Title: **Final Report: Model-Free Operating Envelopes**

Synopsis: This is the final report of the "Model-Free Operating Envelopes at NMI Level" Project funded by C4NET. The report presents the final version of the model-free operating envelope (OE) approach developed throughout this project and presents the main findings of the last 6 months of this project.

Data Document ID: UoM-C4NET-Model-Free_OEs-FinalReport_v02

Date: 2nd August 2023

Prepared For: Centre for New Energy Technologies (C4NET)
C/- SBC, Level 2, 520 Bourke St, Melbourne
VIC 3000, Australia

Prepared By: Vincenzo Bassi
Prof Luis(Nando) Ochoa
Department of Electrical and Electronic Engineering
The University of Melbourne

Revised By: Prof Luis(Nando) Ochoa
Prof Tansu Alpcan
Department of Electrical and Electronic Engineering
Prof Christopher Leckie
School of Computing and Information Systems (CIS)
The University of Melbourne

Contact: Prof Luis(Nando) Ochoa
+61 3 9035 4570
luis.ochoa@unimelb.edu.au

Executive Summary

This is the Final Report of the project "Model-Free Operating Envelopes at NMI Level" funded by the Centre for New Energy Technologies (C4NET) and involving the Australian Distribution Network Service Providers (DNSPs) AusNet Services, Jemena, Citipower & Powercor, and United Energy.

The increasing adoption of residential distributed energy resources (DERs), such as solar photovoltaics (PV), batteries, and electric vehicles (EVs), is driving the need for DNSPs to implement Operating Envelopes (OEs), i.e., meter-level time-varying import and export limits for active customers, to ensure network integrity. However, the main challenge for DNSPs is that an accurate determination and assessment of OEs would normally require power flow analyses and, consequently, need detailed three-phase electrical models of low voltage (LV) distribution networks, which, in practice, are rarely available and are costly and time-consuming to produce.

To overcome the lack of electrical models, this project has successfully demonstrated that is possible to exploit the historical smart meter data of customers to capture the physics of three-phase LV networks and create an electrical model-free approach to calculate voltages. These model-free voltage calculations can then be used along with a heuristic algorithm and tailored approximations to assess several import or export values for active customers (i.e., customers engaged with aggregators) and thus, determine its corresponding OEs. This approach represents an accurate, cheap, and extremely quick alternative to traditional model-based approaches, enabling DNSPs to bypass the costly, time-consuming, and error-prone process of producing and validating electrical models.

The report at hand builds on top of our previous reports [1] (foundations and methodology), [2] (extensive performance tests), and [3] (improvements and updates), and presents the final version the model-free OE approach and the main findings of the last 6 months of this project. This includes final improvements to the offline data pipeline designed to cater for customers with extremely low variability, the incorporation of thermal constraints of key assets into the model-free OEs formulation. It also provides a general overview of the model-free OEs implementation architecture and a qualitative analysis of the scalability of the proposed model-free voltage calculations for a high voltage (HV) feeder. Finally, it presents further analysis of partial smart meter data availability. The conclusions of this report can be summarised as follows.

Model-Free Operating Envelopes

Using real smart meter data from 690+ customers across United Energy, Jemena, and AusNet Services we have demonstrated that our model-free voltage calculations can achieve a high accuracy (average deviation below 1.5 Volts in most cases) which, in turn, means accurate calculations of OEs as well as for other applications (e.g., connection requests, hosting capacity, etc.).

The proposed model-free OE approach considers using model-free voltage calculations instead of power flow simulations. From an application perspective, the proposed model-free OE approach can be divided into two main stages: offline (production of the NN) and online (model-free OE calculation). Details from all steps within each stage are presented in this report. First, a neural network (NN) is trained offline to capture the underlying relationships among the historical smart meter data of customers and the corresponding LV network. Then, the trained NN is used online along with a heuristic algorithm to explore different import or export values for active customers and thus, calculate the OEs. Meter-level voltage compliance is assessed through the trained NN, whereas thermal compliance of transformer, conductors, and customers' connection point, is assessed through its respective thermal ratings and customers' demands. Removing, completely, the need for power flow analyses and electrical models. A case study is presented to illustrate the implementation of the proposed model-free OEs, showing that the proposed approach is an accurate, cheap, and fast alternative to traditional model-based approaches to calculate OEs, enabling DNSPs to accurately calculate OEs without the need for the costly, time-consuming, and error-prone process, of producing and validating electrical models.

Offline Data Pipeline: Final Improvements

Improvements to the offline data pipeline were made due to data issues involving customers with extremely low variability (i.e., all or most values in a range of a few Watts).

The improvements were necessary because the historical data of such customers does not provide variety enough for the NN to adequately capture its relationships. Thus, any adequate value for such customers will be far from what has been observed during training and will cause erroneous calculations for all customers. This is critical in the context of OEs as its calculation and assessment is based on pushing customers' limits. The potential impacts of such customers when deploying the proposed approach are shown and to address this, Step 4 of the offline data pipeline is adapted to automatically detect and remove such customers from the historical data and thus, from the proposed model-free voltage and OEs calculations. Once sufficient adequate data for these customers is available, the NN can be updated to cater for these customers.

Thermal Constraints in Model-Free Operating Envelopes

The proposed model-free OEs approach can capture thermal violations of key assets and limit the corresponding OEs accordingly to ensure network integrity, providing a full model-free approach to assess both, voltage, and thermal compliance, removing completely the need for power flow analyses or electrical models.

For simplicity, only voltage constraints and customers' connection point limitations were assessed before this report. However, thermal constraints of key assets can also limit active customers imports or exports and thus, must be considered when calculating OEs. The NN used to calculate voltages and thus, assess meter-level voltage compliance when calculating OEs, cannot cater for asset utilisation as it was trained to calculate voltages only. Therefore, thermal constraints are approximated by using customers net demands and the thermal capacity (in kVA) of the corresponding assets.

Model-Free Operating Envelopes: Implementation Architecture

The proposed model-free OEs approach can be potentially integrated into current systems of DNSPs. This requires two blocks, Model-Free Engine and Model-Free OEs.

The first block is offline, is dedicated to produce the corresponding NN and thus, it requires access to the different data bases where its corresponding inputs are found every time an update is carried out. The second block, on the other hand, is dedicated to use the obtained NN to calculate OEs online, it must be integrated within the Distributed Energy Resources Management System (DERMS) and must have access to operational advanced metering infrastructure (AMI) data or forecast, depending on how OEs are required, i.e., in near real-time or in advance, respectively.

Voltage Calculations per HV Feeder: Qualitative Analysis

Having a single NN per HV feeder is likely to be impractical as HV and LV networks can be subject to changes (e.g., reconfigurations, reinforcements, or connection of new customers) which in turn would require more frequent updates of the NN than when applied to individual LV networks.

The proposed model-free OEs approach is carried out per distribution transformer. However, scaling the approach to a single NN per HV feeder could bring interesting benefits, such as simplicity, accuracy improvements and, potentially, reducing computational times. Nonetheless, while this might seem attractive, it might result impractical as the multiple changes that occur on the HV and LV networks would require the NN used for the entire HV feeder (i.e., catering for multiple LV networks simultaneously) to be updated frequently.

Partial Smart Meter Data Availability

The proposed model-free approach can calculate voltages with an adequate accuracy and, therefore, adequate OEs, even if data from only 20% of the customers is available.

Although Victoria has full deployment of smart meters, the data ($\boldsymbol{P}$, $\boldsymbol{Q}$, and $\boldsymbol{V}$) of commercial and industrial (C&I) customers might not be available to the DNSPs (only kWh) as it is managed by third parties. To understand the implications of limited data availability (which is also relevant for DNSPs outside Victoria), this report investigates decreasing levels of smart meter data availability. Results are promising as adequate model-free voltage calculations and OEs can be obtained even if data from only 20% of customers is available. Nonetheless, the consistency of the OEs, will depend on the location and characteristics of the customers whose smart meter data is not available.

Table of Contents

List of Figures

List of Tables

Acronyms

AMI: Advanced Metering Infrastructure

C&I: Commercial and Industrial

DERMS: Distributed Energy Resources Management System

DERs: Distributed Energy Resources

DNSP: Distribution Network Service Provider

EV: Electric Vehicle

HV: High Voltage

LV: Low Voltage

NN: Neural Network

OE: Operating Envelope

PV: Photovoltaic

1 Introduction

The increasing adoption of residential distributed energy resources (DERs), such as solar photovoltaics (PV), batteries, and electric vehicles (EVs) is opening the door for customers to engage with aggregators and provide system-level services (hereinafter, customers engaged with aggregators will be referred to as active customers). This, in turn, drives the need for distribution network service providers (DNSPs) to implement Operating Envelopes (OEs), i.e., meter-level time-varying import and export limits for active customers, to ensure network integrity. Proof of this are the multiple works in the area that have been recently carried out across Australia [4]-[9].

The challenge for DNSPs is that an accurate determination and assessment of OEs would normally require power flow analyses [10]-[12] and, consequently, need detailed three-phase electrical models of low voltage (LV) distribution networks, which are rarely available and are costly and time-consuming to produce [13].

To overcome the lack of electrical models, this project has successfully demonstrated that is possible to exploit the historical smart meter data of customers to capture the physics of three-phase LV networks and create an electrical model-free approach to calculate voltages. These model-free voltage calculations can then be used along with a heuristic algorithm and tailored approximations to assess several import or export values for active customers and thus, determine its corresponding OEs. This approach represents an accurate, cheap, and extremely quick alternative to traditional model-based approaches, enabling DNSPs to bypass the costly, time-consuming, and error-prone process of producing and validating electrical models.

The foundations and methodology of the proposed electrical model-free voltage calculations along with the initial test were presented in an interim report “Deliverable 0: Concept, Smart Meter Data, and Initial Findings” [1]. The second report, “Deliverables 1-2-3a: Model-Free Voltage Calculations and Operating Envelopes” [2], showed that the proposed approach can carry out accurate multi-LV circuit voltage calculations achieving and average deviation below 1 V in many cases. It was found that the proposed approach can use, at any moment in time, as little as 3 weeks of historical smart meter data to produce a NN that is adequate to calculate voltages for at least 10 months without requiring further updates. The model-free approach to calculate OEs was also presented, giving the foundations for its implementation, and presenting its initial tests and findings. Furthermore, the third report “Deliverables 3b-4: Improved Model-Free Operating Envelopes and Other Considerations” [3] presented improvements to the offline data pipeline, an additional allocation technique to calculate OEs, and assessed up to which extent the proposed approach can cater for voltage regulation devices. Besides, preliminary analyses of partial smart meter data availability are presented, showing that the proposed approach can obtain accurate voltage calculations as well as consistent OEs when smart meter data from all customers is not available.

This report presents the final version the model-free OE approach developed throughout this project (Section 2) and the main findings of the last 6 months of this project (Sections 3 to 7). Section 3 presents final improvements to the offline data pipeline designed to cater for customers with extremely low variability. Section 4 shows the incorporation of thermal constraints of key assets into the model-free OEs formulation. Section 5 presents a general overview of the model-free OEs implementation architecture. Section 6 shows a qualitative analysis of the scalability of the proposed model-free voltage calculations per high voltage (HV) feeder, and Section 7 presents further analysis of partial smart meter data availability.

2 Model-Free Operating Envelopes

This section builds on top of the previous reports of this project [1]-[3], as well as on top of the latest findings presented in subsequent sections of the report at hand and presents the complete model-free OEs approach developed throughout this project.

The proposed model-free OE approach adapts the model-based OE calculation approach defined in project EDGE in [14] to use model-free voltage calculations [1], [15] instead of power flow simulations. To achieve this, first, a NN is trained to capture the underlying relationships among the historical smart meter data of customers and the corresponding LV network. Then, the trained NN is used along with a heuristic algorithm to explore different import or export values for active customers and thus, calculate the OEs. Meter-level voltage compliance is assessed through the trained NN, whereas thermal compliance of transformer, conductors, and customers' connection point, is assessed through its respective thermal ratings and customers' demands. Removing, completely, the need for power flow analyses and thus, enabling DNSPs to bypass the costly, time-consuming, and error-prone process, of producing electrical models.

From an application perspective, the proposed model-free OE approach can be divided into two main stages: offline (production of the NN) and online (model-free OE calculation). The inputs and outputs of the entire process are detailed below. Further details of each stage are presented in the following subsections.

Inputs:

- Historical AMI data (i.e., $\boldsymbol{P}$, $\boldsymbol{Q}$, and $\boldsymbol{V}$) of customers associated with a given distribution transformer.
- Active and reactive power of passive customers (no need for OE calculation). This can come either from operational AMI data or forecasts, depending on how OEs are being calculated, i.e., in near real-time or hours, days ahead, respectively.
- Reactive power of active customers (those needing OE calculation). This can come either from operational AMI data or forecasts, depending on how OEs are being calculated.
- Voltage statutory limits of residential customers (e.g., 216V and 253V).
- Rated capacity (in kVA) of customers' service cables (or connection agreements).
- Rated capacity (in kVA) of conductors at the head of the LV network.
- Rated capacity (in kVA) of the distribution transformer.

Outputs:

- OEs for active customers (in kW), i.e., maximum imports and exports values according to the adopted allocation technique and subject to network constraints.

2.1 Offline Stage: Neural Network Production

The first stage occurs offline, and its objective is to produce a NN that will be then used to assess meter-level voltage compliance when calculating OEs in the online stage. This process is carried out per distribution transformer and, thus, requires meter to transformer relations. Ideally, historical per-phase smart meter data from all customers connected to the corresponding transformer is used. However, it has been shown in [3] and in Section 7, that the proposed approach can produce accurate voltage calculations even when smart meter data from only a reduced number of customers is available. This is also discussed by the end of this section (section 2.4).

The historical per-phase smart meter data of customers required for this stage is given by the data sets $\boldsymbol{P}$, $\boldsymbol{Q}$, and $\boldsymbol{V}$ in (1)-(3). Specifically, $\boldsymbol{P}$, $\boldsymbol{Q}$, and $\boldsymbol{V}$ contain instances, P_t, Q_t, and V_t, comprised of active power, reactive power, and voltage magnitude, respectively, from all customers in C (set of customers indexed by c) at instance $t \in T$ (set of instances). Note that instances are collected at regular intervals (e.g., 5, 30 minutes) throughout the considered period (e.g., a month, a year), $|T|$ is to represent the total number of instances in T, and $|C|$ denotes the total number of customers (per-phase) in C.

$$\boldsymbol{P} = \begin{bmatrix} P_1 \\ \vdots \\ P_{|T|} \end{bmatrix} = [P_t] \in \mathbb{R}^{|T|\times|C|} \tag{1}$$

$$\boldsymbol{Q} = \begin{bmatrix} Q_1 \\ \vdots \\ Q_{|T|} \end{bmatrix} = [Q_t] \in \mathbb{R}^{|T|\times|C|} \tag{2}$$

$$\boldsymbol{V} = \begin{bmatrix} V_1 \\ \vdots \\ V_{|T|} \end{bmatrix} = [V_t] \in \mathbb{R}^{|T|\times|C|} \tag{3}$$

The proposed model-free voltage calculations approach contemplates to use a NN trained to capture the underlying relationships among its inputs (i.e., customers' active and reactive power, $\boldsymbol{P}$ and $\boldsymbol{Q}$, respectively) and its outputs (i.e., customers' voltages $\boldsymbol{V}$), as a proxy of power flow simulations. This is shown in (4), where $\boldsymbol{W_{NN}}$ comprise all the parameters (i.e., weights and biases) of the NN, which dimensions will vary depending on the structure of the NN. Additionally, $\Delta^{\boldsymbol{V}}{}_{\boldsymbol{NN}}$ comprise the corresponding approximation errors (per-phase) of all instances $t \in T$, i.e., $\Delta^{\boldsymbol{V}}{}_{\boldsymbol{NN}} = \left[\Delta^{V}{}_{NN_t}\right] \in \mathbb{R}^{|T|\times|C|}$. Further details on the historical smart meter data required and on the complete model-free voltage calculations approach can be found in [1],[15].

$$\boldsymbol{V} = f_{NN}(\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{W_{NN}}) + \Delta^{\boldsymbol{V}}{}_{\boldsymbol{NN}} \tag{4}$$

While producing a NN for such task can be relatively straightforward using available open-source tools if the corresponding smart meter data is available, there are several application-specific considerations that must be considered for an adequate model-free OEs implementation. These are described below.

- **Smart meter data:** It is a well-known issue that any machine learning model is as good as the data that is fed, and dealing with real smart meter data is not exempt from challenges. Smart meter data is noisy (due to the inherent noise of smart meter devices) and can be subject to missing or erroneous measurements due to error in communications or smart meter malfunctions. Therefore, the historical smart meter data of customers must be pre-processed before being used. For further details, refer to [1], [3], and Section 3.

- **Data requirements:** While from a data analytics perspective it would be ideal to use as much data as possible to produce the corresponding NN, this is unfeasible in the context of LV networks. HV and LV networks are often subject to changes, e.g., reconfigurations, connection of new customers, reinforcements, etc. Hence, extensive data requirements (e.g., a year of data) will likely cover different conditions for the studied LV network (e.g., different topologies) which will ultimately impact the accuracy of the voltage calculations. Therefore, data requirements must be minimised. For further details, refer to [2].

- **Diversity in LV networks:** Different LV networks have different characteristics (e.g., number of circuits, overhead/underground conductors, number of customers, PV penetration, etc.). However, NNs that are found to be the most suitable for a LV network with certain characteristics will not necessarily perform well in another LV network with different characteristics. For example, throughout this project it has been noted that higher learning rates tend to be more effective for smaller LV networks, whereas, on the other hand, smaller learning rates tend to produce better results when considering larger LV networks. Therefore, a NN selection process adaptable to any kind of LV network must be considered to produce the most suitable NN in each case. For further details, refer to [1]-[3], [15].

- **Scalability:** DNSPs can manage thousands of LV networks within their service areas and OEs can be required in near real-time (e.g., calculated every 5 minutes). While deep NNs (i.e., with multiple hidden layers) have shown great performance in different tasks such as natural language processing and image recognition, larger NN structures have bigger production and deployment times. Therefore, to be scalable, the NN structure must be designed to minimise production times (and, thus, minimise the associated times and costs of the offline process) and, more importantly, minimise deployment times (OE calculation must lie within a control cycle). For further details, refer to [15].

This study proposes a model-free voltage calculations approach designed with all these considerations. A schematic of the proposed approach is presented in Figure 1, where it can be observed that the approach can be divided into three main stages: Data processing, NN selection, and Model-Free Voltage Calculations (which occurs online). Thus, first, historical smart meter data is processed to build an adequate training data set. Then, the obtained training data set is used in a NN selection process to train several NN with different hyperparameters as in (5), where the subindex u correspond to the $u-th$ NN in set U (set of trained NNs). With all these NNs, the most suitable NN for the studied LV network, i.e., NN_{u^*}, is defined. By the end of this process, NN_{u^*} is ready to carry out voltage calculations. Inputs and outputs of the entire approach are presented below. For further details, refer to [1], [15].

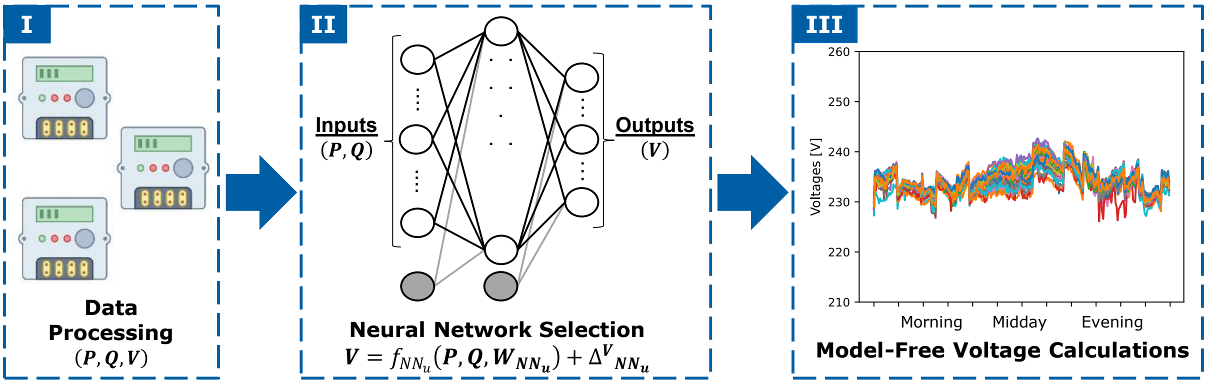

Figure 1. Model-Free Voltage Calculations Approach

$$\boldsymbol{V}=f_{NN_u}\left(\boldsymbol{P},\boldsymbol{Q},\boldsymbol{W}_{\boldsymbol{NN_u}}\right)+\Delta^{\boldsymbol{V}}{}_{\boldsymbol{NN_u}} \tag{5}$$

- **Inputs:** Historical per-phase smart meter data of customers (i.e., $\boldsymbol{P}$, $\boldsymbol{Q}$, and $\boldsymbol{V}$)
- **Outputs:** NN_{u^*} ready to carry out voltage calculations

2.1.1 Data Processing

To maximise the performance of the proposed approach in capturing the physical relationships of a given LV network, and adequate training data set i.e., with correct values for all customers at each instance, is required [1]-[3]. The training data set is denoted by $\{\boldsymbol{P^t},\boldsymbol{Q^t},\boldsymbol{V^t}\}$, and corresponds to a continuous subset of instances from the set of per-phase historical smart meter data $\{\boldsymbol{P},\boldsymbol{Q},\boldsymbol{V}\}$ (e.g., the latest weeks). However, as mentioned before, dealing with real smart meter data is not exempt from challenges. Therefore, to build an adequate training data set first, the following data pipeline is implemented. For further details, refer to [1], [3], and Section 3.

1. **Step 1: Collect raw historical per-phase smart meter data**
   Retrieve historical per-phase smart meter data from all customers in the studied LV network as presented in [1]. Note that smart meters are designed to measure energy consumption in a wide range of formats and thus, the convention adopted for a specific smart meter needs to be properly defined to ensure the correct interpretation of the measurements. For more details, refer to [1].

2. **Step 2: Derive P and Q from raw data**
   While $\boldsymbol{V}$ is measured by smart meters, $\boldsymbol{P}$ and $\boldsymbol{Q}$ must be derived from other available measurements. Such available measurements are specific for each DNSPs and, thus, $\boldsymbol{P}$ and $\boldsymbol{Q}$ derivation will also be specific for each DNSP. Further details of $\boldsymbol{P}$ and $\boldsymbol{Q}$ derivation for each DNSP can be found in [1].

3. **Step 3: Remove invalid and unfeasible instances**
   From the perspective of model-free voltage calculations, it is essential that the training data set is comprised only of valid and feasible instances [1], [2]. Thus, if there are gaps or unfeasible values in the data for a single time instance (e.g., a missing or zero voltage value for one customer) then all customers data for that time instance must be removed and deemed unusable as inputs to the proposed approach. For more details, refer to [3].

   It is important to note that if a customer is found to have invalid (i.e., missing values) or unfeasible values (e.g., V = 0) throughout the entire data set, this customer must be removed from the historical data first and then invalid/unfeasible instances for the remaining customers must be removed. If not, such customer will lead to remove the complete data set.

4. **Step 4: Remove customers whose relationships are not represented**
   Customers with extremely low values and variability can cause erroneous voltage calculations when considering adequate values for such customers as its underlying relationships are not properly represented in the historical data used to train the corresponding NN. Therefore, such customers must be removed from the historical data and, thus, from the voltage calculations. Once adequate data for such customers is available, the NN can be updated to cater for such customers. For more details, refer to [3] and Section 3.

Once the complete data pipeline is implemented, the training data set $\{\boldsymbol{P^t}, \boldsymbol{Q^t}, \boldsymbol{V^t}\}$ is build considering a continuous subset of instances from the data set of per-phase historical data $\{\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{V}\}$. As shown in [2], the proposed approach aims to minimise data requirements and, thus, considers only three weeks of historical smart meter data at 5 minutes resolution. Hence, $\{\boldsymbol{P^t}, \boldsymbol{Q^t}, \boldsymbol{V^t}\}$ is built with the last 6,048 instances (equivalent to 3 weeks at 5 minutes resolution) of $\{\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{V}\}$. Note that this may end up being longer than 3 calendar weeks due to the implemented data pipeline (e.g., if there are gaps in the data).

2.1.2 Neural Network Selection

To cope with diversity in LV networks, a NN selection process adaptable to any kind of LV network based on its corresponding per-phase number of customers (i.e., $|C|$) is implemented. Furthermore, to enhance scalability of the proposed approach in both, production and deployment, this study considers NNs with a single hidden layer only. Further studies were carried out with deeper NN architectures in [15], where it is observed that while deeper NN architectures could bring some minor benefits in terms of accuracy, such benefits come at the expense of longer production and deployment times.

The first step to select the most suitable NN for a given LV network is to define those hyperparameters and NN characteristics that are determined by the nature of the problem. These are presented in Table 1 and are fixed for all NNs that will be trained in this stage. For more details, refer to [1], [2], [15].

Table 1. Hyperparameters and NN characteristics given by the nature of the problem

NN Inputs	NN Outputs	Output Activation function	Scaler	Loss function	Optimiser	Regularisation
$2\|C\|$	$\|C\|$	Linear	[0,1]	Mean Squared Error (MSE)	Adam	L2

To define the remaining hyperparameters, i.e., number of neurons in the hidden layer and its activation function, learning rate, L2 regularisation factor, batch size, and epochs, a K-fold cross validation process is implemented. Thus, search spaces for each hyperparameter are defined (Table 2) and, for each plausible hyperparameter combination, K NNs are trained with K-fold and assessed using its corresponding $RMSE_{kfold}$ (given by the average root mean squared error -RMSE- among the K NNs). For further details on the K-fold cross validation process refer to [1], [2], [3], [15]. Once the most suitable hyperparameter combination is found (lowest $RMSE_{kfold}$), 10 NNs are trained from scratch with those hyperparameters, the NN with the lowest training RMSE correspond to NN_{u^*} and is ready to carry out voltage calculations.

Table 2. Hyperparameters search spaces

N° of Hidden Layers	Number of Neurons	Activation functions	Learning rate	L2 regularisation	Batch size	Epochs
1	$0.5\|C\|$ $1\|C\|$ $2\|C\|$ $3\|C\|$ $4\|C\|$ $5\|C\|$ $6\|C\|$ $7\|C\|$ $8\|C\|$ $9\|C\|$ $10\|C\|$	Tanh ReLu Swish	1×10^{-3} 1×10^{-4} 1×10^{-5}	1×10^{-3} 1×10^{-4} 1×10^{-5}	72 (eq. 6 hrs) 144 (eq. 12 hours) 288 (eq. 24 hours)	500 1,000 2,000
Hyperparameters combinations	2,673					

From an application perspective, it is of interest to reduce the search spaces as this will directly reduce the times and costs associated with the production of the corresponding NN. Based on the results in [2], it can be noted that hyperbolic tangent (Tanh) is selected in all studied LV networks. Thus, the total number of hyperparameter combinations presented in Table 2 can be reduced in 2/3 by fixing the activation function to tanh. Further search space reductions can be explored.

2.1.3 Model-Free Voltage Calculations

Once the most suitable NN for the studied LV network, i.e., NN_{u^*}, is obtained, it can be then used to calculate customers' voltages, i.e., V^{calc}, for any kind of customers what if-scenarios as in (6), where $\bar{P}$, $\bar{Q}$ are to represent a single instance of active and reactive power values for all customers as in (7), (8). For further details, refer to [1], [15].

$$\bar{V}^{calc} = f_{NN_{u^*}}\left(\bar{P}, \bar{Q}, \boldsymbol{W}_{NN_{u^*}}\right) \tag{6}$$

$$\bar{P} = [\bar{p}_1 \quad \cdots \quad \bar{p}_{|C|}] = [\bar{p}_c] \in \mathbb{R}^{|C|} \tag{7}$$

$$\bar{Q} = [\bar{q}_1 \quad \cdots \quad \bar{q}_{|C|}] = [\bar{q}_c] \in \mathbb{R}^{|C|} \tag{8}$$

This capability is further exploited to calculate OEs in the online stage. From an application perspective, it is important to note that NN_{u^*} will require active and reactive power values from all customers considered during training to carry out the corresponding voltage calculations. Therefore, if adequate data for some customer is not available, these values must be replaced by 0. This will ultimately impact the accuracy of the approach as the relationships related to those data points will not be represented. However, as shown in [3] and Section 7, the proposed approach can achieve accurate voltage calculations even when smart meter data from all customers connected to the corresponding LV network is not available.

Finally, the offline stage must be carried out periodically (e.g., every quarter, semester) to update the NN and cater for changes that can occur in LV networks that have an impact in the underlying relationships that are being captured by the NN, e.g., reconfigurations, reinforcements, connection of new customers, etc. While updates are required in the case of a change in the LV, it has been shown in [2] that, if no changes in the LV network occur, the proposed approach can produce a NN that is able to carry out accurate voltage calculations for 10 further months without requiring or benefit significantly from further updates.

2.2 Online: Model-Free Operating Envelopes

In the online stage, the most suitable NN obtained in the offline stage is used to explore several import or export values for active customers and, thus, determine its corresponding OEs. This can be done in near real-time or in advance (e.g., hours, days ahead). When calculating OEs in near real-time, operational AMI data is required. On the other hand, when calculating OEs in advance, forecast data is required.

The proposed model-free OEs approach adapts the model-based OEs calculation approach defined in project EDGE in [14] to use model-free voltage calculations instead of power flow simulations. As in [14], the model-free OEs approach is carried out per instance and assumes that active customers can provide services related to active power only. Additionally, since the exact service required is not known in advance, i.e., either imports or exports, both limits for the OE must be calculated separately. The export limit calculation assumes that all active customers provide export services, and the calculation of the import limits assumes all active customers provide import services. Thus, the two worst-case scenarios, i.e., all active customers simultaneously exporting or importing active power are considered. The proposed model-free approach to calculate OEs is shown in Figure 2 and detailed in the following subsections. For further details, refer to [2], [3].

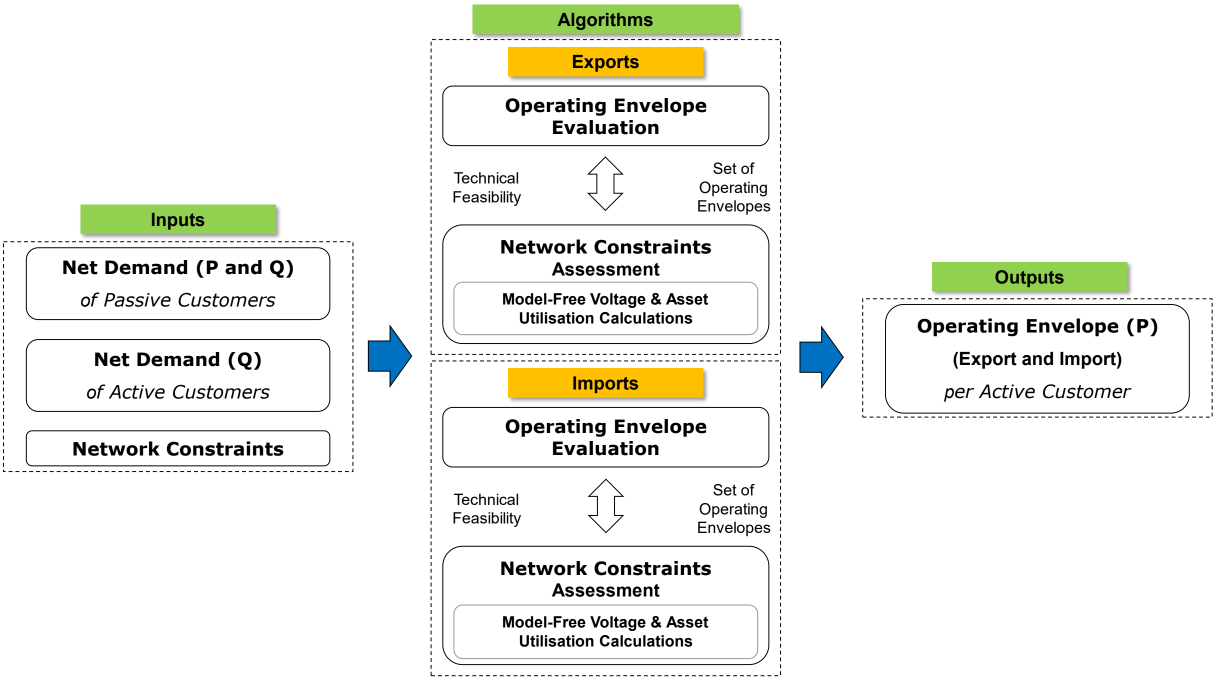

Figure 2. Model-Free OEs Approach

2.2.1 Inputs

Net Demand (P and Q) of Passive Customers: This is the net active and reactive power demand (i.e., P and Q) of passive customers (those not engaging with aggregators) at the instance of interest. This is at the point of connection, i.e., where the smart meter is located.

Net Demand (Q) of Active Customers: This is the net reactive power demand (i.e., Q) of active customers at the instance of interest. This is at the point of connection, i.e., where the smart meter is located.

The inputs above can come either from operational AMI data or forecast, depending on how OEs are being calculated, i.e., in near real-time or in advance, respectively.

Network constraints: These are the technical limits of the studied LV network.

- Voltage statutory limits
- Rated capacity (in kVA) of:
  - The distribution transformer.
  - The conductor at the head of the LV respective circuits.

- Connection point of customers, based on service cable or contracts.

2.2.2 Algorithms

There are two main blocks in the model-free OEs algorithm. OE evaluation and Network Constraints Assessment. Both blocks interact in an iterative process used to determine the maximum exports or imports according to the adopted allocation technique (i.e., how the OEs are allocated among customers).

OE Evaluation: In the first iteration, an initial set of OEs for active customers is built and sent to the Network Constraints Assessment block. There, meter-level voltage compliance and thermal compliance of key assets is assessed. If everything is OK, a new set of OEs with increased values according to the adopted allocation technique is produced and assessed. This process is repeated until the available network capacity is fully utilised, i.e., the set of OEs cannot be increased any further without exceeding the voltage and/or thermal limits of the LV network. The algorithm considered in this project starts with an initial set of OEs with $P = 0\ kW$ for all active customers and increases its value according to the adopted allocation technique in a fixed value, e.g., 0.1 kW each iteration, until network constraints are breached.

- **Allocation technique:** a key feature of the algorithm is the allocation technique to be considered. This will naturally impact the obtained OEs as define how OEs are allocated among the active customers. Two allocation techniques are explored in this project, equally distributed OEs [2] and maximise exports OEs [3]. In the context of equally distributed OEs, all active customers are assigned the same OE at each iteration, whereas maximise exports OEs seeks to maximise the overall imports or exports of the entire LV network by limiting the OEs from those customers that are more sensitive to voltages (i.e., those that are located further from the distribution transformer and thus, subject to larger voltage rise, drops).

Network Constraints Assessment: This block assesses the technical feasibility of each set of OEs sent by the OE evaluation block. In here meter-level voltage compliance is assessed using NN_{u^*}, whereas utilisation of transformer, conductors, and customers' connection point is evaluated using its corresponding thermal ratings and customers' demand. There is a total of 4 network constraints that are assessed within this block. For more details in terms of voltage assessments, refer to [2], [3]. On the other hand, for further details in terms of thermal assessments, refer to Section 4.

1. **Meter-level voltages:** In here, NN_{u^*} is used as in (6) to calculate voltages for all customers, i.e., V^{calc}, and thus, assess voltage compliance with Australian statutory limits as in (9), with $\mathrm{V}_{max} = 253\,V$ and $\mathrm{V}_{min} = 216\,V$ as per Australian regulation [16].

$$\mathrm{V}_{min} \leq \bar{V}^{calc} \leq \mathrm{V}_{max} \tag{9}$$

2. **Connection point:** Here, the apparent power of each customer, i.e., s_c with $c \in C$, is calculated as in (10). Thus, active customers imports or exports are limited by the maximum capacity of its corresponding connection point ($s_{cp,c}$) as in (11).

$$s_c = \sqrt{(P_c)^2 + (Q_c)^2} \tag{10}$$

$$s_c \leq \mathrm{s}_{cp,c}\ \forall c \ in\ C \tag{11}$$

3. **LV conductor:** As detailed in Section 4, LV conductors' constraints are assessed at the head of the LV network (i.e., at the secondary side of the distribution transformer). Thus, the aggregated apparent power from all customers, i.e., s^{Σ}, is calculated as in (12) and active customers are allowed to import or export until the per-phase apparent power, approximated as $\frac{s^{\Sigma}}{3}$, reaches the per-phase capacity of the conductor at the head of the LV network ($s_{conductor}$) as in (13).

$$s^{\Sigma} = \sum_{c \ in \ C} s_c \tag{12}$$

$$\frac{s^{\Sigma}}{3} \leq s_{conductor} \tag{13}$$

4. **Transformer:** In here, s^{Σ} is calculated as in (12) and customers are allowed to import or export up to the maximum capacity of the transformer as in (14).

$$s^{\Sigma} \leq s_{transformer} \tag{14}$$

While LV conductor and transformer analyses should be, ideally, implemented per-phase, phase-grouping information is rarely available for DNSPs. Thus, simplifications are made. However, note that data-driven techniques have been presented in the literature [17]-[19] to identify customers' phase grouping in LV networks using only historical voltage measurements. Hence, these approaches could be potentially implemented to identify customers' phase grouping and carry out a per-phase analysis of LV conductor and transformer utilisation. However, this is outside of the scope of this study.

Finally, the steps to calculate model-free OEs with each of the studied allocation techniques are summarised in Table 3. Note that, for compactness, only the case of export limits is presented. However, it is important to highlight that for the case of import limits, the calculation mirrors the one presented in Table 3 but considering active customers importing active power.

Table 3. Model-Free OEs Calculation Steps. Equally distributed and maximise exports

Equally Distributed OEs	Maximise Exports OEs
1. Obtain NN_{u^*}	**1. Obtain NN_{u^*}**
2. OE Evaluation → Initialise input values: • Near Real Time OEs: Operational AMI data • In advance: Forecast • 1st set of OEs: all active customers $P = 0\,kW$	**2. OE Evaluation → Initialise input values:** • Near Real Time OEs: Operational AMI data • In advance: Forecast • 1st set of OEs: all active customers $P = 0\,kW$

3. Network Constraints Assessment → Check:
• Check voltage and thermal compliance of initial set of OEs
  ◦ If Ok, continue
  ◦ If not, active customers cannot export ($OE = 0\,kW$)
	3. Network Constraints Assessment → Check:
	• Check voltage and thermal compliance of initial set of OEs
	  ◦ If Ok, continue
	  ◦ If not, active customers cannot export ($OE = 0\,kW$)

4. OE Evaluation → Increase OEs • Increase in 0.1 kW the OE of all active customers	**4. OE Evaluation → Increase OEs** • Identify the active customer with the lowest voltage and increase in 0.1 kW its OE

5. Network Constraints Assessment → Check:
• Check voltage and thermal compliance of the set of OEs
  ◦ If Ok, back to step 4.
  ◦ If not, OE correspond to the previous iteration
	5. Network Constraints Assessment → Check:
	• Check voltage and thermal compliance of the set of OEs
	  ◦ If all Ok, back to step 4.
	  ◦ If only the service cable constraint of the active customer identified in Step 3 is breached, back to the previous iteration, step 4, remove the active customer from the set of active customers whose OE can be increased and continue. If there are no more active customers whose OE can be increased, OE correspond to the previous iteration.
	  ◦ If any of the other 3 constraints is breached, OE correspond to the previous iteration.

2.2.3 Outputs

Operating Envelope per Active Customer: The final output corresponds to the OEs for active customers, i.e., maximum imports and exports values according to the adopted allocation technique and subject to network constraints.

2.3 Case Study

To illustrate the implementation of the proposed model-free OEs approach, the case of Jemena Substation B is used. Specifically, day ahead OEs are calculated for the 8th of January of 2021 considering both allocation techniques described in Section 2.2.2. As detailed in Section 2.2.1, in practice, this require forecast of passive customers' active and reactive power as well as active customers' reactive power. However, forecast is out of the scope of this project. Therefore, hindsight OEs are calculated, i.e., perfect forecast for these values is considered (i.e., values are obtained from the historical data). Besides, it is important to note that, for compactness, only the case of export limits is presented. However, it is important to highlight that for the case of import limits, the calculation mirrors the one for the case of export limits but considering active customers importing active power. Further details of each stage are presented below.

2.3.1 Offline Stage: Producing the Neural Network

2.3.1.1 Data Processing

As detailed in Section 2.1.1, the first step is to retrieve and process the historical smart meter data of customers. This historical smart meter data of customers is then used to produce an adequate training data set that is used in the subsequent NN selection process to define and train the most suitable NN for the studied LV network.

Jemena substation B corresponds to a LV network comprised of 4 LV circuits with a total of 155 customers (116 single-phase customers and 39 three-phase customers). Originally, 5 minutes resolution historical smart meter data for the period 01-09-2020 to 30-09-2021 was provided by the DNSP. However, later incorporation of customers, which directly impacts the underlying relationships of the LV network, were detected at 05-09-2020, 19-10-2020, and 22-11-2020. Therefore, the longest period without changes in the LV network was used for this project, i.e., 22-11-2020 to 30-09-2021. For further details, refer to [2].

Steps 1-3 of the data pipeline defined in Section 2.1.1 are implemented to the complete data set of historical smart meter data. Details for each step are provided below.

1. **Step 1: Collect raw historical per-phase smart meter data**
   The raw per-phase historical smart meter data received for the case of Jemena substation B correspond to customers' voltage magnitudes, current magnitudes, power factors, and real currents, spanning the period from 22-11-2020 to 30-09-2021. For further details, refer to [1].

2. **Step 2: Derive P and Q from raw data**
   Here, the raw historical smart meter data retrieved in Step 1 is processed to derive the corresponding values of customers' active and reactive power, i.e., $\boldsymbol{P}$ and $\boldsymbol{Q}$. For further details on $\boldsymbol{P}$ and $\boldsymbol{Q}$ derivation for the case of Jemena substation B, refer to [1].

3. **Step 3: Remove invalid and unfeasible instances**
   Here, instances with missing or unfeasible values of active power, reactive power, or voltage magnitudes, for at least one customer are removed from the historical data. After removing such instances, the complete data set of historical smart meter data for Jemena substation B is comprised of 44,632 instances, which is equivalent to $\approx$ 155 days (i.e., $\approx$ 22 weeks), spanning the period among 22-11-2020 and 30-09-2021 For further details, refer to [2].

After Steps 1-3 of the data pipeline are implemented, there is an equivalent to 22 weeks of historical smart meter data for Jemena substation B. To cater for a summer case study (i.e., with high injections from PV systems), this case study considers the first 6,048 instances (i.e., an equivalent of 3 weeks at 5 minutes resolution) of the historical data set as training data set, spanning the period between 22-11-2020 and 07-01-2021. Note that this correspond to more than 3 calendar weeks as instances with invalid or unfeasible values for at least one customer were removed in Step 3.

Once the training data set is defined, Step 4 of the data pipeline presented in Section 2.1.1 is implemented, i.e., customers whose relationships are not represented are removed from the training data set and thus, from the model-free calculation of voltages. Note that Step 4 is applied only once the training data set is defined as customers removed at this stage can be subject to smart meter malfunctions, which may not be persistent throughout the historical data, i.e., adequate data for these customers could be found in other parts of the historical data.

4. **Step 4: Remove customers whose relationships are not represented**
   Here, all customers with 99% of its corresponding active power values below 100 W are removed from the training data set and consequently, from the voltage calculations. For further details, refer to [3] and Section 3. For this case study, 8 single-phase customers, 1 phase of a three-phase customer and a three-phase customer are removed.

After the complete data pipeline is implemented, the training data set is comprised of an equivalent to three weeks at 5 minutes resolution (i.e., 6,048 instances) of historical smart meter data (i.e., $\boldsymbol{P}$, $\boldsymbol{Q}$ and $\boldsymbol{V}$) for 146 customers; 108 single-phase customers, 1 two-phase customer, and 37 three-phase customers, i.e., per-phase number of customers $|C| = 108 + 2 + 37 \times 3 = 221$. The obtained training data set spans the period between 22-11-2020 and 07-01-2021 and is presented in Figure 3.

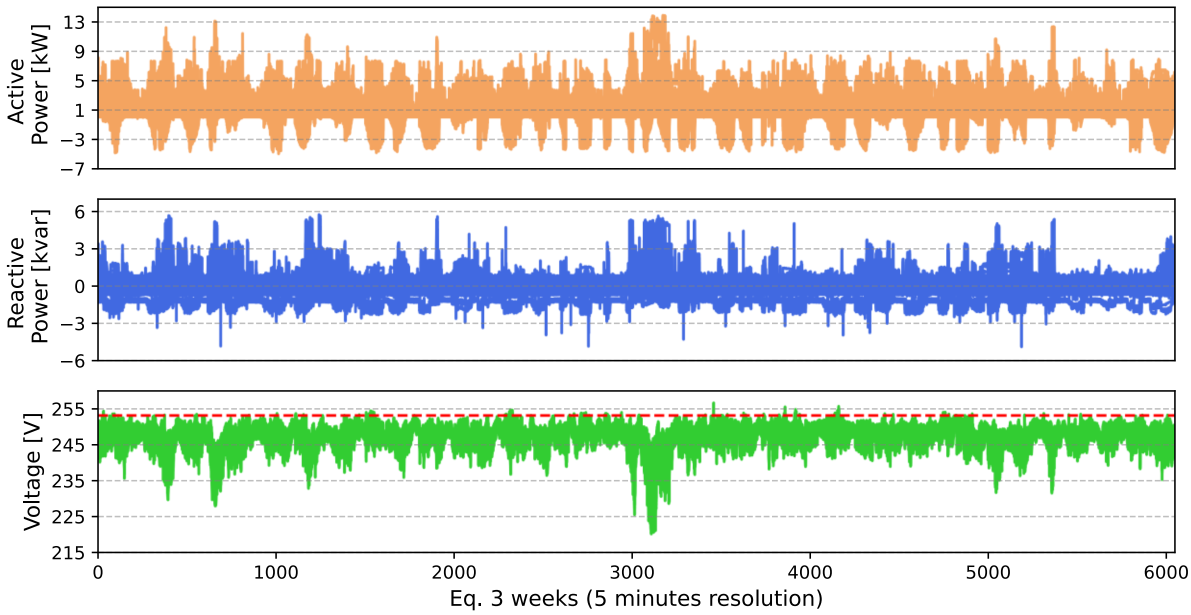

Figure 3. Training Data Set

From Figure 3, in terms of active power, it can be noted that PV generation is observed. In fact, a total of 37 PV systems are identified in the training data set. From these PV systems, 11 are found to be connected to single-phase customers and 26 are found to be connected to the different phases of 10 three-phase customers. Thus, 21 of 146 customers are found to have at least one PV system connected, accounting for a ≈ 14 % of PV penetration. Besides, in terms of reactive power, it can be noted that both inductive and capacitive behaviour can be observed throughout the training data. Finally, in terms of voltages, it can be noted that the LV network is already constrained in terms of voltages as customers' maximum voltages are close to or beyond the upper statutory limit of 253 V through large portions of the training data set.

2.3.1.2 Neural Network Selection

As detailed in Section 2.1.2, after the training data set is obtained is then used within a NN selection process to define and train the most suitable NN for the studied LV network. The first step to define the most suitable NN for the studied LV network is to fix those hyperparameters and NN characteristics that are given by the nature of the problem and are fixed for all NNs that will be trained. These are presented in Table 4 and further details can be found in [1], [15].

Table 4. Hyperparameters and NN characteristics given by the nature of the problem

NN Inputs	NN Outputs	Output Activation function	Scaler	Loss function	Optimiser	Regularisation
442	221	Linear	[0,1]	MSE	Adam	L2

Then, a K-fold cross validation process is implemented to define the remaining hyperparameters, i.e., number of neurons, activation function, learning rate, l2 regularisation factor, batch size and epochs. The search spaces considered for this process are presented in Table 5, where the value selected for this case study is highlighted in bold in each case.

Table 5. Hyperparameter search spaces and obtained values

N° of Hidden Layers	Number of Neurons	Activation functions	Learning rate	L2 regularisation	Batch size	Epochs
1	$0.5\|C\|$ $1\|C\|$ $2\|C\|$ $3\|C\|$ $4\|C\|$ $\mathbf{5\|C\|}$ $6\|C\|$ $7\|C\|$ $8\|C\|$ $9\|C\|$ $10\|C\|$	**Tanh** ReLu Swish	1×10^{-3} $\mathbf{1 \times 10^{-4}}$ 1×10^{-5}	1×10^{-3} 1×10^{-4} $\mathbf{1 \times 10^{-5}}$	**72 (eq. 6 hrs)** 144 (eq. 12 hours) 288 (eq. 24 hours)	500 **1,000** 2,000

Once all hyperparameters are defined, 10 NNs are trained from scratch using the complete training data set. The one found with the lowest RMSE in training corresponds to NN_{u^*} and is ready to carry out model-free voltage calculations for any kind of customers' net demand what-if scenarios.

2.3.1.3 Model-Free Voltage Calculations

To assess the performance of NN_{u^*} obtained in Section 2.3.1.2, historical smart meter data is used. Hence, a test data set is built with 3 weeks of historical smart meter data (i.e., 6,048 instances, consecutive to those considered in the training data set). Thus, the test data set spans the period between 07-01-2021 and 13-03-2021 and it is presented in Figure 4. Note that this correspond to more than 3 calendar weeks as instances with invalid or unfeasible values for at least one customer were removed in Step 3 of the data pipeline.

From Figure 4 it can be noted that the test data set is of similar characteristics as the training data set presented in Figure 3. Hence, in terms of active power, PV generation is observed. In terms of reactive power, inductive and capacitive behaviour are observed throughout the data, whereas, in terms of voltages, it can be noted that the LV network is constrained in terms of exports.

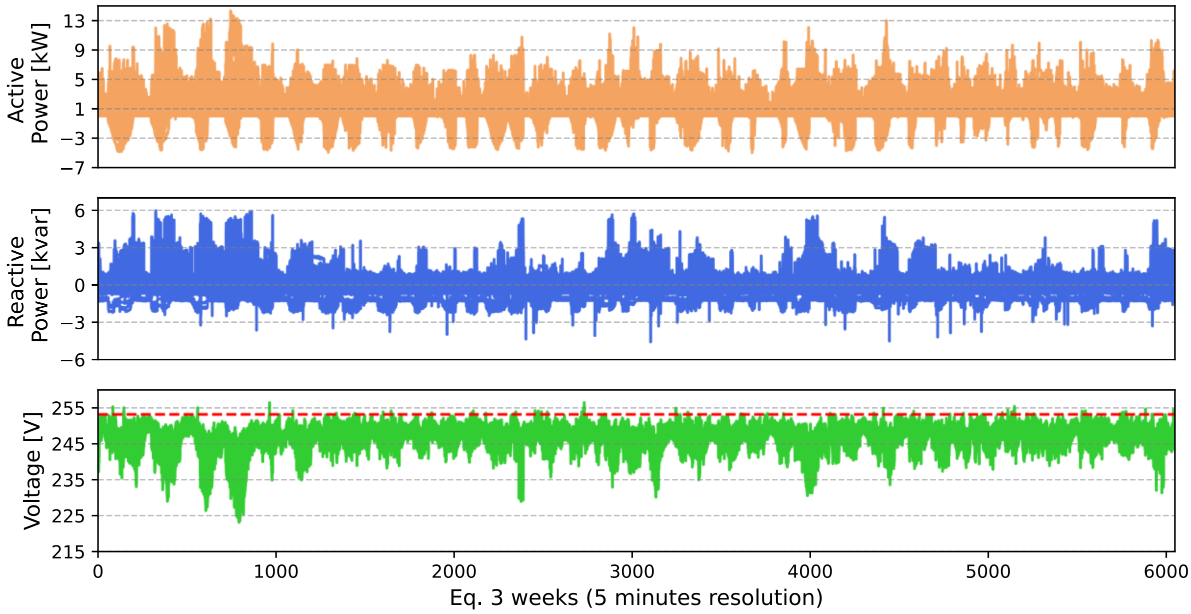

Figure 4. Test Data Set

Thus, the NN obtained in Section 2.3.1.2 is used to calculate voltages as in (6) for all instances in the test set. NN performance is assessed by comparing the obtained voltages against the actual voltage values in the test set. Results are presented numerically in Table 6. Time-series results are presented in Figure 5, where the coloured regions are to represent the difference between the maximum and the minimum voltage value among all customers for each instance. Besides, voltage calculations for all customers at all instances are plotted in the scatter plot in Figure 6, where the y-axis is to represent the voltage values calculated by NN_{u^*} and the x-axis is to represent the corresponding actual voltage values. Furthermore, deviations for all customers at all instances are presented as violin plot in Figure 7.

Table 6. Model-Free Voltage Calculations Results

Training Initial Instance	Training Final Instance Test Initial Instance	Test Final Instance	RMSE V	Av. Dev. V	Max. Dev. V
22-11-2020	07-01-2021	13-03-2021	0.6992	0.5369	9.3516

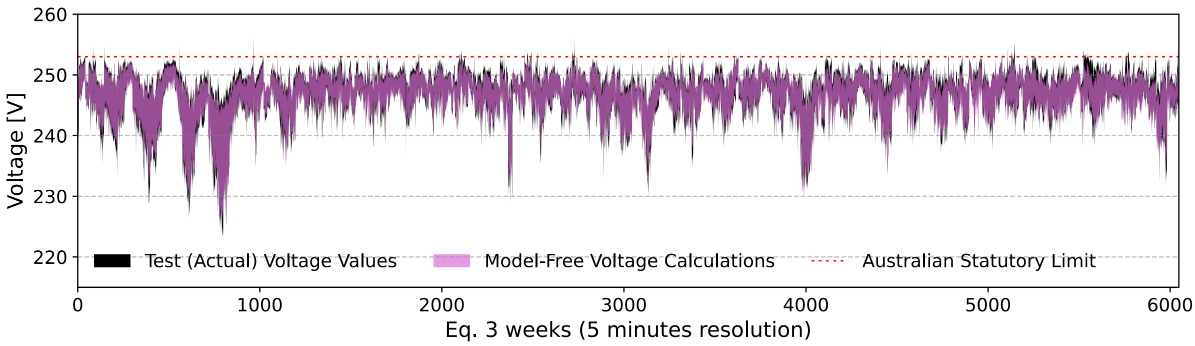

Figure 5. Model-Free Voltage Calculations: Time-Series Results

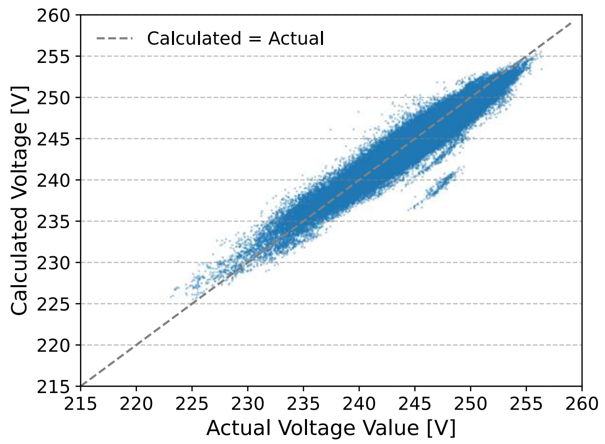

Figure 6. Model-Free Voltage Calculations: Calculated vs Actual Voltage Values

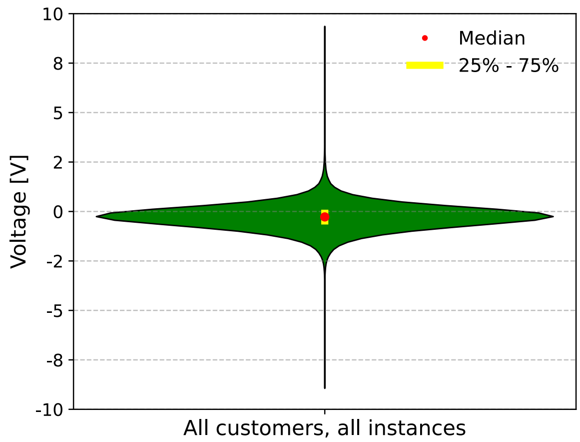

Figure 7. Model-Free Voltage Calculations: Deviation Violin Plot

From Table 6, it can be observed that accurate voltage calculations are obtained in the test data set, achieving a RMSE of 0.6992 V, an Av. Dev. of 0.5369 V, and a Max. Dev. of 9.3516 V. From Figure 5 it can be noted that accurate voltage calculations are obtained throughout the entire data set, even in those instances where large voltage drops (e.g., below 235 V) are observed. Furthermore, from Figure 6 it can be noted that NN_{u^*} accurately captures the underlying relationships of Jemena substation B as the obtained voltage calculations resemble the identity line of the scatter plot. Finally, from Figure 7 it can be highlighted that although Max. Dev. can be high (up to 9 V), such deviations occurs in a reduced number of instances, leaving 99%, 90%, 75%, and 50% of deviations below 2.0792 V, 1.1335 V, 0.7494 V, and 0.4238 V, respectively. As described in [2], instances with high deviations are easy to flag and avoid in operation by using a simple post-processing filtering process. Furthermore, such large deviations are normally caused by sudden peaks in the data of some customers (normally, with low values/variability) that goes well beyond the values observed for these customers during training. Therefore, the thresholds adopted in Step 4 of the data pipeline to define customers with low values/variability can be revised. For further discussions on this, refer to Section 3.

2.3.2 Online Stage: Calculating Operating Envelopes

As discussed in Section 2.2, once NN_{u^*} is obtained it can be then used to explore voltage compliance under several scenarios of import or export values for active customers and, thus, determine its corresponding OEs. This case study considers all single-phase customers with PV systems as active customers (i.e., 11 active customers) and calculates day ahead OEs for the 8th of January of 2021 considering both allocation techniques described in Section 2.2.2. The day (8th of January of 2021) is chosen for three main reasons: 1) It corresponds to a summer day (i.e., high PV injections) close to the training data set; 2) it is one of the days with the largest availability of data for Jemena substation B during summer 2021, 234 over 288 instances are considered valid after the data pipeline is implemented (rest is deemed unused due to gaps); and 3) it presents both, large and small voltage headroom throughout the day, enabling to assess the approach in both situations. Time-series model-free voltage calculations obtained for the 8th of January using NN_{u^*} are presented in Figure 8.

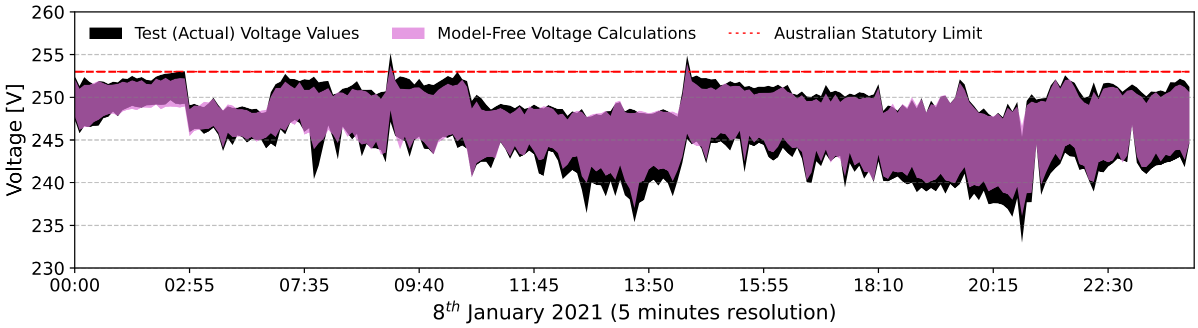

Figure 8. Model-Free Voltage Calculations: Time-Series Results 8th of January 2021

2.3.2.1 Inputs

Day ahead OEs are calculated. This, in practice, would require forecast. However, forecast is outside of the scope of this project and therefore, hindsight OEs are calculated, i.e., perfect forecast is considered.

Net Demand (P and Q) of Passive Customers and Net Demand (Q) of Active Customers: values are obtained from the historical data.

Network constraints:

- **Voltages:** $\mathrm{V}_{max} = 253\ V$ as per Australian regulation [16]
- As thermal capacity of the corresponding assets was not available, assumptions are made:
  - **Connection point:** A maximum capacity of 14 kVA is considered, i.e., $s_{cp,c} = 14\ kVA, \forall c\ in\ C$
  - **LV Conductors:** Conductors of 230 AMPs per phase are considered at the head of the LV network. Therefore, $s_{conductor} = 52.9 \frac{kVA}{phase} x\ N°\ LV\ circuits = 52.9 \frac{kVA}{phase} x\ 4 = 211.60 \frac{kVA}{phase}$
  - **Transformer:** Capacity is obtained by multiplying the per-phase number of customers by an ADMD of 3.5 kVA and then approximating to the closest *traditional* transformer size. Therefore, $s_{transformer} = |C| \times ADMD = 221 \times 3.5\ kVA = 773.5\ kVA \approx 750\ kVA$

2.3.2.2 Algorithms

OEs are calculated considering both allocation techniques described in Section 2.2.2, i.e., equally distributed OEs and maximise exports OEs. Thus, at each instance, OEs are calculated following the steps presented in Table 3 according to the adopted allocation technique. At each export level for active customers explored by the algorithm, network constraints are assessed using (9) – (14) and the limits presented in Section 2.3.2.1.

2.3.2.3 Outputs: Model-Free Operating Envelopes

OEs calculated with both allocation techniques, i.e., equally distributed and maximise exports, are presented in stacked plots in Figure 9 and Figure 10, respectively, where each coloured region is to represent an active customer. By comparing the obtained OEs with the actual voltage profile of Jemena substation B for the day of interest presented in Figure 8, it can be noted that the proposed approach produces consistent OEs in both cases. In this context, it can be noted that larger OEs are obtained at times when significant voltage headroom is available, whereas, on the other hand, more constrained OEs are obtained at times when little to no voltage headroom is available. In this context, it can be highlighted that OEs goes to 0 kW at 09:10 and 14:30 hrs, which is consistent with the two voltage peaks that goes beyond the voltage limit during the 8th of January. Similarly, OEs can go as high as its maximum (i.e., only limited by service cables) around 19:00 hrs or 20:30 hrs when large voltage headroom is available. Besides, it can be noted in Figure 10 that overall, larger exports from active customers can be obtained throughout the entire day by constraining the OEs of those customers that presents higher sensitivity to voltages (maximise exports OEs).

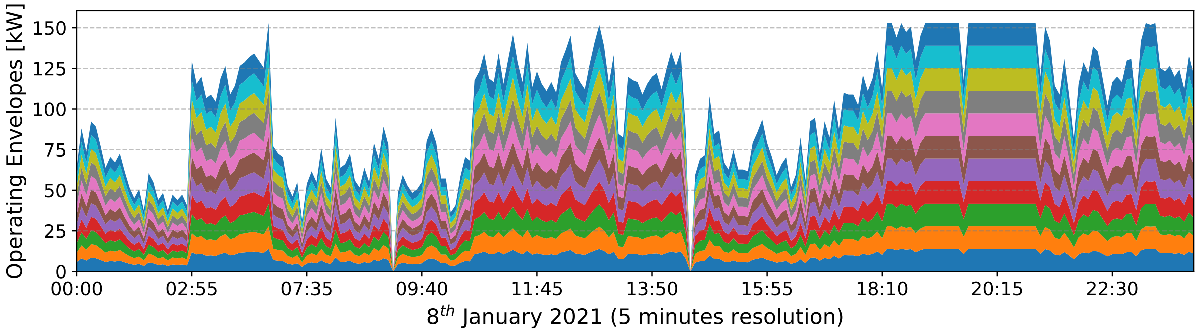

Figure 9. Model-Free OEs: Equally Distributed

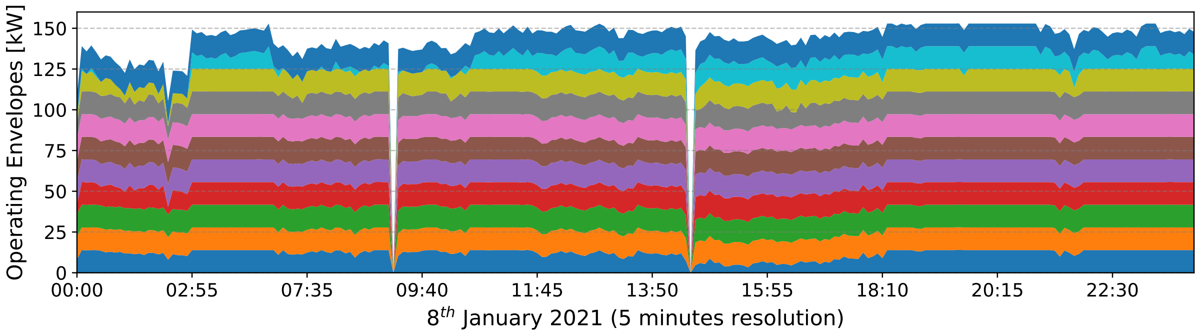

Figure 10. Model-Free OEs: Maximise Exports

Finally, it is worth noticing that equally distributed OEs were calculated in an average of 1.3 s throughout the whole day, whereas maximise export OEs were calculated in an average of 45 s. All calculations were carried out in a standard commercial computer with an Intel i7 processor of 2.80 GHz and 16 GB of ram (DDR4). This difference is caused because while the equally distributed allocation technique increases the OE of all customers simultaneously in each iteration, the maximise exports allocation technique increases the OE of the customer with the lowest voltage only, and thus, naturally involves a larger number of iterations.

2.4 Partial Smart Meter Data Availability

While smart meters are installed in almost every house, business, and factory in Victoria, the smart meter data of some customers might be managed by third parties and therefore, unavailable for DNSPs. On the other hand, for DNSPs outside Victoria, smart meter deployment is still ongoing. Thus, it is of crucial interest to assess the performance of the proposed approach in calculating OEs when smart meter data from all customers connected to the corresponding LV network is not available. For further details, refer to [3] and Section 7.

To assess the effectiveness of the proposed model-free OEs approach under partial smart meter data availability, model-free OEs are calculated as in Section 2.3 but considering that the smart meter data of all two and three-phase customers connected to Jemena substation B is managed by third parties and therefore, unavailable (hereinafter referred to as partial observability case). For further analysis on partial smart meter data availability, refer to Section 7. Note that as the data from such customers is considered unavailable, NN production, assessment, and OE calculations are carried out considering the historical data of single-phase customers only.

Thus, a NN is trained from scratch using the same hyperparameters as those defined with full observability in Section 2.3 and the training data of single-phase customers only. Hence, the per-phase number of customers is $|C| = 108$. It is important to note that from the perspective of the NN, three-phase customers contain more information than single-phase customers (as considered as three single-phase customers). Therefore, while the data of ≈74 % of customers is available, from the NN perspective this represents only ≈50 % of the data.

Once the NN is trained its performance is assessed using the test data set. Results are presented numerically in Table 7 and as time-series in Figure 11, where it can be observed that even though a performance decrease with respect to the case of full observability (Section 2.3) is obtained (as expected), accurate voltage calculations are produced, achieving a RMSE of 0.9385 V, an Av Dev. of 0.7122 V and a Max. Dev. of 9.3323 V.

Table 7. Model-Free Voltage Calculations Results (Partial Observability)

Training Initial Instance	Training Final Instance Test Initial Instance	Test Final Instance	RMSE V	Av. Dev. V	Max. Dev. V
22-11-2020	07-01-2021	13-03-2021	0.9385	0.7122	9.3323

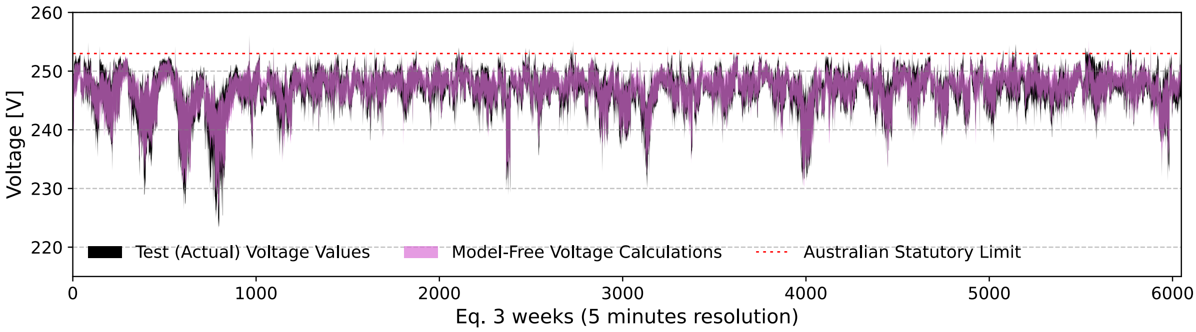

Figure 11. Model-Free Voltage Calculations (Partial Observability): Time-Series Results

To calculate OEs, the same day used in Section 2.3 is considered. The time-series model-free voltage calculations results for this day are presented in Figure 12, where the accuracy decrease with respect to the case of full observability (Figure 8) can be observed.

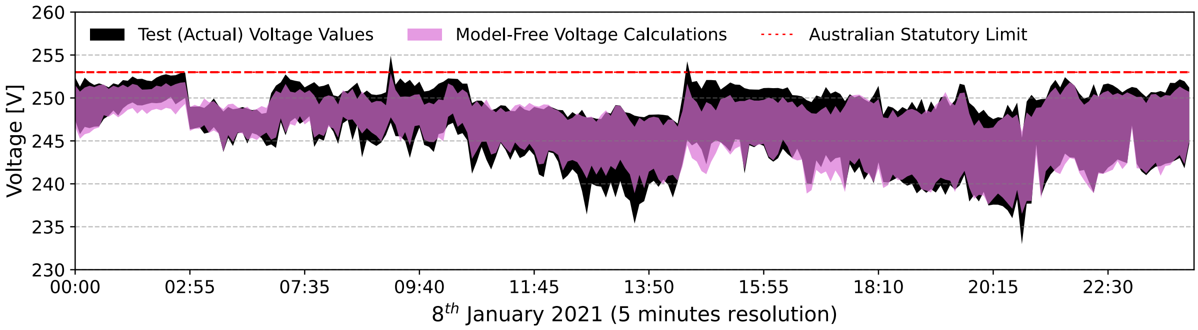

Figure 12. Model-Free Voltage Calculations (Partial Observability): Time-Series Results 8th of January 2021

OEs calculated with both allocation techniques, i.e., equally distributed and maximise exports, are presented in Figure 13 and Figure 14, respectively. As for the case of full observability, it can be observed that the obtained OEs are consistent with the actual voltage profile of Jemena substation B for the day of interest presented in Figure 12, as larger OEs are obtained at times when significant voltage headroom is available, whereas, on the other hand, more constrained OEs are obtained at times when little to no voltage headroom is available. Furthermore, the obtained OEs are also consistent with those calculated in Section 2.3 (Figure 9 and Figure 10). However, as expected, there are differences in the obtained OEs. For instance, it can be noted that while OEs are calculated as 0 kW at 09:10 and 14:30 hrs in the case of full observability, for the case of partial observability small OEs are calculated using both allocation techniques.

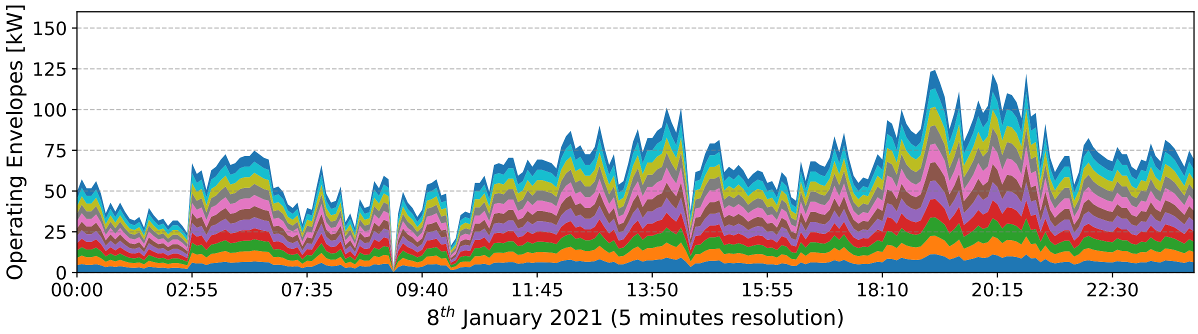

Figure 13. Model-Free OEs (Partial Observability): Equally Distributed

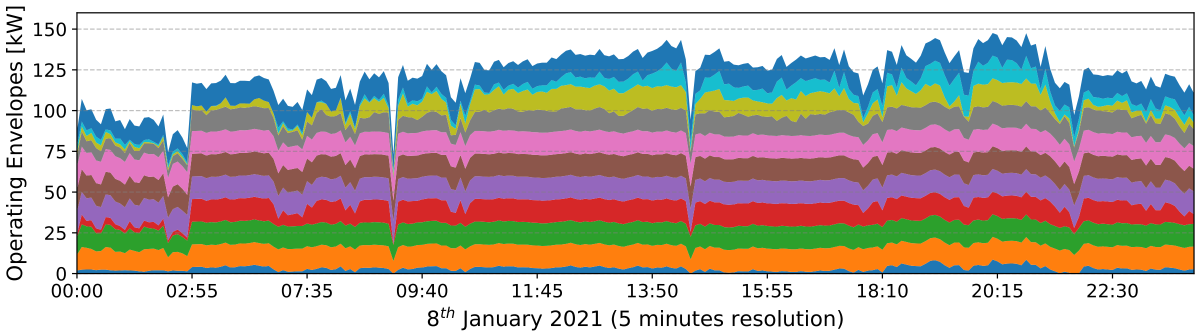

Figure 14. Model-Free OEs (Partial Observability): Maximise Exports

It can be noted that, for this case study, OEs calculated with partial observability are generally lower than those obtained in the case of full observability throughout the day of interest. While this can be seen as *conservative* OEs, as it will further restrict exports and thus ensure network integrity. It can be noted that this might not be always the case when dealing with partial smart meter availability. This will be strictly dependent on the operation and location of those customers whose smart meter data is not available. For further analyses, refer to Section 7. Finally, note that the present study is limited in the context of utilisation of transformer and LV conductors as active customers are allowed to export up to the full capacity of the corresponding assets. As detailed in Section 7, in the context of partial smart meter data availability, the capacity of transformer and LV conductors must be reduced to cater for the utilisation of those customers whose data is not available, which could further reduce the obtained OEs.

2.5 Model-Free Calculations for Planning Purposes

The capability of the proposed model-free OEs approach in assessing both, thermal and voltage compliance, for any kind of customers net demand what-if scenarios can also be exploited by DNSPs for planning applications. For instance, the proposed model-free approach could be adapted to:

- Estimate the hosting capacity of different DER technologies, from solar PV to EVs, or to assess different measurements that could be used to increase it (e.g., control of voltage regulation devices such as the on-load tap changer at the zone substation or capacitor banks).

- Quickly and efficiently assess DER connection request, evaluating its respective impacts into the existing LV network to determine if further actions are required.

2.6 Key Remarks

This section presents the complete model-free OEs approach developed throughout this project. Thus, builds on top of the previous reports of this project [1]-[3], as well as on top of the latest findings presented in subsequent sections of the report at hand. The effectiveness of the proposed model-free OEs approach is demonstrated with a case study considering Jemena substation B, which corresponds to the substation with the largest number of LV circuits, i.e., 4, used in this study.

The full approach is implemented to calculate OEs for the 8th of January 2021. Hence, the historical smart meter data is processed to build an adequate training data set which is then used to define and train the most suitable NN for the corresponding LV network. All this process occurs online. Once the NN is obtained, its corresponding accuracy is assessed with historical data, achieving accurate results. Then, the NN is used along with a heuristic algorithm and tailored approximations to explore different export values for active customers and thus, calculate its OEs. Two different allocation techniques, equally distributed and maximise exports, are considered. Results shows consistency between the obtained OEs and the actual voltage profile for the day of interest. Additionally, it is shown that the proposed approach can adequately capture the underlying relationships and enable larger overall exports from active customers by constraining the OE of those customers that present higher sensitivity to voltages (maximise exports allocation technique). Furthermore, it is shown that if the data of all three-phase customers is considered unavailable, the proposed approach can still produce accurate voltage calculations and consistent OEs.

Overall, it is shown that the proposed approach corresponds to an accurate, cheap, and fast alternative to traditional model-based approaches to calculate OEs, enabling DNSPs to bypass the costly, time-consuming, and error-prone process of producing electrical models. Furthermore, it can be highlighted that the capabilities of the proposed model-free OEs approach in assessing both, thermal and voltage compliance, for any kind of customers net demand what-if scenarios can also be exploited by DNSPs for planning applications, e.g., evaluate DER hosting capacity or connection request.

3 Offline Data Pipeline: Final Improvements

This section introduces final improvements to the offline data pipeline defined in [3] based on latest findings. Note that this section is focused on refining the offline data pipeline of the proposed model-free OEs approach and by no means intends to provide a full overview of the complete offline data pipeline. This can be found in Section 2.

As shown in [3], customers with extremely low values of demand (zero or a few Watts) throughout all or most of the training data set can cause issues when deploying the NN with adequate values for such customers. This is because the relationships of these customers are not properly represented in the historical data used to train the NN and thus, any adequate value for such customers will be far of what has been observed in training and cause erroneous calculations when assessing what-if scenarios. Thus, improvements to the offline data pipeline were presented in [3] such that these customers are automatically removed from the training data set and therefore, from the model-free voltage calculations. Once sufficient data for these customers is available, the NN can be updated to cater for these customers.

Further studies have shown that customers with extremely low variability (i.e., all or most values in a range of a few Watts) will cause similar issues and therefore, must also be removed. This is because the historical data of such customers do not count with enough variety for the NN to accurately capture the underlying relationships. Essentially, any feature that does not provide meaningful information to the NN can create wrong voltage calculations when deploying the NN to calculate voltages. Specifically, if adequate values for such features are considered, which is key when calculating OEs.

An example of two customers with low variability (hereinafter C1 and C2) can be found in United Energy (UE) substation A during the period between 03-05-2021 and 25-05-2021 (equivalent to 3 weeks at 5 minutes resolution, group 36 in [2]). The active and reactive power of C1 and C2 during this period is presented in Figure 15 and Figure 16, respectively. Further specifications of the data of each customer are presented in Table 8.

Table 8. UE substation A - Customers with low variability

Customer with low variability	Max. *P* [W]	Min. *P* [W]	99% values	Max. Q [var]	Min. Q [var]	99% of values
C1	270	30	70	2	-1	3
C2	195	8	90	-30	-70	40

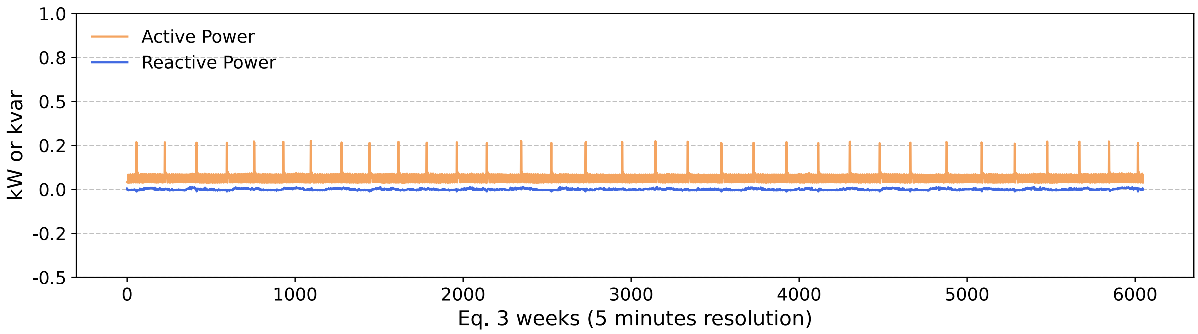

Figure 15. Customer with low variability – C1

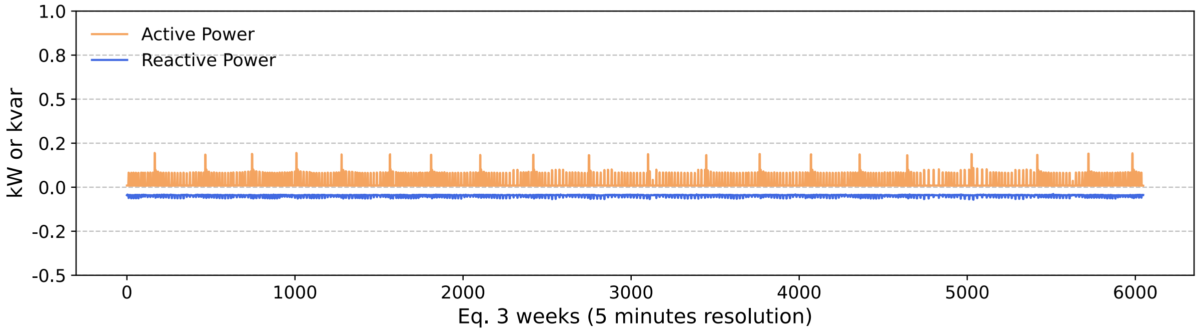

Figure 16. Customer with low variability – C2

After further discussions with DNSPs it is noted that these demands could correspond to smart meter devices that are installed in a construction site, but effective demand is not observed yet. To illustrate the effects of these customers over the proposed approach, a NN is trained considering the period of 03-05-2021 and 25-05-2021 (equivalent to 3 weeks at 5 minutes resolution, group 36 in [2]) as training data set. Thus, the offline data pipeline presented in [3] is implemented as shown in Table 9. It is important to highlight in here that, in Step 4 all customers whose relationships are not represented are identified and removed from the training data set (and, consequently, from the voltage calculations). As in [3], only customers with extremely low values are considered, which correspond to customers whose 99% of active power values are found to be lower than 10 W.

Table 9. Offline Data Pipeline

Step 1	Collect raw historical per-phase smart meter data
Step 2	Obtain customers’ active and reactive power from raw historical smart meter data
Step 3	Remove invalid and unfeasible instances
Step 4	Remove customers whose relationships are not represented

Within Step 4, a single-phase customer is identified to have extremely low values (i.e., 99% of their active power values below 10 W) and thus, removed from the training data set. However, C1 and C2 are still part of the training data set. Specifically, 99% of their values are within a range of 70 W and 90 W, respectively, and therefore, are not removed in Step 4.

After the offline data pipeline is implemented, the training data set is comprised of historical data for 61 single-phase customers and 1 three-phase customer. A NN is trained using the hyperparameters defined in [2] for UE substation A for the case of 3 weeks of data. These hyperparameters are presented in Table 10. Note that L2 regularisation was introduced at a later stage in [3] and thus, was not part of the selection process carried out in [2]. For this study, L2 regularisation is considered as 0.1 times the learning rate.

Table 10. UE substation A - NN Hyperparameters

$\lvert C \rvert$	Number of Neurons	Activation functions	Learning rate	L2 regularisation	Batch size	Epochs
64	$4\lvert C \rvert$	Tanh	1×10^{-4}	1×10^{-5}	144 (eq. 12 hours)	2,000

Model-free voltage calculations results for all customers are presented numerically in Table 11 and as time-series in Figure 17, where it can be observed that accurate calculations are obtained, achieving a RMSE of 1.0711 V, an Av. Dev. of 0.8253 V, and a Max. Dev. of 5.5125 V.

Table 11. Model-Free Voltage Calculations Results

Training Initial Instance	Training Final Instance Test Initial Instance	Test Final Instance	RMSE V	Av. Dev. V	Max. Dev. V
03-05-2021	25-05-2021	16-06-2021	1.0711	0.8253	5.5125

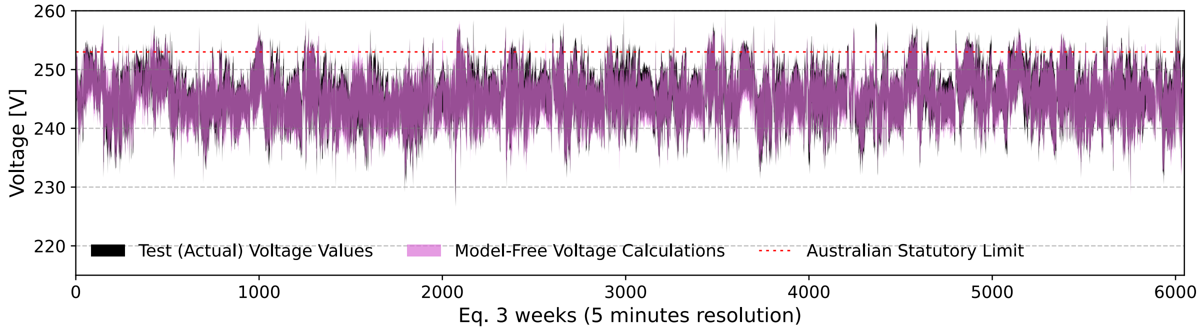

Figure 17. Model-Free Voltage Calculations: Time-Series Results

As observed in Table 11 and Figure 17, accurate voltage calculations are obtained and there is no clear indication of issues caused by the data of C1 and C2 as it was observed for the cases of customers with extremely low values in [3], where the NN calculates erroneous values simultaneously for all customers. However, it must be noted that this is only because the test data of C1 and C2 maintain the behaviour observed for these customers during training as shown in Figure 18 and Figure 19. But, if adequate values for such customers are considered, the NN will calculate erroneous values for all customers as in [3], which is critical as the calculation of OEs is all about pushing customers’ limits.

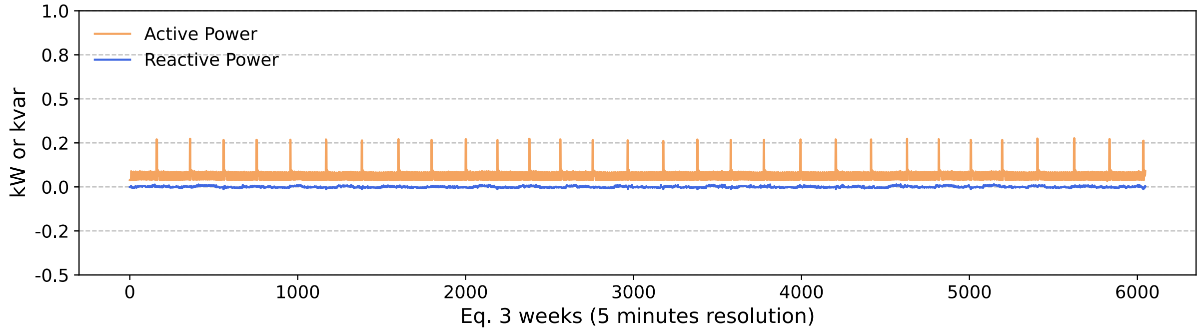

Figure 18. UE Sub A: Customer with low variability – C1 (Test Data)

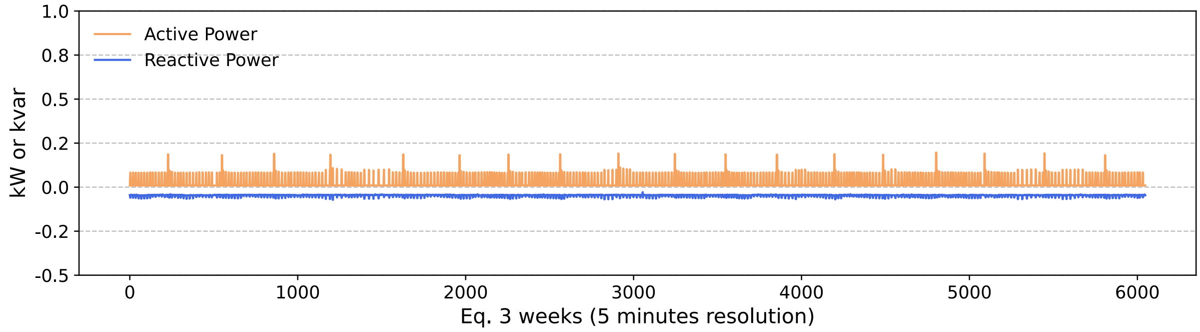

Figure 19. UE Sub A: Customer with low variability – C2 (Test Data)

To illustrate the issues that C1 and C2 can cause when deploying the NN to calculate OEs, the first instance at noon of the test set is considered (i.e., 25-05-2021 12:00 hrs). Thus, the NN is used to assess an OE of 5 kW (exports) for each customer at the selected instance. Hindsight OEs are considered, i.e., for each customer, active and reactive power of the remaining customers plus its corresponding active power are obtained from the historical data. The obtained results are presented in Figure 20, where each boxplot correspond to the voltage calculations obtained for all customers when assessing an OE = 5 kW for the corresponding customer.

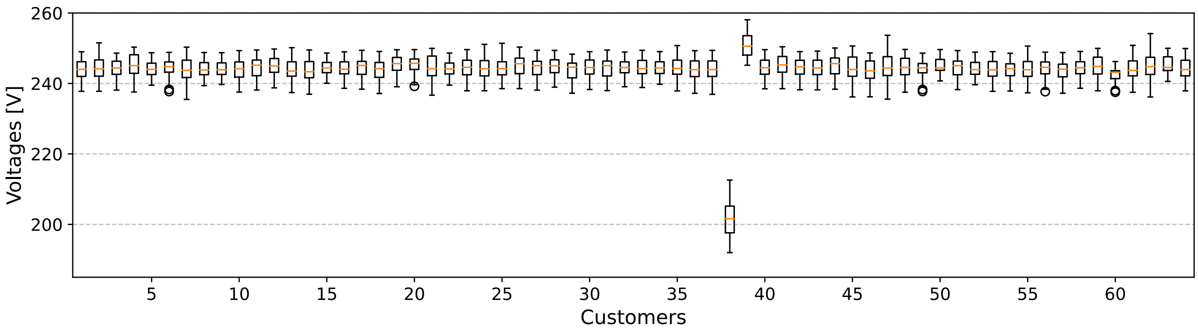

Figure 20. Assessment of OE = 5 kW per customer at 25-05-2021 12:00 hrs

From Figure 20 it can be observed that when assessing an OE = 5 kW for C1 and C2, voltage calculations for all customers goes far from what it is expected and from what it is observed for the rest of the customers. This resembles what was observed in [3] when adequate values were considered for those customers found with extremely low values. This is because the relationships for C1 and C2 are also not properly represented (as observed in Figure 15 and Figure 16) and 5 kW of exports goes far away from what it is observed for these customers during training. This, in turns, end up causing erroneous calculations for all customers. In fact, any what-if scenario that considers adequate values for C1 or C2 will be subject so such erroneous calculations.

Therefore, the offline data pipeline must be improved to automatically detect and remove such customers. In this context, Step 4 is adapted to cater for both, customers with extremely low values and variability, and consequently, automatically detect all customers whose relationships are not properly represented. Thus, instead of removing all customers with 99% of their values below 10 W, which only caters for customers with low extremely low values, the criteria is modified to remove all customers with 99% of their values within a range of 100 W.

To show the effectiveness of the proposed improvements, the improved data pipeline is implemented in the same case of UE substation A presented above. After the improved data pipeline is implemented, the training data set is comprised of historical data for 59 single-phase customers and 1 three-phase customer, as C1 and C2 has been detected and removed in Step 4.

With this data, A NN is trained using the same hyperparameters as before (Table 10). Results are presented numerically in Table 12 and as time-series in Figure 21, where it can be observed that a slight accuracy decrease with respect to the previous case is observed, achieving a RMSE of 1.1184 V, an Av. Dev. of 0.8606 V, and a Max. Dev. of 5.8766 V.

Table 12 Improved Offline Data Pipeline - Model-Free Voltage Calculations Results

Training Initial Instance	Training Final Instance Test Initial Instance	Test Final Instance	RMSE V	Av. Dev. V	Max. Dev. V
03-05-2021	25-05-2021	16-06-2021	1.1184	0.8606	5.8766

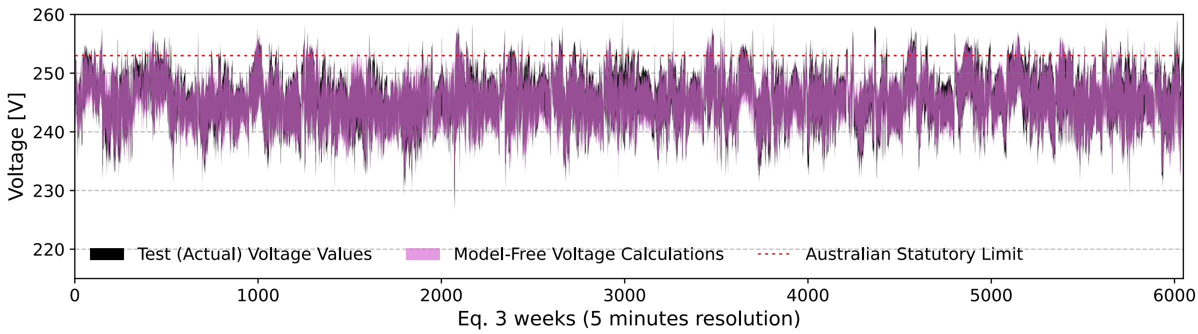

Figure 21. Improved Offline Data Pipeline - Model-Free Voltage Calculations: Time-Series Results

Same as before, the first instance at noon of the test set is considered (i.e., 25-05-2021 12:00 hrs) and OEs of 5 kW (exports) for each customer at the selected instance are assessed. The obtained results are presented in Figure 22, where it can be observed that calculations for all customers lies within similar and expected values. Thus, it can be noted that relationships for all the remaining customers are properly captured, and no issues are to be expected when deploying the NN to calculate OEs.

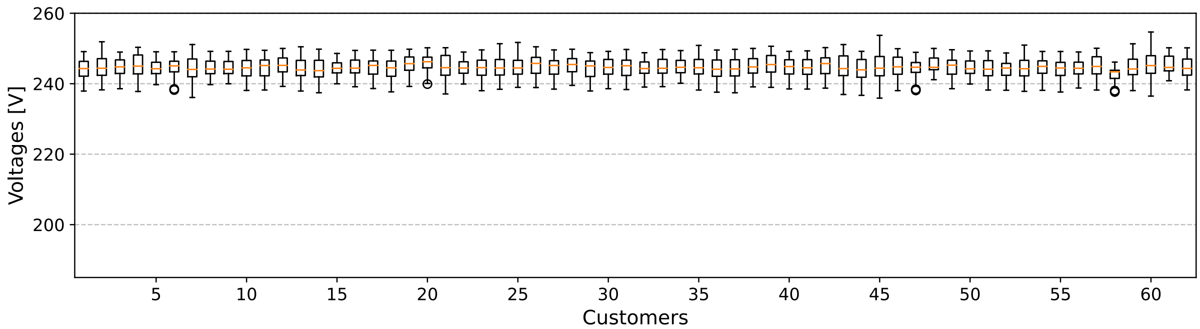

Figure 22. Improved Offline Data Pipeline - Assessment of OE = 5 kW per customer at 25-05-2021 12:00 hrs

Therefore, it can be noted that the adaptations to Step 4 have shown effective in automatically detecting and removing customers with extremely low values and variability for the case of UE substation A, enabling the proposed approach to avoid considering such customers, which can lead to erroneous calculations for all customers when deploying the NN. Specifically, when adequate values for such customers are considered.

Note that, for simplicity, this study identifies customers whose relationships are not properly represented by using thresholds (specifically, if 99% of their active power data is within a range of 100 W). However, the definition of such thresholds such that they are adequate to all cases is not exempt from challenges. On one hand, while extending such thresholds leads to a more *conservative* approach as it will ensure customers whose relationships are not properly represented are removed, it could also lead to remove other customers unnecessarily, which impacts accuracy. On the other hand, limiting such thresholds will likely lead the approach to consider customers whose relationships are not properly represented, which can create issues when deploying the NN. A more thorough approach will be to incorporate a post-processing NN check. Thus, after the NN is trained, the analyses carried out in this section, i.e., OE = 5 kW for each customer at a given instance, could be implemented to directly detect customers whose relationships are not properly represented and remove them from the historical data.

Finally, it is important to note that the proposed offline data pipeline consider removing customers whose relationships are not properly represented. However, an alternative is to flag such customers and consider values that are part of the training data set of such customers (e.g., the median) when deploying the NN. This will prevent the NN to calculate erroneous values when assessing what-if scenarios but does not allow analyses for these customers. Besides, it will require a further post-process of NN check and increments complexity of the approach as any calculation will need to carefully consider observed values for these customers.

3.1 Key Remarks

It was shown in [3] that customers with extremely low values can cause issues when deploying the NN with adequate values for such customers. This is because the relationships of these customers are not properly represented in the historical data used to train the NN and thus, any adequate value for such customers will be far of what has been observed in training and will cause erroneous calculations when assessing what-if scenarios. Thus, improvements to the offline data pipeline were proposed in [3] such that these customers are automatically detected and removed from the historical data. However, further studies have shown that the same issues can occur in the case of customers with extremely low variability (i.e., all values or most values within a range of a few Watts). Thus, Step 4 of the offline data pipeline is adapted to automatically detect and remove these customers.

To illustrate the case of customers with low variability and its corresponding impacts over the proposed approach when deployed with adequate values for such customers, the case of UE substation A is used. In here, it can be noted that although for this case study the impacts of customers with low variability are not directly observed in the obtained model-free voltage calculations as the test data set for these customers resembles the behaviour observed during training, a simple analysis can show that the NN will calculate erroneous values for all customers when considering adequate values for these customers. To overcome this issue, the offline data pipeline presented in [3] is improved to automatically detect and remove these customers.

4 Thermal Constraints in Model-Free Operating Envelopes

The proposed model-free voltage calculations approach can be used along with a heuristic algorithm and tailored approximations to assess several import or export values for active customers and thus, determine its corresponding OEs. For simplicity, only voltage and customers' connection point constraints were initially assessed when calculating model-free OEs in [2], [3]. However, the utilisation of key assets, i.e., transformer and conductors, can also constraint active customers' imports or exports and, therefore, must be considered. This section provides the foundations to assess conductor and transformer thermal constraints within the model-free OEs algorithm in [2] and presents a case study that illustrates such assessment. Note that this section is focused on incorporating thermal constraints within the model-free OEs algorithm presented in [2] and by no means intends to provide a full overview of the complete model-free OEs approach. This can be found in Section 2.

As described in [2], [3], the core of the model-free OEs algorithm is comprised of two main blocks, OE evaluation and Network Constraints. A schematic of this approach is presented in Figure 2 (Section 2). The OE algorithm defined in this project starts by producing an initial set of OEs with $P = 0\,kW$ for the active customers in the OE Evaluation block. This set of OEs is then assessed within the Network Constraints Assessment block, if constraints are not breached, a new set of OEs with increased values is produced according to the adopted allocation technique within the OE Evaluation block. This process is repeated iteratively until network constraints are breached.

Within the Network Constraints Assessment block, meter-level voltage compliance of each set of OEs produced in the OE Evaluation block is assessed using the proposed model-free voltage calculations approach as detailed in [2], [3]. While thermal constraints can also impact customers maximum imports or exports and thus, must be considered, the corresponding NN was trained to calculate voltages only and cannot cater for asset utilisation. Therefore, approximations must be considered. In this project, thermal constraints are approximated by using customers net demands and the thermal capacity (in kVA) of the corresponding assets. Thus, for each set of OEs to be analysed, thermal constraints are considered as follows.

1. **Customers' connection point:** The capacity of the connection point of each customer, which can be given by the capacity of the corresponding service cable or connection agreements, will limit its imports or exports. Hence, the apparent power of each customer is calculated as in (15) and each customer is allowed to import or export up to the maximum capacity of its corresponding connection point, i.e., $s_{cp,c}\ \forall c \ in\ C$ (set of customers indexed by c).

$$s_c = \sqrt{(P_c)^2 + (Q_c)^2} \tag{15}$$

$$s_c \leq s_{cp,c}\ \forall c\ in\ C \tag{16}$$

2. **LV conductor:** The capacity of the different LV conductors throughout a LV network will limit the imports or exports of active customers that are connected downstream the correspondent segment. However, a thorough analysis of LV conductor constraints for each segment throughout the LV network will inherently require knowledge from the corresponding electrical model, i.e., LV network topology, phase groupings, and the capacity of each segment, which are rarely available or accurate in practice. Therefore, assumptions must be made.

   This project assesses LV conductor capacity at the head of the LV network. This minimises the data requirements (only thermal capacity of this conductor is required) and provides a good proxy for LV conductor utilisation as this point will aggregate the demands from all customers connected to the LV network. Note that this hold for most LV networks as they are normally built considering tapering, i.e., the cross-section area of LV conductors is gradually decreased as distance from the distribution transformer increases. However, if a thinner conductor is found in the middle of the LV network, this will not represent a good proxy and further information is required to adequately assess LV conductor constraints in such case.

   Thus, the aggregated apparent power from all customers, i.e., s^{Σ}, is calculated as in (17) and active customers are allowed to import or export until the per-phase apparent power, approximated as $\frac{s^{\Sigma}}{3}$, reaches the per-phase capacity of the conductor at the head of the LV network ($s_{conductor}$) as in (18).

$$s^{\Sigma} = \sum_{c\ in\ C} s_c \tag{17}$$

$$\frac{s^{\Sigma}}{3} \leq s_{conductor} \tag{18}$$

4. **Transformer:** The capacity of the distribution transformer will limit the imports or exports of all active customers connected to it. Thus, s^{Σ} is calculated as in (17) and customers are allowed to import or export up to the maximum capacity of the transformer as in (19).

$$s^{\Sigma} \leq s_{transformer} \tag{19}$$

Ideally, LV conductor and transformer analyses should be implemented per phase. Nonetheless, phase grouping information is rarely available for DNSPs in practice and thus, simplifications are made in this project. However, note that data-driven techniques have been presented in the literature [17]-[19] to identify customers' phase grouping in LV networks from historical voltage measurements only. Hence, these approaches could be potentially implemented to identify customers' phase grouping and carry out a per-phase analysis of LV conductor and transformer utilisation. However, this is outside of the scope of this study.

4.1 Case Study

To illustrate the implementation of thermal constraints within the proposed model-free OEs approach, the case of Jemena substation B is used. Specifically, the NN produced in Section 2 is considered to calculate OEs for 40 active customers for a given instance of the 8th of January 2021 (noon, 12:00 hrs). Note that this case study has been specifically designed to illustrate the implementation of thermal constraints within the model-free OEs approach. Therefore, the instance of interest is chosen to cater for a large voltage headroom (as observed in Figure 8) which will allow the exploration of larger OEs, and a high number of active customers is considered to stress asset utilisation.

OEs are calculated as in Section 2 considering both allocation techniques, i.e., equally distributed and maximise exports. For each allocation technique, two scenarios are considered, one where only voltage and customer connection point constraints are considered (i.e., S1) and other one where LV conductor and transformer constraints are incorporated (i.e., S2). The constraints considered for this case study are as in Section 2 and are detailed below.

- **Voltages:** $V_{max} = 253\ V$ as per Australian regulation [16]
- As thermal capacity of the corresponding assets was not available, assumptions are made:
  - **Connection point:** A maximum capacity of 14 kVA is considered, i.e., $s_{cp,c} = 14\ kVA, \forall c\ in\ C$
  - **LV Conductors:** Conductors of 230 AMPs per phase are considered at the head of the LV network. Therefore, $s_{conductor} = 52.9 \frac{kVA}{phase} x\ N°\ LV\ circuits = 52.9 \frac{kVA}{phase} x\ 4 = 211.60 \frac{kVA}{phase}$
  - **Transformer:** Capacity is obtained by multiplying the per-phase number of customers by an ADMD of 3.5 kVA and then approximating to the closest *traditional* transformer size. Therefore, $s_{transformer} = |C| \times ADMD = 221 \times 3.5\ kVA = 773.5\ kVA \approx 750\ kVA$

OEs calculated using the equally distributed allocation technique are presented in Figure 23 and Figure 24. In here it can be noted that in S1 (i.e., if only voltage and connection point constraints are considered) active customers exports are limited by voltage constraints and the obtained OE is 13.6 kW per active customer. However, in S2 (i.e., if LV conductor and transformer constraints are also considered), it can be observed that in fact, active customers cannot export more than 12.3 kW due to LV conductor limitations.

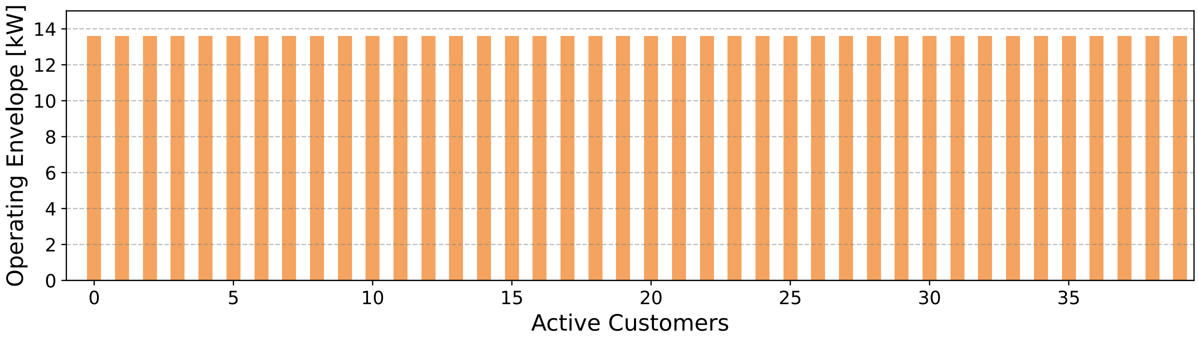

Figure 23. Model-Free OEs: Equally Distributed – S1

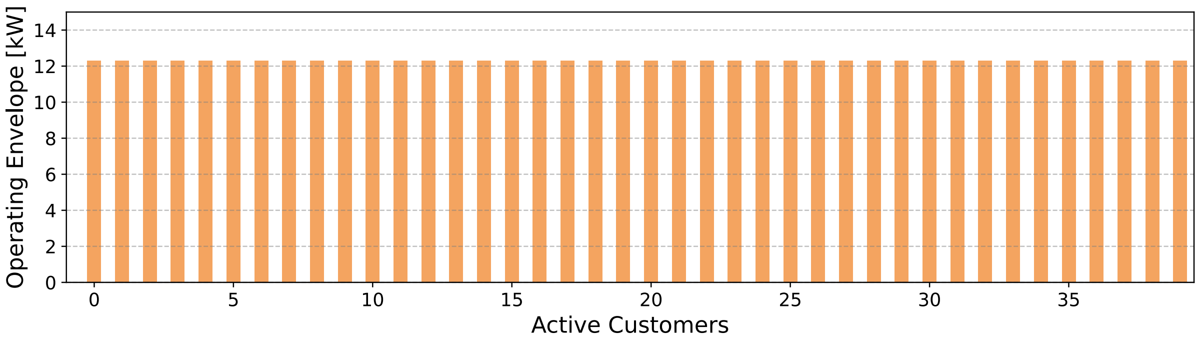

Figure 24. Model-Free OEs: Equally Distributed – S2

Similarly, OEs calculated using the maximum exports allocation technique are presented in Figure 25 and Figure 26. In here, it can be highlighted that in S1 active customers can export a total of 555.1 kW due to voltage constraints. However, in S2 can be observed that active customers can only export up to 495.5 kW due to LV conductor limitations. In fact, some active customers that are allowed to export up to its maximum capacity in S1 are highly constrained in S2 due to LV conductor limitations (e.g., active customers 6, 20 and 25 in the graphs).

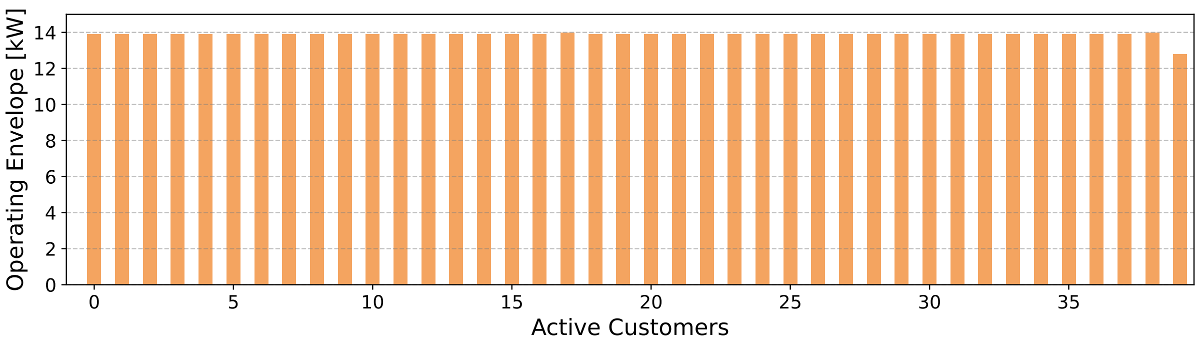

Figure 25. Model-Free OEs: Maximise Exports – S1

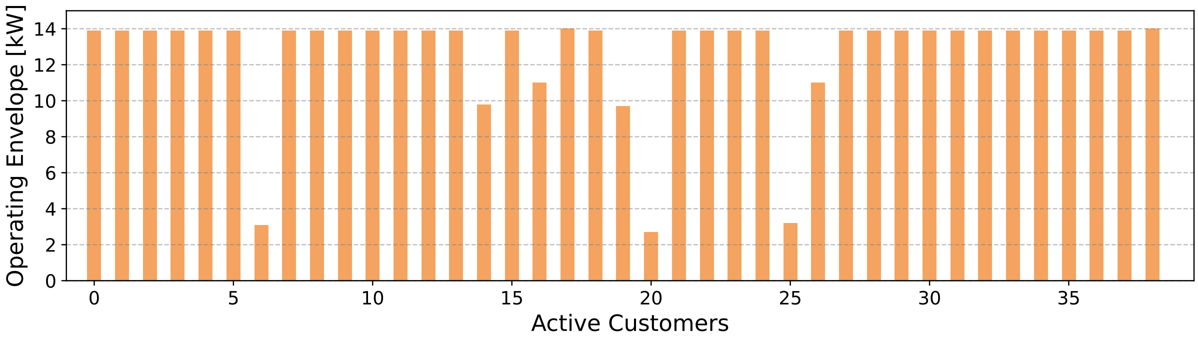

Figure 26. Model-Free OEs: Maximise Exports – S2

Therefore, it can be noted that the proposed model-free OEs approach can capture thermal violations and limit the corresponding OEs accordingly to ensure network integrity, providing a full electrical model-free approach to calculate OEs that enables the assessment of both, voltage, and thermal compliance, without the need for electrical models or power flow analyses.

4.2 Key Remarks

Thermal constraints of key assets can also limit active customers imports or exports and thus, must be considered when calculating OEs. However, the NN used to calculate voltages and thus, assess meter-level voltage compliance when calculating OEs, cannot cater for asset utilisation as it was trained to calculate voltages only. Therefore, approximations must be made. In this project, thermal constraints are approximated by using customers net demands and the thermal capacity (in kVA) of the corresponding assets.

To illustrate the effects of considering transformer and LV conductors constraints when calculating OEs, the case of Jemena substation B is used. Hence, OEs are calculated considering two scenarios, one where only voltage and customers’ connection point constraints are considered (as in [2], [3]) and one where LV conductor and transformer constraints are also considered, showing that the proposed model-free OEs approach can capture thermal violations and limit the corresponding OEs accordingly to ensure network integrity. It can be noted that by incorporating thermal constraints into the model-free OE formulation, the proposed approach enables the assessment of both, voltage, and thermal compliance, removing completely the need for power flow analyses or electrical models.

5 Model-Free Operating Envelopes: Implementation Architecture

This section presents a general overview on how the proposed model-free OEs approach could be potentially implemented within DNSPs' current systems. Note that the proposed architecture is built based on interactions with AusNet Services and thus, some parts can be specific to AusNet and could differ for other DNSPs.

As discussed in Section 2, the inputs required for the proposed model-free OEs approach are detailed below.

- Historical AMI data (i.e., $\boldsymbol{P}$, $\boldsymbol{Q}$, and $\boldsymbol{V}$) of customers associated with a given distribution transformer.
- Active and reactive power of passive customers (no need for OE calculation). This can come either from operational AMI data or forecasts, depending on how OEs are being calculated, i.e., in near real-time or hours, days ahead, respectively.
- Reactive power of active customers (those needing OE calculation). This can come either from operational AMI data or forecasts, depending on how OEs are being calculated.
- Voltage statutory limits of residential customers (e.g., 216V and 253V).
- Rated capacity (in kVA) of customers' service cables (or connection agreements).
- Rated capacity (in kVA) of conductors at the head of the LV network.
- Rated capacity (in kVA) of the distribution transformer.

Thus, for each distribution transformer, the approach must be able to access the inputs mentioned above from existing data bases. It is important to note that different inputs will be required at different stages of the entire process. For instance, historical AMI data (i.e., $\boldsymbol{P}$, $\boldsymbol{Q}$, and $\boldsymbol{V}$) is required to produce the corresponding NN within the offline stage only. Thus, historical AMI data must be accessed only when NN updates are carried out, which occurs periodically (e.g., every quarter). However, on the other hand, operational AMI data is required within the online stage if OEs are to be calculated in near real-time. This, in turns, means that the implementation architecture must be designed such that the approach can access to AMI data from all devices in near real-time.

The proposed model-free OEs implementation architecture is presented in Figure 27. As detailed in Section 2, the first step is to produce the corresponding NN, which occurs offline. This process requires historical AMI data (i.e., $\boldsymbol{P}$, $\boldsymbol{Q}$, and $\boldsymbol{V}$) of customers associated with a given distribution transformer. Thus, historical AMI data is retrieved and processed by the Model-Free Engine as detailed in Section 2.1.1. Once processed, it is then used to produce the most suitable NN as in Section 2.1.2. This process can be carried out periodically (e.g., every quarter, semester) to update the NN and cater for changes that can occur in LV networks and have an impact in the underlying relationships captured by the NN, e.g., reconfigurations, reinforcements, connection of new customers, etc. Alternatively, automated algorithms can be implemented to carry out periodical checks and trigger updates if changes are detected. Note that although thermal capacity of key assets is not required until the online stage, the corresponding data is retrieved at the offline stage as it will not vary unless significant changes in the LV network occurs, which will ultimately require a corresponding update.

Once the NN is trained, it is then used by the Model-Free OEs block within the online stage. This block is located within the distributed energy resources management system (DERMS), which is the software platform or suite of technologies designed to manage distribution energy resources in distribution systems. In here, the Model-Free OEs block uses the obtained NN to calculate OEs as detailed in Section 2.2. This process requires either forecast, which must be produced, or operational AMI data, which can be retrieved from the control and data transport platform. This platform corresponds to a bi-directional secure interface between the DERMS and customers devices, e.g., IEEE 2030.5 utility server. Thus, on one direction, settings that are calculated in the DERMS are sent to this platform to be then broadcasted through the internet to customers devices, this can occur directly to the device, though a gateway or through the cloud. On the other direction, this platform collects measurements from the different devices which are then sent back to the DERMS to be used for the different purposes.

Once OEs are calculated they can be then broadcasted to customers devices through the control and data transport platform in the form of flexible import or export limits or they can be broadcasted to aggregators for them to manage their corresponding portfolios according to such limits. The later will require an additional layer of communication among DNSPs and aggregators. Due data and security restrictions, all calculations are most likely to be implemented on-premises. However, it is important to note that the proposed approach could also be implemented on cloud. This decision will be case specific for each DNSP.

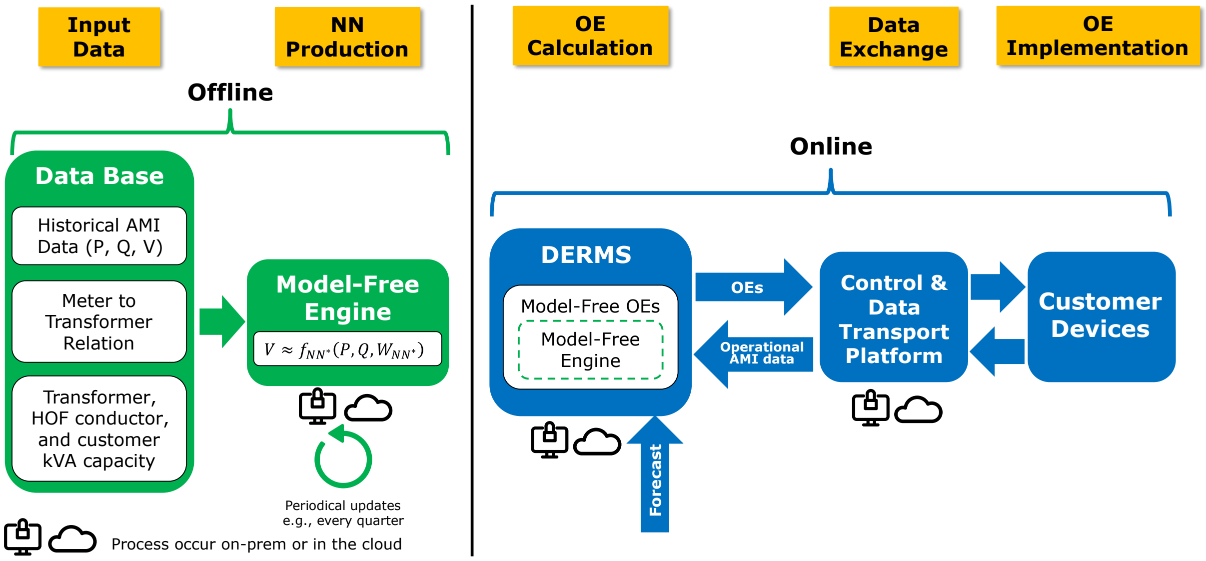

Figure 27. Model-Free OEs Implementation Architecture

5.1 Key Remarks

The proposed model-free OEs approach can be integrated into current systems of DNSPs. This requires two blocks, Model-Free Engine and Model-Free OEs. The former works offline, is dedicated to produce the corresponding NN and thus, it must have access to the different data bases where its corresponding inputs are found every time an update is carried out. The later, on the other hand, is dedicated to use the obtained NN to calculate OEs online, it must be integrated within the DERMS and must have access to operational AMI data or forecast, depending on how OEs are required, i.e., in near real-time or in advance, respectively.

6 Voltage Calculations per HV Feeder: Qualitative analysis

As detailed in Section 2, the proposed model-free OEs approach is based on the proposed model-free voltage calculations, which are carried out per distribution transformer, i.e., one NN is trained per distribution transformer, as presented in Figure 28. From an application perspective, it is of interest to explore if the proposed model-free voltage calculations can be potentially scaled to a single NN per high voltage (HV) feeder as shown in Figure 29. Thus, a single NN could be potentially used to calculate OEs for the thousands of customers supplied by the different LV networks connected to the same HV feeder.

It is important to note that, in this context, a quantitative analysis would require historical data (3 weeks at 5 minutes resolution) from the thousands of customers connected to the same HV feeder, which was not available for the present study. Therefore, this section presents a qualitative analysis and discuss the practical implications of scaling the proposed approach to a single NN per HV feeder.

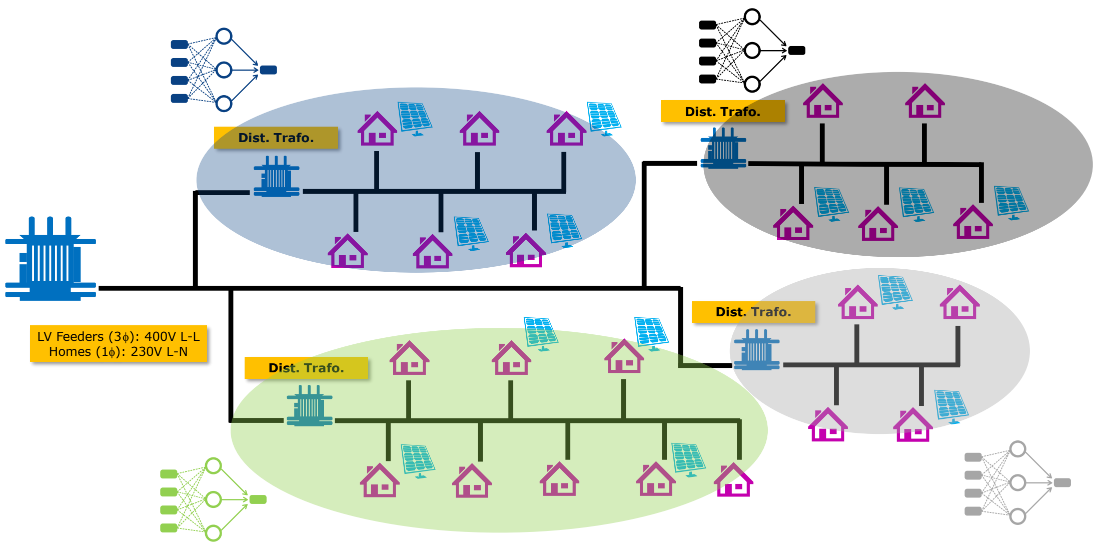

Figure 28. Schematic of per Distribution Transformer Calculations

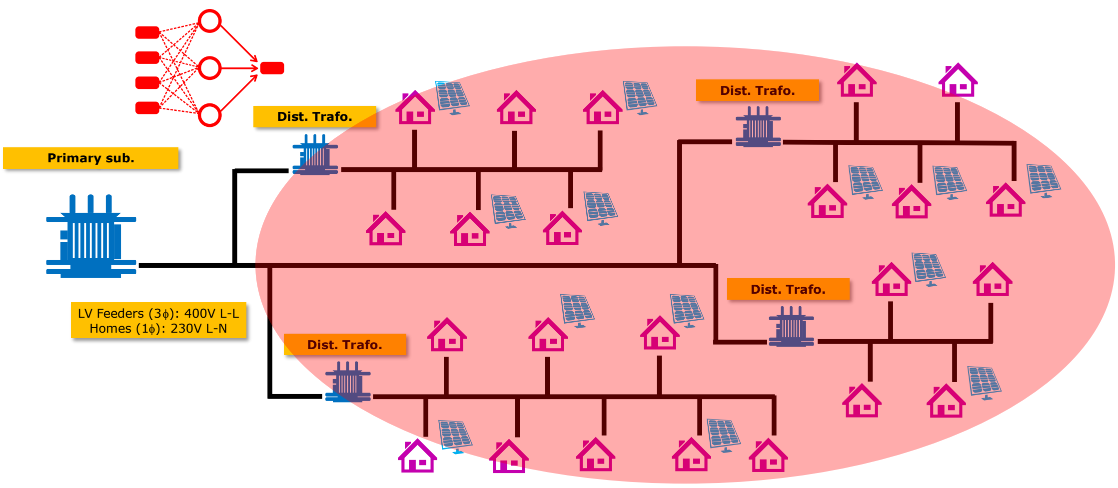

Figure 29. Schematic of per Distribution HV Feeder Calculations

The potential advantages of using a single NN per HV feeder are listed below.

- **Simplicity:** Only 1 NN is produced and only 1 NN is deployed per HV feeder.
- **Accuracy:** Customers' voltages are a result of the interactions among customers' net demands and the physical infrastructure of each LV network (commonly known as a *voltage drop*) plus a voltage reference. This voltage reference corresponds to the voltage at the head of the LV network and is mainly driven by the upstream interactions among the HV feeder and other LV networks connected to it. In fact, it has been demonstrated in [1] that the incorporation of such reference as input to the proposed model-free voltage calculations improves, significantly, the accuracy of the obtained calculations.

  When a NN is trained per distribution transformer, the historical data of its customers is not representative of the major drivers of such voltage reference as they occur outside of the corresponding LV network. However, when a single NN is trained per HV feeder, this information will be embedded in the historical data of the thousands of customers connected to the corresponding HV feeder which could potentially improve accuracy.
- **Computational time:** DNSPs can manage hundreds of HV feeders with dozens to hundreds of distribution transformers each. Depending on the NN architecture (a NN for thousands of customers will likely require a deeper architecture than the one presented in Section 2) and on the implementation of parallel computing, the use of a single NN per HV feeder could potentially reduce the computational times and costs associated with the production and deployment of the proposed approach.

With all the potential advantages presented above, the use of a single NN per HV feeder might seem attractive from an application perspective. However, its technical feasibility must be evaluated before pursuing further analyses. As discussed in Section 2, HV and LV networks are subject to changes, e.g., reconfigurations, reinforcements, or connection of new customers, that impacts the relationships captured by the NN. If such a change occurs, an accuracy decrease is expected until the NN is updated with sufficient data (e.g., three weeks at 5 minutes resolution) representative of the latest condition.

To illustrate the concepts, a simple example of an HV feeder with 100 LV networks supplying 50 customers each is used. Within each LV network, only one change, e.g., reconfiguration, reinforcement, or connection of new customer, per calendar year is considered. The exact date of each change is randomly assigned, and it is considered that all changes are immediately detected. Once each change occurs, sufficient smart meter data representative of the latest condition must be collected (e.g., 3 weeks) and an accuracy decrease is expected until the NN is updated (i.e., throughout those 3 weeks). In this context, if a NN per distribution transformer is considered, a total of 100 NNs (with $|C| = 50$) are trained. In this case, each NN will be updated 1 time throughout the calendar year and thus, accuracy decrease is expected for each customer for 3 weeks only. However, if a NN per HV feeder is considered, a single NN (with $|C| = 5{,}000$) is trained and will need to be updated 100 times throughout the calendar year. This, in turn, means that accuracy decrease is expected throughout most of the year for each customer (355 days in this simple example).

Therefore, by considering a simple qualitative analysis, it can be noted that changes in HV and LV networks could make it impractical to use a single NN per HV feeder. However, further research is needed to adequately assess the impacts of changes in one LV network on the voltage calculations of other LV networks connected to the same HV feeder. Furthermore, it is important to note that if a single NN per HV feeder is considered, the NN structure required for such task will be likely deeper than the single layer NN considered for a single LV network in this project and therefore, the NN structure and its corresponding hyperparameter selection process would need to be revised.

6.1 Key Remarks

The proposed model-free OEs approach is based on the proposed model-free voltage calculations, which are carried out per distribution transformer, i.e., a NN is trained per each distribution transformer. However, from an application perspective, scaling the approach to a single NN per HV feeder could bring interesting benefits, such as simplicity, accuracy improvements and, potentially, reducing computational times. While this might seem attractive, it is noted by a simple qualitative analysis that it might result impractical in the context of distribution networks as HV and LV networks are often subject to changes, e.g., reconfigurations, reinforcements, or connection of new customers, which, in turn, means that the NN used for the entire HV feeder (i.e., catering for multiple LV networks simultaneously) would be constantly subject to accuracy decrease and the need for updates as its underlying relationships will be constantly changing.

7 Partial Smart Meter Data Availability

In Victoria, even though full deployment of smart meters is observed, the smart meter data of some customers might not be available to the DNSPs. This is the case, for example, of some commercial and industrial (C&I) customers whose data is managed by third parties and, hence, unavailable for DNSPs. On the other hand, for DNSPs outside Victoria, smart meter deployment is still ongoing. Hence, it is important to assess if the proposed model-free voltage calculations (and the subsequent model-free OEs) can be used when smart meter data from all customers is not available as this could be the scenario faced by DNSPs across Australia when implementing the proposed approach. Preliminary analyses presented in [3] showed that it is possible to obtain accurate voltage calculations as well as consistent OEs when smart meter data from all customers is not available. However, it is noted that some areas could have C&I predominance, i.e., data from only a reduced number of customers would be available for DNSPs (e.g., 20%). Additionally, further studies have shown that the proposed model-free voltage calculations and its corresponding OEs are impacted by the location and operation of those customers whose smart meter data is not available. This section presents a case study that cater for these aspects.

To illustrate this point, AusNet Site A is used. AusNet Site A corresponds to a LV network with a single LV circuit that supplies a total of 28 customers, 26 single-phase customers and 2 three-phase customers. Let consider AusNet Site A as an area with C&I predominance as depicted in Figure 30. Thus, ≈80% of customers are considered as C&I customers whose data is managed by third parties and therefore, unavailable to the DNSP. In this context, only the data from A6, A16, A5, A4, A20, and A27 will be available to produce the corresponding NN and carry out voltage calculations. However, the remaining customers (≈80%) will have an effect over the voltages of customers A6, A16, A5, A4, A20, and A27 that is not being represented in the historical data used to train the corresponding NN. Therefore, an accuracy decrease is expected for the case of partial smart meter data availability. The magnitude of such accuracy decrease will be associated to the location and operation of such customers.

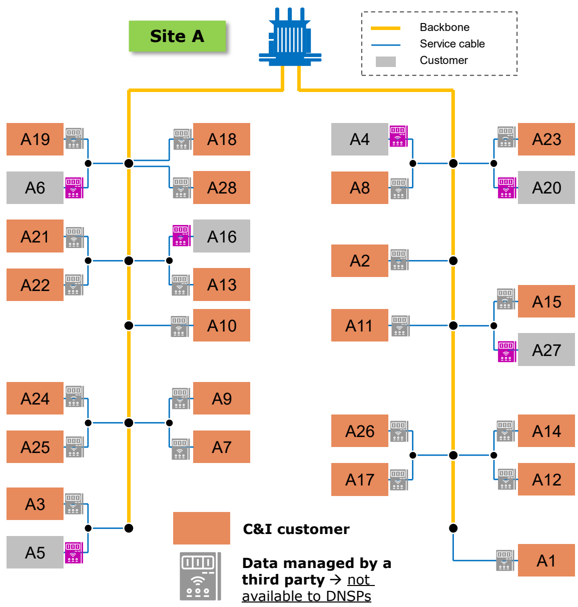

Figure 30. Example: AusNet Site A as an area of C&I predominance

7.1 Case Study

To assess the effectiveness of the proposed approach under partial smart meter data availability, the case of UE substation B is considered. Thus, the offline data pipeline is implemented as presented in Section 2 and training and test data sets are built. After the offline data pipeline is implemented, a total of 72 single-phase customers and 3 three-phase customers are considered. The training data set corresponds to an equivalent to 3 weeks at 5 minutes resolution in the period between 01-12-2020 and 24-12-2020, whereas the test data set is consecutive and of the same length, i.e., between 24-12-2020 and 16-01-2021.

With this data, a progressive decrease in smart meter data availability is considered in steps of 5%. For each smart meter data availability level, 10 cases with different availability of customers are considered. Therefore, for instance, at 50% of smart meter data availability, 10 NNs are produced with data from only 50% of customers. Each NN considers different single-phase customers as C&I customers (i.e., with its corresponding data unavailable). The hyperparameters for all NNs trained in this study correspond to those defined for UE substation B in [2] for the case of 3 weeks of data. These hyperparameters are presented in Table 14. Note that L2 regularisation was introduced at a later stage in [3] and thus, was not part of the selection process carried out in [2]. For this study, L2 regularisation is considered as 0.1 times the learning rate.

Table 13. UE substation B - NN Hyperparameters

$\lvert C\rvert$	Number of Neurons	Activation functions	Learning rate	L2 regularisation	Batch size	Epochs
81	$7\lvert C\rvert$	Tanh	1×10^{-4}	1×10^{-5}	144 (eq. 12 hours)	500

Once all NNs are produced, two studies are carried out. First, the accuracy of the proposed model-free voltage calculations is assessed in the corresponding test data set and then, the produced NNs are used to calculate model-free OEs as in Section 2 considering both allocation techniques, i.e., equally distributed and maximise exports.

7.1.1 Model-Free Voltage Calculations

To establish a benchmark, the case of full observability, i.e., when the historical smart meter data from all customers connected to the LV network is available, is used. The obtained model-free voltage calculations results are presented numerically in Table 14, as time-series in Figure 31, and as scatter plot in Figure 32, where it can be observed that if the historical data from all customers is available, the proposed approach can accurately capture the underlying relationships and produce accurate voltage calculations, achieving a RMSE of 1.1222 V, an Av. Dev. of 0.8600 V, and a Max. Dev. of 7.0467 V.

Table 14. Model-Free Voltage Calculations Results (Full Observability)

Training Initial Instance	Training Final Instance Test Initial Instance	Test Final Instance	RMSE V	Av. Dev. V	Max. Dev. V
01-12-2020	24-12-2020	16-01-2021	1.1222	0.8600	7.0467

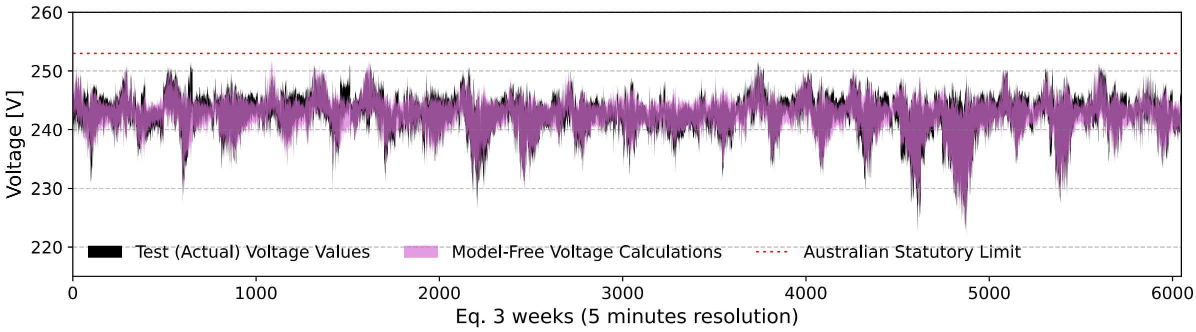

Figure 31. Model-Free Voltage Calculations: Time-Series Results (Full Observability)

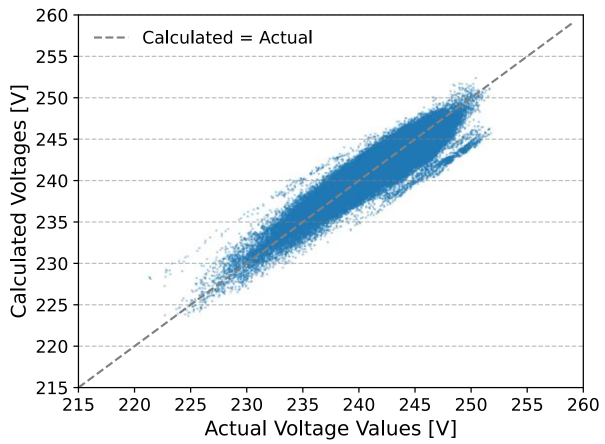

Figure 32. Model-Free Voltage Calculations: Calculated vs Actual Voltage Values (Full Observability)

Results from all smart meter data availability levels considered in this study are presented in boxplots in terms of average deviation and maximum deviation in Figure 35 and Figure 36, respectively. Note that each boxplot is to represent the 10 cases with different availability of customers' data assessed at each smart meter data availability level. Besides, note that the level of 100% is comprised of a single value only as this case comprise data from all customers. Similarly, the level of 20% is also comprised of a single value only as this study removes single-phase customers only and the 20% level is comprised mostly of three-phase customers and the single-phase customers which are to be considered as active customers in the subsequent calculation of OEs (and thus, cannot be removed).

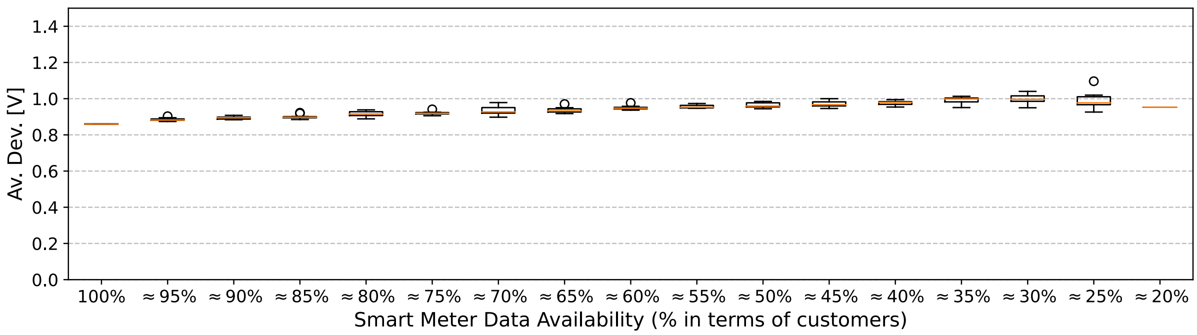

Figure 33. Model-Free Voltage Calculations: Av. Dev. per Smart Meter Data Availability

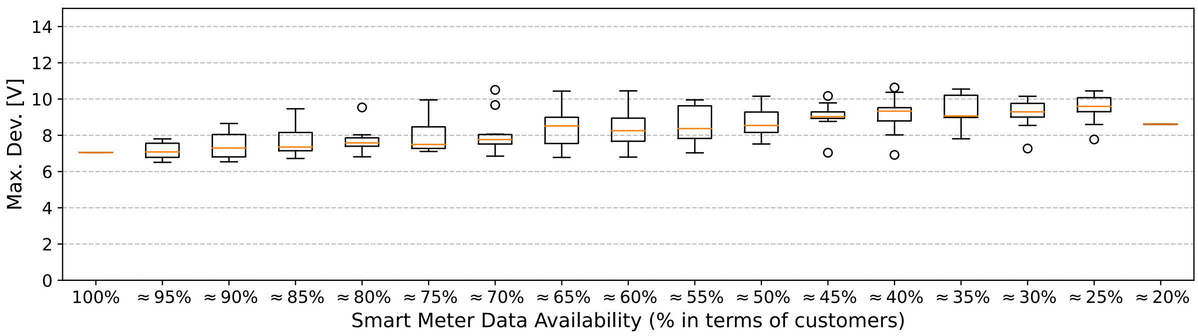

Figure 34. Model-Free Voltage Calculations: Max. Dev. per Smart Meter Data Availability

From Figure 35 it can be observed that similar average deviation is obtained in all cases throughout all smart meter data availability levels, achieving a minimum value of 0.8600 V in the case of full observability (i.e., 100%) and a maximum value of 1.0964 V in one of the cases of 25% of smart meter data availability. Furthermore, from Figure 36 it can be observed that the maximum deviation can present higher fluctuations, achieving a minimum value of 6.5075 V in one of the cases of 95% availability and a maximum value of 10.6358 V in one of the cases of 40% availability. It is important to note that as detailed in Section 2, such high deviations occur only in a reduced number of instances in each case. Time-series model-free voltage calculations results for the worst cases (i.e., worst Av. Dev.) of 90%, 80%, 70%, 60%, 50%, 40%, 30%, and 20% of smart meter data availability are presented in Figure 35 to Figure 42, where it can be observed that, although an accuracy decrease is observed as less data is considered (as expected), accurate voltage calculations are obtained in all cases.

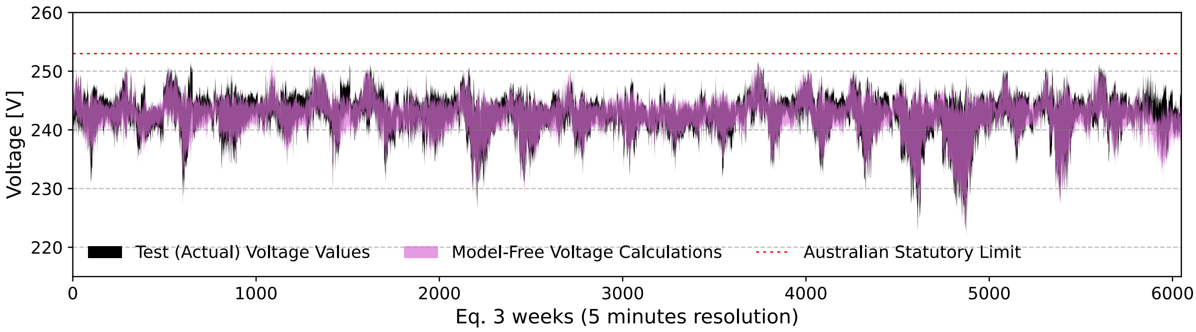

Figure 35. Model-Free Voltage Calculations: Time-Series Results (90%)

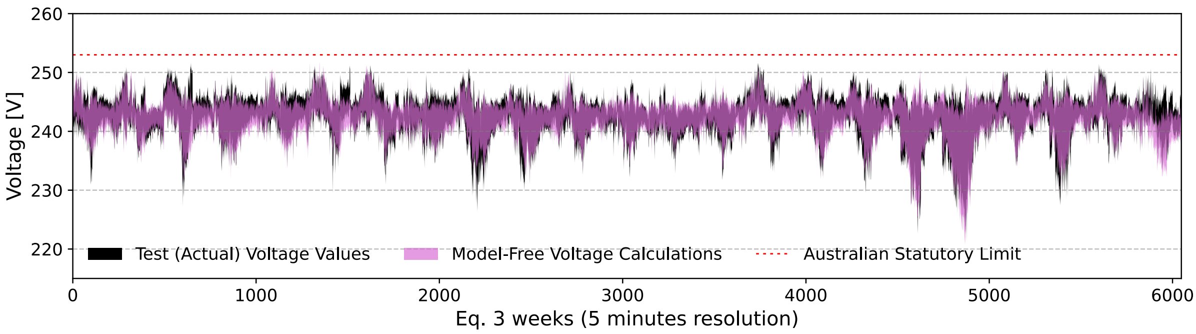

Figure 36. Model-Free Voltage Calculations: Time-Series Results (80%)

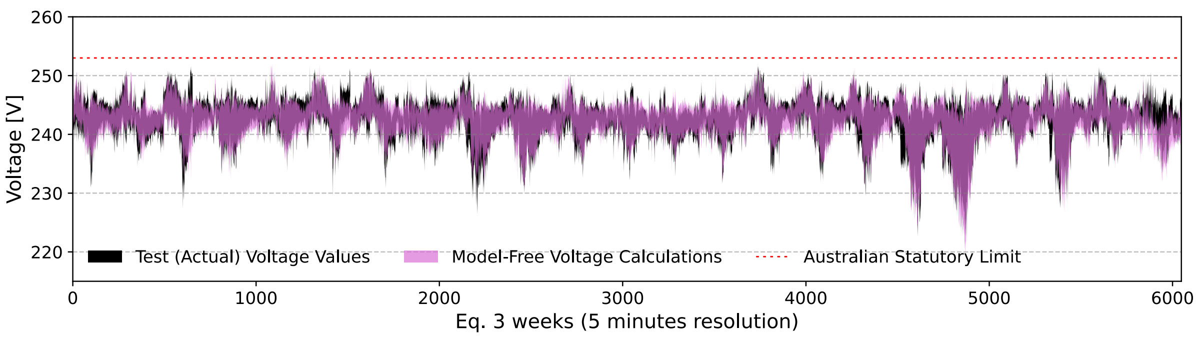

Figure 37. Model-Free Voltage Calculations: Time-Series Results (70%)

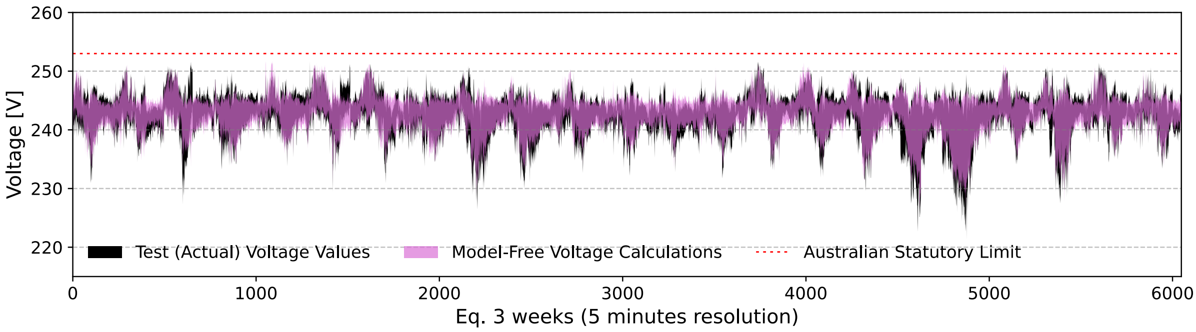

Figure 38. Model-Free Voltage Calculations: Time-Series Results (60%)

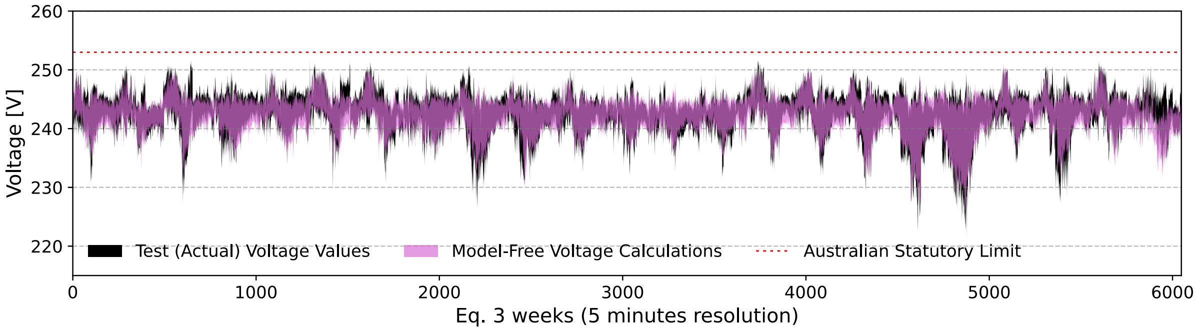

Figure 39. Model-Free Voltage Calculations: Time-Series Results (50%)

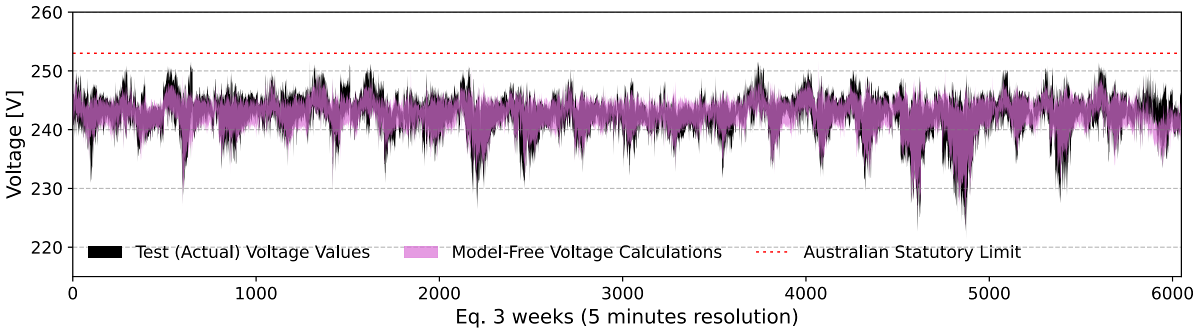

Figure 40. Model-Free Voltage Calculations: Time-Series Results (40%)

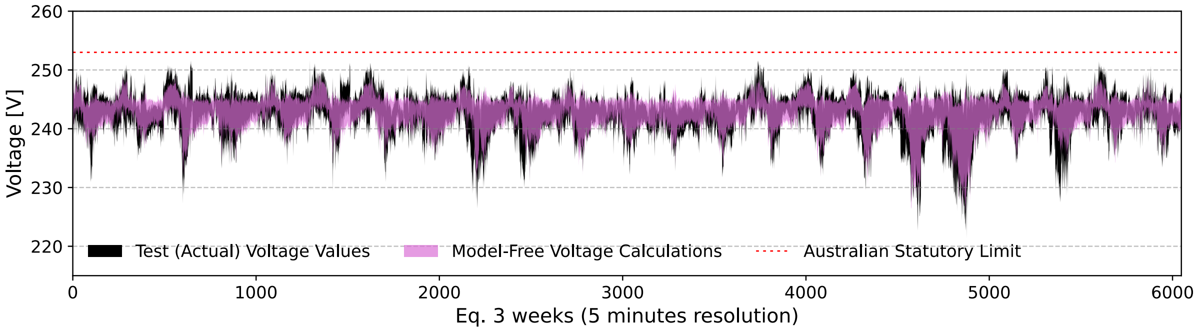

Figure 41. Model-Free Voltage Calculations: Time-Series Results (30%)

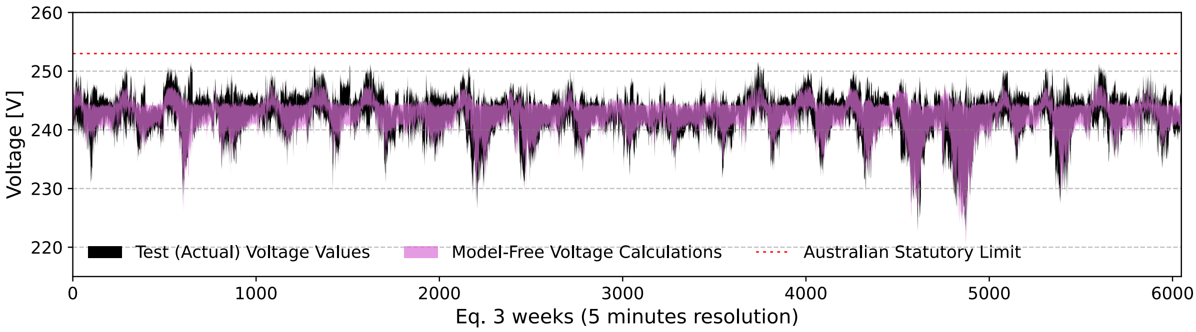

Figure 42. Model-Free Voltage Calculations: Time-Series Results (20%)

Therefore, it can be concluded that, although an accuracy decrease is observed when data from a reduced number of customers is considered (as expected), the proposed approach can produce accurate voltage calculations even when data from only 20% of customers is available.

7.1.2 Model-Free Operating Envelopes

Even though accurate model-free voltage calculations can be obtained in all cases throughout all the considered smart meter data availability levels, the final application of the proposed model-free voltage calculations considered in this project is to be used to calculate OEs. Thus, OEs for the first instance at noon of the test data set are calculated as in Section 2. In this context, equally distributed and maximise exports OEs are considered, and half of the single-phase customers detected with PV system are randomly allocated as active customers (i.e., 9 active customers). Active and reactive power of passive customers as well as reactive power of active customers are obtained from the historical data.

For compactness, this study assesses only the case of export limits. However, it is important to highlight that for the case of import limits, the calculation mirrors the one for the case of export limits but considering active customers importing active power.

Note that, in the context of partial smart meter data availability, thermal assessments of transformer and LV conductors cannot consider its full capacity to define the corresponding limits, i.e., $s_{conductor}$ and $s_{transformer}$ detailed in Section 4. This is because the capacity of these assets will be shared with those customers whose data is not available. Thus, imports or exports must be restricted to what the observed customers can actually utilise, considering that the unobserved customers will be utilising the same assets. In this study, $s_{conductor}$ and $s_{transformer}$ are reduced according to the corresponding smart meter data availability as in (20) and (21), respectively. Note that this is a rather conservative approach as active customers are expected to utilise the most the corresponding assets. Furthermore, for this study, the capacity of the corresponding assets was not available. Therefore, values are approximated from a LV network from AusNet with similar characteristics and correspond to $s_{conductor\ at\ 100\%} = 100\ \frac{kVA}{phase}$ and $s_{transformer\ at\ 100\%} = 220\ kVA$. Customer's connection point and voltage constraints are as in Section 2 (i.e., 14 kVA and 253 V, respectively).

$$s_{conductor\ at\ X\%} = s_{conductor\ at\ 100\%} \times X\% \tag{19}$$

$$s_{transformer\ at\ X\%} = s_{transformer\ at\ 100\%} \times X\% \tag{20}$$

Aggregated OEs obtained with the equally distributed allocation technique are presented for each case at each smart meter data availability level in Figure 43. In here, it can be observed that although different, consistent OEs are obtained in all cases throughout all availability levels. Thus, it can be noted that OEs do not deviate much among the different cases in the each of the availability levels, among availability levels, and with respect to the case of full observability (i.e., 100%) where an aggregated OE of 36.9 kW (i.e., 4.1 kW per active customer) is obtained. Specifically, a minimum value of 25.2 kW (i.e., 2.8 kW per active customer) is obtained in one of the cases for 90% availability and a maximum value of 48.6

kW (i.e., 5.4 kW per active customer) is obtained in cases with 80%, 55% and 35% of smart meter data availability.

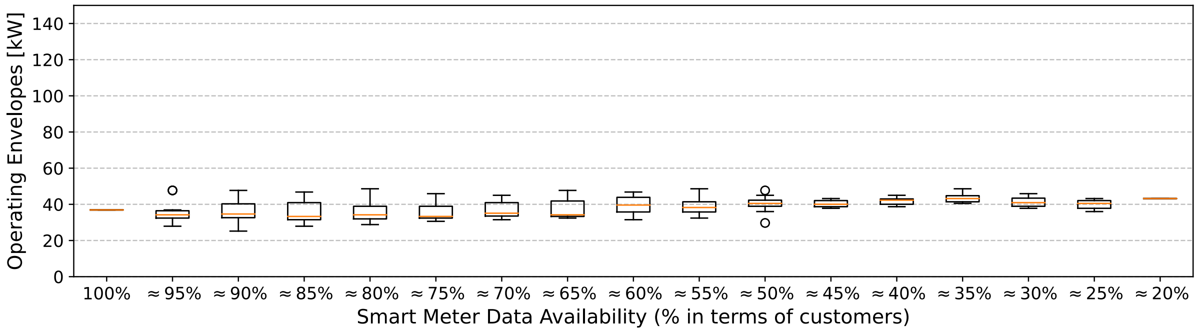

Figure 43. Model-Free OEs: Equally Distributed Operating Envelopes per Smart Meter Data Availability

Aggregated OEs obtained with the maximise exports allocation technique are presented in Figure 44. In here, on one hand it is important to note that OEs after the level of 50% of availability are heavily constrained because of the reductions considered in the capacity of the transformer and LV conductors beyond such level. On the other hand, it can be observed that before this level, the obtained OEs are not as consistent as for the case of equally distributed OEs within each availability level and with respect to the to the case of full observability (i.e., 100%) where an aggregated OE of 70.6 kW is obtained. It can be noted that in several availability levels (e.g., to 95% to 60%) an aggregated OE above 100 kW is obtained in a reduced number of cases, this can be noted in Figure 45, where instead of boxplots, aggregated OEs for each case in each availability level are presented.

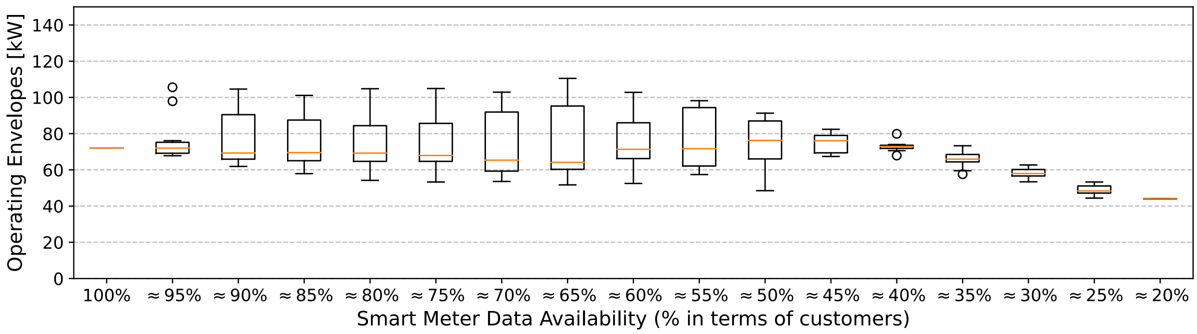

Figure 44. Model-Free OEs: Equally Distributed OEs per Smart Meter Data Availability

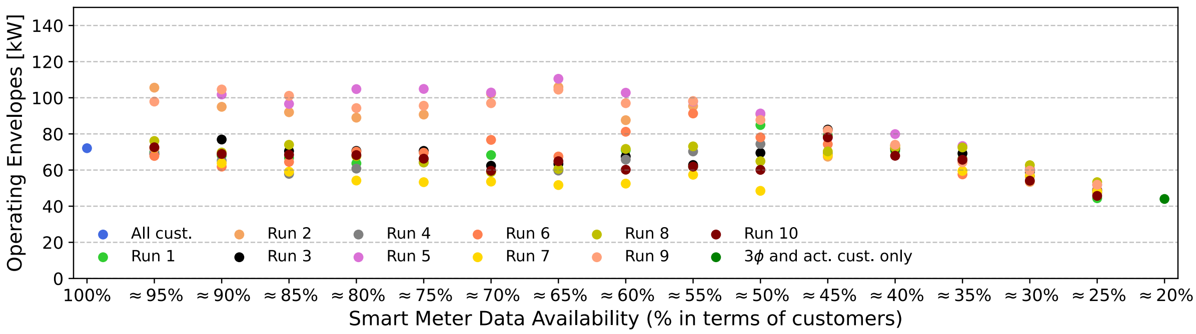

Figure 45. Model-Free OEs: Equally Distributed OEs per Smart Meter Data Availability (detail)

It is important to note that for this case study, OEs obtained when using data from all customers (i.e., 100% availability) are limited by voltage constraints. Specifically, a single-phase customer with a PV system that is not considered as active customer limits the OE for all active customers. This can be observed in Figure 46, where the voltages from all customers for the OEs calculated using the maximise exports allocation technique are presented, note that active customers are highlighted with a dashed grey line and the red dashed line is to represent the upper statutory limit of 253 V.

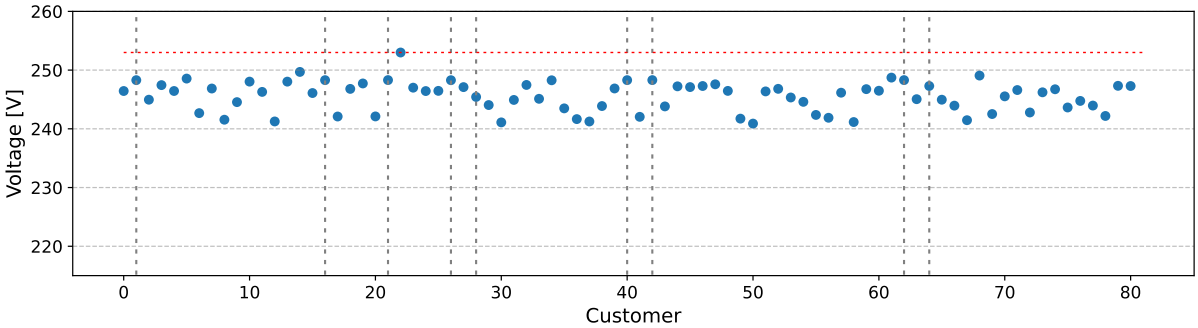

Figure 46. Model-Free Voltage Calculations at the corresponding OE: Maximise Exports – Full Observability

As observed in Figure 46, for this study, the OEs from all active customers are being constrained by the voltage of a specific single-phase customer, probably located far from the distribution transformer and thus, subject to a high sensitivity with respect to voltages (hereinafter, customer 22). This, in turn, means that the OEs obtained when data from all customers is not available will be dependent on the observability of customer 22. On one hand, if the historical data of customer 22 is available to the DNSP, the NN will capture its relationships and thus, the model-free OEs approach will constraint the OEs from active customers such that voltage of customer 22 is complaint with statutory limits (as in Figure 46). Producing OEs that are consistent with the case of full observability. However, on the other hand, if the data of customer 22 is unavailable for the DNSP, the historical data used to train the NN will not be representative of the actual constraint that limits active customers exports. Therefore, the model-free OEs approach will exploit the large voltage headroom observed for the remaining customers and will end up allowing larger exports, overestimating the corresponding OEs.

This is demonstrated in Figure 47, where aggregated OEs obtained with the maximise exports allocation technique are detailed as in Fig 45, but with those cases where the historical smart meter data of customer 22 was considered not available to DNSPs highlighted with a red cross. In here, it can be observed that all cases that show inconsistency among the same level of smart meter data availability and with respect to the case of full observability before the level of 45% of availability (where reductions in thermal constraints starts to limit the corresponding OEs) correspond to cases where the data from customer 22 is not available and thus, the proposed model-free OEs end up exploiting the larger voltage headroom of the observed customers and calculating larger OEs. On the other hand, it can be observed that in those cases where the data from customer 22 is considered available, the obtained OEs are consistent among each availability level and with respect to the case of full observability.

Note that the impacts of the observability of customer 22 are higher when using the maximise exports allocation technique. This is because this allocation technique has been specifically designed to make the most of the observed voltage headroom. However, it can be highlighted that cases were larger OEs are obtained with the equally distributed allocation technique (Figure 43) also correspond to cases were the data of customer 22 is considered unavailable.

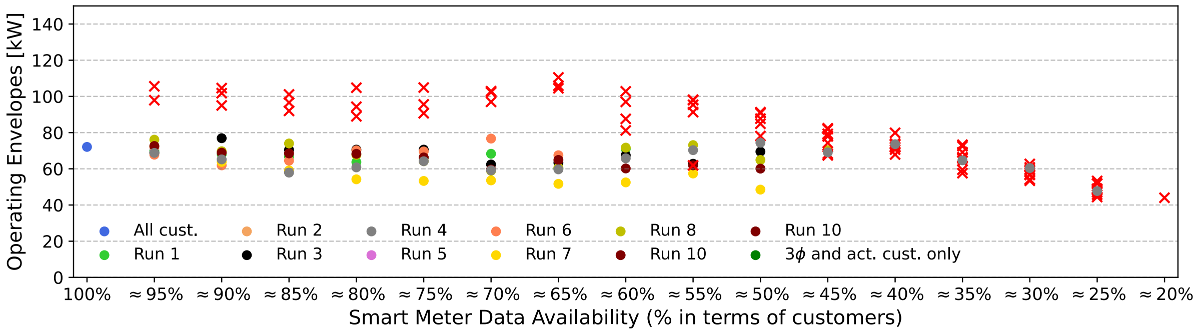

Figure 47. Model-Free OEs: Equally Distributed OEs per Smart Meter Data Availability (detail – cases without customer 22)

Finally, it is important to note the case study presented in this section has limitations in adequately representing the case of Victoria and C&I customers. The demand of the single-phase customers that were removed to assess smart meter data unavailability might not necessarily represent the demand of C&I customers (which is higher and has a different behaviour). Nonetheless, it is also important to note that the case study presented in this section can be considered representative of DNSPs outside Victoria and other places around the world where there is partial deployment of residential smart meters.

7.2 Key Remarks

Although full deployment of smart meters is observed in Victoria, the smart meter data of some customers might not be available to the DNSPs. This is the case of some C&I customers whose data is managed by third parties and therefore, not available to DNSPs. While preliminary analyses in [3] showed that it is possible to obtain accurate voltage calculations and consistent OEs when data from all customers is not available, it is noted that some areas could have C&I predominance, i.e., data from only a reduced number of customers would be available for DNSPs (e.g., 20%). Additionally, further studies have shown that the proposed model-free voltage calculations and its corresponding OEs are impacted by the location and operation of those customers whose smart meter data is not available. This section presents a case study that cater for these aspects.

The case of UE substation B is considered. Thus, several NN are trained with different levels of smart meter data availability, considering 10 different cases of customers' availability in each level. Results shows that accurate model-free voltage calculations can be obtained even when data from only 20% of customers is available. Similarly, consistent OEs are produced in all cases among all levels of smart meter data availability when calculated using the equally distributed allocation technique. However, some inconsistency can be observed for the case of maximise exports OEs. Further analyses shows that this is strictly related to the availability of the smart meter data of a single-phase customer (customer 22) that constraints the OEs from all active customers. Thus, if the historical data of customer 22 is available to the DNSP, the NN will capture its relationships and thus, OEs that are consistent with the case of full observability are produced. However, on the other hand, if the data of customer 22 is unavailable for the DNSP, the historical data used to train the NN will not be representative of the actual constraint that limits active customers exports. Therefore, the model-free OEs approach will exploit the large voltage headroom observed for the remaining customers and will end up allowing larger exports, overestimating the corresponding OEs.

Results presented in this section are promising as adequate model-free voltage calculations and OEs can be obtained even if data from only 20% of customers is available. Nonetheless, the consistency of the OEs, will depend on the location and characteristics of the customers whose smart meter data is not available.

8 Conclusions

The increasing adoption of residential distributed energy resources (DERs), such as solar photovoltaics (PV), batteries, and electric vehicles (EVs), is driving the need for distribution network service providers (DNSPs) to implement Operating Envelopes (OEs), i.e., meter-level time-varying import and export limits for active customers, to ensure network integrity. However, the main challenge for DNSPs is that an accurate determination and assessment of OEs would normally require power flow analyses and, consequently, need detailed three-phase electrical models of low voltage (LV) distribution networks, which, in practice, are rarely available and are costly and time-consuming to produce.

To overcome the lack of electrical models, this project has successfully demonstrated that is possible to exploit the historical smart meter data of customers to capture the physics of three-phase LV networks and create an electrical model-free approach to calculate voltages. These model-free voltage calculations can then be used along with a heuristic algorithm and tailored approximations to assess several import or export values for active customers (i.e., customers engaged with aggregators) and thus, determine its corresponding OEs. This approach represents an accurate, cheap, and extremely quick alternative to traditional model-based approaches, enabling DNSPs to bypass the costly, time-consuming, and error-prone process of producing and validating electrical models.

The report at hand builds on top of our previous reports [1] (foundations and methodology), [2] (extensive performance tests), and [3] (improvements and updates), and presents the final version the model-free OE approach and the main findings of the last 6 months of this project. This includes final improvements to the offline data pipeline designed to cater for customers with extremely low variability, the incorporation of thermal constraints of key assets into the model-free OEs formulation. It also provides a general overview of the model-free OEs implementation architecture and a qualitative analysis of the scalability of the proposed model-free voltage calculations for a high voltage (HV) feeder. Finally, it presents further analysis of partial smart meter data availability. The conclusions of this report can be summarised as follows.

Model-Free Operating Envelopes

Using real smart meter data from 690+ customers across United Energy, Jemena, and AusNet Services we have demonstrated that our model-free voltage calculations can achieve a high accuracy (average deviation below 1.5 Volts in most cases) which, in turn, means accurate calculations of OEs as well as for other applications (e.g., connection requests, hosting capacity, etc.).

The proposed model-free OE approach considers using model-free voltage calculations instead of power flow simulations. From an application perspective, the proposed model-free OE approach can be divided into two main stages: offline (production of the NN) and online (model-free OE calculation). Details from all steps within each stage are presented in this report. First, a neural network (NN) is trained offline to capture the underlying relationships among the historical smart meter data of customers and the corresponding LV network. Then, the trained NN is used online along with a heuristic algorithm to explore different import or export values for active customers and thus, calculate the OEs. Meter-level voltage compliance is assessed through the trained NN, whereas thermal compliance of transformer, conductors, and customers' connection point, is assessed through its respective thermal ratings and customers' demands. Removing, completely, the need for power flow analyses and electrical models. A case study is presented to illustrate the implementation of the proposed model-free OEs, showing that the proposed approach is an accurate, cheap, and fast alternative to traditional model-based approaches to calculate OEs, enabling DNSPs to accurately calculate OEs without the need for the costly, time-consuming, and error-prone process, of producing and validating electrical models.

Offline Data Pipeline: Final Improvements

Improvements to the offline data pipeline were made due to data issues involving customers with extremely low variability (i.e., all or most values in a range of a few Watts).

The improvements were necessary because the historical data of such customers does not provide variety enough for the NN to adequately capture its relationships. Thus, any adequate value for such customers will be far from what has been observed during training and will cause erroneous calculations for all customers. This is critical in the context of OEs as its calculation and assessment is based on pushing customers’ limits. The potential impacts of such customers when deploying the proposed approach are shown and to address this, Step 4 of the offline data pipeline is adapted to automatically detect and remove such customers from the historical data and thus, from the proposed model-free voltage and OEs calculations. Once sufficient adequate data for these customers is available, the NN can be updated to cater for these customers.

Thermal Constraints in Model-Free Operating Envelopes

The proposed model-free OEs approach can capture thermal violations of key assets and limit the corresponding OEs accordingly to ensure network integrity, providing a full model-free approach to assess both, voltage, and thermal compliance, removing completely the need for power flow analyses or electrical models.

For simplicity, only voltage constraints and customers’ connection point limitations were assessed before this report. However, thermal constraints of key assets can also limit active customers imports or exports and thus, must be considered when calculating OEs. The NN used to calculate voltages and thus, assess meter-level voltage compliance when calculating OEs, cannot cater for asset utilisation as it was trained to calculate voltages only. Therefore, thermal constraints are approximated by using customers net demands and the thermal capacity (in kVA) of the corresponding assets.

Model-Free Operating Envelopes: Implementation Architecture

The proposed model-free OEs approach can be potentially integrated into current systems of DNSPs. This requires two blocks, Model-Free Engine and Model-Free OEs.

The first block is offline, is dedicated to produce the corresponding NN and thus, it requires access to the different data bases where its corresponding inputs are found every time an update is carried out. The second block, on the other hand, is dedicated to use the obtained NN to calculate OEs online, it must be integrated within the Distributed Energy Resources Management System (DERMS) and must have access to operational advanced metering infrastructure (AMI) data or forecast, depending on how OEs are required, i.e., in near real-time or in advance, respectively.

Voltage Calculations per HV Feeder: Qualitative Analysis

Having a single NN per HV feeder is likely to be impractical as high voltage (HV) and LV networks can be subject to changes (e.g., reconfigurations, reinforcements, or connection of new customers) which in turn would require more frequent updates of the NN than when applied to individual LV networks.

The proposed model-free OEs approach is carried out per distribution transformer. However, scaling the approach to a single NN per HV feeder could bring interesting benefits, such as simplicity, accuracy improvements and, potentially, reducing computational times. Nonetheless, while this might seem attractive, it might result impractical as the multiple changes that occur on the HV and LV networks would require the NN used for the entire HV feeder (i.e., catering for multiple LV networks simultaneously) to be updated frequently.

Partial Smart Meter Data Availability

The proposed model-free approach can calculate voltages with an adequate accuracy and, therefore, adequate OEs, even if data from only 20% of the customers is available.

Although Victoria has full deployment of smart meters, the data ($\boldsymbol{P}$, $\boldsymbol{Q}$, and $\boldsymbol{V}$) of commercial and industrial (C&I) customers might not be available to the DNSPs (only kWh) as it is managed by third parties. To understand the implications of limited data availability (which is also relevant for DNSPs outside Victoria), this report investigates decreasing levels of smart meter data availability. Results are promising as adequate model-free voltage calculations and OEs can be obtained even if data from only 20% of customers is available. Nonetheless, the consistency of the OEs, will depend on the location and characteristics of the customers whose smart meter data is not available.

References

[1] V. Bassi, D. Jaglal, L. F. Ochoa, T. Alpcan, and C. Leckie, “Deliverable 0 – Concept, Smart Meter Data, and Initial Findings”, February 2022. PDF available on ResearchGate

[2] V. Bassi, D. Jaglal, L. F. Ochoa, T. Alpcan, and C. Leckie, “Deliverables 1-2-3a: Model-Free Voltage Calculations and Operating Envelopes”, July 2022. PDF available on ResearchGate

[3] V. Bassi, L. F. Ochoa, T. Alpcan, and C. Leckie, “Deliverable 3b-4: Improved Model-Free Operating Envelopes and Other Considerations”, January 2023. PDF available on ResearchGate

[4] ARENA, Project EDGE. Available

[5] ARENA, Composing a distributed energy symphony in Western Australia’ largest energy grid. Available

[6] ARENA, Electricity networks gear up to manage electric vehicle demands on the grid. Available

[7] ARENA, SA Power Networks Flexible Exports for Solar PV Trial. Available

[8] DEIP, “Dynamic Operating Envelopes Workstream: Outcomes Report”. Available

[9] The University of Melbourne, Assessing the Benefits of Using Operating Envelopes to Orchestrate DERs across Australia. Available

[10] K. Petrou et al, "Ensuring Distribution Network Integrity Using Dynamic Operating Limits for Prosumers," in IEEE Transactions on Smart Grid, vol. 12, no. 5, pp 3877-3888, September 2021. PDF available on ResearchGate

[11] M. Z. Liu, L. F. Ochoa, P. K. C. Wong and J. Theunissen, "Using OPF-Based Operating Envelopes to Facilitate Residential DER Services," in IEEE Transactions on Smart Grid, vol. 13, no. 6, pp. 4494-4504, Nov. 2022. PDF available on ResearchGate

[12] M. Z. Liu, L. F. Ochoa, T. Ting and J. Theunissen, "BOTTOM-UP SERVICES & NETWORK INTEGRITY: THE NEED FOR OPERATING ENVELOPES," CIRED 2021 - The 26th International Conference and Exhibition on Electricity Distribution, Online Conference, 2021, pp. 1944-1948. PDF available on ResearchGate

[13] M Liu, V. Bassi, L. F. Ochoa, T. Alpcan, C. Leckie, P. Wong, K. Chew, and J. Theunissen, “From Model-Driven to Model-Free: Comparisons Using Real Smart Meter Data and Real Distribution Network from Australia,” International CIGRE Symposium, Cairns, Australia, 2023

[14] M. Z. Liu, and L. F. Ochoa, “Project EDGE - Deliverable 1.1: Operating Envelopes Calculation Architecture”, PDF available on ResearchGate

[15] V. Bassi, L. F. Ochoa, T. Alpcan and C. Leckie, "Electrical Model-Free Voltage Calculations Using Neural Networks and Smart Meter Data," in IEEE Transactions on Smart Grid, vol. 14, no. 4, pp. 3271-3282, July 2023. PDF Available on ResearchGate

[16] Essential Services Commission, “Electricity Distribution Code”, January 2021, Available

[17] A. Simonovska and L. F. Ochoa, “Using Smart Meter Data for Phase Grouping in PV-Rich LV Feeders,” 26th International Conference on Electricity Distribution CIRED 2021, Geneva, Switzerland, 2021, ResearchGate

[18] A. Simonovska and L. F. Ochoa, "Phase Grouping in PV-Rich LV Feeders: Smart Meter Data and Unconstrained k-Means," *2021 IEEE Madrid PowerTech*, Madrid, Spain, 2021, pp. 1-6, ResearchGate

[19] E. Karunarathne, A. Simonovska, L. F. Ochoa, and Tansu Alpcan, “A GMM-Based Phase Group Identification for Residential Low Voltage Networks,” IEEE General Meeting, Orlando, Florida, USA, 2023, ResearchGate