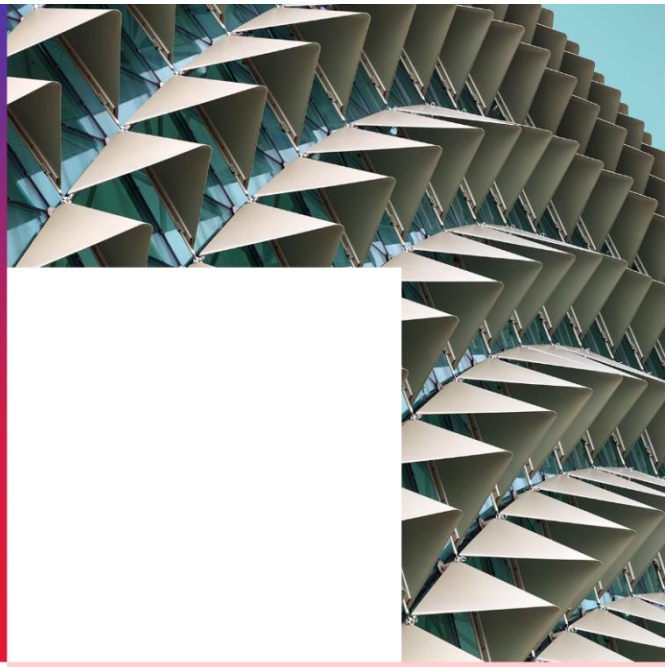


MPAN to Feeder Association Analysis Report (WP5- D3)

SMITN

31/08/2023

Prepared by

Iro Psarra, Nadim Al-Hariri, Alex Gardner

Amendment History

Date	Issue	Status
23/02/2022	0.1	Initial version
29/03/2023	0.2	Revised version
31/03/2023	2.0	Final version
11/05/2023	3.0	Minor update
23/06/2023	4.0	Minor update
31/0/2023	4.1	Minor Update – Clock offset clarification

Contents

1 Introduction

1.1 Background

Smart Meter (SM) data opens up opportunities for Distribution Network Operators (DNOs) to improve their database records and develop a low-voltage (LV) model which will be useful for network planning, fault detection and phase balancing. The adoption of low carbon technologies, such as photovoltaic panels or electric vehicles, is expected to significantly increase in the near future. For this reason, it is important for DNOs to understand the impacts that these technologies may have, particularly, on LV networks.

The SMITN (Smart Meter Innovations and Test Network) project is an innovation project that will investigate how the use of SM data can be used to support network operations. The fourth use case, “Feeder Allocation”, is intended to resolve anomalies in the CROWN or Electric Office records whereby some MPANs may be recorded as having an incorrect distribution substation (SS) or LV feeder. Previous projects, such as the Losses Investigation, have highlighted the issues with the data such as customers being associated with the wrong LV feeder or sometimes the wrong distribution substation.

Data quality and completeness of the LV networks is increasingly important to support planning and operations. Incorrect property assignments to the wrong substation or feeder may have a significant impact on network operations such as in load profiles estimation and the SM aggregation groups derived from CROWN will become unreliable.

Our analysis is focused on 46 secondary substations (SS) in the Milton Keynes area with installed GridKey monitoring devices. Anomalies identified by the SMITN algorithms will therefore be validated against the results of a detailed survey using the HAYSYS Feeder Finder Unit which has been developed in this project.

1.2 Document Purpose

This document presents a high-level description of the algorithms selected for the “Feeder Allocation” use case and will present and discuss the results, using validation data from HAYSYS.

The purpose of this analysis is to:

- detect outlier customers who appear not to be connected to the SS shown in the CROWN database,
- detect outlier customers who appears not to be connected to the LV Feeder shown in the CROWN database,
- propose new SS and LV feeder for each of the identified customer outliers,
- assess the performance of the selected algorithms,
- assess the CROWN MPAN to Feeder connectivity data,
- assess the EO MPAN to Feeder connectivity data.

This will provide useful information to enable National Grid Electricity Distribution (NGED) and other DNOs to implement Business As Usual (BAU) processes to validate existing LV connectivity records and backfilling missing data.

1.3 Overview

Section 2 of this document presents an overview of the SMITN selected algorithms and a breakdown into its different variations.

Section 3 presents the data pre-processing and the summary results of the selected algorithms.

Section 4 discusses the key learning points of this analysis.

1.4 Abbreviations

Term	Definition
BAU	Business As Usual
CB	Circuit Breaker
CIM	(IEC 61970/61968/62325) Common Information Model for the electricity industry.
CROWN	NGED's Enterprise Asset Management system.
CSV	Comma-Separated Value
DCC	Data Collection Company (for SM data)
DD	Data Dictionary
DER	Distributed Energy Resource
DMS	Distribution Management System (such as GE PowerOn)
DNO	Distribution Network Operator
EAC	Estimated Annual Consumption
EAM	Enterprise Asset Management (such as ARM, SAP or Oracle)
EO	Electric Office (NGED's GIS)
ERD	Entity-Relationship Diagram
ETL	Extract, Transform and Load
IEC	International Electrotechnical Commission
GIS	Geographic Information System (such as ESRI or GE Electric Office/Smallworld)
GUID	Globally Unique Identifier
HH	Half-Hourly
HV	High Voltage
LV	Low Voltage
INM	Integrated Network Model
JSON	JavaScript Object Notation
LCT	Low-Carbon Technology e.g. heat pumps, electric vehicles or photovoltaic generation
MAPE	Mean Average Percentage Error
MC	Measurements Calculator
MDM	Master Data Management
MPAN	Meter Point Administration Number (core – the 13-digit format)
NGED	National Grid Electricity Distribution
NHH	Non-Half-Hourly
NOP	Normally Open Point
ODS	Operational Data Store
OGC	Open Geospatial Consortium
RDBMS	Relational Database Management System
RMS	Root Mean Square (= $\sqrt{\frac{1}{n}\sum x^2}$)
RMU	Ring Main Unit
SCADA	Supervisory Control and Data Acquisition
SM	Smart Meter
SMITN	Smart Meter Innovations and Test Network
SS	Secondary Substation
SQL	Structured Query Language
TK	Unique INM record identifier
Tx	Transformer

Term	Definition
URI	Uniform Resource Identifier
UUID	Universally Unique Identifier
VIPQ	Voltage, Current, Active and Reactive Power
WKT	Well-Known Text (an OGC spatial geometry representation format)
XML	Extensible Markup Language

2 Feeder Allocation Data Usage

2.1 Trial area

The core trial area comprises 46 SS with installed GridKey monitoring devices from the Milton Keynes area.

2.2 Data Requirements

The following datasets were used for the feeder allocation use case:

1. Smart meter voltage data, with 1-minute time resolution.
2. Smart meter voltage data, with 30-minute time resolution.
3. Distribution substation voltage monitoring data, from GridKey loggers (1-minute resolution).
4. Premises to distribution substations connectivity data from CROWN.
5. Electric Office mapping data to provide connection phases of customers where this is already known.
6. Electric Office network data.
7. Feeder Validation data from HAYSYS.

2.3 Pre-Calculation Validity checks

2.3.1 Voltage Data

1. Data Anomalies

Low SM and GridKey voltage readings (e.g. 0V) exist in the time-series voltage data due to reading failures. The presence of very low values may have a negative effect in the model, for this reason records with unreasonable voltage readings were assessed and removed or corrected.

2. Clock Offsets

Another issue that may impact the results is the clock offsets between the SM voltage data, i.e. a discrepancy in voltage data time logging between the SMs. This issue is particularly present in voltage data with lower resolution (i.e., average 1-minute voltage data). An algorithm that identifies the offset difference between the SM and GridKey was created and applied. The aim is to synchronise the voltage curves.

3. Measurement Interval Synchronisation

The half hourly (HH) voltage readings from SMs use intervals with an arbitrary starting time within clock half hours. Similarly, voltage data with higher time resolutions (i.e. 1 minute voltage data) may start in the middle of a clock minute. Time synchronisation is important as there is a direct comparison of the voltages from SMs and GridKey for each time period.

As a result, higher-resolution voltage data provides not only an improved visibility of the short-term voltage variations, but also a reduction in the time offsets between the measurement intervals used by each SM.

Two approaches have been developed to deal with this issue:

- Timestamp rounding

For the 1-minute voltage data, the timestamp seconds have been rounded to the closest minute, while for the HH data, the minutes have been rounded to the closest 30-minute period.

- Linear Interpolation

Linear interpolation is helpful when searching for a value between a given set of points. It is a method to construct new data points within the range of a discrete set of known data points. In other words, Linear Interpolation allows us to estimate voltage values at times that we do not have them at, given that we have readings soon before and soon after that time.

2.3.2 GridKey and SM Load Measurements

2.3.2.1 SM Errors

There are occasional errors in the collected kW values from smart meter aggregation groups. These are identified and recorded for each feeder/substation-day so that those feeder/substation-days can be excluded from any analysis.

Smart Meter data errors are identified as a day:

- with any period with a value greater than 500kW,
- average kW for a day is less than -10kW,
- average kW for a day is greater than 500kW

2.3.2.2 GridKey Errors

GridKey Feeder measurements are provided at a time granularity of 1 minute. These are averaged to create 30 min data (HH period).

GridKey errors are identified for any day for the feeder/substation:

- where there are less than 20 minutes in any HH period
- where there are less than 1360 minutes in a day
- Where the measurement is less than 1 W and greater than -1 W (this identifies feeder measurements for a feeder where fuses are removed)

2.4 Feeder Allocation Approaches

2.4.1 Results from Phase Identification

2.4.1.1 Outliers from Phase Identification

In the Phase Identification use case, we initially assumed that the connectivity between premises and distribution SSs obtained by CROWN was correct. However, during the SMITN analysis it became clear that there are errors in the mapping between customers and LV feeders and between customers and SS records in CROWN.

In the “Feeder Allocation” use case, we will use the results obtained from the “Phase Identification” use case. It was known from previous work by Scottish and Southern Electricity Networks and Scottish Power Energy Networks that meters with very low correlation results may be fed from a different substation.

In the first approach (see “Phase Identification” report), the correlation used time-series voltage data from a single-phase SM paired with SS monitoring data. For each single-phase SM, three correlation results were obtained, one for each phase measurement at the substation. Outliers were identified and tested to see if they belong to other distribution substations using topology proximity.

Ideally, the outliers that are identified would be tested against their neighbour SS, but this is out of the scope of this project because of the absence of SS monitoring data and SM data outside of the core area.

In the second approach, “clustering SMs into 3 groups”, the proposed clustering algorithm clusters the SM into three groups, one for each phase. However, meters that are fed from different substations may be assigned to the wrong cluster. Outliers from the clustering techniques will be identified and pinpointed for review.

It is noted that this approach is only effective when identifying meters that are fed from a different substation, while it is not able to distinguish the meters that fed from the same substation but have been assigned to the wrong feeder.

2.4.1.2 Correlation with neighbouring SS

During the SMITN analysis we identified that properties could have a strong voltage correlation result with the GridKey data of a wrongly assigned SS in cases where the two SS are fed from the same primary substation. In this case, even if the properties are fed from the neighbour SS, the previous algorithm will not be able to identify them as outliers.

In this case, the algorithm repeats the “Phase Identification” approaches using the data from both examined SS if they are available in the core area. Then, the SMs are assigned to the SS which have higher correlation results.

2.4.2 Clustering SMs into N groups

This approach is focused on identifying SMs that have been assigned to the wrong feeder. Prior knowledge of phase information is needed for this approach. For this reason, the results from “Phase Identification” will be used to assign a phase to each SM.

The algorithm is outlined as follows:

1. Calculation of voltage correlation between each pair of SMs. This is a matrix of dimensions N_mxN_m, where N_m is the number of SMs in the analysis.

2. For each phase, clustering the SMs into N_f groups using unsupervised machine learning – clustering techniques, where N_f is the number of feeders in each substation.
3. Repeat steps 1 and 2 using multiple variations of the input data (i.e., magnitude of voltage time-series, voltage step changes) and assess their performance.

Hierarchical clustering was applied, as it was proven to be the most effective clustering technique in the “Phase Identification” use case. Moreover, we have tested again the algorithm’s performance of using the time-series voltage data and the step changes. Finally, we tested the algorithm’s sensitivity to SM coverage.

2.4.3 Voltage correlation with aggregated feeder demand

In this approach, feeders are identified by correlating the smart meter voltage drops with the aggregated load on each feeder. This aggregated load data is provided by substation monitoring.

The algorithm is outlined as follows:

1. For each phase, the algorithm subtracts each SM voltage curve from the GridKey voltage reference.

2. It correlates the voltage differences with the aggregated load on each feeder.

3. For each SM, the algorithm assigns the feeder with the strongest correlation as the predicted feeder.

This approach requires accurate phase information. The results from “Phase Identification” use case were used.

2.4.4 Discrepancies between EO and CROWN

The network topology and geospatial data will be obtained from EO. In EO, meter details have been added recently in many places in the LV network topology but there are still many places where meter information is not available.

During the SMITN analysis, we identified that there are discrepancies between CROWN and EO data. To store the data from EO and compare them with CROWN, we used CGI’s INM (Integrated Network Model) data model which can be utilised to enhance and improve the LV model.

In areas where meter information has been populated in EO as well as their service connections to the main cables, we were able to assign substation and feeder number to each meter using the information that exist in the EO cables (“connectivity approach”). The feeder and the substation number for each meter which obtained using the “connectivity” approach will be compared with the CROWN data.

In areas where the meter information is not available in EO, we used the meter coordinates from CROWN data. A distance function was used to find the distance for each meter to its closest LV feeder. Each meter will be assigned a feeder using the “proximity” approach, and this information will be compared with the CROWN data.

2.4.5 Feeder Profiles Outliers

In this approach, we will use the initial results from “Feeder Profiles” use case. The Feeder Profiles use case calculates an estimate of the Feeder load and then compare this with the measured data at the SS. Different algorithms are proposed to calculate the feeder load (see “Profile Planning” report). In this report we will use the NHH1 estimation algorithm. (see Appendix A.8).

It has been confirmed that the algorithms used in “Feeder Profiles” use case are sensitive to customer-to-feeder connectivity. Wrong customer assignments lead to significant errors in feeder load estimation. The

proposed algorithm identified feeder allocation errors by comparing the difference between expected demand and the measured demand. First, the load in SS and feeder level is estimated using the CROWN connectivity data. Subsequently, the feeder load was estimated again using EO data (see section 2.4.3). The daily Mean Absolute Percentage Error (MAPE) is calculated for each feeder (see Appendix A.7). The purpose of this approach is to identify feeders with big EO and CROWN discrepancies and assess the connectivity in two systems using the results from the load profiles.

2.4.6 Validation Data

2.4.6.1 HAYSYS

Identifying the Feeder cables connected to the substation to be surveyed is achieved through the use of the HAYSYS Feeder Finder. The Feeder Finder achieves this identification by placing its current injector unit which induces a current into each of the substation Feeder neutral cores (typically the armouring of the Feeder cable). These currents are at a frequency of 5MHz and modulated with a specific code, that is different for each of the substation feeders being surveyed. The Feeder Finder can inject into up to four feeders. The Feeder Finder detector is a mobile device, that is battery powered and used to detect the magnetic fields surrounding the feeders as a result of the currents injected at the substation. The detector differentiates between the magnetic fields through the identification of the codes modulated onto the 5MHz injected signals. Through the Feeder Finder display software, the codes are processed and the measured feeder value determined, displayed and recorded.

3 Results

3.1 Data Preparation

3.1.1 Core Area

In the core area of the SMITN project, there are 8809 properties, from which 47% of them have installed a SM up to the end of November 2022. This is slightly higher than the overall average for NGED, which is nearer 40%, as substations with a higher proportion of smart meters were preferentially selected for inclusion in the SMITN test network.

Initially, HH voltage data was requested from the devices for the period 1st May 2022 until 31st August. Around 2200 devices gave a response, which gives a success rate of 57%. A request was then issued to all devices to amend the Average Voltage Measurement Period to 1 minute from the default of 30 minutes. 1810 devices gave a response. 1 minute voltage data was requested and from the beginning of October until 30th November 2022.

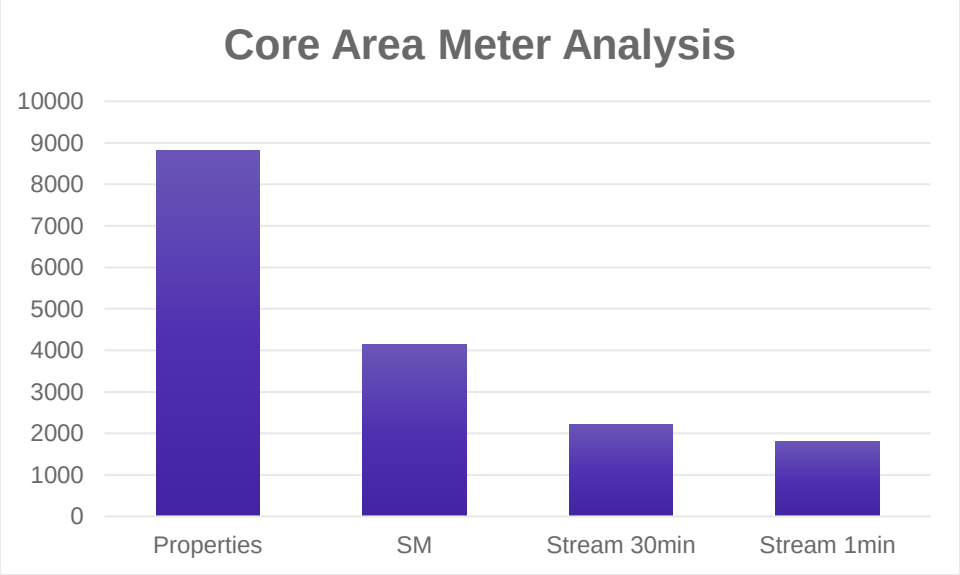


Figure 1: Core Area Device Analysis

3.1.2 Sample Voltage Data

Figures 2 and 3 show an example of SM time-series voltage data and the time-series SS voltage data with 1 minute and HH resolution. The SS monitoring data with one minute time resolution is post-processed to derive HH samples that are aligned with the timing of the 30-minute voltages of SM's.

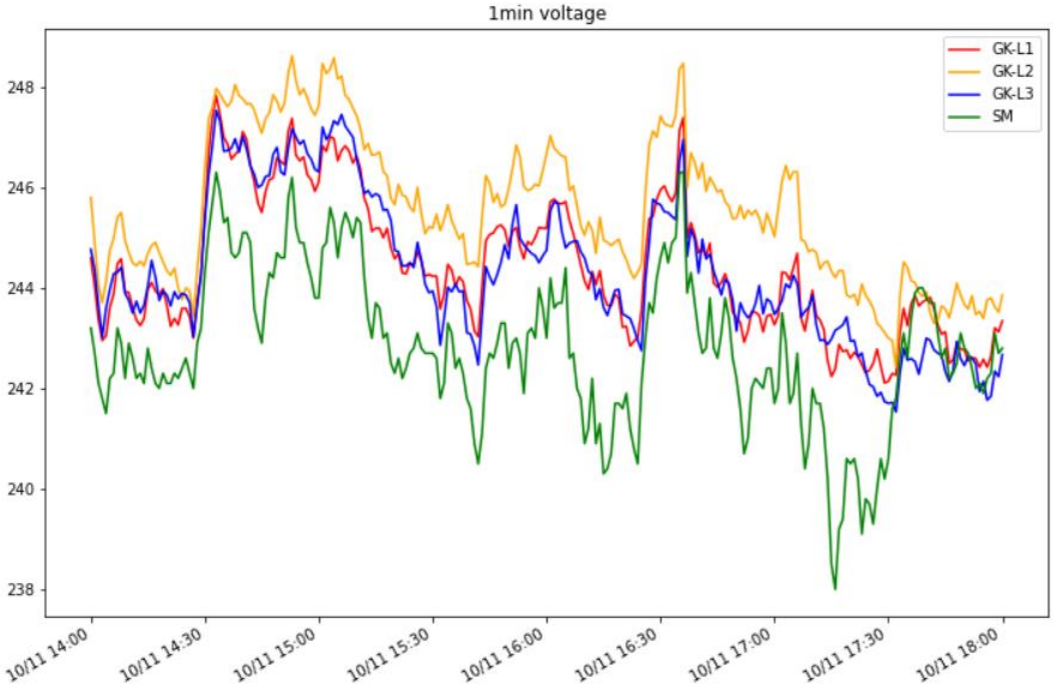


Figure 2:Sample 1-minute voltage data from a SM and its Distribution Substation

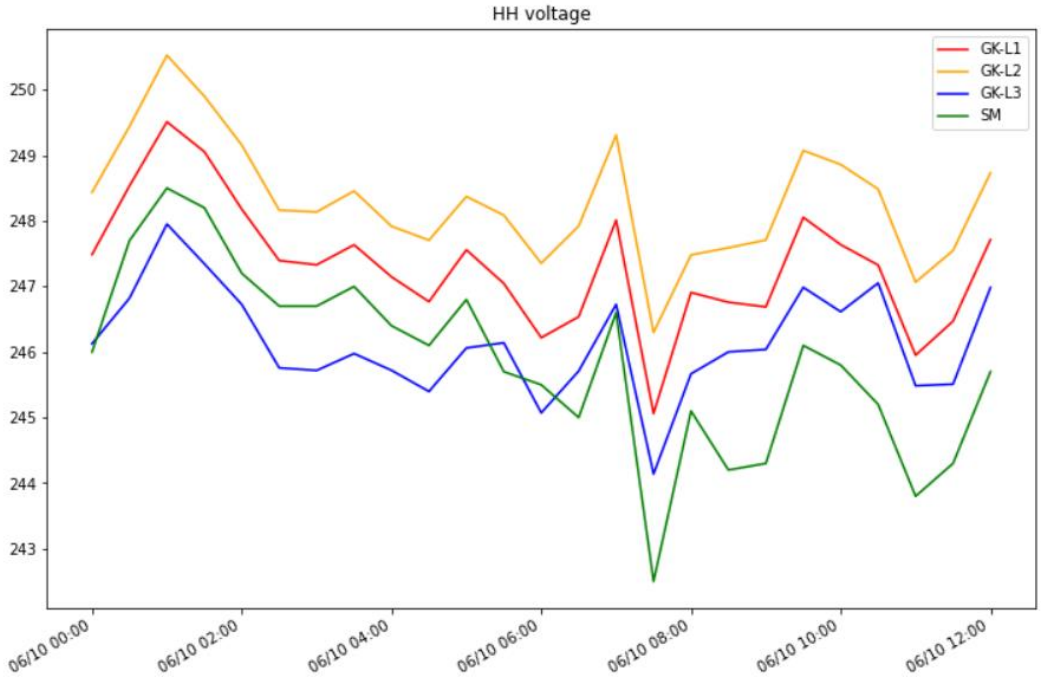


Figure 3: Sample HH voltage data from a SM and its Distribution Substation

3.1.3 Data Cleansing and Meter Delays

Analysis has been undertaken to identify anomalies in the time-series voltage data. In particular, unusually low voltage readings in both 1-minute and 30-minute data. These data quality issues could degrade the performance of the analysis and to the model, so data cleansing was required. Voltage readings lower than 200 V were identified and removed.

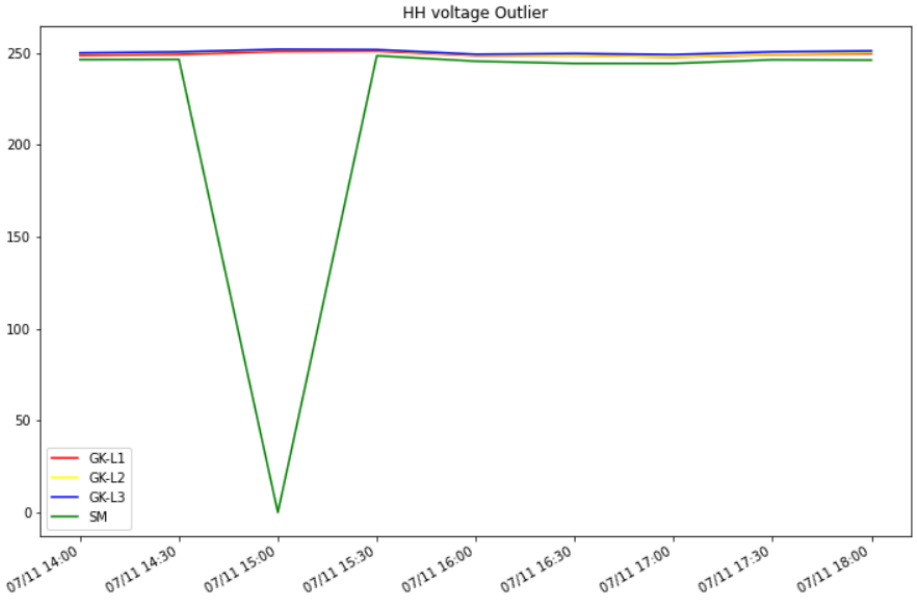


Figure 4:Sample "0" reading in SM data

A second data issue that was identified was SM voltage reading clock offsets relative to substation logger. The timing of the smart meter voltage data was generally well-aligned to the substation monitoring data. However, 56 SMs out of 1810 found to have a clock synchronisation issue relative to substation logger. It is unclear from this data if the offset is introduced from the SM or from the substation logger.

While the voltages from GridKey and the voltages from the SM's have the same timestamps, there are a number of SM's whose voltage profile curves lead the secondary SS's voltage, and others which lag significantly. The presence of this issue is more often in higher resolution data, in our case in 1-minute data. Figure 5 shows an example of a SM voltage having a 5-minute offset. The graph shows a period of 1 hour.

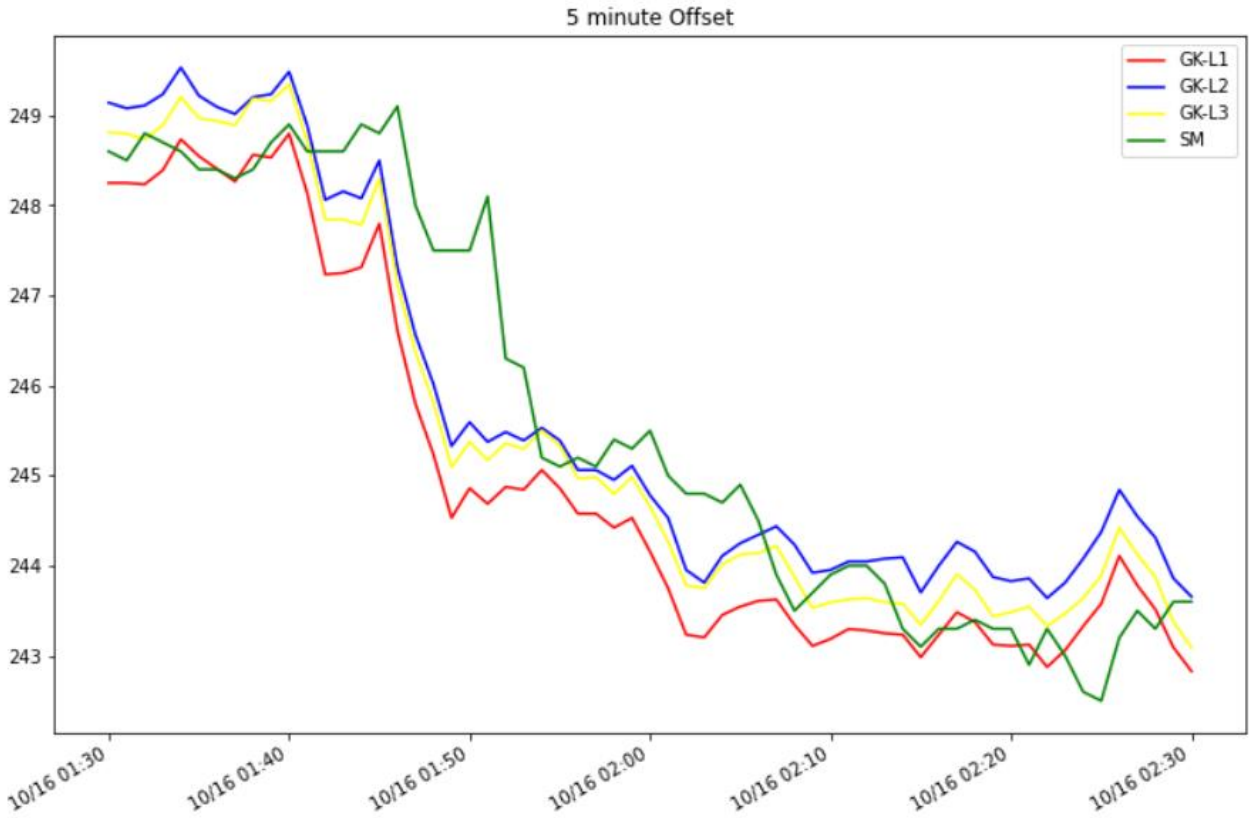


Figure 5: A SM with 5-minute offset relative to substation logger

To deal with this issue, an algorithm was created to identify potential offsets. The offset checking algorithm checks offsets between -5 and +5 minutes with a 15 second resolution. If the algorithm doesn't find an improvement in voltage correlation by a threshold in that range, then it does a broader search, checking between -60 and +60 minutes with a resolution of 60 seconds. The voltage correlation is calculated by using the magnitude of voltage step changes. If the correlation result was improved by more than a set threshold (i.e. 0.1), then the offset was applied to the voltage data. The voltage curve is then corrected. 56 SM devices were identified to have a clock offset relative to substation logger. Table 1 presents the recommended minute offset from the algorithm for sample devices and Figure 6 shows the distribution of the clock offsets for the 56 SMs.

Substation	Device	Approximate minute offset (m)
942757	Device 1	1.75
942086	Device 2	7
945237	Device 3	-0.75
942756	Device 4	1
945113	Device 5	2.25
942756	Device 6	-0.75
942774	Device 7	-2.5
942774	Device 8	-0.75

Table 1:Sample devices with the recommended minute offset

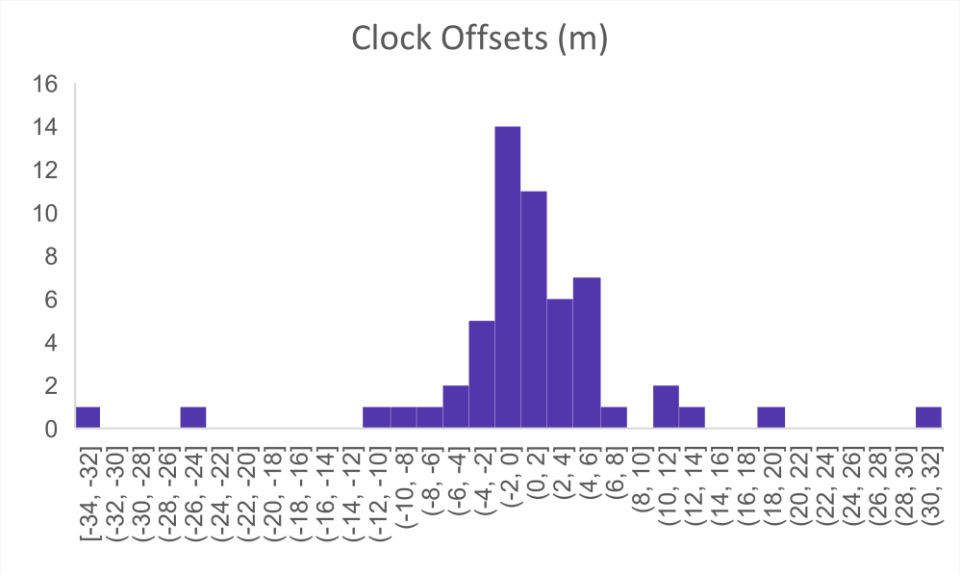


Figure 6:Distribution of clock offsets relative to substation logger (for the 56 SM)

3.2 Results from Phase Identification

3.2.1 "Outliers from Phase Identification"

3.2.1.1 Direct correlation between SM and GridKey

The voltage correlation between the SMs and the GridKey data was expected to be high. However, there are many factors that could impact the correlation results, such as the time synchronisation between meters, the frequency of sampling etc. Moreover, low correlation results between the SMs and GridKey may suggest that these SMs have been assigned to a wrong SS.

To apply the proposed algorithm, each property needs to have:

- Available time-series voltage data,
- Available GridKey time-series voltage data.

Figure 7 illustrates the distribution of the correlation results using as input the time-series of the voltage step changes (step changes) as this was proved to be the most efficient approach in the "Phase Identification" report. We can observe that there are a number of SMs where the correlation values are very low and these have been identified as outliers and pointed out for review. A reasonable threshold is around 0.2 and 0.3, but the threshold needs further investigation when the results from the HAYSYS feeder survey will become available.

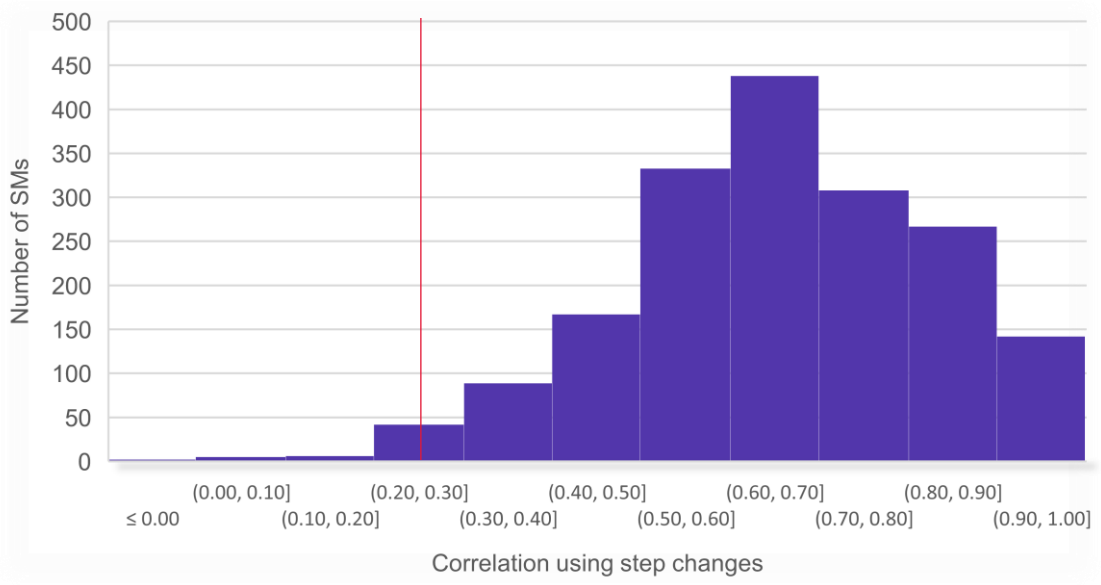


Figure 7: Distribution of correlation results using voltage step changes

Figure 8 shows an example of low correlation results in “942756” SS. The SMs with a yellow colour have been identified by the algorithm as outliers. EO network data has suggested that these SMs are fed from the neighbour substation ‘942351’.

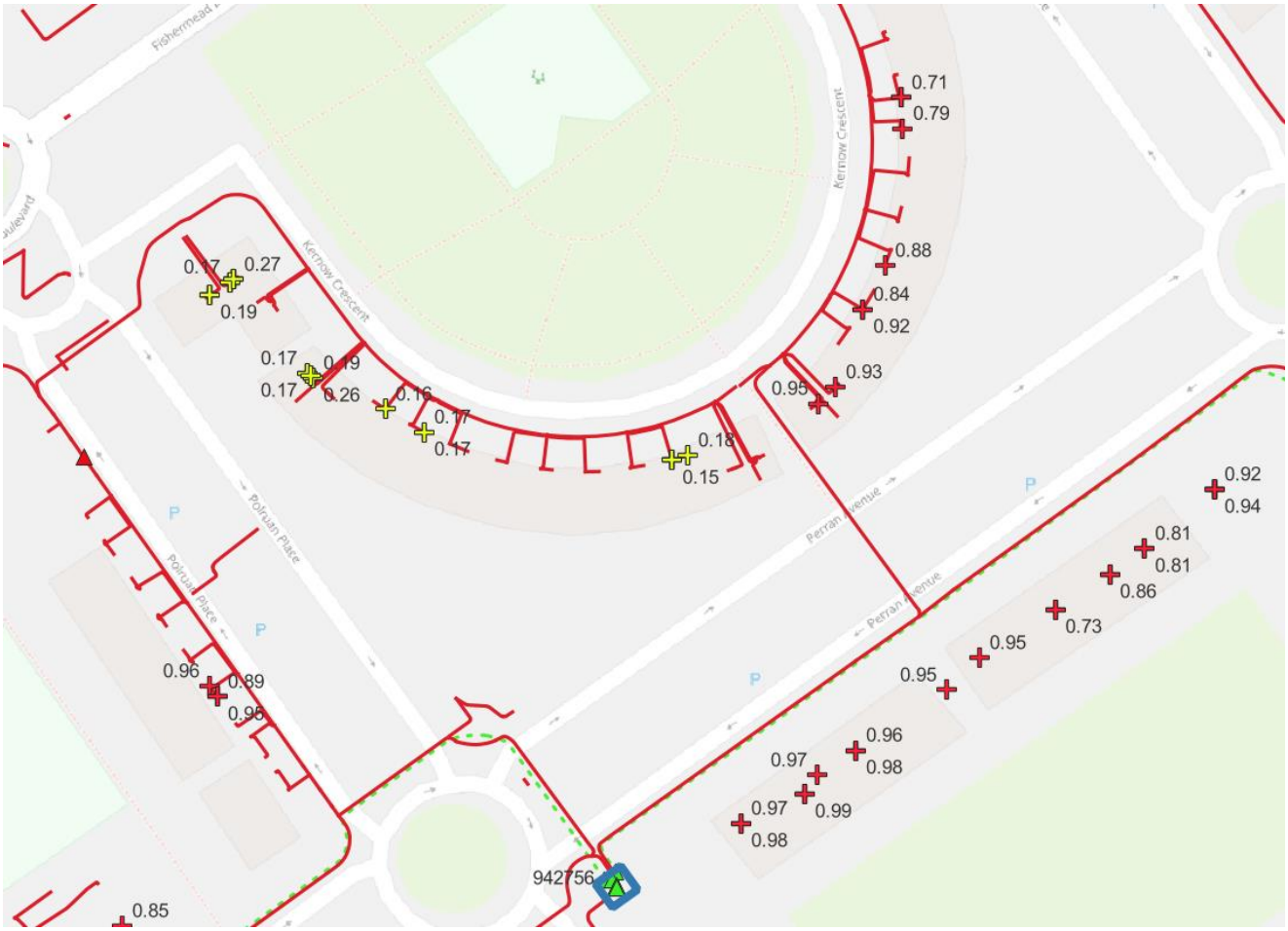


Figure 8: Outliers from Phase Identification for ‘942756’

3.2.2 Correlation with neighbouring SS

In the case of where the examined SS and the neighbouring SS are fed from the same primary feeder, the correlation may be high even if the property has been assigned to the wrong SS. Figure 9 illustrates an example of two neighbouring substations (‘945487’, ‘942840’) fed from the same HV feeder (‘940045/06’). The feeders from each substation have been colour-coded, with blue the feeders from ‘945487’ SS, with red the feeders from ‘942840’ SS and with green the HV feeder from the primary substation ‘940045’. All the properties in this figure (23,24,25,26,27,28,29,30) are fed from the substation ‘942840’ according to CROWN records while there is no information about these properties in EO data. Their correlation result with ‘942840’ GridKey data is very high (correlation results with red)

Figure 9: Wrong MPAN assignments between two SS which are fed from the same primary feeder

Investigating the EO LV network data and the distance of each SM to its closest LV feeder (see approach 3.4), suggests that the SMs 27,28,29,30 are fed from ‘945487’ SS. However, as their correlation results are very high, the algorithm discussed in the previous section will not be able to pick these SMs as outliers.

In these cases, the algorithm identified the SSs that are fed from the same primary feeder, and recalculated the correlation results using the GridKey from the neighbouring SS. If the correlation result increases, the algorithm identifies these properties as outliers and highlights them for review. In the absence of GridKey data, the process can work using SM voltage data from both SS in which we have high confidence that the allocation in CROWN data is correct.

In the example in Figure 9, the correlation has been re-calculated using the GridKey data from ‘945487’ SS. The correlation for SMs 27,28,29,30 has been increased (blue colour correlations) which indicates that the 4 SMs are fed from the ‘945487’ SS and the initial CROWN assignments are incorrect. In contrast, the correlation decreased for SMs 23,24,25,26 which means that the initial CROWN assignments are correct.

3.3 Clustering SMs into Feeder groups

To identify SMs that have been allocated to the wrong feeder, unsupervised machine learning – clustering techniques – will be used. For each phase, the algorithm will create N_f groups, where N_f is the number of feeders in each substation. The correct phase information is very important in this approach and the phase information from the phase identification use case will be used.

Figure 10 shows an example of the clustering results for “940337” SS and phase ‘L2’. On the right, there is a heatmap, which shows the correlation results for each pair of SM. The green colours represent the pairs of SMs that are highly correlated, while the red colours are the pairs of SMs with very weak correlation. In this figure, the SMs have been ordered based on the clustering group (feeder group) that the algorithm assigned them (1,2,3,4). We can observe, how the algorithm managed to create four distinguished clusters, one for each feeder.

To assign the actual feeder number to each cluster, the CROWN data was used. The label of each cluster is determined based on the greatest number of MPANs belonging to a feeder within the cluster. Here we can see that Device 8 and Device 7 have feeder numbers in CROWN that do not agree with the clustering results.

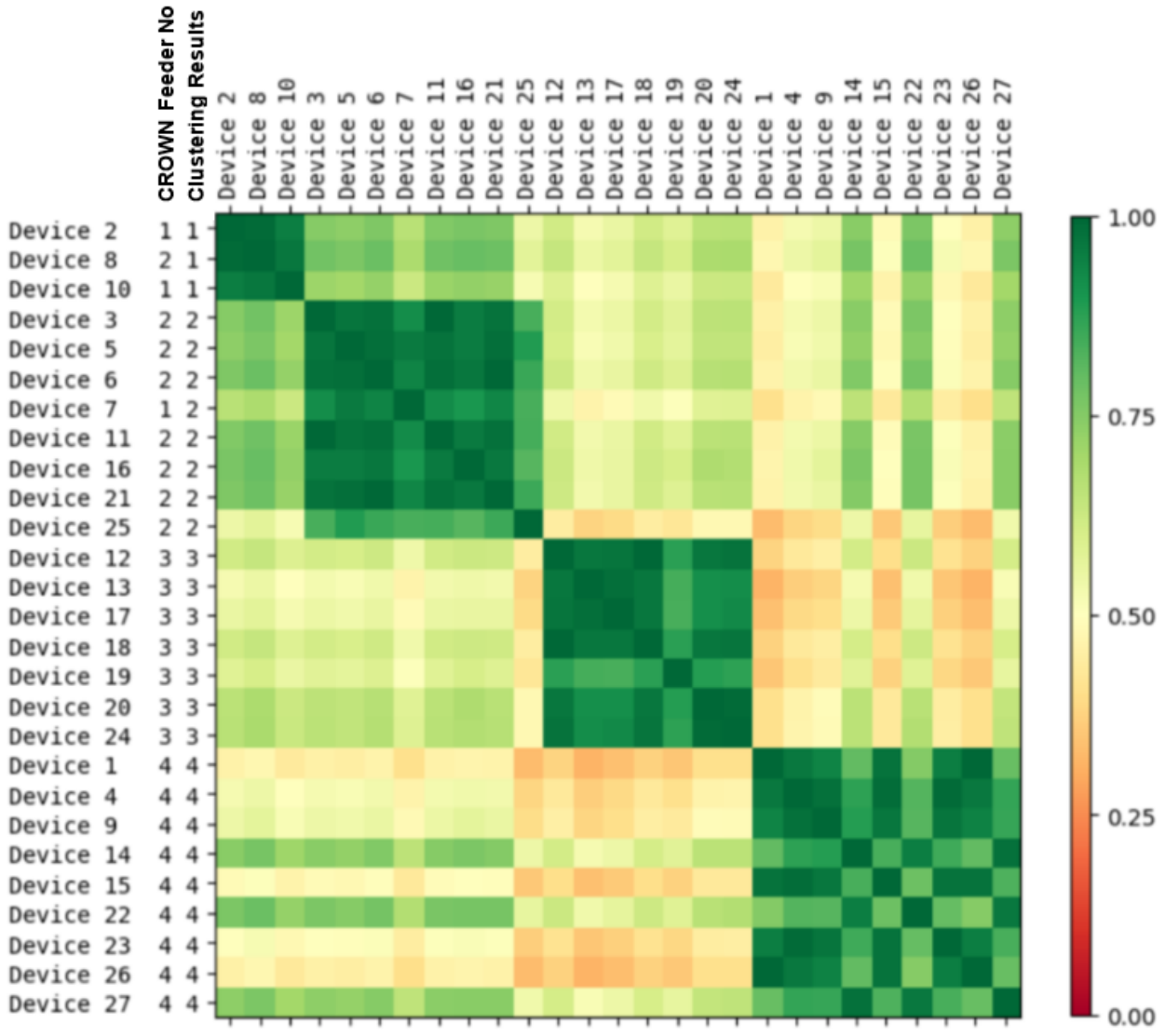


Figure 10: Correlation Heatmap and Clustering results for 940337 phase L2.

The same process repeats for phases L2 and L3. In Figure 11 we can observe the clustering results for the substation '940337'. We colour-coded the number of feeders. Green represents feeder no 1, pink represents feeder no 2, brown feeder no 3 and blue feeder no 4. The squares represent the properties and their colour represent their feeder number in CROWN records. On top of the squares, there are small circles, which

represent the clustering results. When the squares and circles have the same colour, it means that CROWN records and clustering results agree.

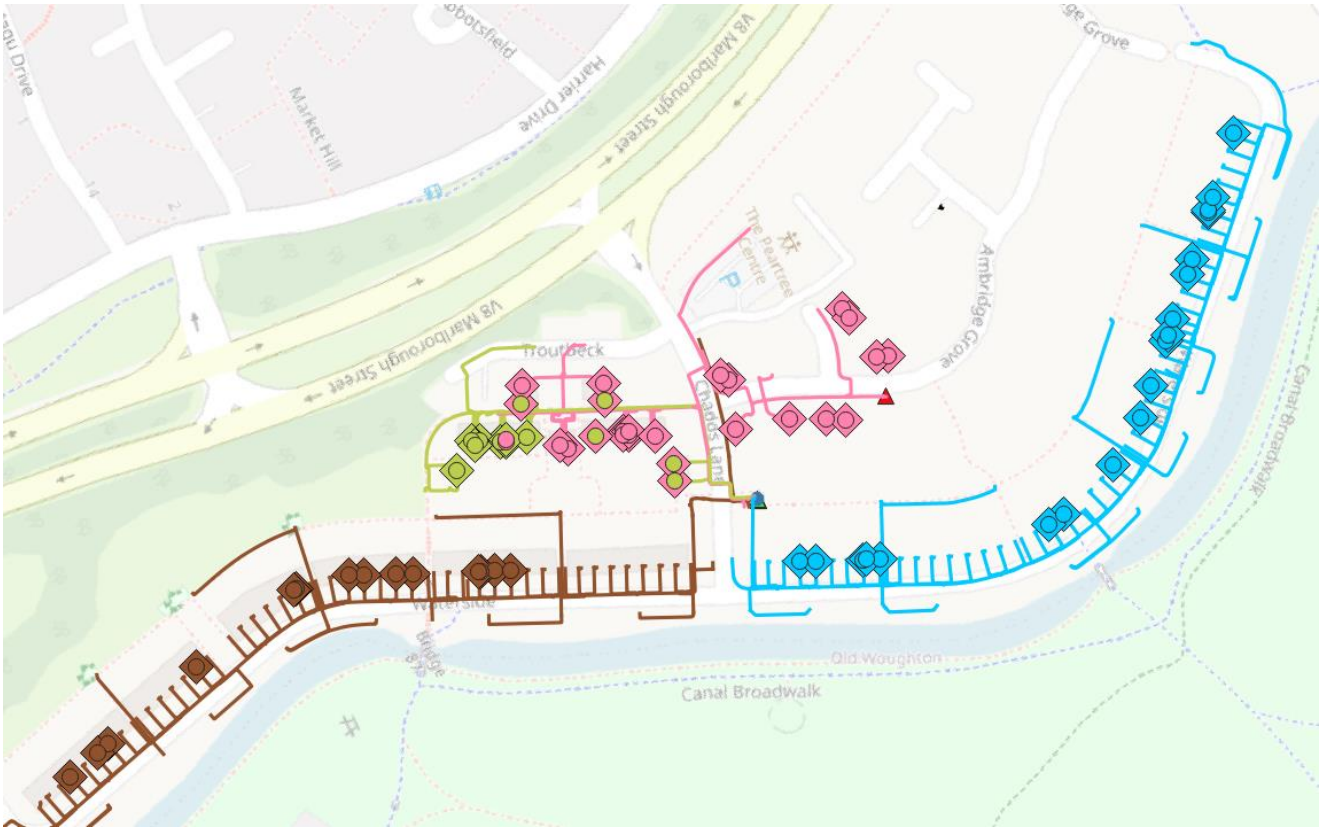


Figure 11: Clustering results for '940337'

The clustering managed to separate feeder 3 and feeder 4 into two separate clusters without any changes to CROWN records. However, feeders 1 and 2 are parallel feeders and the algorithm identified potential incorrect assignments in CROWN records (see Figure 11).

It is noted that SMs that are very close to the SS have very similar voltage curves and the result of the clustering is less reliable. For this reason, the algorithm excludes the SMs that are in a distance of 50 meters from the substation. Moreover, Silhouette Score is a metric used to calculate the quality of a clustering technique. Its value ranges from -1 to 1. 1 means cluster are well apart from each other and clearly distinguished. In order to increase the reliability of the algorithm, SMs with very low Silhouette Score (less than 0.3) has been removed from the analysis and not highlighted for review.

3.4 Voltage correlation with aggregated feeder demand

In this approach, we first calculated the voltage difference between each SM voltage curve and the GridKey voltage reference. Then the algorithm correlates the voltage differences with the aggregated load on each feeder. The algorithm assigns the feeder with the strongest correlation as the predicted feeder.

We then run the correlation algorithm on consecutive 4-hour periods. The feeder demand that correlates best with each property is recorded for each 4-hour block, and then we take the most common result for each property as the feeder assignment. We then calculate a confidence score as the percentage of times that this

correlation algorithm gave the same result as the most common result for that property over all these 4-hour blocks.

Figure 12 illustrates the results of this approach for the substation “940337” that we discussed in the previous section. This approach gave the same results with the previous approach for all the 4 feeders.

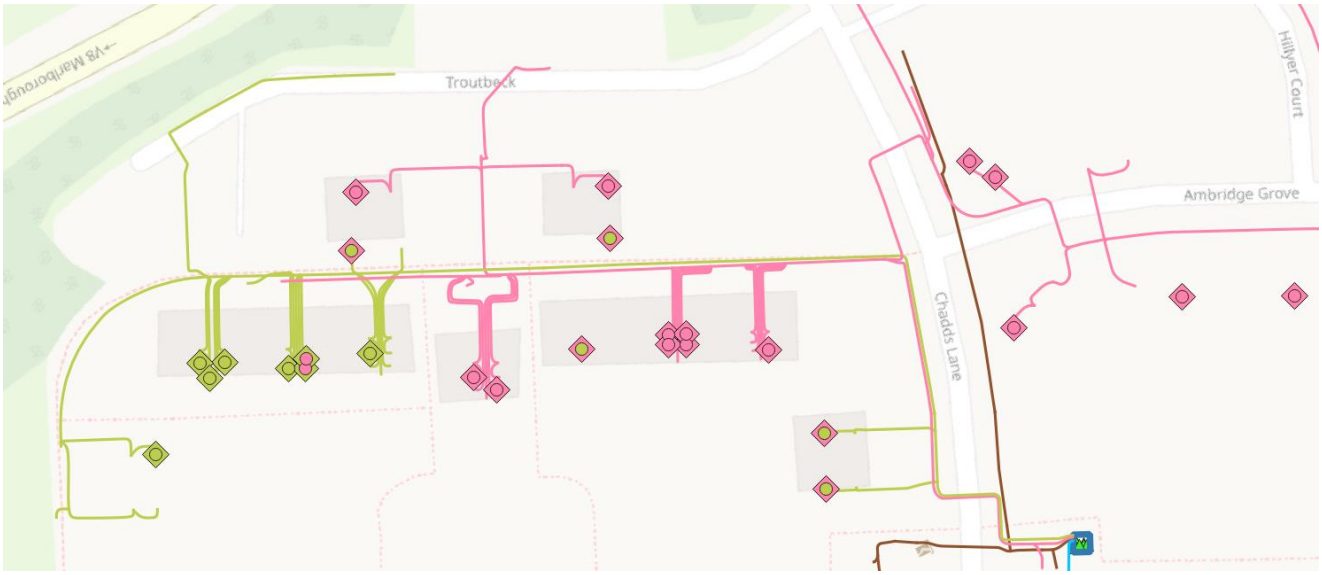


Figure 12: Wrong feeder assignments

3.5 Discrepancies between EO and CROWN

EO data is another source that can be used to identify the MPAN connectivity. To identify the feeder for each MPAN in EO, two methods were used:

1. Where the service connection and premise details exist in EO, the feeder information was obtained from the EO LV feeders with the “connectivity” approach.

2. Where the service connection and MPAN information does NOT exist in EO, a proximity algorithm is used to identify the closest LV feeder for each MPAN using the coordinates obtained from CROWN.

Then, the algorithm compares the information from CROWN and EO and flags the properties where the LV connectivity is different in the two systems.

We need to note that the “proximity” approach may not be representative of the correct MPAN to feeder connectivity in cases where parallel cables exist. For this reason, the algorithm excluded the properties that are close to parallel feeders from the analysis.

Figure 13 illustrates the discrepancies between EO and CROWN per substation. Out of 8809 MPANs, there are 958 MPANs with suspect feeder ids.

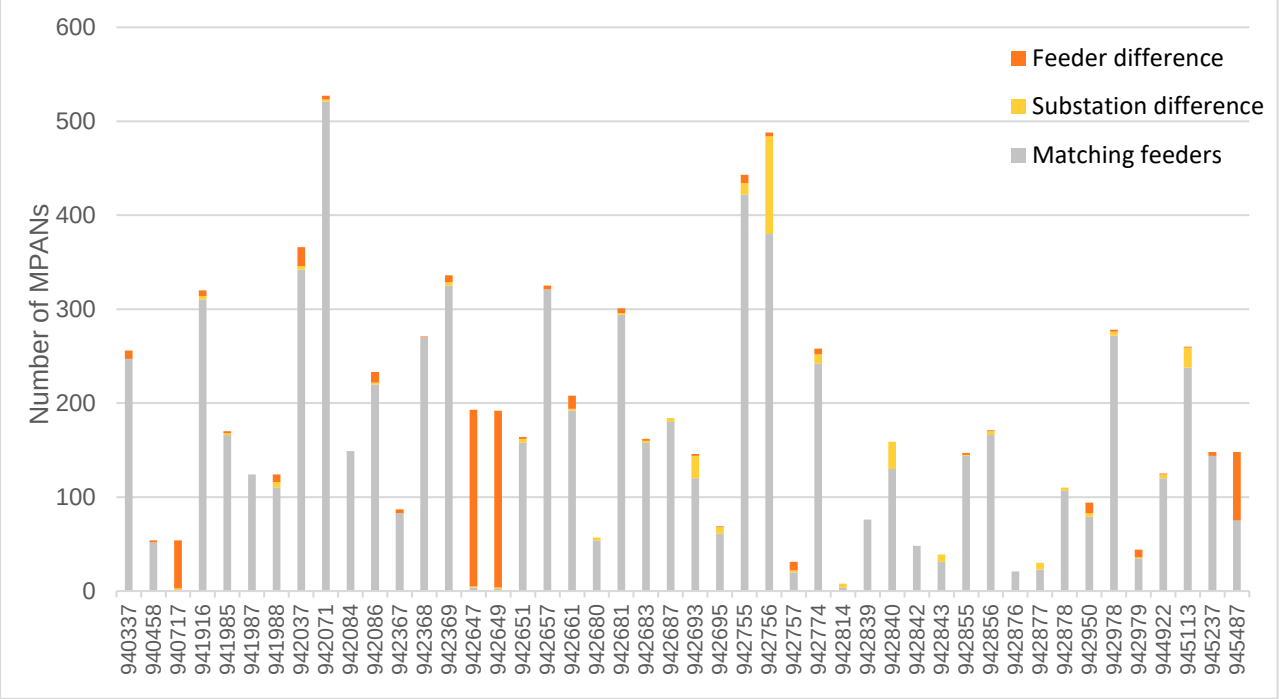


Figure 13: Discrepancies between EO and CROWN

There are various explanations for these differences that would not necessarily mean CROWN is wrong:

- The “proximity” approach can lead to wrong EO assignments as the MPAN might just be connected to a more distant feeder.
- The circuit ids are missing on some cables so a connection to the assigned feeder might not be found.
- The circuit id numbering in EO could be wrong, rather than the numbering in CROWN. The 940717, 942647, 942649, 945487 SSs have a numbering issue.

3.6 Feeder Profiles Outliers

In the “Feeder Profiles” use case, we have discussed that the results from the proposed algorithms are very sensitive to MPAN to Feeder connectivity. The algorithms used the connectivity data both from CROWN and EO and the results varied significantly. However, there were cases that the daily MAPE has decreased significantly when the algorithm used the connectivity data from EO instead of CROWN, but there were cases where CROWN data seems to be more accurate than EO. In this section, we will discuss how we can use the “Feeder Profile” NHH1 approach to validate the CROWN and EO discrepancies which will provide another way to validate the connectivity of the two sources.

3.6.1 Substation Changes

The calculation of the SS load profiles can be used to identify if the MPANs have been allocated to the wrong substation. The algorithm identified the substations where the daily MAPE is very low when we use the data from the one source, but the error increases significantly when the other source has been used. Table 2 includes the daily MAPE from the calculation of the SS profiles using CROWN and EO. The “Out” column

represent how many MPANs are recommended to move out from the substation based on EO data, while the "In" represents how many MPANs are recommended to be included in the substation based on EO data. Then, the algorithm Accepts or Rejects these transfers.

	MAPE		MPAN Transfers		
Substation	CROWN	EO	Out	In	Status
940458	5	46	3	1	Rejected
941988	11	5	6	0	Accepted
942037	3	10	4	0	Rejected
942647	30	4	3	4	Accepted
942680	31	11	3	1	Accepted
942756	33	3	107	0	Accepted
942774	17	9	11	0	Accepted
942814	18	46	4	0	Rejected
942840	19	5	31	0	Accepted
942855	7	47	1	2	Rejected
942877	13	23	8	0	Rejected
942950	20	55	5	3	Rejected
944922	9	17	4	2	Rejected

Table 2: MAPE using the connectivity data from CROWN and EO

An example of wrong SS assignments has been discussed in section 3.2.1 for the '942756' SS. A number of SMs have been identified as outliers from the voltage correlation. As the "SS Profiles" approach works for MPANs with and without SMs, this algorithm has identified 107 wrong assignments in CROWN data for the SS '942756'.

In figure 14, the LV feeders cables from '942756' SS have been coloured in red, while the feeders from the neighbouring SSs are in blue. The properties that are fed from '942756' based on CROWN data are the red dots. We can observe that there are three sections of the feeders where CROWN data looks inconsistent with EO data. From the SS profile estimation results, we can confirm that the use of the EO data decreased the

average percentage of the ME from 33% to 3%. The algorithm accepts the MPAN transfers and EO source labelled as correct.

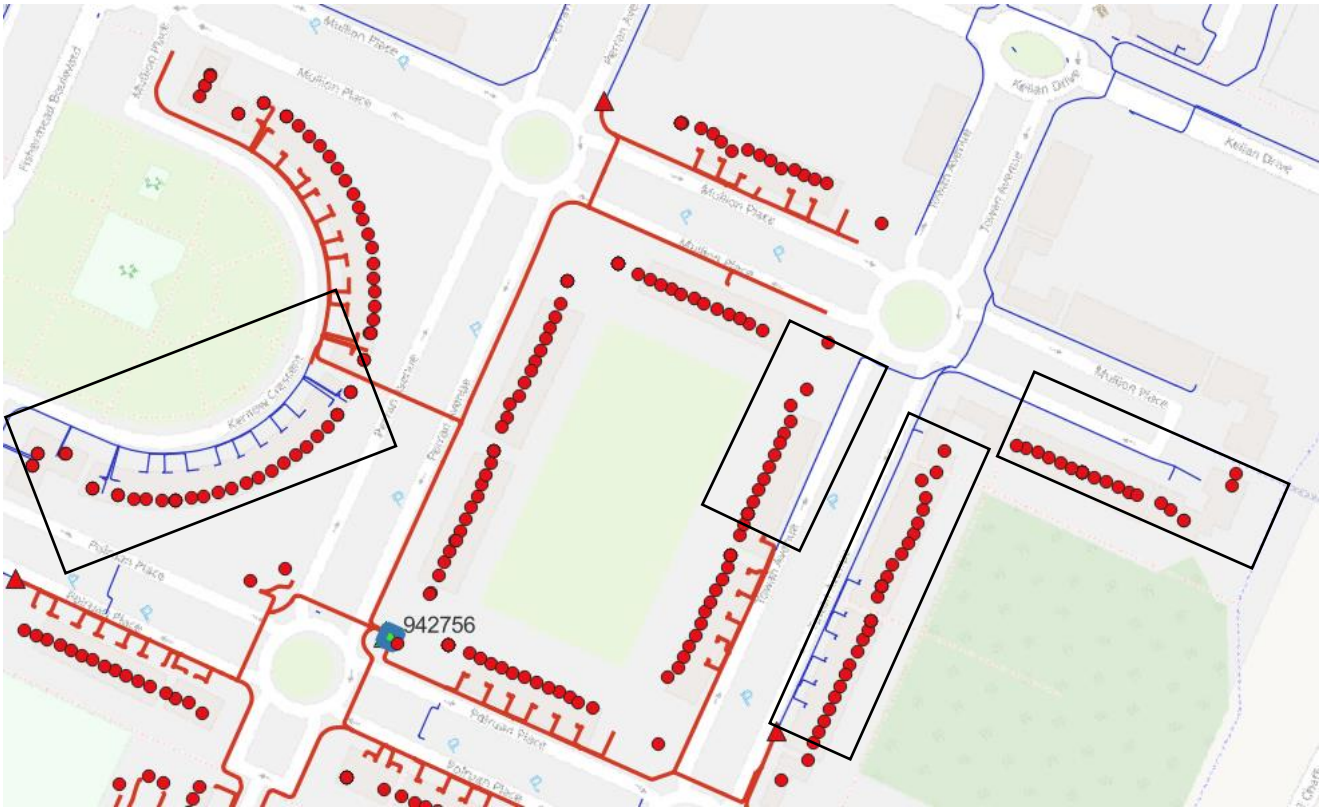


Figure 14: Wrong MPAN assignments in CROWN

3.6.1.1 Feeder Changes

Similar to the previous approach, the estimated feeder profiles can be used to identify wrong MPAN assignments on the feeder level. The algorithm compares the performance of the estimated feeder profiles using MPAN connectivity from CROWN and EO and accepts or rejects the initial proposed reassignments in CROWN data.

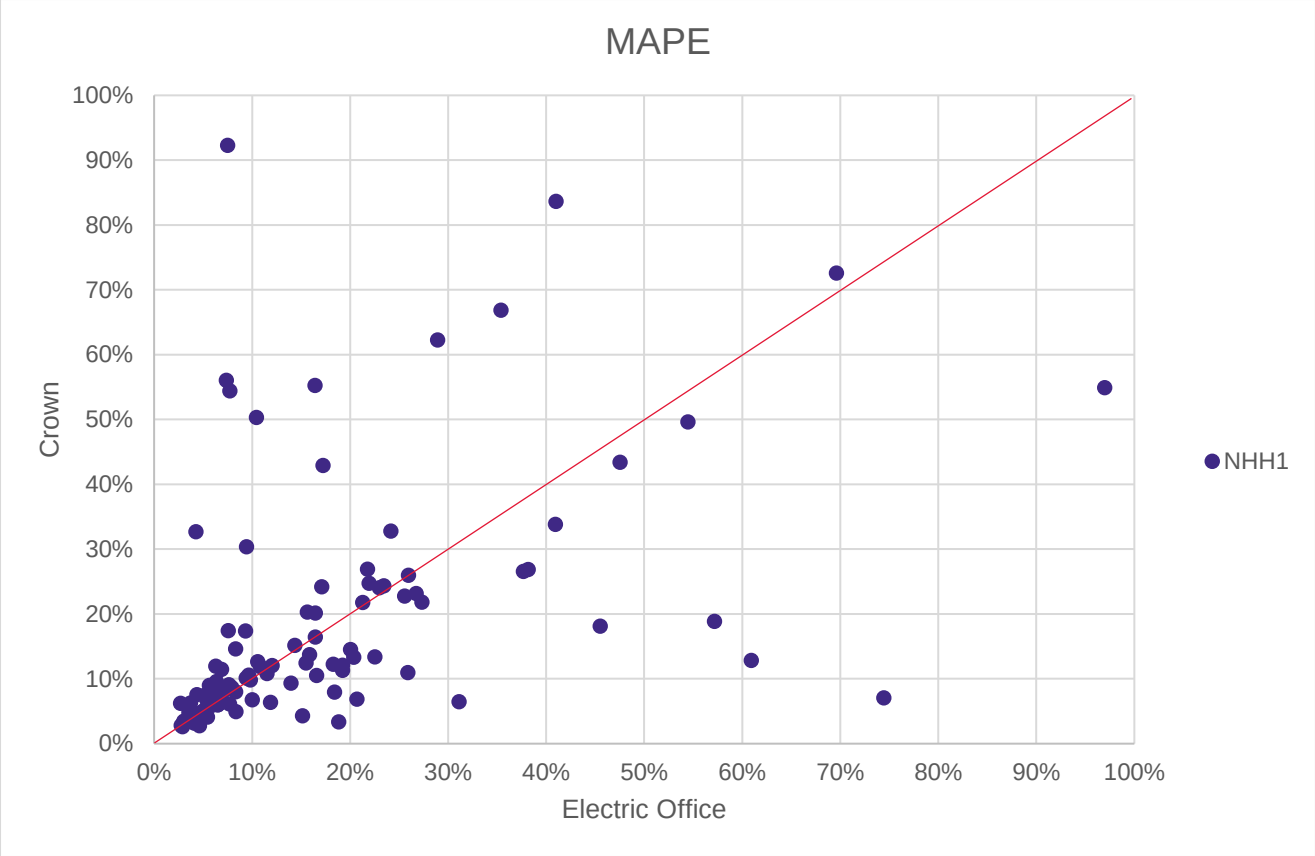


Figure 15: Comparison between EO and CROWN MAPE

Figure 15 illustrates a scatter plot which demonstrates the differences in MAPE between CROWN and EO. The feeders that are close to the red diagonal line have similar MAPE using both CROWN and EO, above the red line are the feeders that perform better using EO data while below are the feeders that CROWN connectivity data seem more accurate.

The algorithm excluded the feeders from the analysis when:

- the MAPE is similar when using CROWN and EO (difference lower than 5%) – 62 feeders,
- the MAPE didn't decrease lower than 20% even if either CROWN or EO were used – 18 feeders. These feeders need further investigation to understand the big daily MAPE.

Table 3 shows the results for the proposed feeder changes.

Feeder	MAPE		Movements from Neighbour SS		Movements between feeders from the same SS		Status
	EO	CROWN	out	in	out	in	
CIRC:940337:1	10	29	0	0	0	10	Accepted
CIRC:940337:2	16	42	1	0	10	0	Accepted
CIRC:941985:3	21	13	4	2	0	2	Rejected

CIRC:942037:5	22	15	3	0	0	0	Rejected
CIRC:942086:1	18	8	0	0	15	0	Rejected
CIRC:942086:2	19	4	2	0	0	15	Rejected
CIRC:942368:4	19	13	0	0	0	1	Rejected
CIRC:942647:2	19	416	1	2	77	11	Accepted
CIRC:942647:3	9	56	0	0	27	71	Accepted
CIRC:942647:4	17	24	2	0	11	33	Accepted
CIRC:942649:2	16	158	3	0	142	70	Accepted
CIRC:942651:3	31	7	1	7	0	2	Rejected
CIRC:942681:1	27	11	1	3	1	1	Rejected
CIRC:942695:2	9	18	7	1	1	1	Accepted
CIRC:942755:2	22	9	0	0	1	8	Rejected
CIRC:942755:4	17	56	5	0	6	1	Accepted
CIRC:942756:1	5	33	49	0	0	0	Accepted
CIRC:942756:2	8	92	58	0	4	0	Accepted
CIRC:942757:1	58	18	3	0	7	0	Rejected
CIRC:942774:2	9	16	10	0	6	5	Accepted
CIRC:942855:1	78	9	1	2	3	0	Rejected
CIRC:942855:4	13	8	0	0	0	1	Rejected
CIRC:942877:1	24	16	8	0	0	0	Rejected
CIRC:942978:1	10	18	0	0	3	0	Accepted
CIRC:942978:3	14	21	5	1	2	0	Accepted
CIRC:944922:2	17	7	4	1	0	0	Rejected
CIRC:944922:4	17	11	0	1	1	1	Rejected
CIRC:945237:1	22	16	0	0	0	6	Rejected
CIRC:945487:1	12	51	0	1	1	73	Accepted
CIRC:945487:5	8	54	0	20	73	1	Accepted

Table 3: Suggested MPAN to Feeder changes for each feeder

Feeders '940337:1' and '940337:2' are two feeders running in parallel on the network (see Figure 12). Using the CROWN connectivity data, the estimated profile for feeder 1 was considerably underestimated, while the feeder 2 was overestimated. Based on the EO "connectivity" and "proximity" approaches, the algorithm suggested 10 MPANs to move from the feeder 2 to feeder 1. Making these changes, the daily MAPE decreased from 29% to 10% for feeder 1 and from 42 to 16% in feeder 2. However, feeder 1 is still underestimated and feeder 2 is overestimated. Figure 16,17 illustrates an example of the estimated profiles for '940337' feeder 1 and 2.

Similar observations were received from the clustering algorithms in section 3.3. Five SMs were suggested by the algorithm to be moved from feeder 2 to 1 and two to move from feeder 1 to 2. Applying the additional changes from the clustering algorithms and calculating again the estimation for feeder profiles, the daily MAPE decreased to 8.7% for feeder 1 and to 13.5% for feeder 2. It is noted that the clustering algorithm can only work for the properties with SMs.

Combining the estimated profiles from feeder 1 and feeder 2 and comparing them with the real measurements led to the daily MAPE decreasing to 5% (see Figure 18), which indicates that there are probably more properties that are still misallocated between the feeder 1 and 2.

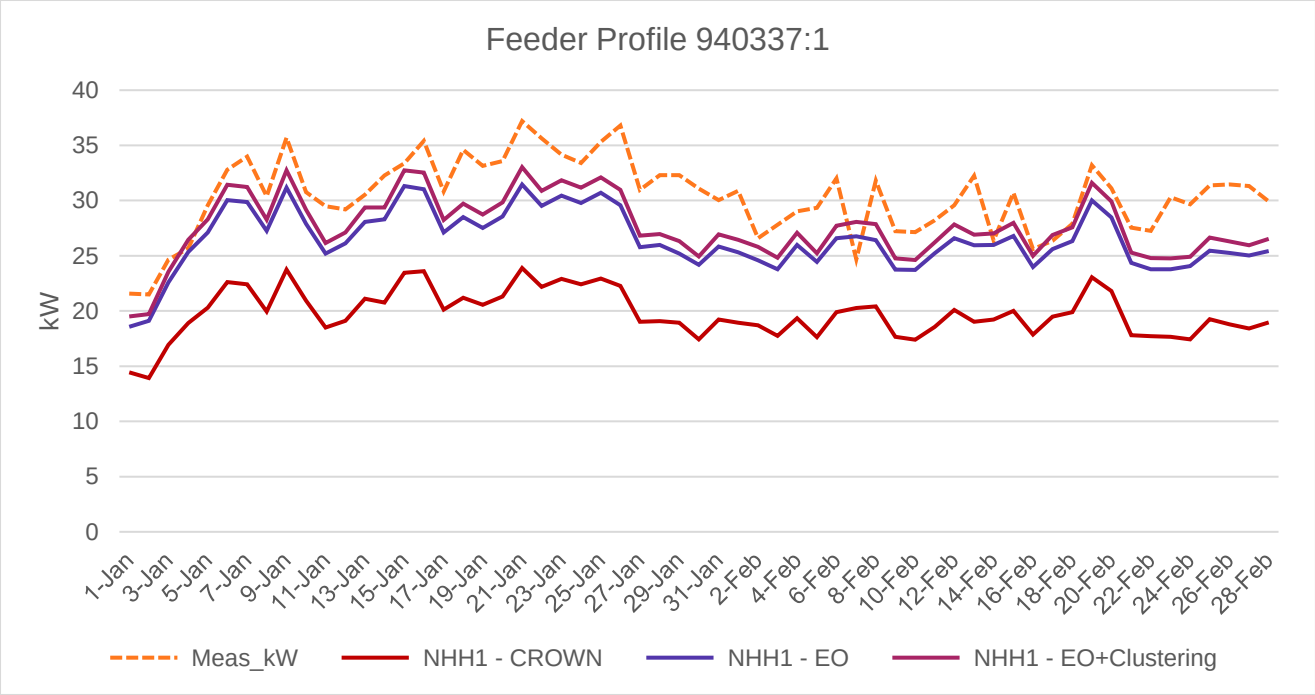


Figure 16: Measured Vs Estimated load profile for 940337:1

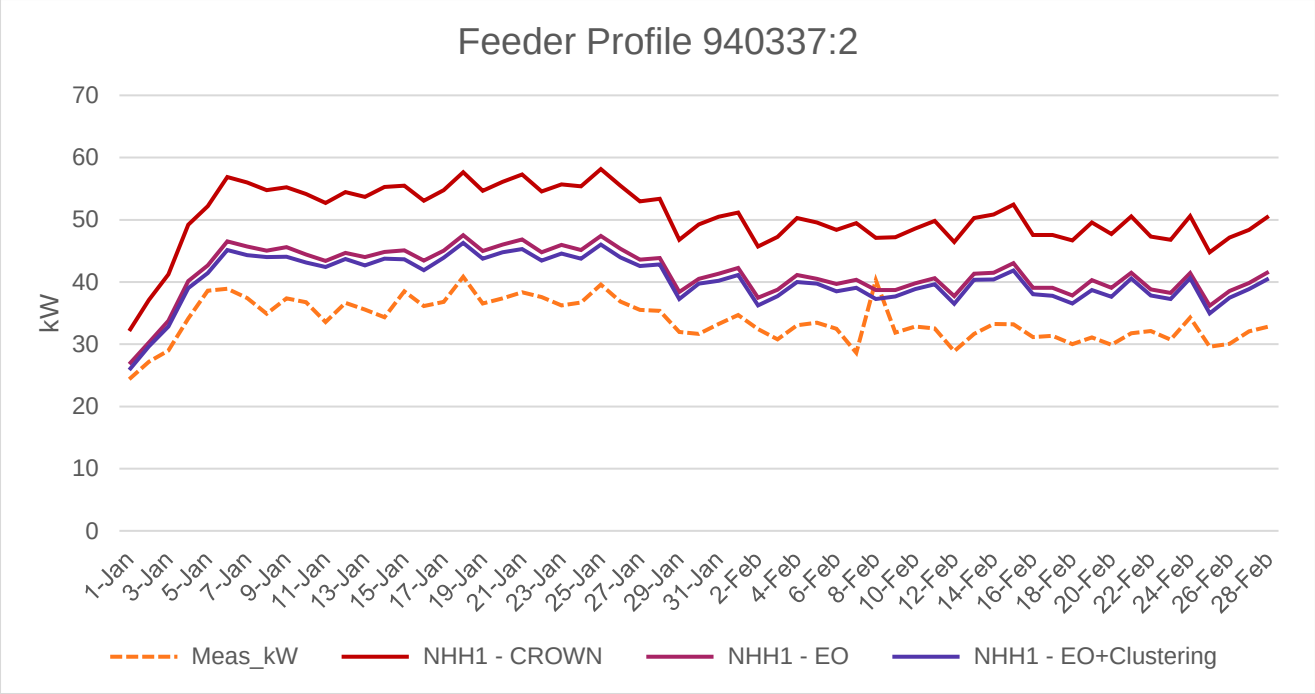


Figure 17: Measured Vs Estimated load profile for 940337:2

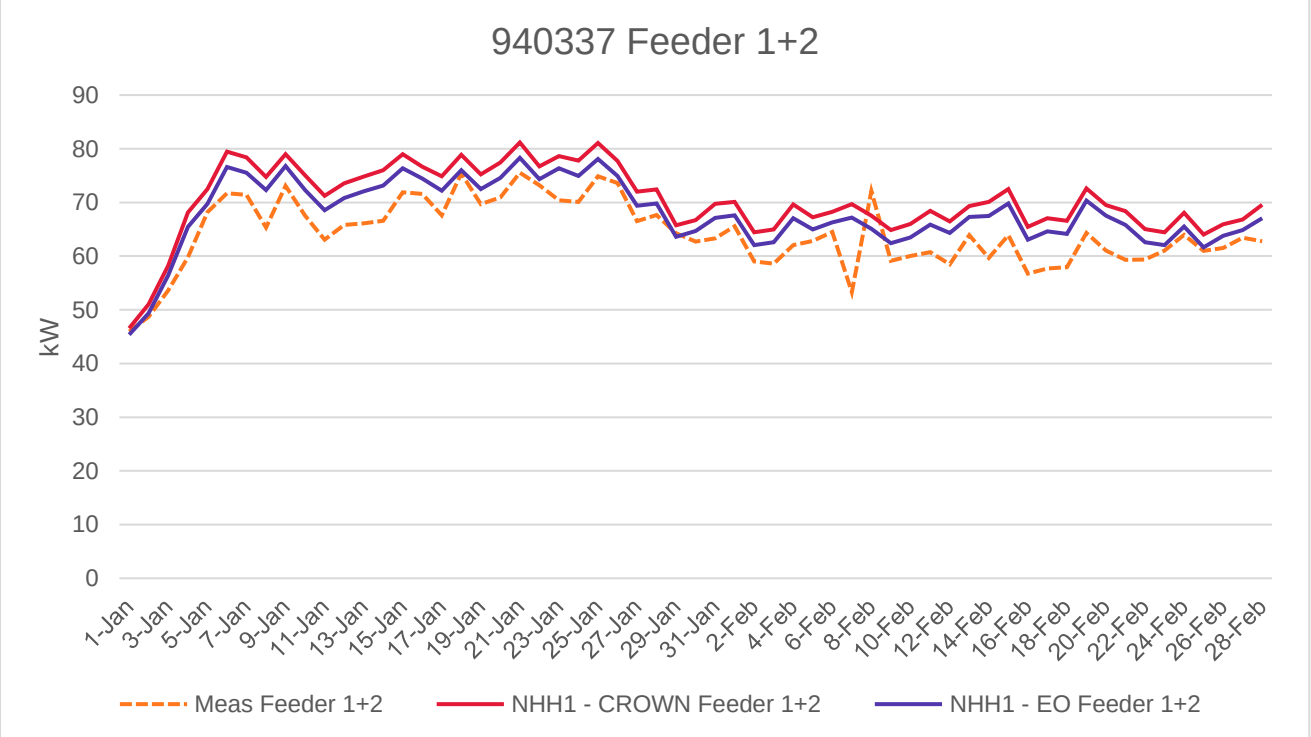


Figure 18: Measured Vs Estimated load profile for 940337:1 + 940337:2

3.7 Ensemble Modelling and BAU Adoption

Some of the proposed algorithms can work without the existence of SS monitoring, while many algorithms need SS monitoring. Moreover, most of the proposed algorithms require the phase to be already known. Figure 19 illustrates the high-level proposed methodology in both cases for BAU adoption.

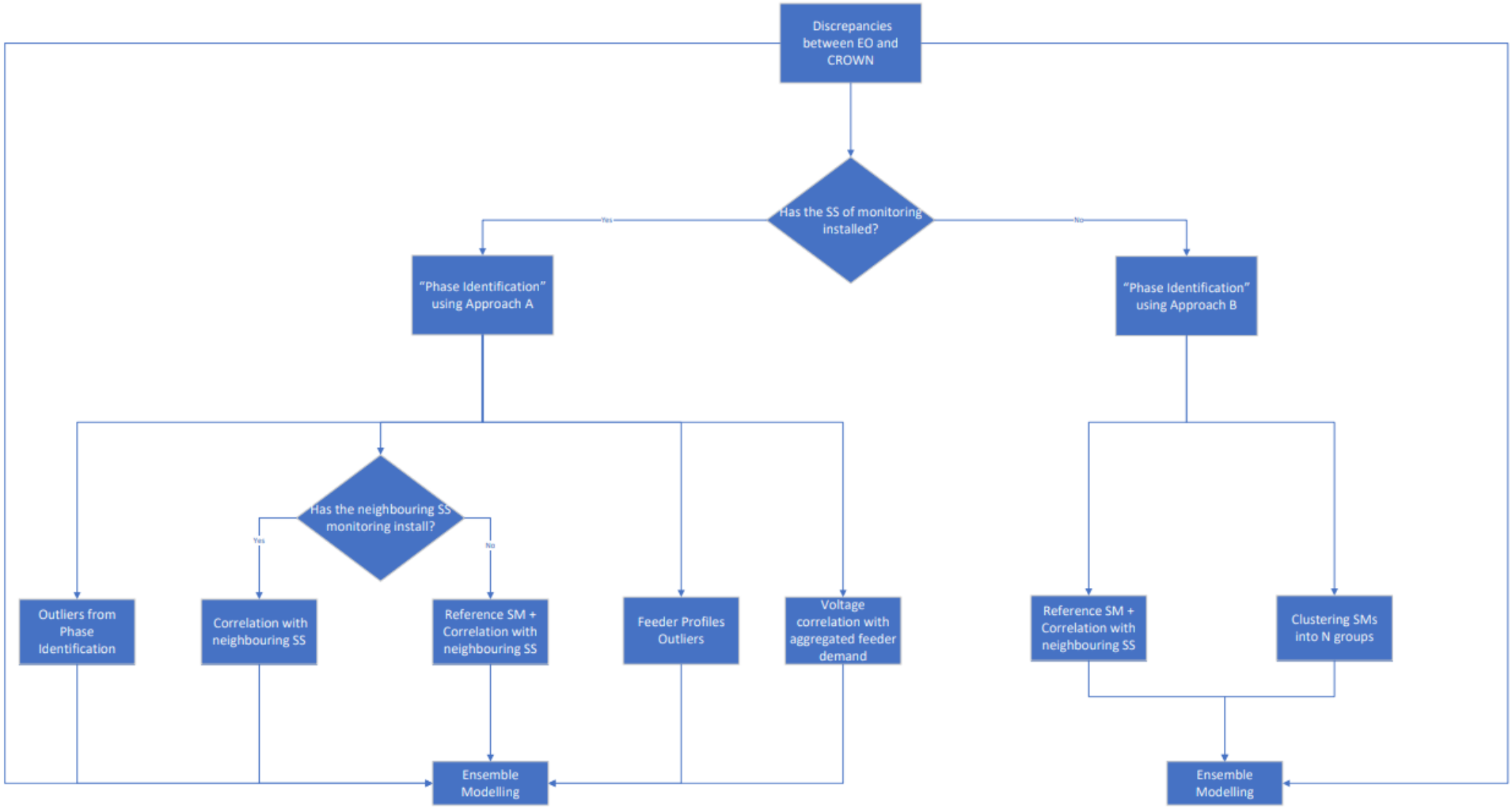


Figure 19: High-level proposed methodology

3.8 Summary Results using HAYSYS validation data

HAYSYS initially tested 218 properties. The feeder survey results for these 218 properties have an accuracy based on CROWN records of about 68%, while using the EO as main source the accuracy increased to 81%.

Then, HAYSYS tested more properties that have been highlighted from the algorithms and the results are the following.

3.8.1 Discrepancies between EO and CROWN - Results

3.8.1.1 Properties that fed from neighbouring SS

There are 337 properties that CROWN and EO disagree regarding the substation that are fed from. 37 properties have been tested from HAYSYS. For 5 properties, HAYSYS agree with CROWN data and for 1 with EO. For 31 properties, HAYSYS did not receive signal. This means that the properties are fed from a different substation and not from the one that CROWN suggests. As a result, the EO data could be the correct source of truth, but we don't have confirmation from HAYSYS feeder survey.

As a result, **86%** of the properties that have been highlighted for review because of the discrepancies between EO and CROWN are not fed from the SS that CROWN records.

3.8.1.2 Properties that fed from different feeder

There are 795 properties that CROWN and EO disagree regarding the feeder number that they are fed from. For 85 properties, we have HAYSYS feeder validation data. For 76 properties, HAYSYS agree with EO data and for 3 with CROWN. Moreover, for 6 properties, HAYSYS did not agree either with CROWN or EO.

As a result, **96%** of the tested properties that have been highlighted for review because of the discrepancies between EO and CROWN are not fed from the feeder that CROWN suggests. **90%** of those are fed from the feeder that EO suggests.

3.8.2 Outliers from Phase Identification - Results

38 SM have been highlighted from the "Outliers from Phase Identification" approach that are fed from different SS. We have HAYSYS validation data for 10 properties. HAYSYS did not receive a signal for these 10 properties, which means that these properties are not fed from the SS that CROWN suggests, and the algorithm correctly highlighted them for review.

As a result, the algorithms accuracy is **100%** for the tested properties.

3.8.3 Correlation with neighbouring SS - Results

26 SM have been highlighted from the algorithm that are fed from the neighbouring SS. We received HAYSYS validation data for 14 SM. For 13 SM, HAYSYS agree with the approach results, while 1 wrongly has been identified as outlier from the algorithm based on HAYSYS validation data.

As a result, the algorithms accuracy is **93%** for the tested properties.

3.8.4 Clustering SMs into Feeder groups

1. Out of 1559 SMs, the algorithm managed to assign a feeder number to 1220 with high confidence. We excluded the SM that are very close to SS (less than 50m) as the voltages are very similar. Also, we excluded the SM where the silhouette score (see Appendix A.3) of the clustering is less than 0.3.

2. For 57 SM, the algorithm did not agree with CROWN and highlighted for review.

3. From the 57 SM, the 11 SM agreed with EO while the other 46 did not agree with either EO or CROWN.

4. 15 out of 1220 SMs have been reviewed by HAYSYS. 1 of them agreed with the algorithms result and other 7 has been allocated to the correct cluster but due to the numbering issue between CROWN and EO, the feeder numbers are different. 6 SMs do not agree either with CROWN or EO or HAYSYS. Most of them belong to "942774". We need to review why the clustering in this substation did not perform well. However, the HAYSYS results for this substation did not agree with EO or CROWN either.

substation	Device	New cluster	EO Feeder	CROWN Feeder	HAYSYS
940337	Device 1	1	1	2	F1
940337	Device 2	1	1	2	F3
942647	Device 3	4	2	4	F2
942647	Device 4	4	2	4	F2
942647	Device 5	4	2	4	F2
942647	Device 6	4	2	4	F2
942647	Device 7	5	1	5	F1
942647	Device 8	5	1	5	F1
942647	Device 9	5	1	5	F1
942756	Device 10	1	3	3	F3
942774	Device 11	4	3	3	F5
942774	Device 12	4	3	3	F5
942774	Device 13	4	3	3	F5
942774	Device 14	4	3	3	F5
942774	Device 15	4	3	3	F5

Table 4: Clustering Results

3.8.5 Voltage correlation with aggregated feeder demand

1. Out of 1559 SM, the algorithm managed to assign a feeder number to 1114 SM with confidence > 0.8.

2. For 300 SM, the algorithm did not agree with CROWN and highlighted them for review.

3. From the 300 SM, the 65 agree with EO while the other 235 did not agree either with CROWN or EO.

4. 15 SMs have been reviewed from HAYSYS. 13 SM agreed with HAYSYS. Accuracy = **87%**

3.8.6 Feeder Profiles - Results

3.8.6.1 Properties that fed from neighbouring SS

In this approach, 58 properties have highlighted for review. The algorithm suggests that for 26 properties the accepted source is EO, while for the rest 32 the accepted source is CROWN. We have only 3 HAYSYS results which agree with the algorithm results. Accuracy = **100%**

3.8.6.2 Properties that fed from different feeder

In this approach 410 properties have been highlighted for review. The algorithm suggests that for the 291 the accepted source is EO while for 119 properties the accepted source is CROWN. 28 properties have been reviewed by HAYSYS. 20 of them agree with algorithms' results. Accuracy = **71%**

4 Summary of Learning Points

- Feeder allocation errors in the database occur on most feeders.
- On some feeders, there are significant numbers of errors due to numbering swaps.
- Feeder identification is less straightforward than phase identification.
- Many approaches requires that phase identification is already correct.
- Voltages for connections near the substations are similar on each feeder and the clustering approach cannot easily separate them.
- EO and CROWN are the two sources that can be used as a starting point to assess the MPAN to feeder connectivity.
- Voltage data are very promising to identify outliers and misallocations.

Future Work

- Identify missing customers that do not exist either in EO or CROWN.
- Investigate the integrated use of a combination of algorithm types (ensemble modelling) to increase the probability of data accuracy.
- The “Voltage correlation with aggregated feeder demand” assigned many SMs in a feeder where neither CROWN nor EO agree. Further investigation is needed to assess and improve the approach.
- Define methodology for BAU adoption.

5 Appendix

A.1 Pearson Correlation

Pearson Correlation measures the strength of the linear relationship between two variables. It has a value between -1 and 1 with a value of -1 meaning a total negative linear correlation, 0 being no correlation and +1 meaning a total positive correlation.

In other words, if the value is in the positive range, the relationship between variables is positively correlated, and both values decrease or increase together. On the other hand, if the value is in the negative range, it shows that the relationship between variables is negatively correlated, and both values will go in the opposite direction.

Pearson's correlation formula is as follows.

$$\frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

A.2 Hierarchical Clustering

Hierarchical clustering methods are divided in two categories, agglomerative and divisive. Agglomerative hierarchical clustering is a "bottom up" approach. In this method, each object is a different cluster in the beginning. Then one pair of clusters is merged at a time and the method continues until all clusters are merged into one cluster. On the other hand, divisive hierarchical clustering is a top-down approach. In this approach, all objects are initially in one big cluster and then split into smaller clusters until each object forms an individual cluster. A dendrogram is used in order to present the results of hierarchical clustering. Figure 13 illustrates an example of the way the agglomerative and divisive hierarchical clustering work.

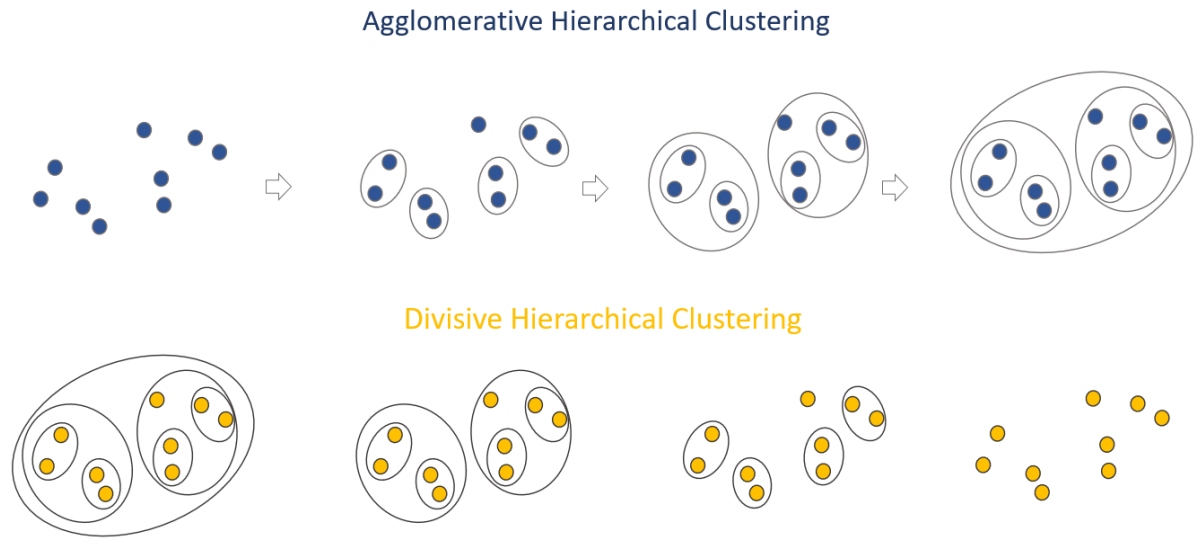


Figure 20: Agglomerative vs Divisive Hierarchical Clustering

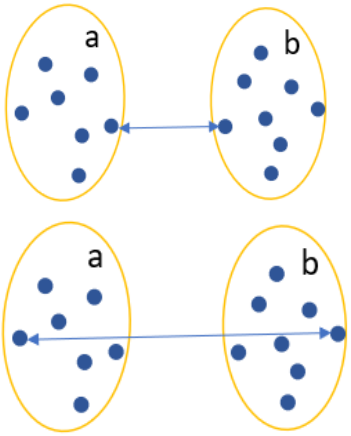
Squareform and pdist are two different methods of calculating distances in hierarchical clustering. Squareform is used when all the data points are arranged in a square matrix and is more efficient in hierarchical clustering as it uses a single distance matrix to calculate all the distances between all the points, while pdist requires separate calculation for each pair of points. On the other hand, pdist is more accurate as it takes into account the actual distances between each pair of points.

A.2.1 Linkage Function between clusters

The creation of the hierarchical cluster tree requires a distance metric which calculates the distance between clusters. There are multiple linkage methods, the single, complete, average, centroid and Ward's linkage method.

Single Linkage

In the single linkage method, the distance between two clusters is defined as the shortest distance between two points in each cluster. In Figure 7, we can see that the distance between two clusters, a and b is the length of the arrow that unites the two clusters' closest points.



$$L(a,b) = min\left(D\left(x_{ai}, x_{bj}\right)\right) \quad (2.3)$$

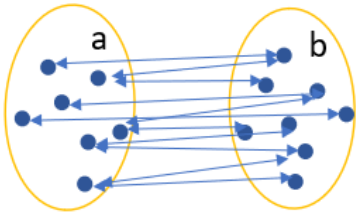
Complete Linkage

On the other hand, in complete linkage hierarchical clustering, the distance between two clusters a and b is the length of the arrow that connects the two points in the two clusters which are as far away as possible.

$$L(a,b) = max\left(D\left(x_{ai}, x_{bj}\right)\right)$$

Average Linkage

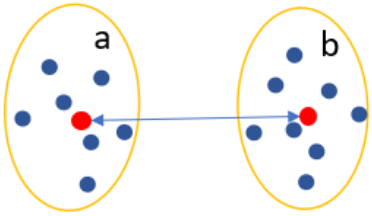
In average linkage method, the distance between two clusters is defined as the average of all pairwise distances across the two clusters.



$$L(a,b) = \frac{1}{n_a n_b}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b} D\left(x_{ai}, x_{bj}\right) \quad (2.5)$$

Centroid Linkage

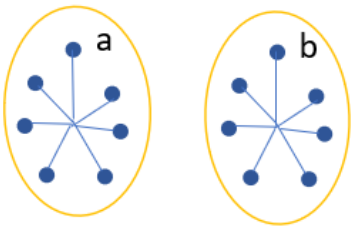
In the centroid linkage method, the distance between two clusters is the distance between the two mean vectors of the clusters. At each stage, the two clusters with the smallest centroid distance are merged.



$$L(a,b) = D\left(\left(\frac{1}{n_a}\sum_{i=1}^{n_a} x_{ai}\right), \left(\frac{1}{n_b}\sum_{j=1}^{n_b} x_{bj}\right)\right) \quad (2.6)$$

Ward's minimum variance method

The Ward's linkage method is the only method that is based on sum-of-squares criterion. At each stage, two clusters merge that provide the smallest increase in the combined error sum of squares.



$$L(a,b) = D\left(\{x_{ai}\}, \{x_{bj}\}\right) = \left\|x_{ai} - x_{bj}\right\|^2$$

A.2.2 Hierarchical Clustering using SMs

The clustering algorithm starts with each device in its own singleton cluster. The algorithm takes the clusters with the smallest distance between them (from the linkage method) and merges them into a single cluster. Then the distances between this new combined cluster and all the other clusters are recalculated, and then the algorithm repeats until all devices are in one cluster. This algorithm then returns the history of all its merges and what distances were between each of the clusters it merged.

From this history, we can pick a threshold where we ignore all merges that occurred where the distance between the merged clusters was above that threshold. This will give us a variable number of clusters depending on the threshold and the distance matrix. We can also choose a fixed number of clusters, by not merging any more clusters once it has made the number of clusters that we want. In phase identification, we pick 3 clusters.

A.2.2.1 An example of Hierarchical Clustering

In Table 5, we can observe an example of a correlation matrix between device's voltage streams.

	Device 1	**Device 2**	**Device 3**	**Device 4**	**Device 5**	**Device 6**
Device 1	1	0.2	0.8	-0.1	0.1	0
Device 2	0.2	1	0	0.9	0.8	0.1
Device 3	0.8	0	1	0	0	0.2
Device 4	-0.1	0.9	0	1	0.9	-0.2
Device 5	0.1	0.8	0	0.9	1	0.2
Device 6	0	0.1	0.2	-0.2	0.2	1

Table 5: correlation matrix

From the correlation matrix, the distance matrix is calculated. Using $D_{i,j} = 1 - C_{i,j}$ the distance between devices 1 and 2 is 0.8, between devices 3 and 3 would be 0, between devices 2 and 4 would be 0.1, etc.

Starting the clustering algorithm, each device is in its own cluster. So, device 1 is in cluster 1, device 2 in cluster 2, etc. The algorithm checks the distance between each pair of clusters and picks the minimum. At the start, this linkage method is the same as the distances between the points (because in $d(u, v) = \max\left(\mathrm{dist}(u_i, v_j)\right)$ there is only one point in each cluster, so it is the maximum over only one distance), so it will just merge the clusters of the closest points. It is noted that the example uses the 'complete' linkage method for simplicity.

Iteration 1: The closest clusters, clusters 2 and 4, have distance 0.1. Clusters 4 and 5 also have distance 0.1. We will pick only one pair of clusters to merge, 2 and 4. Let the new cluster containing devices 2 and 4 be called cluster 7.

Iteration 2: Cluster 4 and 2 got merged into cluster 7, and so now we need to check the distances between this cluster and the other clusters. The distance between cluster 5 and cluster 7 is now

$$\max(\text{dist}(Device\ 5, Device\ 2), \text{dist}(Device\ 5, Device\ 4)) = \max(0.2, 0.1) = 0.2.$$

After calculating the new distances between the new cluster 2 and the other clusters, the closest clusters are now 7 and 5 (linkage distance 0.2), and 1 and 3 (0.2). We'll randomly pick 2 and 5, so now devices 2, 4, 5 are now in the new cluster 8.

Iteration 3: After recalculating the distances between clusters, clusters 1 and 3 have the smallest distance (0.2) so they will be merged into the new cluster 9.

Iteration 4: The distance between clusters 8 and 9 is 1.1, between 9 and 6 it is 1, and between 8 and 6 it is 1.2. So, the clusters 9 and 6 get merged, and now they become cluster 10 containing 1, 3, 6.

Iteration 5: There are only 2 clusters left, now they merge into a single cluster.

From this history, if we want 3 clusters (e.g., one for each phase), then we will take the state after iteration 3 as our clusters, (1,3), (2,4,5), and (6). This looks like a good clustering because the correlations between the devices within each cluster are high, and the correlations of devices in different clusters are low.

Here is a dendrogram diagram that illustrates the process. Note that the numbers along the bottom are the distances between the clusters when they merged.

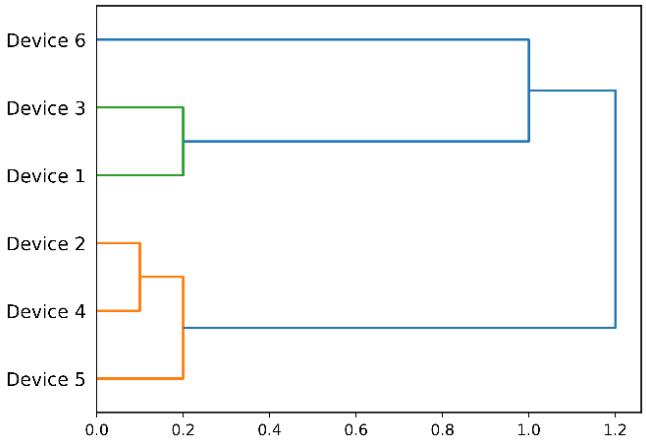


Figure 21: Dendrogram using complete linkage function

Here is the same dendrogram but made with Ward's linkage method instead of complete. Note that the distances at which clusters merge are different.

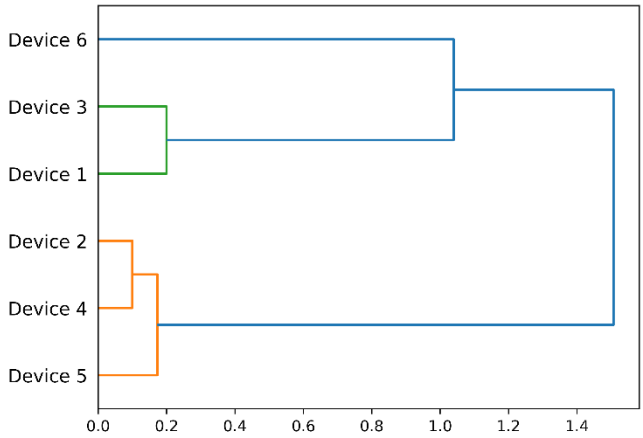


Figure 22: Dendrogram using complete linkage function

A.3 Silhouette Score

Silhouette Coefficient or Silhouette Score is a metric used to calculate the quality of a clustering technique. Its value ranges from -1 to 1. 1 means cluster are well apart from each other and clearly distinguished. 0 means clusters are indifferent, or we can say that the distance between clusters is not significant. -1 means clusters are assigned in the wrong way.

The formula for the calculation of the Silhouette Score is:

$$Silhouette\ Score = (b - a) / \max(a, b)$$

Where,

a = average intra-cluster distance i.e., the average distance between each point within a cluster.
b = average inter-cluster distance i.e., the average distance between all clusters.

A.4 Accuracy

The accuracy of a machine learning classification algorithm is one way to measure how often the algorithm classifies a data point correctly. Accuracy is the number of correctly predicted data points out of all the data points.

Accuracy has the following definition:

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ Number\ of\ predictions}$$

For binary classification, accuracy can also be calculated in terms of positives and negatives as follows:

$$Accuracy = \frac{(TP + TN)}{TP + TN + FP + FN}$$

Where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

A.5 Rand Index & Adjusted Rand Index

Rand Index is a clustering metric that computes a similarity measure between two clusters by considering all pairs of samples and counting pairs that are assigned to the same or different clusters in the predicted and true clustering.

The formula for the Rand Index is:

$$RI = \frac{Number\ of\ Agreeing\ Pairs}{Number\ of\ Pairs}$$

The RI can take values from 0 (completely dissimilar) to 1 (a perfect match).

The Rand Index score is then adjusted for chance using the formula below:

$$ARI = \frac{(RI\ -\ Expected\ RI)}{Max(RI)\ -\ Expected\ RI}$$

This is the adjusted Rand Index, with 0 having a Rand Index the same as an average random labelling, and 1 when the clusters are identical.

A.6 Fowlkes-Mallows Scores

Like the Rand Index, the Fowlkes Mallows scores measure the correctness of the cluster assignments using pairwise precision and recall. A higher score signifies higher similarity.

Fowlkes-Mallows Index (FMI) is a geometric mean of pairwise precision and recall, using True Positive (TP), False Positive (FP) and False Negative (FN).

Fowlkes-Mallow's score does not take into account True Negative (TN), it will not be affected by chance adjustments, unlike Rand Index.

The formula for Fowlkes-Mallows is:

$$FMI = \frac{TP}{\sqrt{(TP + FP\ \ ) \times (TP + FN)}}$$

A.7 Mean Absolute Percentage Error

The mean absolute percentage error (MAPE), also known as mean absolute percentage deviation (MAPD), is a measure of prediction accuracy of a forecasting method in statistics. It usually expresses the accuracy as a ratio defined by the formula:

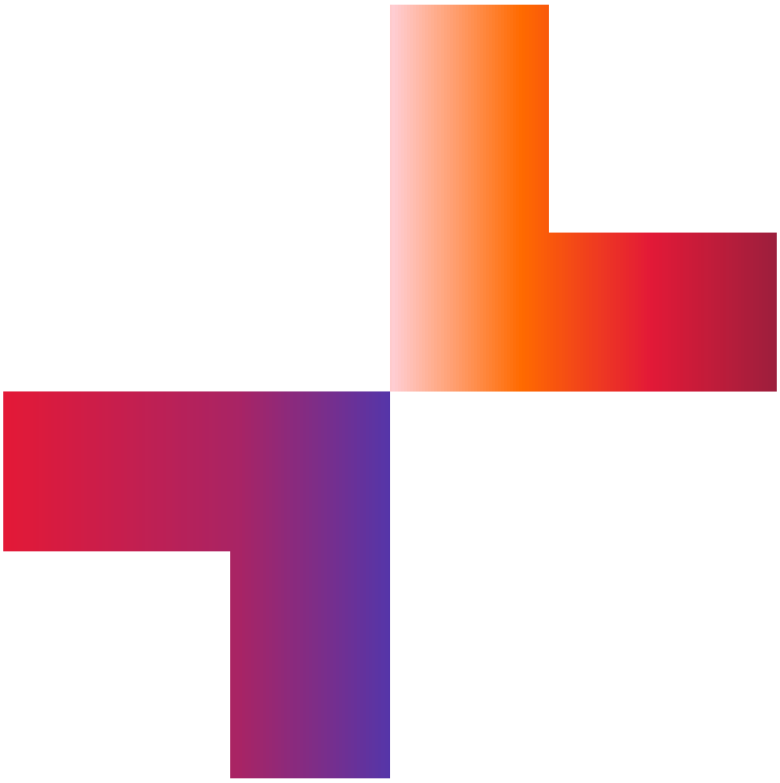
$$MAPE = \frac{100\%}{\mathrm{n}} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

where A_t is the actual value and F_t is the forecast value. Their difference is divided by the actual value A_t. The absolute value of this ratio is summed for every forecasted point in time and divided by the number of fitted points n.

A.8 NHH1

Algorithm Name	HH Customer Demand	Smart meter aggregated demand	NHH customer estimation	Feeder/SS Connectivity
NHH1-Crown	An accurate reflection of the demand taken (Subject to errors in allocation customers to substations and the accuracy of the metered volumes).	An accurate reflection of the demand taken (Subject to errors in allocation customers to substations and the accuracy of the metered volumes).	Utilise EAC and day specific profile coefficients for each day profile class, SSC, TPR.	Crown
NHH1-EO	As above	As above	Utilise EAC and day specific profile coefficients for each day profile class, SSC, TPR	EO

For more information see Combination Load Profiles for Planning Report (WP3- D1).



cgi.com