VALIDATING REAL LV FEEDER MODELS USING SMART METER DATA: A PRACTICAL EXPERIENCE FROM PROJECT EDGE

Michael Z. LIU, Angela SIMONOVSKA, Luis F. OCHOA
The University of Melbourne – Australia
m.liu@ieee.org, a.simonovska@ieee.org,
luis.ochoa@unimelb.edu.au

Peter K.C. WONG, Kenneth CHEW, John THEUNISSEN
AusNet Services – Australia
peter.wong@ausnetservices.com.au,
kenneth.chew@ausnetservices.com.au

ABSTRACT

The proliferation of residential distributed energy resources is driving the need for advanced and proactive approaches to ensure the integrity of distribution networks. Common techniques accessible to distribution companies often involve power flow analysis, which entail the availability of adequate electrical models (an accurate and complete set of data on customer phase grouping, network topology and line impedances). However, in low-voltage (LV) residential areas, these models are often inaccurate and/or incomplete, creating significant challenges and limitations for distribution companies when trying to apply any model-based approach in practice. Exploiting the availability of smart meter data, this paper presents a three-step, data-driven methodology to validate electrical models. The proposed methodology has been successfully demonstrated in Project EDGE (a high-profile, government-funded trial in Australia) to produce accurate, power flow-ready models of real-world feeders. This work also highlights the key insights and learnings from working with real data and LV feeders.

1. INTRODUCTION

The proliferation of residential distributed energy resources (DER), such as rooftop solar PV, batteries, and electric vehicles, is creating significant challenges for distribution companies. This is because the existing distribution networks have not been designed to withstand the diverse behaviours of DERs. For instance, excessive reverse power flows can cause severe voltage rise issues around noon [1] and the highly coincidental charging of electric vehicle can exacerbate voltage drop issues in the evening [2]. Therefore, more advanced and proactive approaches are necessary to ensure network integrity (i.e., customer voltages are within statutory limits and network asset are not overloaded).

In this context, the common techniques accessible to distribution companies (such as heuristic algorithms [3] and optimal power flow-based optimisations [4, 5]) often rely on power flow analysis, which entail the availability of adequate electrical models (i.e., accurate and complete information on customer phase grouping, network topology and line impedances). Although, to a certain extent, modelling data do exist within distribution companies' databases, they are often incomplete and erroneous, which is particularly prevalent for low-voltage (LV) residential feeders [6]. Therefore, being able to produce accurate electrical models of LV feeders is extremely crucial for distribution companies so that they can effectively tackle the challenges of rapid DER growth.

Thanks to the deployment of smart meters, with certain jurisdictions reaching as close as 100% installations for residential and small commercial/industrial customers (e.g., in the State of Victoria in Australia as well as nationwide across Spain, Denmark and Finland), data-driven techniques have emerged in recent literature to tackle the fundamental challenge of inadequate modelling data. Some works focus on tackling one aspect of the electrical model (e.g., phase grouping identification [7], topology construction [8] or line impedances estimation [9]) while others have investigated ways of reproducing the entire model simultaneously [10]. Nonetheless, a key limitation of these works is that they have not been tested for real-world feeders and with real smart meter data.

This paper presents a three-step methodology to produce accurate, power flow-ready electrical models of real-world, three-phase LV feeders. Starting with the existing (and potentially incomplete/erroneous) data from distribution companies, several smart meter data-driven techniques are exploited, along with technician site visits, to correct and/or validated customer phase groupings, feeder topology and line impedances. The proposed methodology leverages the authors' prior work in [7, 9] while also incorporating further adaptations to cater for real smart meter data and feeders. The proposed methodology has been applied in Project EDGE [1] to produce validated electrical models of two LV feeders in Victoria, Australia. This paper also shares the key challenges and learnings from working with real-world smart meter data and feeders.

2. METHODOLOGY

2.1 Overview of a Three-Phase LV Feeder

An illustration of a (residential) three-phase LV feeder is shown in Figure 1. The start of a feeder is named head-of-feeder which corresponds to the secondary terminal of the distribution transformer. Each customer (i.e., a house) is connected to the head-of-feeder through the common backbone (three-phase, as depicted by the three coloured lines) and individual service cables (either single-phase or

1 Project EDGE (link) is a high-profile, government-funded trial in Victoria, Australia that aims to create a proof-of-concept platform for DER to provide services through aggregators. Here, the production of accurate electrical models is essential in enabling the distribution company (AusNet Services) to calculate meter-level operating envelopes (time-varying power export and import limits at the point of connection) so as to ensure network integrity while facilitating DER services.

three-phase, depending on the connection requirements).
From a modelling perspective, the key information required to produce a complete electrical model consist of:

1. *customer phase groupings*, i.e., which of the three phases each customer is connected to;
2. *feeder topology*, i.e., how the backbone and service cables are interconnected with each other; and
3. *line impedances*, i.e., the resistance and reactance for each segment of the backbone and each service cable. NB. Each backbone segment is defined by the section between service cables of adjacent customers.

In this work, the neutral conductor (typically found in LV feeders) is not explicitly considered. This is achieved through the Kron reduction technique [11] which is adequate for LV feeders with sufficient earthing points (such as in Australia, where the neutral conductor is earthed at every customer's meter).

2.2 Data-Driven Model Validation

To a certain extent, modelling data can be extracted from the various databases of distribution companies, such as the geographic information system (GIS) and SCADA system. However, and particularly for LV feeders, there are often data quality issues such as missing information (e.g., phase grouping not being recorded when the customer connection was first established) and erroneous information (e.g., the location of a customer is incorrected recorded). As a result, the existing data is often insufficient to produce an adequate model for the purpose of accurate power flow analysis. Therefore, the three-step process discussed next (phase grouping identification, topology verification and impedance correction) is developed to resolve potential issues with existing modelling data. This methodology leverages historical measurements (voltage magnitude, active power, and reactive power) from both smart meters (installed at the customers' premises) and the head-of-feeder monitor (installed on the secondary side of the transformer). While head-of-feeder monitors are not necessarily common in LV feeders, they have been deployed in Project EDGE to capture the necessary data for modelling and other trial related purposes.

Step 1: Phase Grouping Identification

The first step is to identify the phase grouping of all customers, which uses the clustering algorithm proposed in [7]. Here, two statistical analysis techniques (namely principal component analysis and unconstrained k-means clustering) are exploited to extract the inherent correlation in the time-series voltage profiles of customers belonging to the same phase group. This algorithm is applied to both smart meter data and head-of-feeder data in order to obtain phase group matching of customers with respect to the three-phases of the distribution transformer.

Step 2: Topology Verification

The second step is to resolve any issues with the existing information on feeder topology from the GIS system. This is an iterative process between carrying out desktop

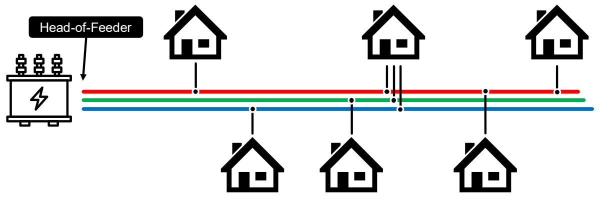

Figure 1. Illustration of a three-phase LV feeder.

analysis (to identify potential issues with the existing topology) and physical site visits by a technician (to physically inspect the infrastructure and verify the topology). Given that site visits are time-consuming and costly for distribution companies (e.g., a technician needs to be scheduled, the site can be very far away, and the inspection itself is manual), the initial desktop analysis can help to identify key areas of concern, reducing the effort required by a technician during a site visit. In certain cases, this could even eliminate the need of a site visit altogether.

In terms of the desktop analysis, a voltage sensitivity-based technique is used to identify potential issues with the existing topology data. This technique involves analysing the effect of a given customer's power imports (or exports) on other customers' voltages, and thus identify their relative position in a feeder. With respect to Figure 1, the sensitivity of the last customer's voltage due to the first customer's powers will be minimal as the only common electrical path between them is the first backbone segment. On the other hand, the sensitivity of the last customer's voltage due to the second last customer's power will be much higher due to sharing almost the entire backbone as the common electrical path.

Step 3: Impedance Correction

The third, and final, step is to correct the conductor impedances. This step uses the linear regression-based algorithm proposed in [9] to estimate the impedances. This process is done successively, starting from the head-of-feeder, for each backbone segment and service cable. It is worth noting that this step relies on adequate results from the previous two steps (i.e., having validated phase grouping and topology information).

2.3 Performance Assessment with Scattered Plots

The performance of an electrical model is assessed using a scattered plot technique, as illustrated in Figure 2. This assessment is carried out by comparing two sets of voltages, namely the *measured* voltages (on the x-axis) and the *calculated* voltages (on the y-axis). The measured voltages are the actual values recorded by smart meters. The calculated voltages are obtained by running power flows using the electrical model and the powers recorded by the smart meters.

In a perfect scenario, the measured voltages should equal to the calculated voltages, which is depicted by the black line $y = x$. However, given that real world measurements are inherently imperfect (attributed by noise, synchronisation issues, etc.), the key criterion of a satisfactory model is one that aligns well with the black line; this is depicted by the green dots in Figure 2 where the calculated voltages are highly correlated with the measured voltages. On the other hand, an unsatisfactory model is one that misaligns with the black line; this is depicted by the red dots in Figure 2.

3. CASE STUDY

3.1 Real Feeders and Data

Two three-phase LV feeders, namely Feeder 1 and Feeder 2, in a residential neighbourhood of Victoria, Australia are considered in this case study. The two feeders are part of the trial sites in Project EDGE and managed by the local distribution company AusNet Services. Feeder 1 is relatively larger in size, servicing 28 customers in total (26x single-phase and 2x three-phase). Feeder 2, in contrast, has 8 customers in total (5x single-phase and 3x three-phase). Both feeders have a nominal supply voltage of 400V line-to-line which means customers have a nominal voltage of 230V line-to-neutral.

For Feeder 1, the unvalidated modelling data available from the distribution company consists of the feeder topology (extracted from the GIS system), partial phase grouping information (for ~75% of all customers) and estimated line impedances (derived from the estimated conductor length using GIS data and manufacturers' specifications). The same set of unvalidated data is also available for Feeder 2, except that the existing data contains phase grouping information for all customers.

For both feeders, three weeks of historical smart meter data and head-of-feeder monitor data (installed on the secondary side of the distribution transformer) are considered. The measurements are collected per phase and consist of voltage magnitude, active power, and reactive power. These measurements are instantaneous readings with five-minute resolution. Given that each phase of the three-phase customers is independently monitored, from a modelling perspective, they are treated as equivalent single-phase customers.

The rest of this section will present the validation process for each feeder as well as the key insights and learnings. For clarity, the electrical model produced using original data from the distribution company is termed the *unvalidated model* and the final model after the validation process is termed the *validated model*.

3.2 Model Validation of Feeder 1

Customer Phase Grouping. As shown in Table I, the unvalidated model has missing phase grouping information for 7 (out of 28) customers. This is resolved after carrying out the Phase Grouping Identification step presented in Section 2.2, also shown in Table I. Apart from the missing data of 7 customers, no further issue was identified with the original data. The validity of phase grouping information can be straightforwardly verified using both smart meter data and head-of-feeder data. This is shown in Figure 3 (for a 12-hour window) which compares, per phase, the active power recorded at the head-of-feeder and the sum of active power from all customers. As illustrated by the time-series plot in Figure 3, the power at the head-of-feeder (solid lines) is closely matched with the sum of customers (dashed lines), for each phase and at all times; this confirms that the phase grouping information of the validated model (as shown in Table I) is indeed correct.

Feeder Topology. With the validated phase grouping information, the next step is to produce the sensitivity values and cross-check with the unvalidated topology. A single-line representation of the unvalidated topology is shown in Figure 4. The desktop-based sensitivity analysis is explained using A1, which is one of the mistakes identified in the original data. The sensitivity values of customer A1 (with respect to other customers) is shown in

Table I. Customer phase grouping data for Feeder 1.

	Unvalidated Model	Validated Model
Phase A	A1, A13, A14, A20, A22, A24, A26, A28	A1, **A4**, A13, A14, **A19**, A20, A22, A24, A26, A28
Phase B	A1, A3, A8, A9, A12, A16, A24,	A1, A3, **A6**, A8, A9, A12, **A15**, A16, **A23**, A24,
Phase C	A1, A2, A5, A7, A10, A17, A18, A24, A25, A27	A1, A2, A5, A7, A10, **A11**, A17, A18, **A21**, A24, A25, A27
Unknown	**A4, A6, A11, A15, A19, A21, A23**	-

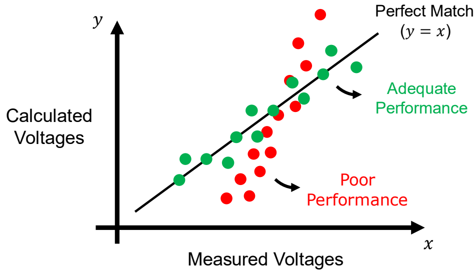

Figure 2. Model assessment with scattered plots.

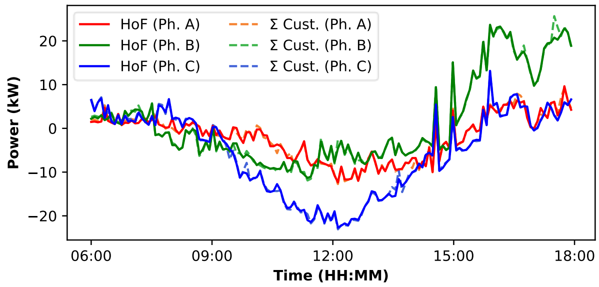

Figure 3. Phase grouping verification with power plots.

Table II. From Table II, it can be verified that A1 has significantly higher sensitivity due to customers A4, A20, A14 and A26. However, this result contradicts the topology data shown in Figure 4 where A1 is shown to be a customer near the head-of-feeder (i.e., have led to negligible sensitivity with respect to any customer). Consequently, this discrepancy is flagged as a potential issue for AusNet Services to verify during a site visit. After going through the process of both desktop analysis and site visit, the validated topology shown in Figure 5 is obtained (key changes are customers A1, A9, A18 and A28).

Line Impedances. With the revised phase grouping and topology data, the Impedance Correction step in Section 2.2 is carried out to produce the full validated electrical model. The overall changes in both positive sequence (R1 and X1) and zero sequence (R0 and X0) impedance values are illustrated in Figure 6. Here, each block represents a backbone segment or service line of Feeder 1 and the accent $*$ is used to denote the validated values. As shown in Figure 6, the validated impedances can be several times larger than the original, unvalidated values, with the aggregated positive sequence resistance and reactance increasing by 170% and 400%, respectively.

Performance Assessment. The overall performance of the unvalidated and validated electrical models are compared using the scattered plot introduced in Section 2.3. For benchmarking purposes, the customers with unidentified phase groups in the unvalidated model are all assigned to phase A. As illustrated by Figure 8, the unvalidated model (red) showed higher discrepancies between the measured voltages and calculated voltages compared with the validated model (green). In contrast, the validated model offers substantial improvement as it is well aligned with the reference line $y = x$.

3.3 Model Validation of Feeder 2

The same model validation process is applied to Feeder 2. For brevity, only key results are highlighted.

For Feeder 2, the original phase grouping and topology data are verified to be correct. However, significant changes are required in the line impedances. The validated topology is shown in Figure 7 and the performance comparison is shown in Figure 9. As illustrated by Figure 9, despite having the correct phase grouping and topology information, extremely poor performance is still observed for the unvalidated model (red), which is attributed by inaccurate line impedances. On the other hand, the performance is significantly improved with the validated model (green).

3.4 Practical Insights and Learnings

Real world data is imperfect. The use of real-world data introduces additional challenges due to the inherent noise/error within these data which can be caused by factors such as synchronization, equipment's accuracy

Table II. Sensitivity analysis of customer A1.

ID	Sensitivity (V/kW)	ID	Sensitivity (V/kW)
A24	0.01	A4	0.09
A13	0.03	A20	0.13
A19	0.04	A14	0.25
A22	0.05	A26	0.27
A28	0.05		
A28	0.05		

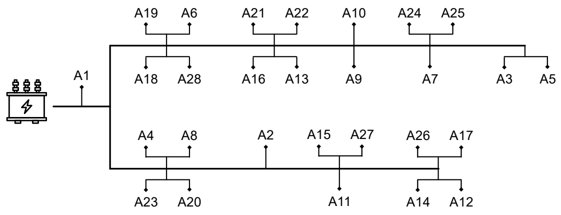

Figure 4. Unvalidated topology of Feeder 1.

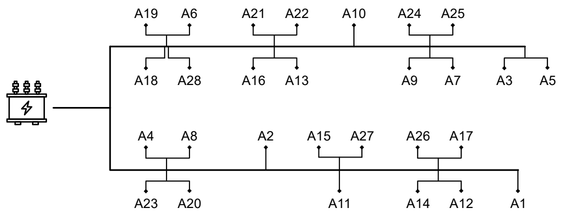

Figure 5. Validated topology of Feeder 1.

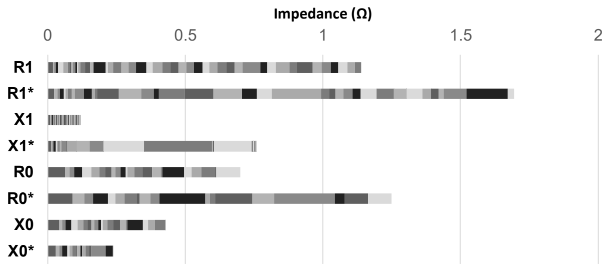

Figure 6. Line impedances of Feeder 1: unvalidated (R1, X1, R0, X0) vs validated ($R1^*$, $X1^*$, $R0^*$, $X0^*$).

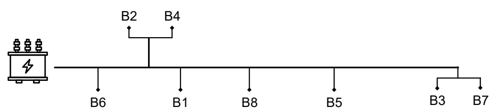

Figure 7. Validated topology of Feeder 2.

class, etc. To this end, the scattered plots (presented in Section 2.3) has proven to be an extremely effective tool as it offers quick and intuitive performance visualisations. It is also worth highlighting that the key criterion should be assessing the overall trend on the scattered plots (i.e., aligned with the line $y = x$), rather than specific points.

Model validation can be labour-intensive and time-consuming. The process to produce validated electrical

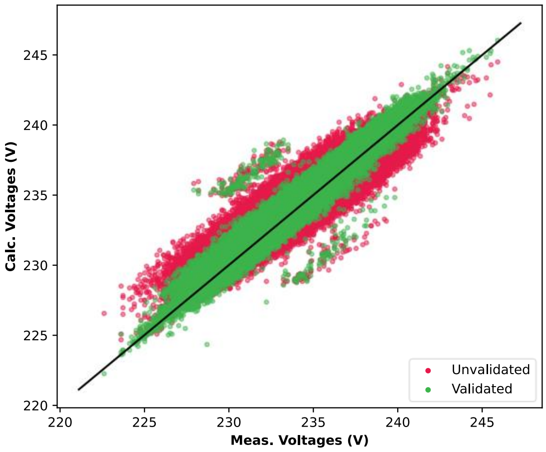

Figure 8. Performance assessment of Feeder 1: unvalidated (red) vs validated (green).

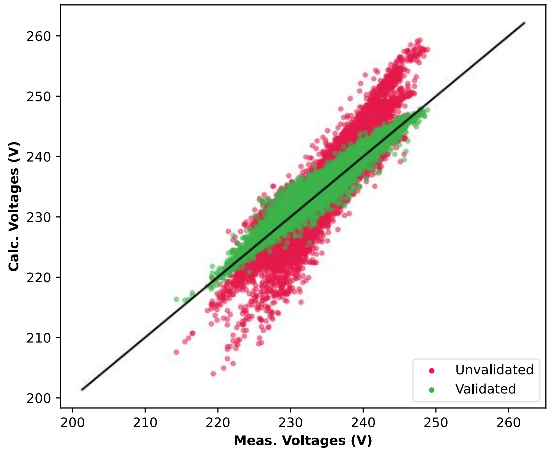

Figure 9. Performance assessment of Feeder 2: unvalidated (red) vs validated (green).

model is non-trivial. Firstly, every site visit requires the scheduling of an available technician, which could introduce delays. Furthermore, the presence of underground cables (increasingly common in urban networks) poses further challenges as the conductors themselves are difficult/impossible to be physically inspected. Finally, given the conductor impedance correction step relies on the topology information, multiple iterations could be necessary to finalize both the topology and the impedances.

4. CONCLUSION

An adequate electrical model is the foundation of many applications in distribution networks. Although, to a certain extent, data for modelling purposes do exist within distribution companies' databases, they are often incomplete and erroneous. This issue is particularly common in low-voltage (LV) residential feeders. By exploiting the availability of smart meter data, this work presents a three-step model validation methodology to produce accurate, power flow-ready electrical models of real-world LV feeders. The proposed methodology is successfully demonstrated using real smart meter data for two three-phase LV feeders in Victoria, Australia.

Acknowledgments

The authors would like to thank AusNet Services for their support and the data used in this paper.

REFERENCES

[1] A. Navarro-Espinosa and L. F. Ochoa, "Probabilistic Impact Assessment of Low Carbon Technologies in LV Distribution Systems", *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2192-2203, 2016

[2] J. Quirós-Tortós, L. F. Ochoa, S. W. Alnaser, and T. Butler, "Control of EV Charging Points for Thermal and Voltage Management of LV Networks", *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 3028-3039, 2016

[3] M. Liu and L. F. Ochoa, "Deliverable 1.1: Operating Envelopes Calculation Architecture", Available: https://electrical.eng.unimelb.edu.au/power-energy/projects/project-edge.

[4] M. Z. Liu, L. F. Ochoa, P. K. C. Wong, and J. Theunissen, " Using OPF Based Operating Envelopes to Facilitate Residential DER Services", *IEEE Trans. Smart Grid*, vol. 13, no. 6, pp. 4494-4504, 2022.

[5] V. Rigoni, D. Flynn, and A. Keane, "Coordinating Demand Response Aggregation with LV Network Operational Constraints ", *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 979-990, 2021.

[6] A. Navarro-Espinosa, L. F. Ochoa, R. Shaw, and D. Randles, "Reconstruction of Low Voltage Distribution Networks: From GIS Data to Power Flow Models", *Proc. 23rd International Conference on Electricity Distribution*, pp. 1-5, 2015.

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

[8] A. B. Pengwah, L. Fang, R. Razzaghi and L. L. H. Andrew, "Topology Identification of Radial Distribution Networks Using Smart Meter Data", *IEEE Systems Journal*, vol. 16, no. 4, pp. 5708-5719, 2022.

[9] Y. Wang, M. Z. Liu, and L. F. Ochoa, "Assessing the effects of DER on voltages using a smart meter-driven three-phase LV feeder model", *Electric Power Systems Research*, vol. 189, pp. 106705, 2020.

[10] V. C. Cunha, W. Freitas, F. C. L. Trindade, and S. Santoso, "Automated Determination of Topology and Line Parameters in Low Voltage Systems Using Smart Meters Measurements", *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 5028-5038, 2020.

[11] W. H. Kersting and R. K. Green, "The application of Carson's equation to the steady-state analysis of distribution feeders", *Proc. 2011 IEEE/PES Power Systems Conference and Exposition*, pp. 1-6, 2021.