high respective low variability, which could provide insights in how the NTL is being controlled.

Temperature corelation

The final characteristics of the NTL analysed where how it correlated to the outside air temperature, which could help in identifying what type of load had gone unmetered. The correlation coefficient between the NTL-load and outside air temperature was calculated as

$$r(A, B) = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{A_i - \mu_A}{\sigma_A} \right) \left(\frac{B_i - \mu_B}{\sigma_B} \right) \quad (1)$$

where A_i is the NTL-load at time instant i , μ_A is the average NTL-load, and σ_A is its standard deviation. Temperature is similarly represented by B_i , μ_B , and σ_B while N is the total number of time instances.

Temperature data was obtained from the closest SMHI temperature metering station to the network location with a similar elevation above the sea. Correlations were calculated for seasonal, daily, and hourly variations. For seasonal variations, a 30-day rolling average filter was applied to both datasets to isolate the seasonal trend. For daily and hourly variations, the data was resampled to the appropriate resolution, and the seasonal component (the 30-day rolling average) was removed to focus on short-term fluctuations.

3.1.2 Localization of NTL

In order to identify where in the network the cause of the NTL is located, the three methods for NTL localization presented in [17] were applied. These methods are based on the principle that an NTL introduces a localized and unexpected voltage drop in the network. The three methods presented are:

Voltage Drop at Measurement Point, the error between measured and simulated voltage can indicate the location of NTL. This error is represented by calculating the Directional Mean Error (DME) over all time instances classified as having unusual high losses. The DME is calculated as:

$$DME = \frac{1}{N} \sum_{i=1}^N (Sim_i - Meas_i) \quad (2)$$

Where N is the number of data points, while Sim_i and $Meas_i$ are the simulated voltage and the measurement voltage at time instant i . A positive DME suggests that the measured voltage is consistently lower than the estimated voltage when losses are high, which may indicate NTL.

Correlation: Voltage - Voltage, the correlation between measured voltage and simulated voltage is calculated. The corelation is calculated according to (1), but with A and B being the measured and simulated voltage instead. Ideally, this correlation should be close to one and deviations indicates that the voltage time series have different behaviour which could be due to NTL.

Correlation: Voltage - Energy, this method examines whether a change in power at an energy meter corresponds to a change in voltage. The measured correlation and the simulated correlation are compared, and any significant deviation may indicate NTL.

Handling Voltage Noise

Since all methods rely on voltage comparisons between measurements and simulations, it is important that the voltage from the simulation is accurate and that the measured voltage is of good quality. However, if as in this project, this was not the case, this section will describe the method used to minimize the uncertainty originating from the voltage signals.

Figure 1 shows the phase voltage at a bus closer to the bottom of a network for both the simulation and the smart meter, the hole in the plot is due to a time period during September and October with missing data. What is notable is that the measurement variability of the voltage is significantly greater than the simulated variability. This variation in the measurements is consistent across all buses and is due to variations in the voltage from the HV side of the transformer. In the simulations, the overlaying HV system has been modelled as a Thevenin equivalent i.e. constant voltage behind an impedance. However, in the measurements, the source is not constant but carries variations induced by other connected networks. Furthermore, the voltage measurement has a resolution of 1 V, and the accuracy of the smart meter is not known but could be as low as 1% (≈ 2.4 V).

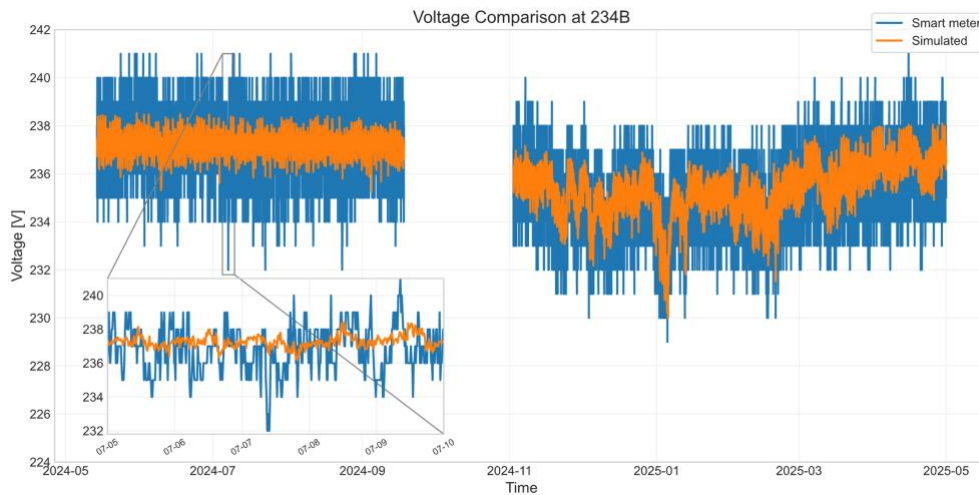


Figure 1: Comparison between measured and simulated voltage at bus 234 and phase B. Data is missing 18/09/2024 – 02/11/2024.

To make the simulation and measurement voltages more comparable, different methods can be applied, either to reduce variations in the measurements or to increase variations in the simulations. A method to increase variation in the simulations is to replace the constant voltage source behind the Thevenin impedance with a voltage profile based on the measured transformer voltage, preferably from the HV side if available. This would replicate the measured

conditions as accurately as possible by accounting for source voltage variations. This method was however not applied in this project as the case studied (see Section 3.2) lacked sufficient voltage data from the transformer meter.

Another method, and the workaround used due to the lack of transformer voltage measurements, is to extract an average variation across all buses. Since these variations originate from the source voltage, they propagate through the network and are therefore present at all buses. To extract them, the difference between the measured voltage and the simulated voltage was calculated for each bus, and the average difference across all load buses was taken as the variation. This can be done per phase if variations differ by phase, but in this case, the variation was consistent across all phases and was therefore calculated as a single average.

This variation was then added to the simulated voltages to create estimated voltages. The resulting comparison between measured, simulated, and estimated voltages is shown in Figure 2. Adding the extracted variation to the simulation produces a voltage profile that is significantly more comparable to the measurements than the original simulation alone. The estimated voltages are therefore of more use in the analyses for localizing the NTL.

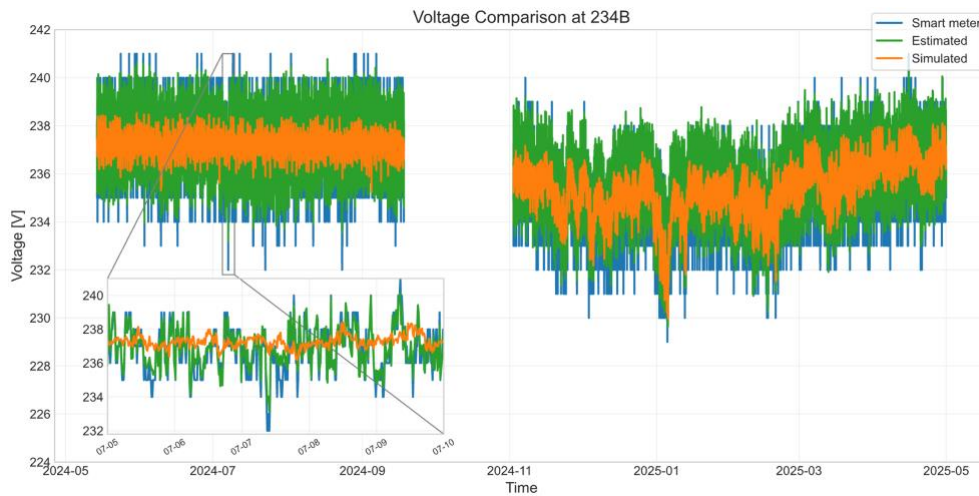


Figure 2: Comparison between measured, simulated and estimated voltage at bus 234 and phase B, where the estimation adds extracted measured voltage variation on the simulation. Data is missing 18/09/2024 – 02/11/2024.

3.2 CASE STUDY

For the case study on detection of NTL, a real distribution-level system configuration experiencing suspicious high losses was shared to the project along with its smart meter measurements. The described methodology was applied to this system in order to identify if NTL were present in the system and if so, what are the characteristics of the NTL and where in the network the NTL are located.

3.2.1 Setup and description

The topology of the shared distribution network used in this project is shown in Figure 3. The system consists, at the top, of the transformer and the point of

connection to the larger grid. The transformer steps down the voltage from 11 kV to 400 V, and the larger grid is modelled as a Thevenin equivalent. The system consists of 58 load buses, 8 of which contain distributed PV installations. The network is separated into five different branches, and the nodes are named accordingly, as shown in Figure 3.

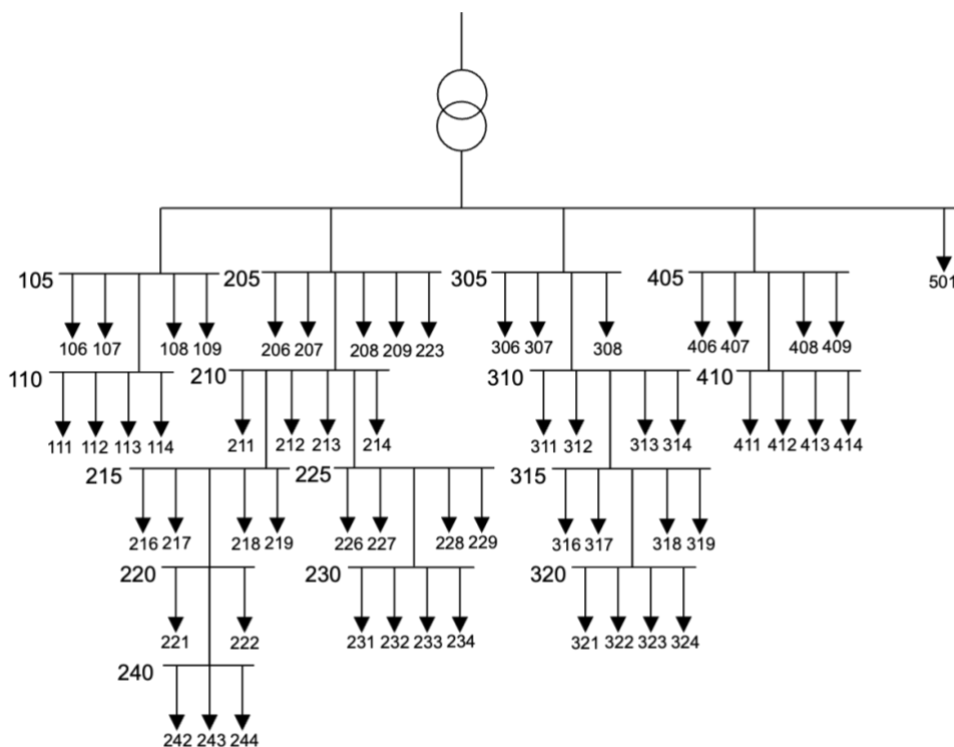


Figure 3: Topology of distribution system network

To avoid being locked into expensive licensing software, the project used OpenDSS to build the model and to simulate the system. To aid with the simulations, the Python library `py-dss-interface` [10] have been used to interface with OpenDSS through Python. The simulations run on the model where quasi-dynamic simulation with a timestep of 15 minutes. The loading in the model system is per phase, with the phase-specific loading for each timestep gathered from the smart meter data. The smart meter data was available from each load bus and the transformer at the top of the network.

Data availability and quality

There were two datasets of smart meter data available, referred to as **consumption logs** and **analytics logs**. The specifications and limitations of each dataset where as follows:

Consumption logs:

- Data Recorded: Three-phase bidirectional energy consumption per load and bidirectional energy flow through the transformer.
- Resolution: 15-minute intervals.
- Time Period: 14 May 2024 – 1 May 2025.

- Missing Data: Five files miss a limited number of data points, at most three hours' worth.
- Limitations / Exceptions:
 - One file reports half of the time period with hourly resolution. This has been interpolated to 15 minutes to fit the rest.
 - Bidirectionality for the transformer was enabled starting in November 2024.

Analytics logs:

- Data Recorded: Voltage and current per phase for load buses and the transformer.
- Resolution: 15-minute intervals.
- Time Period: 14 May 2024 – 1 May 2025.
- Missing Data: 18 September 2024 – 2 November 2024 missing in all files.
- Limitations / Exceptions:
 - No bidirectionality.
 - Transformer data available only until 18 September 2024.
 - One PV load bus has data only until 9 September 2024.
 - Voltage recorded with 1-volt resolution.

From this description, it is apparent that only the analytics logs have three-phase data, while only the consumption logs have bidirectional data. As the analytics logs are the ones with three-phase data, it was necessary to use these to specify the loading in the model. However, as the analytics logs did not have bidirectional data, the bidirectional information in the consumption logs were used to determine when the combined PV-load bus is producing or consuming power. From the consumption logs, it was determined at what data points bus is a net producer and the directionality of the current in the analytics logs where changed for these times. This is not a perfect solution, as during a 15-minute timestep the power can oscillate between production and consumption, resulting in the majority of the power cancelling out leaving, for example, only a small net production. The analytics logs, however, take the average of the current's absolute value during the 15-minute timestep and, in this example, classify all of that current as production. Hence, the model will use a too high a production value for this timestep compared to reality. The same will happen if it was instead a small net consumption, but in the opposite direction. The error this generates in the model is assumed to be small due to low currents when operating close to zero power.

Figure 4 and Figure 5 show stacked plots of the load consumption and PV generation from the consumption logs. The unit used throughout the report is average power in kW over the 15-minute timesteps. Both consumption and production exhibit seasonal patterns as expected with loading being higher in winter, while PV production is higher in summer. They also show daily patterns,

with consumption peaking in the morning and afternoon, while PV production is greatest at midday and absent during the night.

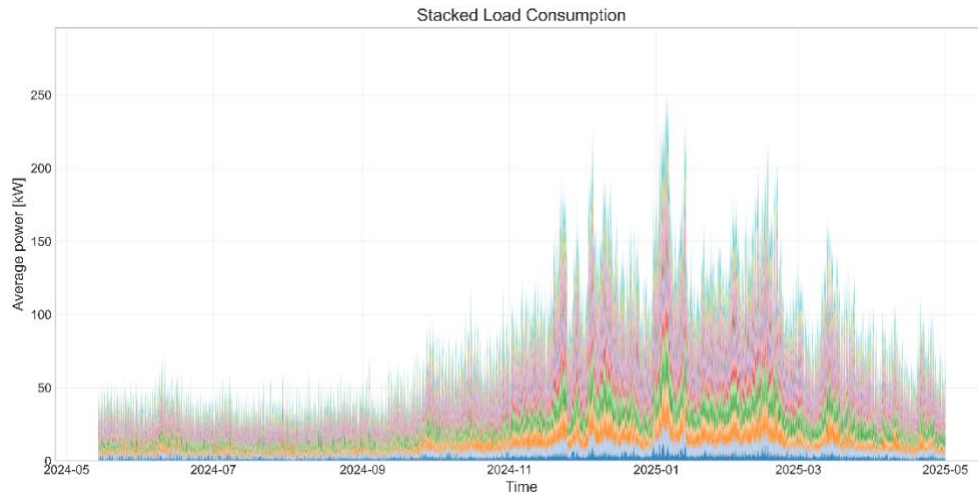


Figure 4: Average total power consumption of the network over 15-minute time steps

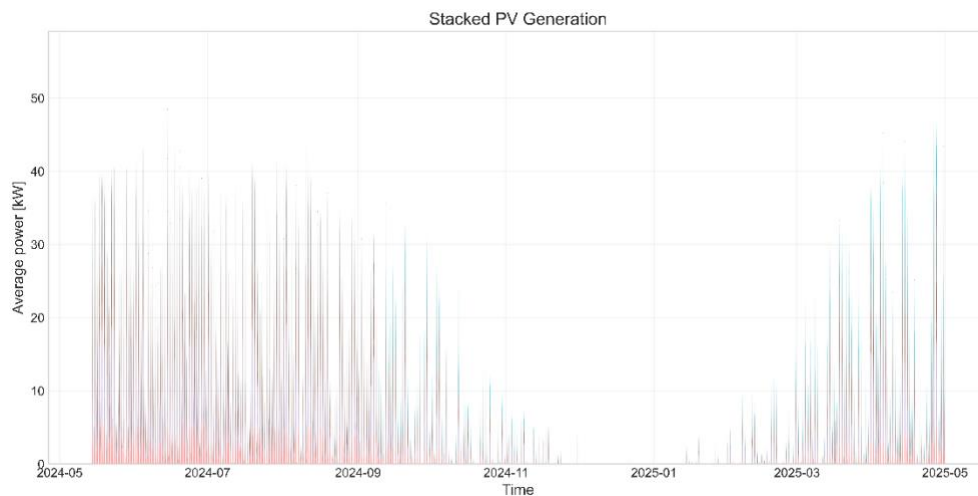


Figure 5: Average power generation of the network over 15-minute time steps

Simulation and smart meter comparison

A power flow comparison between the smart meter measurement and the simulation is only possible at the transformer. This is because smart meters are installed at the transformer and at the load buses, and that by design, the power consumption at the load buses is the same for both the measurements and the simulation. However, as the per phase data at the transformer is limited, it is only possible to do a combined three phase power flow comparison using the consumption logs.

Figure 6 shows the three-phase power flow through the transformer for both the simulation and the smart meter measurements. The figure shows the entire studied time period, and on the top left is a zoomed-in extract of July 2025. Based on the

figure, the smart meter energy flow and the simulated power flow behave similarly although some differences still occur. The major difference during part of September and October is due to a lack of data in the analytics logs, which the loading in the simulations is based on, while the consumption logs which the transformer loading is based on does not have the same information loss. The smaller difference with the blue smart meter measurement exceeding the simulation, most notable during the latter half of the analysed period, are presumed to be due to NTL in the system. How the simulation and the smart meter differ is shown more clearly in Figure 8. However, Figure 6 suggests that there are no major faults with the model and that the separation of PV generation from load consumption in the analytics logs works sufficiently.



Figure 6: Comparison between measured and simulated transformer power flow. Simulation is missing 18/09/2024 – 02/11/2024 due to the missing data in analytics logs.

Figure 7 shows the time series of losses in the network specifically from the smart meter data. The losses are calculated by comparing the total power measured entering the network at the transformer with the total power measured exiting the network at the nodes, accounting for the PVs. The difference between power in and power out represents the losses shown here. What stands out is the shape of the losses over time. There are clearly two distinct levels of losses, with points transitioning between them. Initially, during the summer months with low power flow, losses are generally at the lower level. As winter approaches and power flow increases, losses are at the higher level, then appear to return to the lower level as winter ends and power flow decreases.

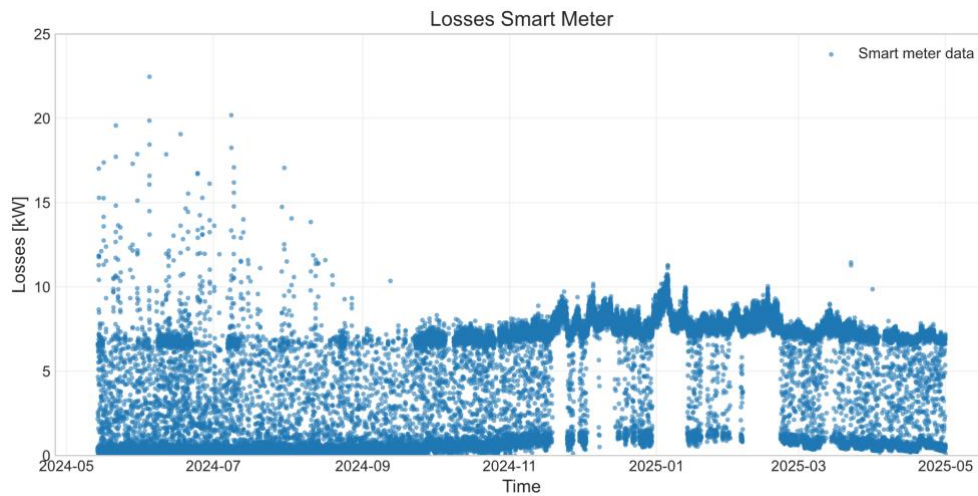


Figure 7: Losses in network as calculated from the measurement.

Figure 8 compares the measured smart meter loss levels to the simulated losses, plotted relative to the power flow through the transformer. Blue dots represent losses derived from measurements, and orange dots represent simulated losses. The simulation aligns well with the lower loss level. The higher loss level is assumed to be caused by NTL in the network. An additional load with a constant power consumption of 6.5 kW was therefore added to the simulation model without a meter measuring its consumption. The resulting losses, including the added load, are shown as red dots in Figure 8. These align closely with the higher measured loss level, therefore making the additional 6.5 kW load a convincing cause of the higher loss level.

One aspect of Figure 8 that should be noted is that the simulation shows losses for negative transformer power flow, while the measurements do not. This is because bidirectional metering was not enabled on the transformer during the initial period (summer), when power flow was often negative. Instead, the transformer measurement during the negative power flow is zero or close to it, causing inflated losses from measurements.



Figure 8: Comparison of losses between measurement, simulation and simulation with additional NTL-load.

3.2.2 Results

The following section will present the result of an in-depth analysis of the differences between the smart meter measurements and the simulation given the assumption that the higher loss level identified from the measurements of the network is due to NTL and specifically an additional load of 6.5 kW connected without being metred. The objective of this analysis is to categorize and localize the NTL using the methodology described in Section 3.1.

Categorization of NTL

The measured loss profile in Figure 7 and Figure 8 follows a clear pattern and by identifying this pattern some characteristics about the additional load can be identified.

NTL-load Profile

The yearly load profile of the assumed load causing the NTL is shown in Figure 9. Outliers outside the range of 120% to -20% of 6.5 kW were removed, and a 6-hour moving average was applied to visualize the pattern, as the raw data would otherwise oscillate rapidly between fully on and off states. As expected, the NTL-load profile resembles the losses in Figure 7, except that the trend caused by the technical losses has been removed. The load profile shows a strong seasonal pattern, with higher values in winter compared to summer. The result of the average daily loading of the NTL-load, shown in Figure 10, indicates a on average constant daily consumption.

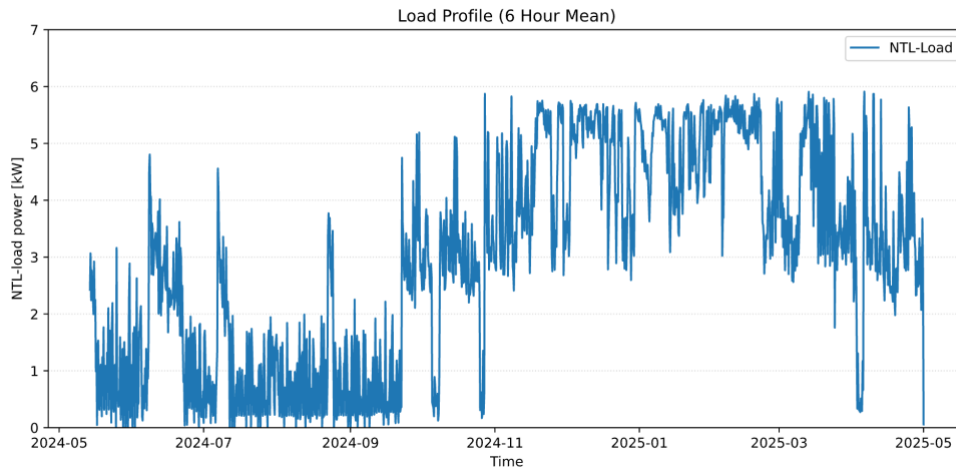


Figure 9: Loading of NTL-load based on comparison of losses between measurement and simulation.

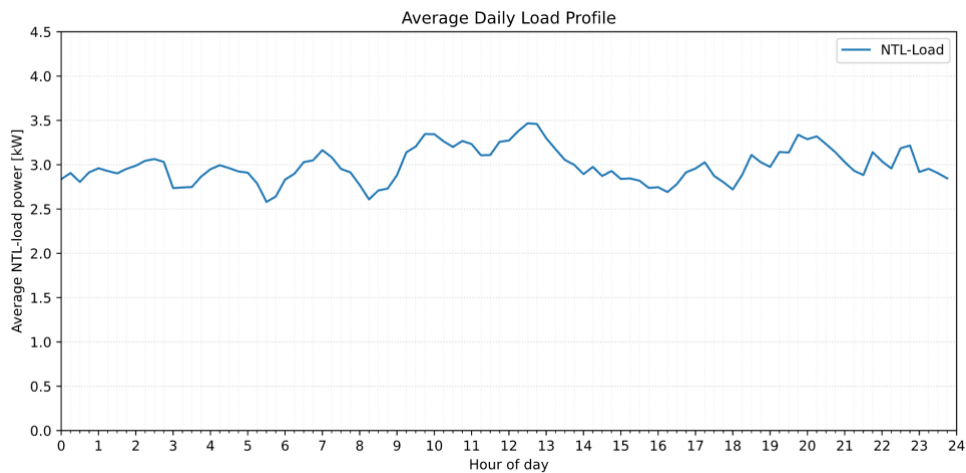


Figure 10: Daly average load profile of NTL-load based on comparison of losses between measurement and simulation.

Transitions

Transition refers to the data points between when the load is fully on or off, where the load is transitioning to a new state. Figure 11 shows that transitions mostly happen around midday, with a slight increase in the afternoon and early evening. There is no clear trend indicating whether transitions are more likely to be upward or downward at specific times of day. There is also no noticeable variation based

on the day of the week. However, a clear seasonal trend emerges with transitions being less frequent during winter.

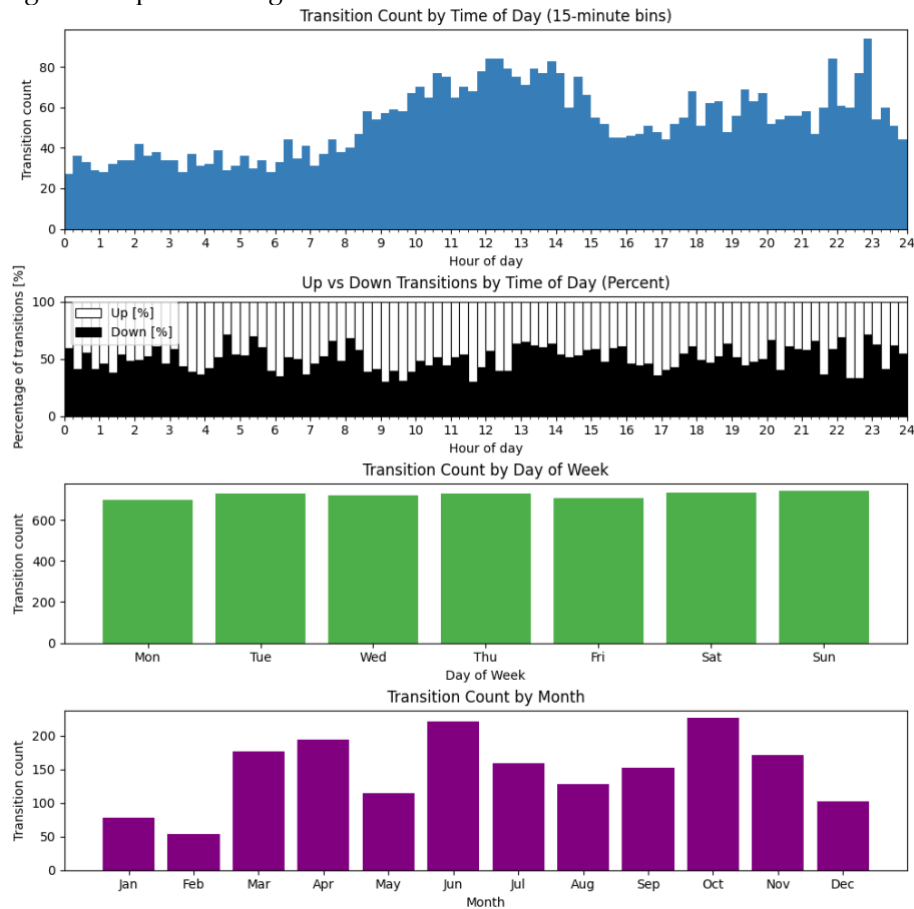


Figure 11: Distribution of number of NTL-load transitions per day, week and month.

Considering the results from Figure 10 and Figure 11, one possible explanation is that the NTL-load is automatically controlled to maintain a constant reference. This is supported by the approximately constant daily profile, the roughly equal number of upward and downward transitions throughout the day, and the consistent weekly distribution. The monthly trend also supports this, as the controlled parameter, likely temperature, is influenced by seasonal changes. During winter, the NTL-load is mostly on, possibly because the controlled parameter does not reach its reference value, explaining the small number of transitions. In summer, the NTL-load is mostly off, but there are still many transitions, indicating active regulation.

Assuming automatic control is applied to maintain a constant reference, the daily distribution suggests that either greater variability around the reference is allowed during approximately 00:00–07:00, or the controlled parameter is stiffer, as indicated by fewer control actions. Conversely, during 09:00–15:00, the higher number of control actions suggests that variability around the reference is more tightly controlled, or the controlled parameter is less stiff.

Temperature correlation

A specific characteristic of the NTL that have been noted is that high losses occur more frequently during winter, when power flow is higher and temperatures are lower, while lower losses are generally observed during summer. Hence, it was interesting to investigate the outside air temperature correlation to the NTL-load for seasonal, day to day and hour to hour variations. Table 1 presents the correlations calculated based on (1) where a value of -1 notes that the two curves follow each other perfectly, but in the opposite direction, where an increase of curve one corresponds to a decrease in curve two.

Table 1: Correlation coefficients between seasonal, day to day and hour to hour variations for the NTL-load and the outside air temperature

Temporal range	Correlation
Seasonal	-0.986
Day to day variations	-0.579
Hour to hour variations	-0.215

The seasonal time series for NTL-load and temperature are almost perfectly negatively correlated, with a correlation coefficient of -0.986. The corresponding time series are shown in Figure 12, where they clearly follow each other very well, but in opposite directions.

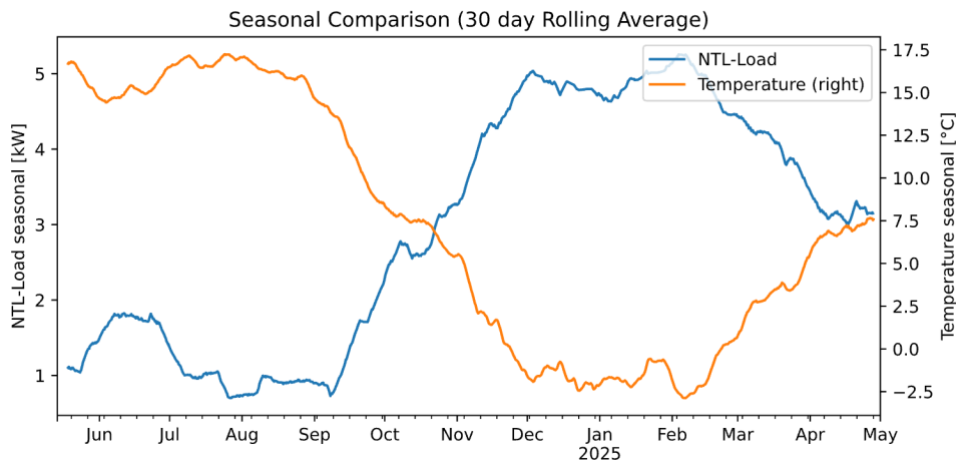


Figure 12: Time series of seasonal trend for NTL-load and outside air temperature.

The correlations for day to day and hour to hour variations are shown through scatter plots in Figure 13 and Figure 14 respectively. The day to day variations show a correlations coefficients of -0.579 suggesting a trend towards higher NTL for lower temperature and vice versa but is significantly less clear then the seasonal correlation. The hour-to-hour variation has a further smaller negative correlation of -0.215 and a visual inspection also suggests that any trends between temperature and NTL is slight.

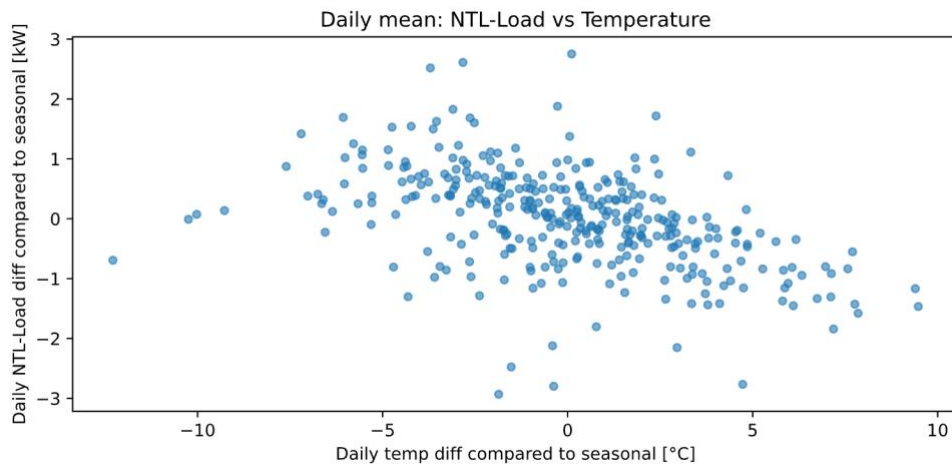


Figure 13: Visual representation of correlation between variations in NTL-load and outside air temperature from day to day.

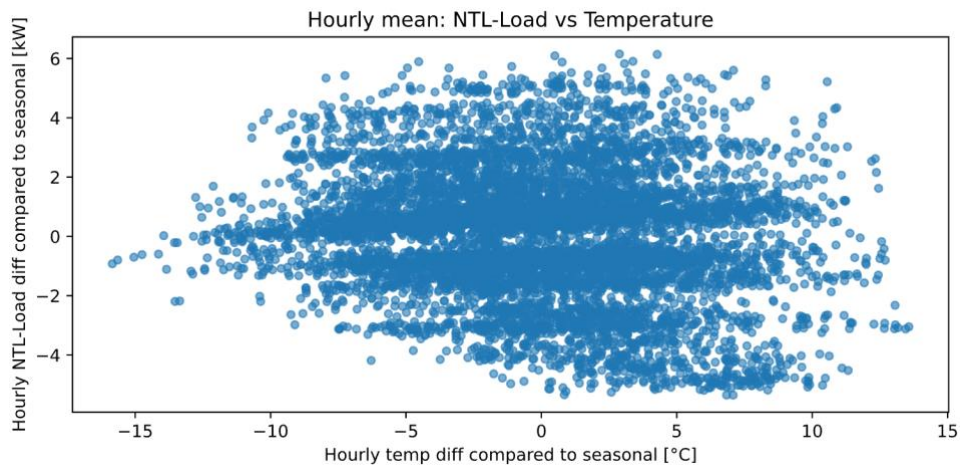


Figure 14: Visual representation of correlation between variations in NTL-load and outside air temperature from hour to hour.

Note that correlation decreases as finer temporal variations are examined. The reduction in correlation at higher temporal resolutions is most likely due to thermal insulation. Assuming the NTL-load controlled relative a temperature, the steady daily profile in Figure 10 suggests that the controlled environment is well insulated. This means short-term temperature fluctuations (hourly or daily) have little effect on the NTL-load, as only sustained changes over longer periods would trigger adjustments. Additional influencing factors for the decreases in correlation are:

- **Load behaviour:** The NTL-load switches between on and off states. When data is aggregated to lower resolutions, the load appears more continuous and better aligned with temperature changes. At higher resolutions, it is more likely to remain at maximum or minimum values.
- **Geographical differences:** The temperature data comes from a station some distance away, and these differences become more significant at shorter timescales.

To summarize, assuming that deviations in smart meter losses compared to expected losses through simulation are due to an additional unmetered load, this NTL-load shows a rated power consumption at 6.5 kW and strong negative correlation with seasonal temperature trends, suggesting it is a heater used for temperature control. The daily load profile and transition distribution of Figure 10 and Figure 11: Distribution of number of NTL-load transitions per day, week and month. indicate automatic control to maintain a constant reference. The lack of correlation at higher temporal resolutions, combined with the steady daily profile, suggests a sufficient thermal insulation to make the controlled environment largely unaffected by short-term changes in the outside air temperature.

Localization of NTL

The localization of the NTL is based on the principle that the NTL will introduce a local voltage drop in the network. The three methods for NTL localization based on [17] and described in Section 3.1 have been applied to the studied network. The methods are applied when both studying each phase individually and when aggregating the phases to a per bus analysis.

Voltage Drop at Measurement Point

Figure 15 shows the voltage DME for each bus and phase calculated according to (2) from the time instances when the measured losses exceed expected losses, what is referred as the high loss category. A high positive DME would indicate that the measured voltage is consistently lower than the estimated voltage when losses are high, which could suggest NTL. From the figure, bus 324b shows the highest DME. However, there is significant variation in DME values across the 320-series buses, both positive and negative. This variability is present across all buses but is most pronounced in the 320-series. The presence of negative DME suggests that these deviations may not be caused by NTL but rather by factors such as, the measurement uncertainty, the low voltage measurement resolution, or a phase mismatch in the model (i.e., what has been noted as phase A of an end node in the model may in the actual network be correlated to phase B whole for a different end node, the model phase A may be correlated to phase C). Due to this variability in the DME, it is not possible to identify the location of NTL based on this analysis.

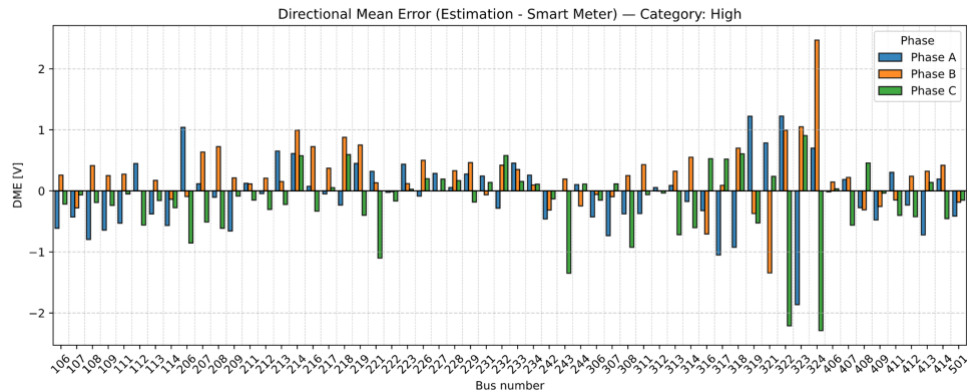


Figure 15: Directional mean error between estimated voltage and measured voltage for the high loss category, by bus and phase.

As one of the potential causes of the high variability and difficulty in locating the NTL is a possible mismatch between phases in the model and the actual network, the method was applied again, but this time, voltages were averaged over the phases beforehand to minimize any effects caused by this mismatch. Figure 16 shows the DME results for the high-loss category. In this case, the result is clearer, with bus 214 showing a high positive DME, suggesting NTL. In addition, several buses close to 214 also show a somewhat increased DME, which would be expected if 214 were the bus with NTL. Complicating the picture, however, is bus 243, which shows a high negative DME, indicating a higher measured voltage than the estimation, which is an unexpected result.

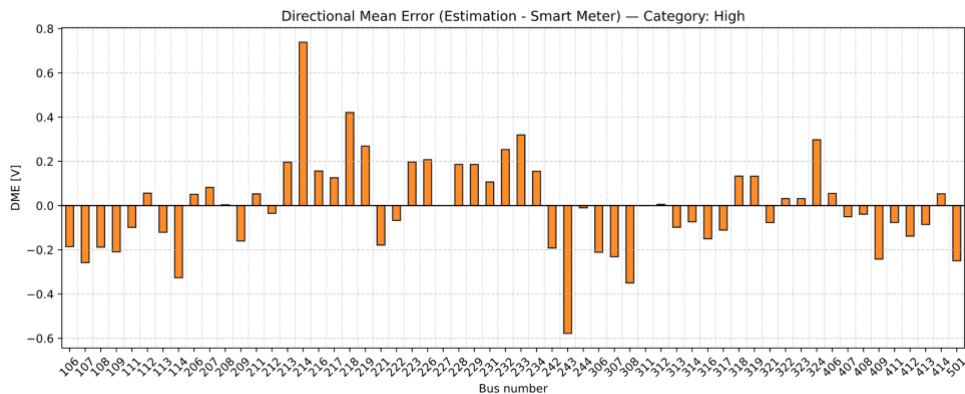


Figure 16: Directional mean error between estimated voltage and measured voltage for the high loss category, by bus.

A further investigation of the DME can be done by comparing the DME between the low-loss category (i.e. when network losses were close to expected based on the simulations) and the high-loss category. The expectation is that during low losses, the DME should be close to zero for all buses as there is no NTL, while for high losses, the bus with NTL should see an increase in DME along with the buses electrically close by. This comparison is shown in Figure 17, and the DME for the low-loss category is consistently close to zero as expected, resulting in a noticeable DME increase for bus 214 and some nearby buses during high losses. However, a significant DME decrease for bus 243 during high losses is also noted.

It should also be noted that this comparison between low and high losses should be interpreted with care, as the network loading characteristics differ between these cases. Low losses mostly occur when the network is lightly loaded, while high losses occur under heavier loading. The difference between the two cases is therefore not only the presence of NTL but also the loading conditions, which could influence how DME compares.

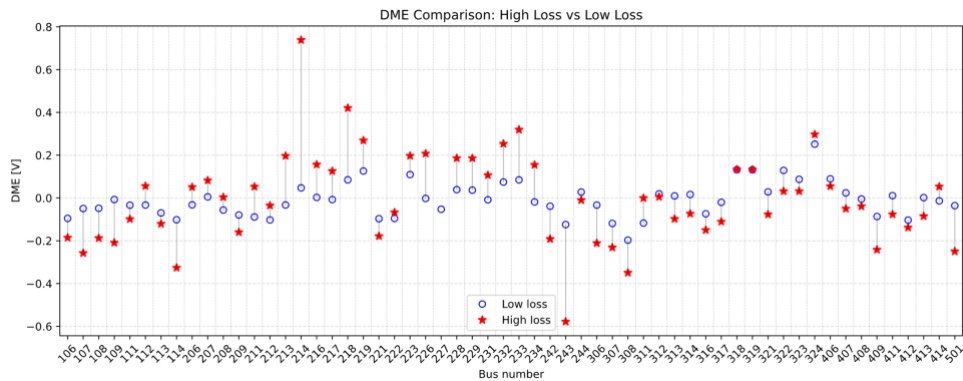


Figure 17: Directional mean error between estimated voltage and measured voltage compared between low loss category and the high loss category, by bus.

Correlation: Voltage – Voltage

The second method compares measured voltage to estimated voltage, where a low correlation could indicate NTL. The correlations for all data points and when looking at phases individually are shown in Figure 18. Results are similar to the previous method, with no clear indication of NTL location when looking at all phases. Again, buses in the 320-series stand out slightly, with 322c and 324c having the lowest correlations. However, high variations in the correlation over all busses makes any result suspect and nothing to clearly stand out, making it again not possible to suggest where the location of the NTL might be.

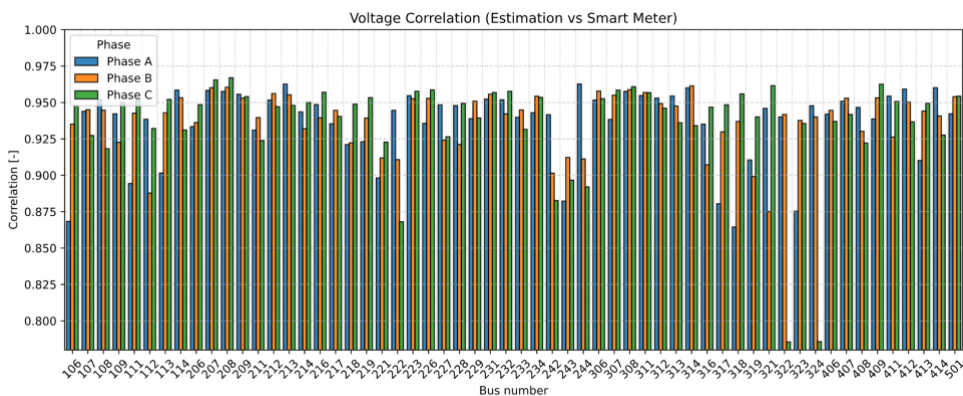


Figure 18: Correlation coefficient between the estimated voltage and the measured voltage, by bus and phase.

As for the voltage drop method, when combining the individual phases an analysing per buss to minimize phase mismatch errors, buses 214 and 243 stand out compared to the rest indicating possible NTL locations, as shown in Figure 19. These are the same buses that stood out in the voltage error results.

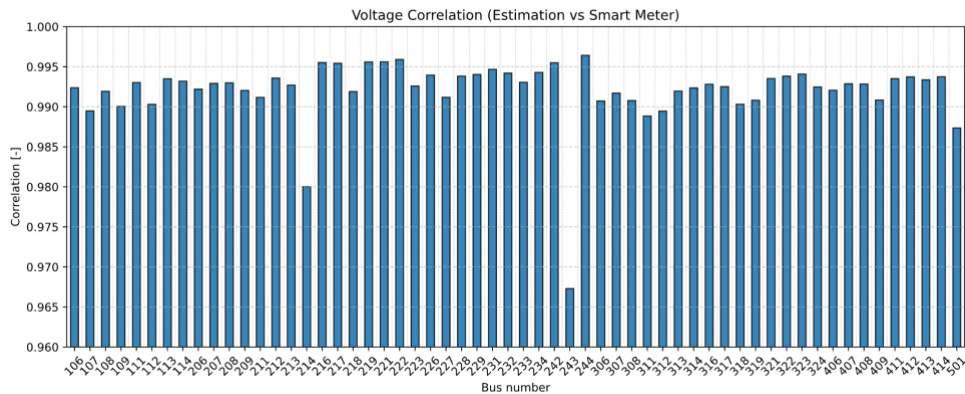


Figure 19: Correlation coefficient between the estimated voltage and the measured voltage, by bus.

Correlation: Voltage – Power

Figure 20 shows the correlation between voltage and power for both measurements and estimated values across all available data points separated by phase. A large difference between the smart meter measurements and the estimated values is indicative of NTL. However, similar to the other per phase analysis, this correlation comparison does not provide a clear indication of NTL location. Buses around 317 to 324 stand out slightly, showing higher correlation for measurements compared to estimations while being electrically close. However, similar patterns appear elsewhere in the network. The direction of the difference is also unexpected, if NTL were present, a lower measurement correlation would be expected. Bus 324c shows an expected difference, but overall variability and inconsistency make these results unreliable.

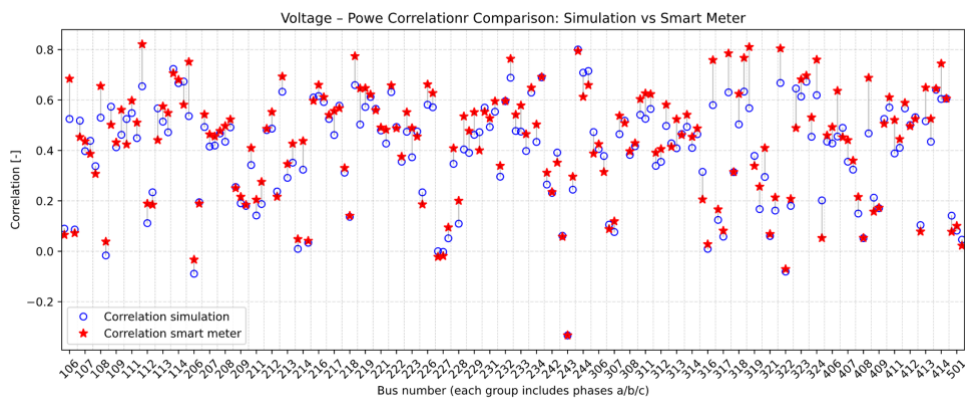


Figure 20: Correlation coefficient between voltage and power compared between the estimation and the measurement, by bus and phase.

Finally, a per bus aggregated analysis for the voltage power correlations was done and is shown in Figure 21. However, different from the voltage drop case and the voltage – voltage correlation case, the voltage – power correlation case does not produce any clearer results when aggregating per bus. Hence the voltage – power correlation did not, for the studied case, produce any location that where indicative for the NTL.

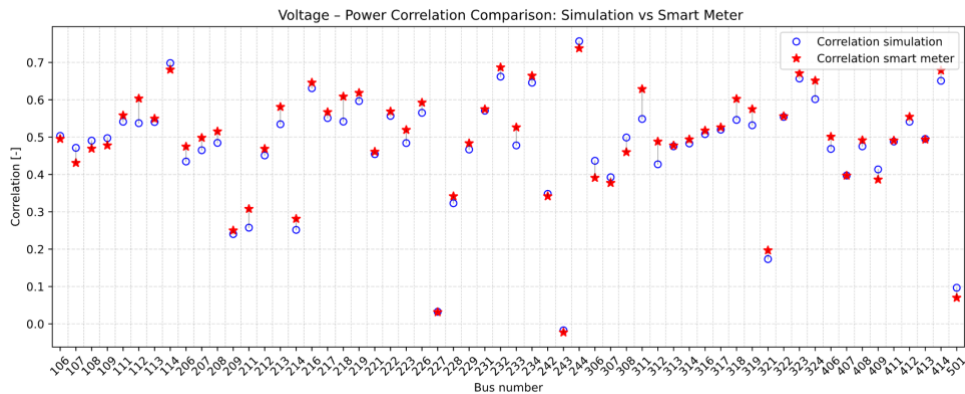


Figure 21: Correlation coefficient between voltage and power compared between the estimation and the measurement, by bus.

Localization summary

Overall, comparing the localization results between the per-phase analysis and the per-bus analysis, the per-bus approach gives a clearer picture of where the NTL might be located. The high variability between phases and buses for all methods when analysing per-phase makes it not possible to suggest a location of where the NTL may originate from in this case. It should be noted that the buses in the 320s did somewhat stand out across all methods in the per-phase analysis, but as the way they stood out is not always consistent with what is expected from NTL, it suggest that this is more likely due to measurement or modelling issues rather than NTL. This suggests that the phases in the model are not aligned with the actual network, which needs to be addressed before any per-phase localization can be attempted.

When instead analysing aggregated per bus, a higher DME as well as a lower voltage - voltage correlation was identified for bus 214. This bus is therefore a likely candidate for NTL. What complicates the picture is the significant negative DME for bus 243, along with its lower voltage - voltage correlation. This suggests NTL-like behaviour for bus 243, but is indicative of an unmetered power injection at the bus. This is the opposite from what is expected as the network were showed to have unusually high losses, suggesting an unmetered power extraction, as was noted for 214.

There are still uncertainties in the aggregated per-bus analysis related to voltage measurement accuracy and resolution. Of note is that the presumed NTL load is 6.5 kW, which is only 2.3% of the network's maximum measured load (280 kW). Hence, if the NTL is located near the transformer, where voltage is stiffer, its impact on voltage may be negligible in relation to the uncertainty and resolution. In the studied case, possible locations of the NTL where identified, but it still is an aspect to keep in mind and highlights the necessity of accurate voltage measurements in order to accurately perform this type of analysis.

3.3 DISCUSSION

The results of the case study show that it is possible to identify, gather insight into and localize NTL in a real network. After conversations with the owner of the network from which the data were obtained, it was confirmed that the localized bus of 214 did have an unmetered load connected, being the cause of the NTL.

The voltage drop at measurement point analysis provided the largest amount of information for the localization of NTL in the studied case, as the calculated DME parameter is directional. Its directional nature proved important in distinguishing bus 214, which exhibited the expected NTL behaviour, from bus 243, which showed anomalous behaviour, although not characteristic of the expected NTL pattern. However, it should be noted that the reason for the unexpected behaviour at bus 243 was not determined. Because the voltage drop analysis is directional, it is also possible to distinguish between unregistered loads and unregistered power injections.

The voltage – voltage correlation analysis was also useful in finding the NTL location, as it clearly highlighted buses with anomalous voltage correlations. Given that the voltage drop analysis and the voltage – voltage correlation analysis identified the same bus, using a combination of these two analyses strengthen the validity of the identified bus. The final method tested, the voltage–power correlation analysis, did not provide any clear results in the studied case.

One aspect the case study highlighted is the importance of accurate and good quality data. The data used were sufficient to achieve the objectives of the study and identify the NTL location. However, several workarounds were necessary due to limitations in the data. Major challenges included the lack of per-phase bidirectional measurements, high uncertainty in the voltage measurement and a large proportion of missing data. These limitations required significant additional effort and introduced uncertainty into the results. It is also important to note that the impact on the voltage from the NTL needs to be greater than the voltage measurement uncertainty for localisation analysis to work. If a small NTL is located at a voltage-stiff point in the network, the accuracy requirements for the voltage measurements become high.

Finally, the case study showed that any per-phase analysis failed to provide any clear results. The suggested reason for why the per-phase analysis failed is due to phase connections in the actual network not lining up with the phase connections in the model. In order to do any per-phase analysis, it is therefore necessary to have a complete and accurate understanding of all phase connections in the network.

4 Impact of the unbalanced operation on the hosting capacity of the system

Hosting capacity studies for PVs are conducted to estimate the maximum additional PV production that can be connected to the distribution system without exceeding its operational limits. The hosting capacity of a network depends on multiple factors such as the network configuration, load patterns, and grid operator limitations among others. There is no specific methodology defined for performing hosting capacity studies in literature and thus, a simple iterative methodology is adopted in this project. The hosting capacity is estimated both on individual bus level and at the system level.

4.1 METHODOLOGY

Estimation of the hosting capacity of the individual system buses

The hosting capacity for individual buses is conducted to evaluate the maximum PV injection that is possible at each bus while retaining the secure operation of the rest of the system. This type of analysis is done to address new connection requests from the customers by evaluating the maximum solar PV production that can be installed [18]. By the secure operation of the system, it is meant that the system operates within its operational limits such as bus voltage limits and line current limits [19]. The lower and upper bus voltage limits of 0.95 and 1.05 pu are assumed as operational limits when calculating the hosting capacity in this project. As PVs are installed on the rooftops in a residential distribution network, only the load buses in the network are considered for PV integration.

The hosting capacity methodology uses an iterative process depicted in Figure 22. The network with existing loads and PVs is considered as the base case network. To the base network a new PV is added at a bus, and the quasi-dynamic simulation is run for the entire duration of the time-series data with the new configuration. The PV is then scaled up by a factor, and the simulation is re-run. This process is repeated till one of the bus voltages in the entire network exceeds the threshold of 5% at any point during the simulation. When the voltage threshold is reached, the corresponding size of the PV is recorded and scaled down to zero, effectively returning to the base network configuration. The recorded size of the PV is then the hosting capacity at that bus. A new PV is added at another load bus, and the iterative process is run again till all the buses are covered.

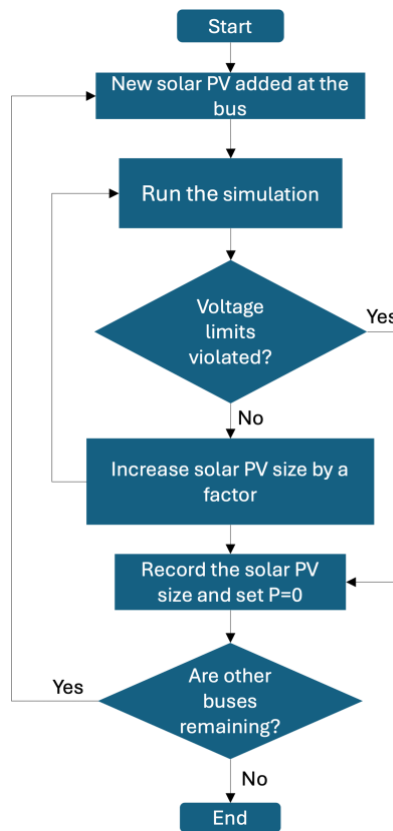


Figure 22: Flow chart of the iterative process of PV hosting capacity calculation for individual buses.

Estimation of the hosting capacity of the entire system

The hosting capacity study of the whole distribution system is conducted to estimate the maximum PV penetration for the whole system when PV panels are gradually, in steps, and over time being installed over multiple buses. PVs are installed on customer requests and so the order in which PVs are being installed at different locations (buses) in the system is not deterministic. To account for different possible orders of locations in which PVs are installed, Monte Carlo simulations are used when determining the hosting capacity of the analyzed distribution system. Using Monte Carlo analysis, a number of system load buses' permutations are created defining scenarios giving the locations' order. Since the total number of permutations is vast and the computation time for each simulation is long, only 51 of the possible permutations are run to illustrate statistical results.

As the number of scenarios run is not enough to get statistically relevant results, two additional deterministic bus orders in which PVs are being installed are considered: *systematic* and *conservative* order. The first *systematic* order considers that locations where PVs are installed start from the electrically closest bus to the electrically furthest bus from the substation. This *systematic* order is expected to lead to the highest system hosting capacity, i.e. optimal order of locations in which PVs should be installed. The second *conservative* order takes a conservative approach assuming that the locations where PV are installed start from the electrically furthest to the electrically closest bus to the substation.

Nowadays in Sweden, PV connections are given on the rule *first come - first served*. This rule is simple, however not entirely fair. If a party who comes first connects large PV installation that limits the remaining hosting capacity for the others filing connection requests later, the others may not be able to connect their PV installations of similar or even smaller size. Secondly, the rule *first come - first served* may sub optimally use the total distribution system hosting capacity therefore resulting in uneconomical use of distribution system infrastructure.

To be able to use the simple *first come - first served* rule and still address the aforementioned problems, some distribution systems may limit the maximum size of the PV connections per location. While this is not currently the case in Sweden, this practice is used by a number of European DSOs, especially in the countries where PV penetration in distribution systems is larger. To analyze how maximum allowed PV connection power affects the total hosting capacity, the sensitivity analysis is done. Maximum PV connection power per location is varied from 5 kW to 20 kW (typical range of Swedish PV installations) with addition of 50 and 100 kW limits accounting for larger stand-alone systems or more commercial/non-household PV plants.

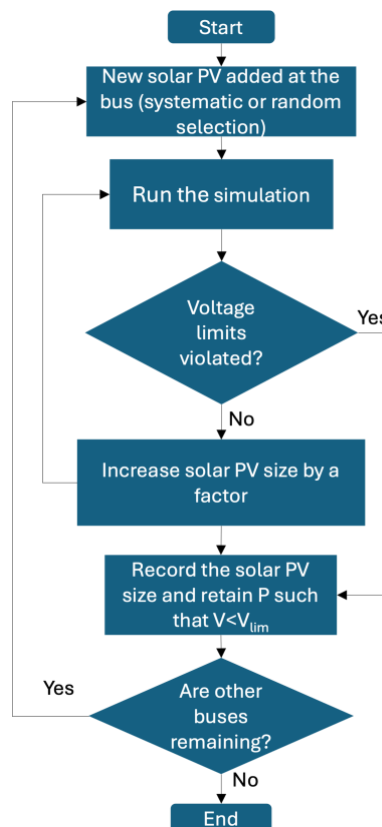


Figure 23: Flow chart of the iterative process for estimating the hosting capacity of the entire distribution network.

The iterative process for the hosting capacity estimation of the network follows a similar approach to the one used for the individual buses and is shown in Figure 23. However, the difference here is that once the hosting capacity of a specific bus

is estimated, the PV power is maintained at that value for the next iteration. Like for the individual buses, the analysis here is done on two different network conditions, one base case scenario and the other unbalanced scenario.

4.2 CASE STUDY

4.2.1 Setup and description

The system analyzed in this case study is the same as the one described in detail in Section 3.2.1. The original system had 8 PV installations. To estimate the total possible PV hosting capacity of the system, additional PVs are sequentially connected to the systems' load buses using algorithms described in Figure 22 and Figure 23. To simulate newly connected PV's output variation during the year, operational statistic available for originally connected 8 PVs is used. Therefore, newly installed PVs follow the same pattern of varying their active power output as originally existing PVs i.e. we assume that all the PVs in the system are affected by the same meteorological conditions.

The overlying grid has been in the study modelled as a slack bus as information on its parameters and configuration or short-circuit capacity of at the point of common coupling were not available. Bus voltage limits of 0.95 and 1.05 pu are considered as the operational limits while, due to a lack of data, the current limits of the cables have not been considered in the study.

For estimating the system's hosting capacity, two operation scenarios are considered:

1. balanced operation (base case) of the system over the three phases and
2. unbalanced operation with 1-phase outage on the branch connecting substation and bus 205. It is assumed that in the case of one phase outage, PV inverters can continue to operate on two other healthy, available phases.

Three approaches of choosing bus orders in which PVs are integrated in the system, described in Section 4.1, are investigated:

- a. Systematic
- b. Randomized (Monte Carlo)
- c. Conservative

Finally, dependency on available hosting capacity of the system is evaluated against 6 different values for maximum power of PV installation per location (bus): 5, 10, 15, 20, 50, 100 kW.

4.2.2 Results

Estimation of the hosting capacity of the individual buses

The hosting capacity estimation for individual buses is shown in Figure 24. The figure on top shows the maximum size of PV that can be connected at the system load buses. The results are given for both balanced (base case) and unbalanced (1-phase) outage operation of the system. The corresponding electrical distances of the load buses are plotted in the bottom figure. The electrical distance is calculated between the bus and the LV side of the substation transformer. Comparing the hosting capacities for balanced and unbalanced system operation, a common trend seen is that the phase outage leads to a decrease in the hosting capacity across all the buses, with an average reduction of about 30 kW. Considering only the base case, there are some variations in how much each load bus can withstand the PV injection into the network. Bus 501, where PV501 is located, has the highest capacity while buses 243 and 244 have the lowest. If we take a look at the electrical distances of the buses, bus 501 is electrically the closest to the substation and therefore its voltage strength is higher. Thus, it can host a large amount of PV before reaching the voltage limits compared to the other buses in the analysed system. In contrast, buses 243 and 244 are the furthest from the substation and so they have the lowest hosting capacity.

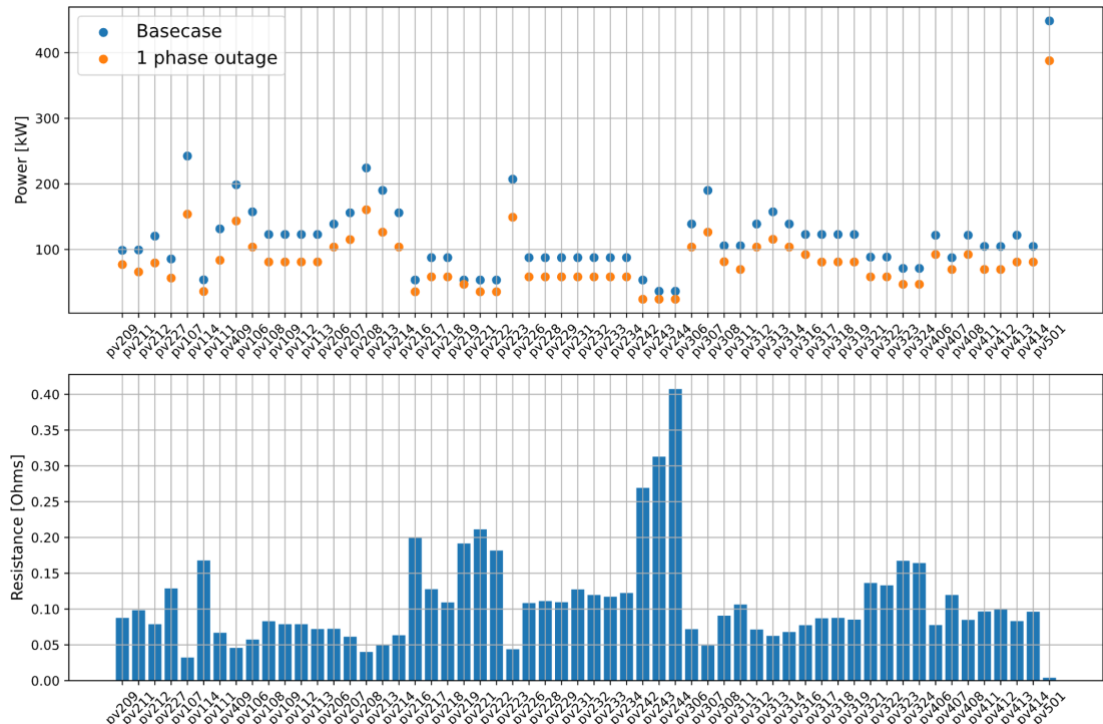


Figure 24: Result for the hosting capacity estimation of individual buses in the network. The figure on top shows the maximum PV power that can be installed at each bus. The electrical distances of the buses from the substation are shown in the bottom figure.

Estimation of the hosting capacity of the entire system

Figure 25 shows the hosting capacity of the system for balanced operation and the *systematic* order of bus selection for integration of PVs. The results are presented for different maximum allowed PV power per connection/bus. If the power limit per connection is set to 5 kW, then all PV installations in the system are reaching their allowed power limit. This means that in the case when DSO limits the maximum power of the connection to 5 kW, there are no need to run additional hosting capacity studies for each individual new PV connection as all of them even with their maximum allowed power of 5kW will not jeopardize the system's operating limits. Furthermore, this also ensures that each connection point has the same connection conditions and fair sharing of the system's hosting capacity.

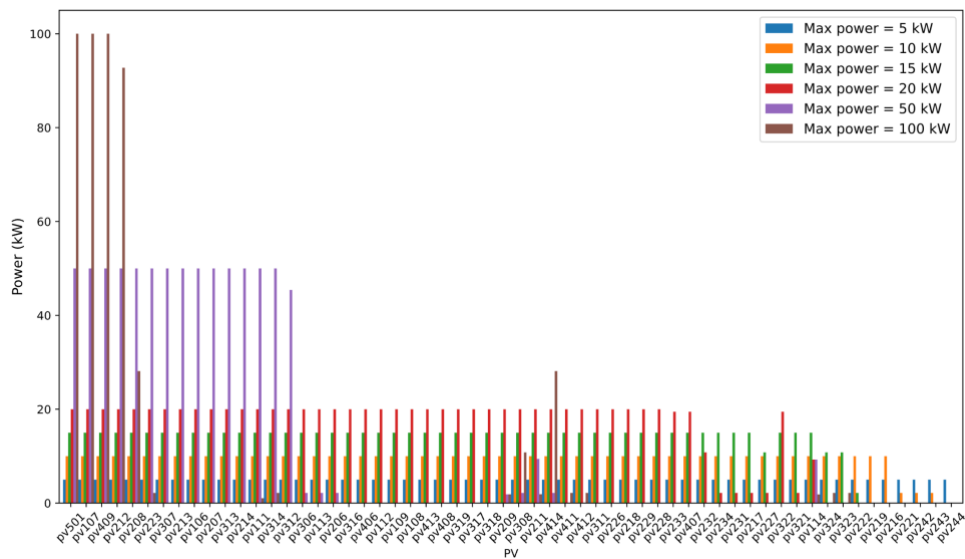


Figure 25: Hosting capacity per load bus and power limit per connection for the systematic selection of PVs with the base case network configuration.

In the case when maximum PV allowable power is higher than 5kW, it can be seen that not all PV connections can be sized to reach that allowed power limit. For example, in case of the maximum allowed power of PV connection of 10 kW, three PV installations will need to be limited more than initially allowed power of the installation in order not to jeopardize the distribution system operation. For these three PV installations, it would not be possible to directly allow 10 kW connection. Instead, separate calculations/studies would need to be carried out for each of these connections to determine maximum power of PV allowed to be connected to the system. From the practical point of view, such procedure would require more work from DSO, lengthier connection process and ultimately, not fair sharing of the grid hosting capacity for at least three PV connections.

While limiting PV connection power can solve the problem of fair sharing of the grid hosting capacity, simplify the connection procedures and decrease burden on DSOs, it also affects how efficiently the hosting capacity of the system is used, i.e. how much power of PV installations can be installed in total in the system. The

first row of Table 2 shows the hosting capacity results for *systematic* order of bus selection for integration of PVs when different power limit of PV connection is applied. It can be seen that the highest utilization of hosting capacity of the system is achieved when limiting the power of PV installation to 20 kW. In this case, 39 locations can be allowed to connect at maximum 20 kW power of PV while for the rest, individual bus hosting capacity studies would need to be carried out. As a conclusion, to efficiently use the distribution system infrastructure when integrating PV, it is important to provide limit to the allowed PV connections but also in the limited number of cases do special hosting capacity calculations.

The utilized hosting capacity of the system also depends on the order in which PVs are connected in the system. For this case study, the impact of three different connection orders, described in 4.1, is assessed: *systematic*, *conservative* and *randomized (Monte Carlo)*. Comparing the results from Table 2, it can be seen that the highest utilization of the system's hosting capacity is achieved with *systematic* order of connecting PVs, while *conservative* order mostly provides the worst results. Average value of the used system's hosting capacity over run Monte Carlo simulations (Randomized order) is between values obtained for *systematic* and *conservative* orders. The minimum values over Monte Carlo runs, illustrating randomized order of integrating PVs, are mostly slightly higher than the value obtained for *conservative* order of integrating PVs while the maximum values are mostly slightly below values obtained when *systematic* order of integrating PVs is used.

In case of randomized order (Monte Carlo) of locations for integration of PVs, maximum hosting capacity is achieved when limiting power of PV connections to 20 kW, similarly as when choosing *systematic* order of connecting PVs. However, if PVs are installed starting first from the electrically furthest buses (*conservative* order), then limiting the power of PV connections to 5 kW yields the highest used system's hosting capacity for PVs.

Table 2: Hosting capacities in kW for each power limit per connection and each case study.

		Power limit per PV connection (kW)					
Bus Order for connecting PVs		5	10	15	20	50	100
Systematic	Balanced	290,00	546,49	756,27	869,27	776,75	479,38
	Unbalanced	279,30	444,09	620,83	650,53	486,30	445,20
Conservative	Balanced	290,00	118,14	139,04	136,92	136,92	136,92
	Unbalanced	117,60	109,20	98,24	100,67	98,95	98,95
Randomized (Monte Carlo) Balanced	Min	290,00	211,67	215,89	198,27	112,23	112,23
	Max	290,00	536,49	661,42	682,43	583,00	561,01
	Avg.	290,00	391,31	415,37	425,69	342,18	257,75

Randomized (Monte Carlo) Unbalanced	Min	204,19	178,55	141,16	188,76	112,23	112,23
	Max	290,00	536,49	661,42	682,43	583,00	561,01
	Avg.	281,71	359,01	390,71	402,66	321,98	249,95

Assuming that PV inverters can continue to operate on two other healthy, available phases, unbalanced operation with 1-phase outage on the branch connecting substation and bus 205 decreases the available PV hosting capacity over all the analyzed scenarios. Comparing the results from Table 2, the used hosting capacity decreases by in some cases only 4% to even up to 60%. Since unbalance is caused only on the part of the network, the decrease of the hosting capacity is strongly influenced by the selected bus order in which PV panels are connected. This can be also seen comparing the results of unbalanced and balanced scenario of the system operation for randomized order of selecting buses for PV connection. Maximum hosting capacity over Monte Carlo runs is the same for both unbalanced and balanced scenario. This is because the PVs in both balanced and unbalanced scenarios are connected in the part of the system that is not affected by 1-phase outage on the branch connecting substation and bus 205.

4.3 DISCUSSION

To simulate the randomized order in which PVs are connected to the different buses in the system, Monte Carlo simulations are run together with assessing the hosting capacity as described in Section 4.1. The results from Table 2 are obtained by running 51 of these simulations, representing 51 different bus orders in which PVs are integrated in the system. The total number of bus permutations (different bus orders in which PVs can be integrated in the system) is $N!$ (! – factorial, where $N = 58$ is the number of buses of the analysed system). Due to the lack of the resources in this project, number of run simulations (51) is much less than the number of simulations needed to get statistical relevance of the results. Therefore, the results from **Error! Reference source not found.** for Monte Carlo simulations should be taken as illustrative and not definite. To get more accurate statistical results for the hosting capacity of the analysed system, more Monte Carlo simulations would need to be run which remains as the future work.

In the limited number of *randomized* Monte Carlo runs, lower hosting capacity than in case of *conservative* order is obtained as well as in some limited number of *randomized* runs higher hosting capacity is obtained than in the case of using *systematic* order of buses for integrating PVs. This means that while *conservative* and *systematic* order of integrating PVs (as described in Section 4.1 illustrate dependence of PV hosting capacity on the order in which PVs are integrated in the system, they do not give theoretically the minimum and maximum available hosting capacity, respectively. To find theoretically minimum and maximum hosting capacity of the system optimization-based methods should be used, which was not the aim in this project.

In the case of unbalanced operation of the system, raising from the loss of one phase in the part of the system due to a fault, the hosting capacity is reduced as compared to the balanced operation of the system. There is no clear trend on how much this reduction is. Since the unbalance is caused only on the part of the system, the decrease of the hosting capacity is strongly influenced by the selected bus order in which PV panels are connected. A future work should analyse multiple causes for unbalanced system operation to get a better and more general picture on how unbalanced operation can affect the hosting capacity of the system.

In calculation of the hosting capacity, some simplifications have been made. The overlaying system has been assumed as “strong” and not having significant impact on the voltage changes in the analyzed LV network. Therefore, it was modelled as slack bus. While LV bus voltage limits have been taken into account, current limits of the cables have not been considered due to the lack of data. The assumption was that in LV distribution systems voltage limits are more constraining especially in the cases when reactive power compensation is not used as a mean to control the voltage. 10-minutes value of bus voltage limits of 0.95 to 1.05 pu are assumed in the study. However, some distribution system operators tend to put even more constraining voltage limits allowing rapid voltage changes of not more than 3% [20]. In case of PV integration, shading of PV panels may introduce such rapid voltage changes. These changes have not been considered in the study and remain as future work. Due to the aforementioned simplifications, hosting capacity results might have been overestimated in this study. The purpose of the study was not to precisely estimate the hosting capacity but to illustrate impacts of unbalanced operation on it.

5 Conclusions

Fast digitalization of distribution systems opens new opportunities to monitor and plan their operation as well as address the challenges with increasing penetration of PVs and EVs in the system. Some of the challenges include proper assessment of the system hosting capacity, ensuring efficient power system operation, resiliency and reliability. Smart meters, as one of the main components of distribution systems digitalization, may provide better observability of the system's current but also possible future planned operation. This project showcases the use of smart-meter data to improve distribution system operation, namely by identifying the non-technical losses and estimating system hosting capacity under various operating conditions.

In this project, smart meter data measurements were compared with simulated expected system operation to detect abnormalities. In this case – non-technical system losses. Different methods for analysing correlation between real smart meter data the simulated one were assessed to identify the discrepancies. From these discrepancies, non-technical losses have not only been identified but they were also characterized in the sense of their location and pattern (behaviour). It has been noted that the more precise measurements and availability of smart meter data from more locations can give better/more reliable results on detection and characterization of the system's nontechnical losses.

Another example of using smart meter data to improve distribution systems' operation and planning was illustrated in this project on estimating PV hosting capacity of the chosen Swedish distribution system. Smart meter data provides better operational statistics of the system allowing for more precise hosting capacity calculations. In the carried case study, the calculations have been done together with couple of sensitivity analysis scenarios. It has been assessed how the hosting capacity of the system is affected by different bus order in which PVs are integrated in the system and different power limits of PV installations per location. The results showed strong dependency of available total hosting capacity of the system on the bus order in which PVs are integrated in the system. Besides simplifying connection requirements and fair allocation of the system capacity, limiting the maximum power per PV connection can also improve total utilization of the system's hosting capacity. The results show that there is optimum value for power limit of PV installations per location which was in the studied case 20kW per connection.

The analyses in this project were carried out using open-source tools (openDSS) and data formats (CIM). The side objective was to test these tools and assess how they perform for distribution system operation and planning studies. No major issues have been identified. Open-source analysis and modelling tools as well as standardized formats for exchanging information are beneficial in the presence of many different distribution system operators and vendors providing equipment. They provide transparent environment for data, information and knowledge exchange as well as wider availability of tools that can be used for improving the system operation and planning.

6 References

- [1] Y. Huang, E. Grahn, C. J. Wallnerström, L. Jaakonanti and T. Johansson, "Smart meters in Sweden – Lesson learned and new regulations," in *3rd AIEE Symposium on Current and Future CHallenges to Energy Security*, Milan, Italy, 2018.
- [2] A. Benzerger, A. Bahmanyar, G. Derval and D. Ernst, "A Unified Definition of Hosting Capacity, Applications, and Review," *IEEE Access*, vol. 13, pp. 7819 - 7836, 2024.
- [3] Y. Hou, J. Zhu, M. Z. Liu, W. J. Nacmanson and L. F. Ochoa, "EV Hosting Capacity and Voltage Unbalance: An Australian Case Study," in *IEEE PES Grid Edge Technologies and Exposition*, San Diego, USA, 2023.
- [4] D. Schwanz, S. K. Rönnberg and M. Bollen, "Hosting capacity for photovoltaic inverters considering voltage unbalance," in *IEEE PowerTech*, Manchester, 2017.
- [5] A. A. Anderson, T. E. McDermott and E. G. Stephan, "Enabling Data Exchange and Data Integration with the Common Information Model: An Introduction for Power Systems Engineers and Application Developers," Pacific Northwest National Laboratory, Richland, Washington, 2023.
- [6] Electric Power Research Institute, "OpenDSS Documentation: Introduction to OpenDSS," 2024. [Online]. Available: <https://opendss.epri.com/IntroductiontoOpenDSS.html>. [Accessed 20 February 2026].
- [7] R. C. Dugan, "OpenDSS Training Workshop: Introduction to OpenDSS," 30 August 2021. [Online]. Available: <https://www.epri.com/events/fb4cf76b-d97d-44cc-9523-4c9243381bd1>. [Accessed 20 February 2026].
- [8] Electric Power Research Institute, "OpenDSS Documentation: Power Flow," 2024. [Online]. Available: <https://opendss.epri.com/PowerFlow.html>. [Accessed 20 February 2026].
- [9] Electric Power Research Institute, "OpenDSS Documentation: Brief History and Objectives," 2024. [Online]. Available: <https://opendss.epri.com/BriefHistoryandObjectives.html>. [Accessed 20 February 2026].
- [10] P. Radatz, "py-dss-interface: A Python package that interfaces with OpenDSS powered by EPRI," 2026. [Online]. Available: https://github.com/PauloRadatz/py_dss_interface. [Accessed 20 February 2026].
- [11] P. Radatz, "py-dss-toolkit: Advanced Python Tools for OpenDSS powered by EPRI," 2025. [Online]. Available: https://github.com/PauloRadatz/py_dss_tools. [Accessed 20 February 2026].
- [12] Sandia National Laboratories, "GridPV Toolbox," [Online]. Available: <https://pvpmc.sandia.gov/tools/gridpv-toolbox/>. [Accessed 20 February 2026].
- [13] J. Deboever and P. Radatz, "OpenDSS Training Workshop - 2021: Advanced topics & Applying DSS in R&D," 2 September 2021. [Online]. Available:

- <https://www.epri.com/events/fb4cf76b-d97d-44cc-9523-4c9243381bd1>.
[Accessed 20 February 2026].
- [14] A. A. Anderson, E. G. Stephan and T. E. McDermott, "Enabling Data Exchange and Data Integration with the Common Information Model: An Introduction for Power Systems Engineers and Application Developers," Pacific Northwest National Laboratory, Richland, Washington, 2022.
 - [15] J. Kirtley Jr., "Introduction to Power Systems," Massachusetts Institute of Technology, 2003. [Online]. Available: https://ocw.mit.edu/courses/6-061-introduction-to-electric-power-systems-spring-2011/7603169f5d8514b769a46fa900b79a57_MIT6_061S11_ch4.pdf.
 - [16] A. H. A. Rashid, M. Abdel-Akher and K. M. Nor, "Improved three-phase power-flow methods using sequence components," *IEEE Transactions on Power Systems*, vol. 20, pp. 1389-1397, 2005.
 - [17] M. Persson and A. Lindskog, "Detection and localization of non-technical losses in distribution systems with future smart meters," in *IEEE Milan PowerTech*, Milan, Italy, 2019.
 - [18] M. B. TAIS TAVERES DE OLIVEIRA, "Acceptansgränsberäkningar i distributionsnät," 2025. [Online]. Available: <https://energiforsk.se/media/34196/2025-1094-acceptansgransberakningar-i-distributionsnat.pdf>.
 - [19] F. H. Math Bollen, *Integration of Distributed Generation in Power System*, Wiley, 2011.
 - [20] H. B. Robles, A. Emanuelsson, C. J. Walnerström, M. Dalheim and M. Tjäder, "Recent updates in the swedish regulation on voltage quality," 2025. [Online]. Available: <https://ei.se/download/18.226ac6e0198bb6d7eda22e6/1755676337517/CIRED-2025-Paper-330-Recent-updates-in-the-Swedish-regulation-on-voltage-quality.pdf>.

THE BENEFITS OF NEW SMART METERS AND UNBALANCED GRID MODELS FOR INCREASED DIGITALIZATION

Digitalization of distribution systems and deployment of new smart meters brings new capabilities to solve the challenges related to the distribution systems' operation and planning. By using open-source tools, this project applies these capabilities on a case study of the Swedish distribution system to better estimate system's PV hosting capacity under unbalanced operating conditions as well as to identify and characterize non-technical system losses.

The study illustrates potential to solve the essential system challenges using license-free open-source tools and available 15-minutes measurements from smart-meters. The project results pave the way for further discussions on benefits of distribution systems digitalization and using of open-source tools and file formats to model, analyse and exchange information about the distribution systems' planning and operation.

Ett nytt steg i energiforskningen

Forskningsföretaget Energiforsk initierar, samordnar och bedriver forskning och analys inom energiområdet samt sprider kunskap för att bidra till ett robust och hållbart energisystem. Energiforsk är ett politiskt neutralt och icke vinstutdelande aktiebolag som ägs av branschorganisationerna Energiföretagen Sverige och Energigas Sverige, det statliga affärsverket Svenska kraftnät, samt gas- och energiföretaget Nordion Energi. Läs mer på energiforsk.se.

