

UNIVERSIDADE FEDERAL DE ITAJUBÁ – UNIFEI
GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

**A FRAMEWORK BASED ON MIXED-INTEGER MODELS TO SUPPORT THE
OPERATION OF ACTIVE DISTRIBUTION SYSTEMS WITH INVISIBLE DERS**

**UMA METODOLOGIA BASEADA EM MODELOS DE PROGRAMAÇÃO INTEIRA
MISTA PARA APOIAR A OPERAÇÃO DE SISTEMAS DE DISTRIBUIÇÃO ATIVOS
COM DERS INVISÍVEIS**

Pedro Henrique Naves Vasconcelos

ITAJUBÁ

2025

UNIVERSIDADE FEDERAL DE ITAJUBÁ – UNIFEI
GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

Pedro Henrique Naves Vasconcelos

**A FRAMEWORK BASED ON MIXED-INTEGER MODELS TO SUPPORT THE
OPERATION OF ACTIVE DISTRIBUTION SYSTEMS WITH INVISIBLE DERS**

**UMA METODOLOGIA BASEADA EM MODELOS DE PROGRAMAÇÃO INTEIRA
MISTA PARA APOIAR A OPERAÇÃO DE SISTEMAS DE DISTRIBUIÇÃO ATIVOS
COM DERS INVISÍVEIS**

Thesis submitted to the Graduate Program in
Electrical Engineering, Federal University of
Itajubá, as part of the requirements for obtaining
the degree of Doctor of Philosophy in Electrical
Engineering.

Field of Electrical Power Systems

Supervisor: Antonio Carlos Zambroni de Souza

Co-Supervisor: Bala Venkatesh

Co-Supervisor: Glauco Nery Taranto

ITAJUBÁ

2025

ACKNOWLEDGMENT

I want to express my deepest gratitude to all those who contributed to completing this work in different ways.

I am grateful to my supervisor, Prof. Antonio C. Zambroni de Souza, for the valuable lessons throughout this academic journey. I also thank Prof. Bala Venkatesh for his guidance and warm reception in Canada, and Prof. Glauco N. Taranto for his support and encouragement throughout this work.

I am thankful to Prof. Walmir Freitas, Prof. Fernanda Trindade, P.Eng. John Penaranda, and all colleagues from the research groups in Itajubá and Toronto, who consistently offered support, insightful suggestions, and friendship.

I extend my heartfelt thanks to my family for their unconditional love and constant encouragement. To Ana, whose presence made the difficult days lighter and the moments of achievement even more special.

I am also grateful to the Brazilian National Council for Scientific and Technological Development (CNPq), the Federal University of Itajubá, and the Centre for Urban Energy at Toronto Metropolitan University for their financial support and opportunities for academic and personal development.

RESUMO

A modernização dos sistemas elétricos de potência, marcada pela ampla integração de Recursos Energéticos Distribuídos (DERs), está transformando redes de distribuição tradicionais em sistemas dinâmicos e ativamente gerenciados. No entanto, persiste um grande desafio operacional: a visibilidade limitada dos DERs não monitorados, denominados DERs invisíveis. Embora sejam fundamentais para a flexibilidade futura da rede, esses recursos introduzem incertezas que prejudicam a observabilidade do sistema, a tomada de decisões operativas e a resiliência. Esta pesquisa propõe uma estrutura metodológica baseada em modelos de otimização inteira mista para apoiar a operação de sistemas de distribuição ativos sob condições de baixa observabilidade dos DERs. A abordagem concentra-se no desenvolvimento de modelos agregados equivalentes para DERs invisíveis, permitindo que operadores de sistema infiram variáveis críticas—como magnitudes de tensão e fluxos de potência em ramos—utilizando dados de medição esparsos ou incompletos. Ao formular o problema como uma tarefa de Programação Não Linear Inteira Mista (MINLP) convexa, a metodologia permite o posicionamento e dimensionamento estratégico de modelos agregados de DERs que reproduzem o comportamento observado do sistema. A pesquisa também introduz uma variante linearizada do modelo, incluindo uma formulação de Programação Linear Inteira Mista (MILP) utilizando relaxações de McCormick, para aprimorar a tratabilidade computacional sem comprometer a precisão das estimativas. Além disso, modelos híbridos de DERs, combinando tecnologias como geração fotovoltaica e sistemas de armazenamento de energia, são incorporados para representar melhor o comportamento em regime permanente de redes de distribuição modernas com DERs invisíveis híbridos. Diversos estudos de caso demonstram que a estrutura proposta é capaz de estimar variáveis de estado ausentes de forma precisa mesmo em cenários de escassez de dados, alcançando baixos erros médios enquanto reduz significativamente os tempos de solução por meio das técnicas de linearização. A integração com o OpenDSS possibilita a validação da implementação em uma ferramenta-padrão da indústria, comprovando a precisão numérica elevada e a redução dos tempos de solução na modelagem de alimentadores equivalentes com DERs agregados representando inúmeros recursos invisíveis. De forma geral, este trabalho avança o estado da arte ao oferecer uma abordagem de modelagem escalável e eficiente em dados, que capacita operadores de sistemas de distribuição a manter operações confiáveis e eficientes mesmo diante do crescente número de DERs invisíveis.

Palavras-Chave: Recursos Energéticos Distribuídos, Rede Equivalente, Cálculo de Fluxo de Carga, Otimização, Sistemas de Distribuição Radiais, Estimação de Estado.

ABSTRACT

The modernization of electric power systems, marked by the widespread deployment of Distributed Energy Resources (DERs), is transforming traditional distribution networks into dynamic, actively managed systems. However, a major operational challenge persists: the limited visibility of unmonitored DERs, hereon referred to as invisible DERs. While critical to future grid flexibility, these resources introduce uncertainty that hampers system observability, operational decision-making, and resilience. This research proposes a novel methodological framework based on mixed-integer optimization models to support the operation of active distribution systems under limited DER observability. The approach focuses on developing equivalent aggregate models for invisible DERs, enabling system operators to infer critical network states—such as voltage magnitudes and branch power flows—using sparse or incomplete measurement data. By formulating the problem as a convex Mixed-Integer Nonlinear Programming (MINLP) task, the methodology allows for strategically placing and sizing equivalent aggregate DER models that best replicate observed system behavior in steady-state. The research introduces a linearized model variant, including Mixed-Integer Linear Programming (MILP) formulation using McCormick relaxations to enhance computational tractability without compromising estimation accuracy. Furthermore, hybrid DER models combining technologies such as photovoltaic generation and battery storage are incorporated to better capture the steady-state behavior of modern distribution networks with invisible hybrid DERs. Comprehensive case studies demonstrate that the proposed framework can accurately estimate unobserved system states even under reduced number of metered buses, achieving low average errors while significantly reducing solution times through linearization techniques. The integration with OpenDSS enables the validation of the implementation in an industry-standard tool of equivalent feeder models with aggregate DER models representing innumerable invisible resources with high numerical accuracy and reduced solution times. Overall, this work advances the state-of-the-art by providing a scalable, data-efficient modeling approach that empowers distribution system operators to maintain reliable and efficient grid operations despite the growing presence of invisible DERs.

Keywords: Distributed Energy Resources, Equivalent Network, Load Flow Calculation, Optimization, Radial Distribution Systems, State Estimation.

LIST OF FIGURES

Figure 1.1. Comparison of traditional and future workflows for power flow scenario analysis in DER-rich distribution feeders.	19
Figure 1.2. An overview of the research reported in this document, along with the corresponding inputs and outputs of the models in each section.	19
Figure 2.1. Diagram of a radial distribution system with hidden sections containing invisible DERs.	23
Figure 2.2. (a) Radial distribution feeder with observed and unobserved sections, including invisible DERs; (b) equivalent system representation with aggregate DER models.	26
Figure 2.3. Power output of equivalent DERs as piecewise linear functions of weather inputs.	29
Figure 2.4. Relative voltage magnitude estimation errors for all metering scenarios combining the NY values of 1 and 9.	36
Figure 2.5. Boxplots of relative branch flow estimation errors for all metering scenarios combining the NY values of 1 and 9.	37
Figure 2.6. Actual and estimated branch active power flows and equivalent DER output for different NY.	38
Figure 2.7. Branch active power flow estimation accuracy (ACC, %) considering the variation of NY and the number of metered buses.	39
Figure 2.8. Branch flow estimation errors for different DER penetration levels.	40
Figure 2.9. Equivalent DER allocation for distinct locations of hidden DERs in the 25% DER penetration scenario.	42
Figure 2.10. Average branch flow estimation and processing time values varying the length of historical data.	43
Figure 2.11. Branch flow estimation error for different SNRs.	44
Figure 2.12. Comparison of results considering branch flow measurements at the substation exit.	45
Figure 2.13. Relative branch flow estimation errors when maximizing and minimizing the squared RMSE metric.	46
Figure 2.14. Example of application for the modified IEEE 69-bus radial test system with a reduced number of equivalent DERs.	47
Figure 2.15. Branch active power flow estimation accuracy (ACC, %) varying the number of metered buses and the choice of NY.	48
Figure 2.16. Estimation accuracy across the seven feeders considering different meter coverage levels.	49
Figure 3.1. Comparison of the power output profiles of the different hybrid PV-BESS models. The BESS charges exclusively through PV output and discharge between hours 16 and 21.	53
Figure 3.2. TOU pricing signal used to schedule BESS DERs.	55
Figure 3.3. Diagram of the test system displaying DER information.	57
Figure 3.4. Relative estimation errors under different meter coverage levels. (a) Observed voltages. (b) Unobserved voltages. (c) Unobserved branch flows.	60
Figure 3.5. Branch flow estimation accuracy (ACC, %) varying the meter coverage level and the maximum number of aggregate DERs (setting of parameter NY).	61

Figure 3.6. Comparisons between the proposed method and other linearized and nonlinear approaches: (a) unobserved branch flow estimation errors, and (b) solution time.	67
Figure 3.7. Branch flow estimation accuracy for all system feeders considering varying meter coverage levels. 50% DER penetration scenario.	68
Figure 4.1. System plots for the 240-bus case with DER locations marked. (a) Detailed system with 128 DERs. (b) Reduced version with 15 equivalent DERs.	73
Figure 4.2. 10-bus test feeder. (a) Original system. (b) Reduced system.	76
Figure 4.3. 873-bus test system. (a) Original system. (b) Reduced system.	77
Figure A.1. Representation of the l -th line segment connecting buses i and j	88

LIST OF TABLES

Table 2-1. Summary of the availability of system and DER information.	23
Table 2-2. Expected inputs and outputs of the proposed MINLP formulation for aggregate modeling of invisible DERs.	32
Table 2-3. Load and DER capacity per bus at the MV level for the 10-bus test feeder.	33
Table 2-4. Load and DER capacity per feeder for the 873-bus distribution test system.	34
Table 2-5. Estimation Performance: 10-bus system, NY = 9, 100% MCL	35
Table 2-6. Absolute error values for the estimation of branch flows for different spatial distribution of PV DERs	41
Table 2-7. Estimation Performance: 69-bus system, NY = 5, 100% MCL	48
Table 2-8. Branch Flow Estimation Performance: 873-bus system, NY = 5 (per feeder), 100% MCL.....	49
Table 3-1. Load and DER capacity per feeder for the 240-bus distribution test system.	58
Table 3-2. Estimation Performance: 11-bus system, NY = 9, 100% MCL, 40% DER.	59
Table 3-3. Estimation Performance: 11-bus system, NY = 9, 100% MCL, Varying DER Penetration	62
Table 3-4. Estimation Performance: 11-bus system, NY = 9, 100% MCL, 40% DER, Noise in Measurements.....	63
Table 3-5. Estimation Performance: 11-bus system, NY = 9, 100% MCL, 40% DER, Branch Impedance Errors	64
Table 3-6. Estimation Performance: 11-bus system, NY = 9, 100% MCL, 40% DER, Total DER Capacity Errors	64
Table 3-7. Estimation Performance: 11-bus system, NY = 9, 100% MCL, 40% DER, Net Load Measurement Errors	65
Table 3-8. Formulation Comparison of the Proposed Model and Linear and Nonlinear Benchmark Approaches	66
Table 4-1. Branch Flow Deviations Under Different System Unbalance Levels	74
Table 4-2. Power flow results obtained using OpenDSS. Balanced three phase models for the 10- and 873-bus test systems.	75
Table 4-3. Solution time after 10,000 power flow calculations in OpenDSS.....	76
Table B-1. Summary of the average monthly 24-hour weather data profile used in simulations.	92

LIST OF ABBREVIATIONS AND ACRONYMS

ACC	Accuracy
AE	Absolute Error
AMI	Advanced Metering Infrastructure
BESS	Battery Energy Storage System
Bonmin	Basic Open-Source Nonlinear Mixed Integer Programming
COIN-OR	Computational Infrastructure for Operational Research
DER	Distributed Energy Resource
DSO	Distribution System Operator
IED	Intelligent Electronic Device
LV	Low Voltage
MAE	Mean Absolute Error
MCL	Meter Coverage Level
MILP	Mixed Integer Linear Programming
MINLP	Mixed Integer Nonlinear Programming
MIP	Mixed Integer Programming
MV	Medium Voltage
MW	Megawatt
PMU	Phase measurement Unit
PV	Photovoltaic
RE	Relative Error
RMSE	Root Mean Square Error
TOU	Time of Use
WT	Wind Turbine

LIST OF SYMBOLS

Indices

i, j, k	A system node. i refers to a bus located at the first end of a line, while k refers to a bus at the second end of a line.
l	A system line.
t	A time interval.
x	A hybrid PV-BESS model.

Parameters

B, G	Shunt susceptance and conductance elements of a line.
$h_t, h^r, \underline{h}$	Instantaneous, rated, and minimum threshold direct solar irradiance values for the calculation of the power output of PV DERs.
K^U	Weight matrix for voltage magnitude measurement residuals in the objective function.
K^{PV}	Coefficient for calculating the power output of PV DERs.
K^{WT}	Coefficient for calculating the power output of WT DERs.
M	Modified bus incidence matrix.
NM	Number of metered buses.
NDT	Number of time intervals in historical data.
NY	Maximum number of buses candidate for equivalent DER allocation.
p_t^{PV}	Per unit active power output of PV DERs at time t
p_t^{WT}	Per unit active power output of WT DERs at time t
R_l, X_l, Z_l	Resistance, Reactance, and Impedance elements of a line, respectively.
ToU_t	Time of use pricing signal to schedule BESS DERs.
t_x	BESS start charging time for the hybrid DER model x .
$\tilde{U}$	Matrix with squared voltage magnitude measurements.
$w_t, w^r, \underline{w}, \overline{w}$	Instantaneous, rated, cut-in, and cut-out wind speed values for the calculation of the power output of WT DERs.

Variables

D_k^L	Share of the total peak feeder load demand allocated at bus k .
$J_{l,t}$	Square of the magnitude of the current flowing in line l at time t .
$P_{l,t}^F, P_{l,t}^S$	Active power flow in the first and second ends of a line at time t , respectively.
P_t^F, P_t^S	Column vectors containing active power branch flows for all system lines at time t .
P_t^L	Column vector containing all load active power injections at time t .
$P_{k,t}^L$	Active power consumption at node k and time t .

P_t^L	Total active power feeder load at time t .
$\mathbf{P}_t^{PV}$	Column vector containing all PV power injections at time t .
$P_{k,t}^{PV}$	Active power output of the equivalent PV DER at bus k , time t .
$\overline{P}_k^{PV}$	Maximum generating capacity of the equivalent PV DER at bus k .
P_{total}^{PV}	Total installed PV capacity at the feeder level.
$\mathbf{P}_t^{WT}$	Column vector containing all WT power injections at time t .
$P_{k,t}^{WT}$	Active power output of the equivalent WT DER at bus k , time t .
$\overline{P}_k^{WT}$	Maximum generating capacity of the equivalent WT DER at bus k .
P_{total}^{WT}	Total installed WT capacity at the feeder level.
$\mathbf{P}_t^{BESS}$	Column vector containing all BESS power injections at time t .
$P_{k,t}^{BESS}$	Active power output of the equivalent BESS DER at bus k , time t .
$\overline{P}_k^{BESS}$	Maximum discharge capacity of the equivalent BESS DER at bus k .
P_{total}^{BESS}	Total installed BESS capacity at the feeder level.
$Q_{l,t}^F, Q_{l,t}^S$	Reactive power flow in the first and second ends of a line at time t , respectively.
$\mathbf{Q}_t^F, \mathbf{Q}_t^S$	Column vectors containing reactive power branch flows for all system lines at time t .
$\mathbf{Q}_t^L$	Column vector containing all load reactive power injections at time t .
$Q_{k,t}^L$	Reactive power consumption at node k and time t .
Q_t^L	Total reactive power feeder load at time t .
$U_{k,t}$	Squared bus voltage magnitude at bus k and time t .
$\mathbf{U}$	Matrix with squared voltage magnitude estimates.
$W_{j,l,t}$	Auxiliary variable for the linearization of the calculation of the squared line current magnitude using McCormick envelopes.
$\mathbf{Y}^{DER}$	Column vector containing all binary variables for the allocation of equivalent DERs.
Y_k^{DER}	Binary decision variable for allocating equivalent DERs.
Y_k^L	Binary decision variable for allocating load models.

CONTENTS

1. INTRODUCTION.....	14
1.1. Motivation.....	14
1.2. Literature Survey.....	16
1.3. Thesis Objective.....	18
1.4. Thesis Structure.....	19
2. A MINLP MODEL FOR THE AGGREGATE MODELING OF INVISIBLE DERS.....	21
2.1. Chapter Introduction.....	21
2.2. Theoretical Background.....	21
2.3. Problem Statement.....	25
2.4. Solution Method.....	30
2.5. Results and Discussion.....	35
2.6. Chapter Summary.....	50
3. A MILP MODEL FOR ENHANCED AGGREGATE MODELING OF INVISIBLE HYBRID DERS....	51
3.1. Chapter Introduction.....	51
3.2. Linearization in Power Distribution Systems Modeling.....	51
3.3. Problem Statement.....	52
3.4. Solution Method.....	56
3.5. Results and Discussion.....	58
3.6. Chapter Summary.....	69
4. INTEGRATION INTO AN INDUSTRY-STANDARD TOOL FOR DISTRIBUTION SYSTEMS MODELING AND ANALYSIS.....	70
4.1. Chapter Introduction.....	70
4.2. Integration with OpenDSS.....	70
4.3. Comparisons Between Detailed and Reduced System Models.....	71
4.4. Chapter Summary.....	78
5. CONCLUSION.....	79

5.1. Chapter-Wise Summary	79
5.2. Thesis Contributions.....	79
5.3. Directions for Future Research.....	80
6. REFERENCES.....	82
APPENDIX A—DEDUCTION OF THE LINE-WISE POWER FLOW EQUATIONS	88
APPENDIX B—TEST SYSTEM DATA USED IN SIMULATIONS	91
APPENDIX C—MAIN PUBLICATIONS RELATED TO THIS RESEARCH	94

1. INTRODUCTION

1.1. Motivation

The advancing modernization and expansion of electrical power systems already brought significant changes to the design and operation of power distribution systems. Adopting grid-edge technologies and distributed generation has set up a transition from a passive distribution system to a more dynamic one, where distribution utilities must actively manage power flows while integrating a growing share of Distributed Energy Resources (DERs) [1]. Although the pace of these transformations varies across different countries—with more developed economies enabling faster shifts in customer consumption patterns—emerging economies tend to gradually follow a similar trajectory, facing their own unique challenges in updating electricity infrastructure while ensuring energy equity and resilience to an increasingly unpredictable climate [2], [3].

These transformations accelerate as electricity replaces fossil fuels in transportation, heating, and industry [4]. The deep electrification of the economy, complemented by the rise of high-consumption, high-reliability loads—such as data centers and electric vehicle (EV) fast-charging hubs—introduces new challenges by increasing peak demand and straining local grid infrastructure [5]. In other words, the ever-growing interdependence between the power sector and key aspects of society further increases the need for improved power system practices, as disruptions in power supply may have cascading effects across multiple sectors. Given this landscape, the role of Distribution System Operators (DSOs) becomes even more critical [6], [7]. Their ability to maintain a reliable power supply will depend on enhanced system modeling, monitoring, control, and operational flexibility [8].

When deployed in a coordinated manner, DERs may prove valuable for DSOs in dealing with voltage regulation and congestion management [9]. However, the impracticality of monitoring and integrating DER data at scale leaves many of these resources invisible to the operation and challenges system observability requirements for effective decision-making. This, associated with the invisible nature of small-scale DERs, can lead to unexpected power flows and potential voltage and grid unbalance problems [10], [11]. Invisible DERs refer to resources typically installed within the distribution utility's service area, being unmonitored due to the limited capability to collect, process, and integrate data from many small-scale resources. DERs might also remain invisible when connected to feeder sections that are managed by third-party entities with limited data exchange with primary utility.

In such cases, granular data related to each DER site remains unknown, such as real-time power output and control settings, introducing challenges related to data availability and integration with power flow and optimal power flow tools DSOs may employ [12]. Although not accurately representing invisible DERs might not be detrimental to the analysis of distribution systems under low DER penetration, some

utility practices are affected in scenarios with increased DER participation [13]. Examples include resource allocation [14], load transfer [15], voltage regulation [16], overcurrent protection [17], and system restoration [18].

Considering the survey results from [19], most distribution utilities tend to only monitor in real time and explicitly model DER units exceeding 250 kW of generating capacity. However, there are cases where only DERs above 1 MW of capacity are included in system models, leaving a larger number of existing DERs invisible to the operation. As a result, while network information and the total installed DER capacity at the feeder level might be known, the individual contributions and collective impact of small-scale DERs are not directly observable. On the other hand, certain feeder sections might be operated by third-party entities that restrict data sharing with the primary utility. These network sections are typically treated as hidden from the utility's perspective, with limited information about the DERs and loads they contain. The only DER information that might be accessible in these scenarios is the peak DER generating capacity for each type of DER present, requiring equivalent models to address their impact in the main trunk feeder.

For example, in Ontario, Canada, power distribution utilities classify feeders into M- and F-class [20]. M-class feeders contain sections managed by third-party companies, such as industrial or commercial facilities or private operators controlling part of the distribution infrastructure. These feeders tend to be more complex, particularly due to lacking system visibility; however, F-class feeders, despite being primarily controlled by the distribution utility itself, may still face challenges in assessing the impact of DERs on grid performance due to the presence of innumerable unmonitored small-scale resources. Examples of distribution feeders facing massive integration of small-scale DERs can be found in Brazil after the country-wide uptake of the solar photovoltaic (PV) market since 2022—after the enactment of Law no. 14.300/2022 [21]. This law established the legal framework for distributed generation. On average, there were two PV installations per minute at the power distribution level until the end of 2024 [22], adding up to three million individual PV installations. Since the average installed generating capacity is around 12 kW per system (total country-wide aggregated total is approximately 34.8 GW [23]—considering distributed generation only), most installations remain unmonitored and thus invisible to local operators.

In addressing the challenge of improving the modeling fidelity of distribution systems in the presence of invisible DERs, much of the existing research literature focuses on developing equivalent models and methods to estimate their impact in steady-state. Many studies have proposed aggregate DER representations, relying on statistical and heuristic approaches to estimate power generation and consumption patterns of invisible resources. Other works have explored state-estimation techniques in the presence of invisible DERs, often using algorithms that can infer the presence and behavior based on real-time network conditions and historical data.

1.2. Literature Survey

The efforts to improve system visibility, considering the presence of invisible DERs, have gained prominence in recent years, and three main classes of methods are identified in the literature. Some methods aim to identify the aggregate power output of invisible DERs downstream to a metering device by disaggregating DER generation from flow measurements [25]-[31]; others aim to improve system visibility by identifying the location of invisible DERs using voltage sensitivity data [32]-[35]. The third class of methods focuses on distribution system state estimation in the presence of invisible DERs [36]-[43].

Ref. [24] employs a linear regression method to extract the peak aggregate PV capacity in a region downstream to a metering device, assuming the availability of accurate weather (e.g., solar irradiance) and line current measurements. Ref. [25] proposes an improved regression method considering the spatial correlation between the aggregate output of hidden PVs and data from neighboring (visible) PV plants. Similarly, regression models are used in [26] to estimate the peak generating capacity of PV and EV charging stations aided by an artificial neural network (ANN) model for feature extraction. The authors in [27] employ a statistical approach for behind-the-meter unit detection and capacity estimation using data-mining heuristics. In [28], a parametric PV model is built based on manufacturer data to approximate simulations to existing field measurements and estimate behind-the-meter PV generation. With equivalent parameters determined at multiple locations, the authors employ a clustering method to obtain aggregate models and estimate their impact at the feeder level. Ref. [29] defines a baseline demand curve from which PV generation is obtained when a significant deviation from historical recordings is identified. In [30], a machine learning model trained with feeder loading data is used to estimate the power output of unmonitored PV sites and is validated under various weather conditions.

The limitation of such methods arises from their pre-definition of areas of interest for analysis (e.g., areas that the system modeler knows/suspects an unmonitored DER may exist) and their reliance on data from visible DER sites to support model training and tuning. These limitations are addressed in [31], where branch flow measurements at the substation exit are disaggregated into net load and DER output data per DER type in the feeder. This accurately estimates the steady-state behavior of existing DERs, eliminating the need for granular data to assess the impact of invisible DERs at the feeder level.

Regarding models based on voltage sensitivity information, the authors in [32] determine buses to which invisible DERs are connected by assuming a linear relationship between node voltage magnitudes and active power injections from unmonitored resources. This approach, however, requires power flow calculations to determine bus voltage sensitivities. Algorithms for voltage magnitude [33], [34] and voltage sensitivity estimation [35] considering the presence of DERs improve the solution. However, these cannot be directly applied when considering that not all DER

data might be available for power output estimation. Hence, equivalent models must be developed to represent the impacts of these resources and yield numerically accurate power flow calculations.

To enable state estimation in low-observable distribution systems with DERs, Ref. [36] employs a weighted least absolute value estimator using pseudo-measurements to update missing system and DER parameters and improve power flow calculations. Pseudo-measurements are generated using average historical load data, while in [37] they are generated based on assumed correlations between existing measurements, the power output of visible DER sites, and unmetered system locations and DERs. The authors discuss the existence of smoothed behavior resulting from aggregating small, less predictable invisible DERs and loads connected at the low-voltage level.

Other state estimation methods employ machine learning models trained using time-series power flow calculations [38], physics-informed neural networks [39], and probabilistic models considering sampling of low-fidelity customer-side data [40]. Other approaches leverage accurate system topology and line impedance data to propagate known voltage values across similar buses and enhance system observability [41], [42]. Ref. [43] proposes a coupled power flow formulation, combining two subsequent snapshots for recovering system states over short-time intervals. A set of determined parameters is introduced by considering that all injections from unmetered nodes are stationary. Then, a state estimation problem is formulated to infer power injections, assuming that the variability of measurements arises from the variability of the power output of invisible DERs.

While these methods integrate low-fidelity information to enhance system visibility, considering the presence of DERs, their performance is limited by increased data requirements, assumptions over the behavior of unmetered nodes, and measurements at DER locations or buses with similar voltage profiles. Refs. [36]-[42] also require full system observability. Even though [43] does not have the same requirements, it relies on consecutive power flow snapshots, and all non-metered buses must be connected to a metered one through vertex-disjoint paths (i.e., paths with no common shared nodes, except for the source and end locations). Moreover, many of these methods are computationally intensive and operate on detailed network models without offering a pathway to reduced-order representations—which is essential for improving scalability and integration into existing utility workflows.

Given this context, there is an opportunity to explore novel methods to establish equivalent aggregate representations of invisible DERs to enhance their visibility and improve the steady-state analysis of power distribution systems with invisible DERs. Equivalent DER models representing innumerable invisible resources can be integrated into traditional power flow and optimal power flow solution algorithms to improve the fidelity of the resulting system models. This research explores Mixed-Integer Programming (MIP) formulations that enable optimally allocating (locating and

sizing) aggregate DER models and estimating unobserved system states. This research also explores how these models can be integrated into utility-standard power systems modeling and analysis tools, such as OpenDSS [44], for enhanced steady-state studies involving scenario analysis.

An important feature of the proposed framework is the control over the number of resulting equivalent invisible DER models, enabling the aggregate representation of these resources and reduction of the resulting system model. This leads to simplified yet accurate models that expedite the solution time required for power flow analysis. This proves beneficial for studies requiring many scenarios and the modeling of larger-scale systems with innumerable DERs.

1.3. Thesis Objective

The research reported in this thesis aims to enhance the steady-state analysis of radial distribution systems in the presence of invisible DERs by developing equivalent aggregate DER models that improve power flow calculations. Specifically, the research explores nonlinear and linearized MIP formulations to allocate aggregate DER models that capture the combined, steady-state impact of many invisible DERs. Another outcome of the proposed optimization formulations is the estimation of unobserved system states (e.g., node voltage magnitude and branch flows) without relying on dense measurement infrastructure. In this sense, the main objectives and highlights of this research can be listed as:

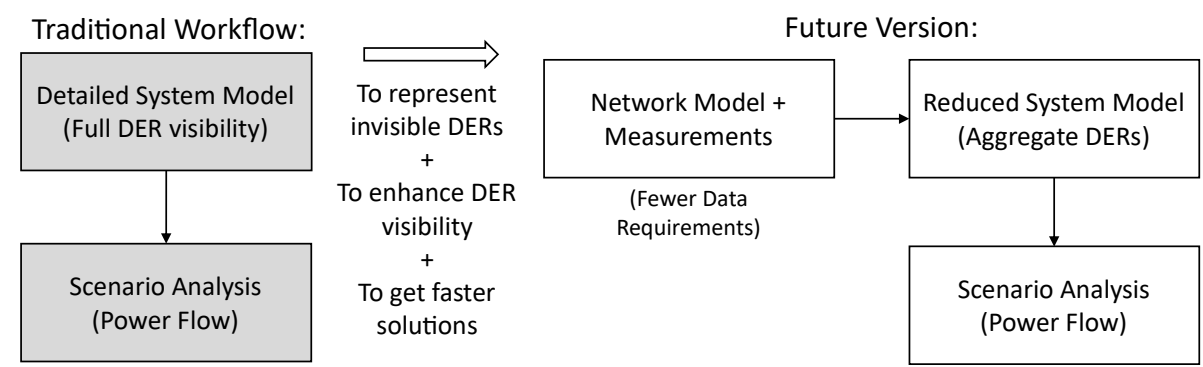
- To develop optimization-based methods to determine the locations and generating capacities of equivalent aggregate DER models representing innumerable invisible DERs.
- To improve the accuracy of steady-state system analyses by integrating the resulting equivalent models into traditional power flow and optimal power flow solution algorithms.
- To enable the estimation of unobserved system states in low-observable radial distribution networks without relying on dense measurement infrastructure.
- To ensure that the resulting outputs are compatible with utility-standard tools for power systems analysis and facilitate practical implementation.

Figure 1.1 contrasts the traditional and future (with the deployment of equivalent aggregate DER models) workflows for utilities to perform scenario analysis (e.g., power flow calculations) with DER-rich distribution feeders. In the traditional approach (left side), utilities rely on a detailed system model that requires full visibility of all DERs. While this method provides high accuracy, it depends on extensive data collection and processing, which becomes increasingly burdensome as more small, often unmonitored DERs are added to the system.

The future version (right side) proposes a more efficient method by using a simplified network model combined with available measurement data to build a reduced system model where DERs are aggregated. This approach lowers DER data

requirements and enables the representation of invisible resources while preserving system behaviors. The reduced model supports faster power flow solutions—due to the reduced number of elements represented—allowing utilities to run scenario analysis more efficiently—a key contribution of this research.

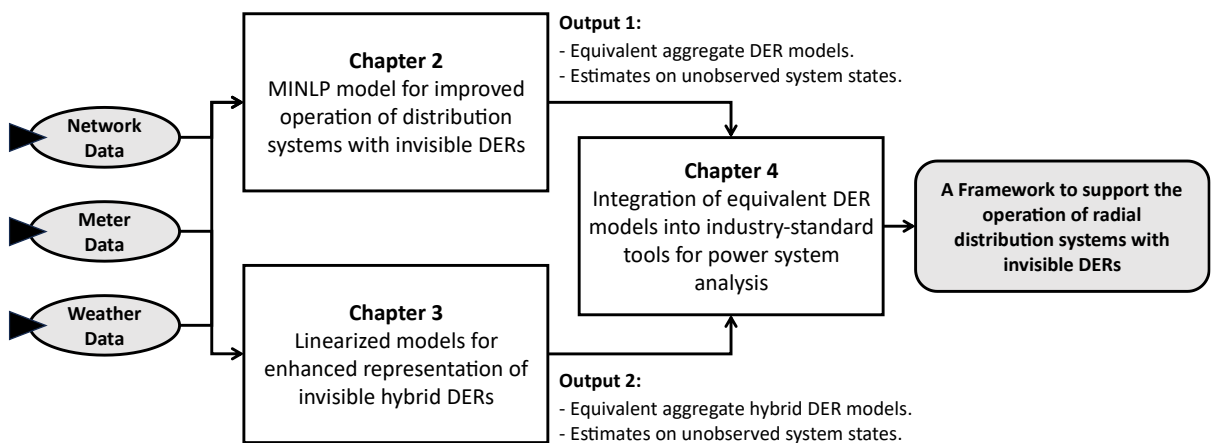
Figure 1.1. Comparison of traditional and future workflows for power flow scenario analysis in DER-rich distribution feeders.



1.4. Thesis Structure

Each chapter of this thesis follows a logical progression from theoretical background to methodology, validation, and practical application—ultimately demonstrating the effectiveness of the proposed framework in addressing the challenge of developing accurate aggregate models for invisible DERs. The interactions between each section, along with the corresponding inputs and outputs, are depicted in Figure 1.2.

Figure 1.2. An overview of the research reported in this document, along with the corresponding inputs and outputs of the models in each section.



Each chapter can be summarized as follows:

- **Chapter 2** presents the main convex mixed-integer nonlinear formulation, its modeling assumptions, and its outcomes in terms of developing equivalent aggregate DER models for radial distribution systems.
- **Chapter 3** explores linearized versions of the original formulation to enhance computational performance while yielding numerically accurate results. The modeling of hybrid DERs (e.g., PV+BESS) is introduced to better represent existing DER configurations.
- **Chapter 4** discusses the practical implications of the proposed methods, including their integration into OpenDSS and potential application cases for distribution utilities.
- **Chapter 5** concludes the thesis by summarizing key findings and suggesting future research directions.
- **Appendix A** provides the deduction of the line-wise power flow model used to represent the steady-state behavior of radial distribution systems. This section was included to inform the potential reader that the power flow equations employed in formulations yield exact solutions.
- **Appendix B** presents all the data used in simulations to allow the replication of results.
- **Appendix C** outlines the publications related to this research and other research outputs realized during this Ph.D. degree.

2. A MINLP MODEL FOR THE AGGREGATE MODELING OF INVISIBLE DERs¹

2.1. Chapter Introduction

This chapter presents the proposed Mixed-Integer Nonlinear Programming (MINLP) approach designed to develop aggregate models of invisible DERs and leverage the resulting equivalent models to enhance the operation and planning of radial distribution systems. The chapter begins by providing minimal introductions to the key concepts serving as the basis for the modeling aspects of the research. Since the proposed optimization approach relies on system measurements and historical data, discussing system visibility and state estimation issues in modern distribution systems is relevant. The rationale behind the chosen power flow model to represent the steady-state behavior of radial distribution systems is also discussed, along with algorithmic considerations regarding the convexity and computational complexity of optimization problems in the power systems context.

Following the theoretical discussion, the chapter details the proposed MINLP formulation, along with a presentation and discussion of the results. The resulting aggregate DER models represent the aggregated effects of many invisible DERs, enabling enhanced steady-state analysis and power flow estimation—offering practical advantages, as they can be integrated with existing distribution system analysis tools such as OpenDSS, as discussed in Chapter 4.

2.2. Theoretical Background

2.2.1. Distribution system visibility and state estimation issues

System visibility in power distribution networks refers to the extent to which a DSO can assess the state of the grid in real-time based on the availability of measurements for bus voltages, power injections, and branch flows. The level of visibility is determined primarily by the placement of metering infrastructure, the frequency of data collection, and integration into supervisory and control systems.

Unlike in transmission systems, the practical implementation of state estimation methods for distribution systems has traditionally been overlooked, as these networks operated predominantly under unidirectional power flows from substations to consumers, relying on simple radial configurations and predetermined operational settings [45]. However, with increasing penetration of DERs, power flows are not strictly unidirectional, and distribution systems have become more dynamic. New forms of flexible loads, demand response programs, and the impending deep electrification further contribute to increased operational complexity and uncertainty.

A common approach for increasing system visibility consists of using pseudo-measurements, which are synthetic or simulated measurements not directly derived from physical sensors but often based on historical data, forecasting, and statistical

¹ This chapter's contents were published in part in the paper: P. N. Vasconcelos, F. C. L. Trindade, B. Venkatesh, W. Freitas, A. C. Zambroni de Souza, and G. N. Taranto, "A Mixed-Integer Nonlinear Model to Support the Operation of Distribution Systems with Hidden DERs," *IEEE Transactions on Power Delivery*, vol. 40, no. 1, pp. 484-496, Feb. 2025. © 2025 IEEE.

assumptions of unobserved quantities. Recently, more researchers have explored state estimation techniques incorporating data from advanced metering infrastructure (AMI), phasor measurement units (PMUs), and intelligent electronic devices (IEDs) [46]. However, practical implementation by distribution utilities has been limited, as discussed in the industry survey results of [47], with most respondents stating that lack of data, data quality, and data integration issues are current limitations for accurate system modeling and estimation. Concerning DERs, most respondents reported that only utility-scale installations are actively monitored—typically above the 250 kW of peak generating capacity but above the 1 MW limit in some cases—leaving small-scale and behind-the-meter resources invisible to DSOs.

As a result, state estimation has become an area of interest for DSOs. However, achieving full observability is often impractical due to the high cost of deploying metering devices at scale [48]. This challenge presents a research opportunity: developing novel methods to leverage existing data (such as feeder-level aggregated information) and enhance steady-state system representation even under low-observable conditions—an outcome of this thesis.

2.2.2. The nature of invisible DERs

In this work, it is considered that invisible DERs in distribution systems stem from two primary conditions: the presence of small-scale DERs below the monitoring threshold and hidden feeder sections where visibility is restricted due to limited data access. Each condition presents unique challenges for real-time monitoring, modeling, and steady-state system analysis.

Small-scale DERs refer to units whose generating capacities fall below utility monitoring thresholds—commonly around 250 kW. These DERs are not individually metered or directly observable in real time. As a result, their power injections and control behaviors are unknown to system operators. While utilities may maintain aggregate information on installed DER capacity at the feeder level, they typically lack the integration of detailed site-specific data such as precise geographical location, generation profiles, fuel types, and control schemes. This lack of granularity creates significant uncertainty in assessing the operational impacts of these resources, especially considering scenarios at high DER penetration levels.

Hidden feeder sections further contribute to DER invisibility in the system. These, include parts of the network—often feeder laterals—that are either outside of the utility’s metering scope or managed by third-party entities. In such cases, data sharing between the third party and the primary utility may be limited or entirely absent. Although voltage magnitude measurements might still be available at specific points within the system (e.g., via protection relays, billing meters, or AMI devices), they often do not provide insight into the DER activity or load conditions within these unobservable sections. Figure 2.1 illustrates a representative radial feeder where voltage magnitude measurements are available along the bulk feeder section at the medium voltage (MV) level, while lateral branches remain unmonitored and potentially

host invisible loads and DERs. Note that, even in cases where voltage measurements are available at nodes with invisible DERs, these measurements might not be dedicated to DER monitoring.

Figure 2.1. Diagram of a radial distribution system with hidden sections containing invisible DERs.

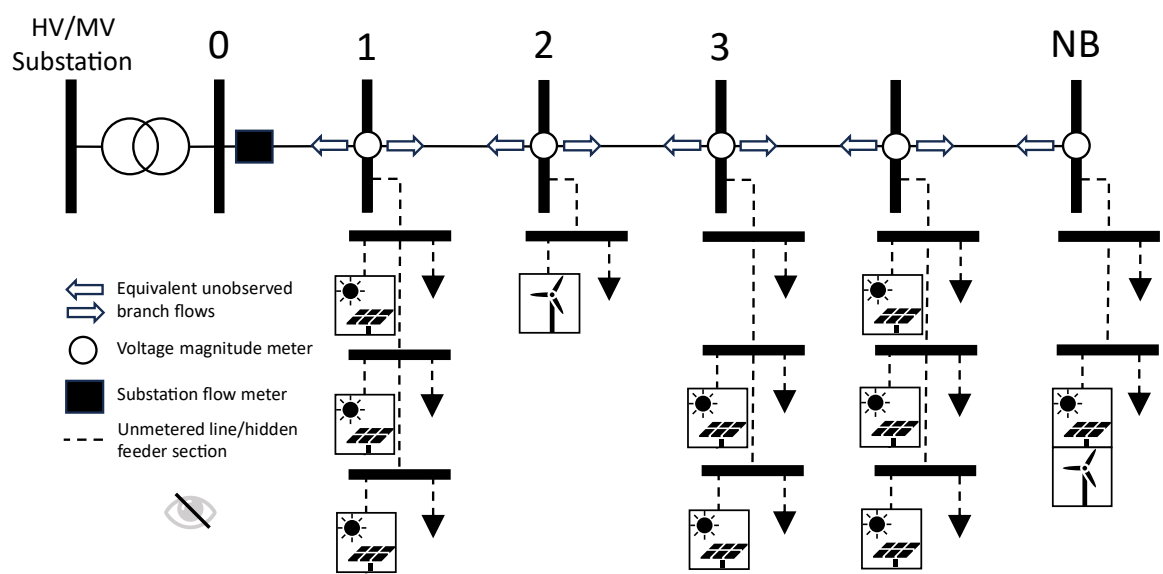


Table 2-1 provides an overview of the network, DER, and measurement information availability for the scenarios with small-scale DERs and hidden feeder sections, as assumed in this research. This distinction informs the modeling assumptions and the structure of the proposed aggregate DER models and their integration with traditional power flow solution methods.

Table 2-1. Summary of the availability of system and DER information.

Case	Network Information	DER Information	Measurement Availability
Small-scale DERs	Known feeder details (line impedances, topology, bus locations).	Utility does not monitor DERs; only aggregated peak DER capacities known.	Bus voltage magnitudes (if available); no direct DER measurements.
Hidden feeder sections	Main feeder details are known; limited or no data of third-party sections.	Utility cannot monitor DERs as data is inaccessible; only aggregated peak DER capacities known.	Bus voltage magnitudes (if available); equivalent representations required for unobservable sections.

2.2.3. Power flow model for radial distribution systems

A power flow model is a mathematical representation of an electrical network describing how power is distributed across the system under steady-state conditions. The model's equations must be sufficient to describe the relationship between bus voltages, power injections, and branch flows across all system components. There are many ways to categorize power flow models, with a frequent approach being separating bus injection and line-wise methods [49]. The former represents the power system regarding nodal power injections, while the latter tracks power flows along individual system lines, where Kirchoff's laws can be applied directly to model branch currents and voltage drops.

Although bus-injection models tend to be advantageous for large-scale power flow analysis since they leverage sparse matrix techniques for computational efficiency—while the Jacobian matrix for line-wise models tends to be relatively denser—the line-wise power flow formulation proposed in [50] leads to a more linearized set of equations. This property supports the integration into classical optimization problems. Such a feature exists due to the absence of voltage angle in the equations, obtained by rearranging the quadratic term referring to the voltage difference between two ends of a line. This step was first addressed in [51], referring back to the *DistFlow* equations for the line-wise model for radial distribution systems proposed by [52].

Appendix A provides a deduction of the line-wise power flow equations used in the foregoing sections to model radial distribution systems with DERs. This appendix was introduced to discuss the model and evidence that no simplifying assumptions are made. As discussed in [53], these equations yield an exact model of radial distribution systems even in the conditions of reverse power flow. The reader is referred to [54] for a line-wise model addressing meshed systems.

2.2.4. Convexity in optimization problems

Convexity is a key concept in optimization that affects the tractability and solvability of optimization problems. An optimization problem is considered convex if its objective function is convex and the feasible region, defined by its constraints, is also convex. This means that any local minimum is also a global minimum—a desirable property for optimization, which ensures the algorithm converges to an optimal solution without getting trapped in suboptimal points. Formally, a function f is convex if, for any two points in its domain, x and y [55],

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y), \quad \forall \alpha \in [0,1].$$

This means that any point on the straight line between $(x, f(x))$ and $(y, f(y))$ is greater than or equal to the function value at the corresponding point between x and y . If f is twice-differentiable, it is only convex if its Hessian matrix (constituted of all its second-order partial derivatives) is positive semidefinite, i.e., $\nabla^2 f(x) \succeq 0$.

In the context of power systems, convexity depends on the power flow equations used to describe the system's behavior. The traditional bus-injection model often leads to nonconvex optimization problems, especially because the power flow equations are inherently nonlinear, involving products of voltage magnitudes and trigonometric functions with phase angles [56]. In contrast, the aforementioned line-wise model yields equations that are more amenable to convex optimization techniques [57]. The transformation of power flow equations into convex formulations is primarily performed using relaxation and linearization techniques. A more detailed discussion of traditional linearization methods applied when addressing radial distribution systems is presented in Chapter 3.

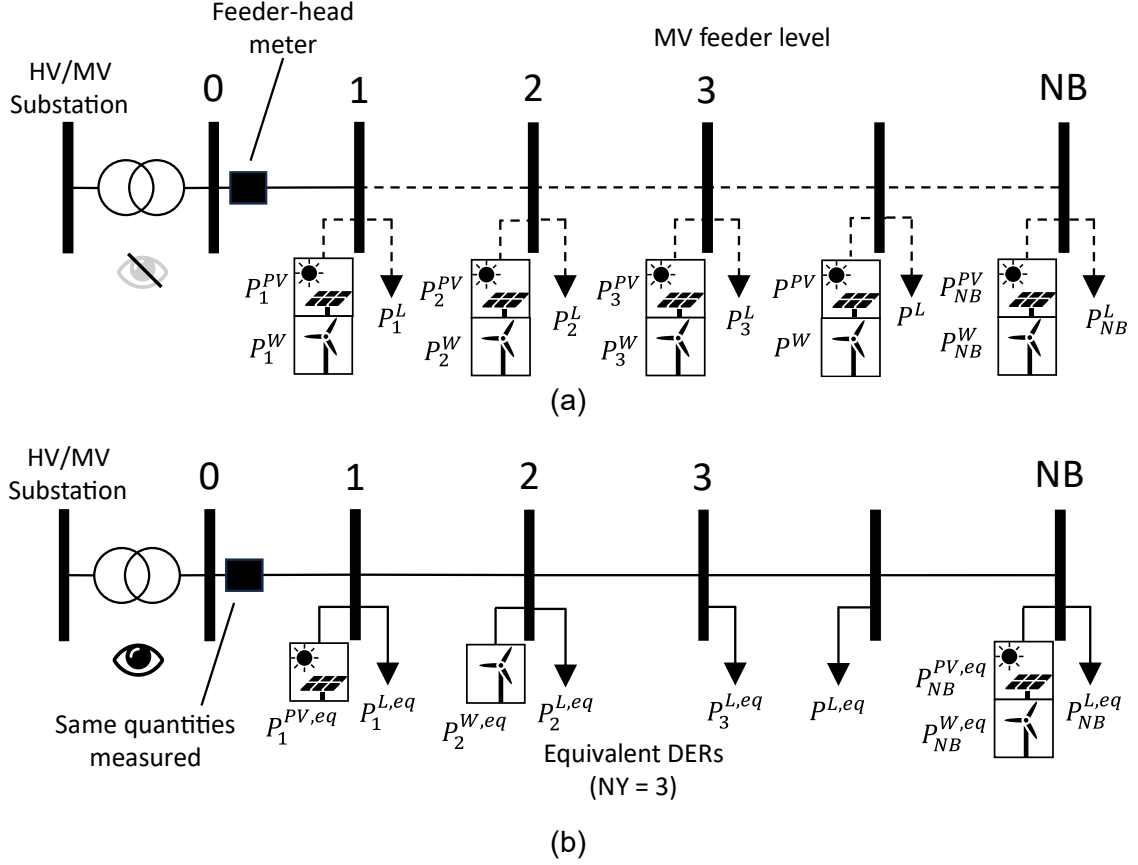
Relaxation techniques are employed to convexify inherently nonconvex problems, enabling the use of efficient convex optimization methods. These techniques often involve approximating or reformulating the original problem to obtain a convex counterpart that is theoretically easier to solve. For the scope of this research, the set of line-wise power flow equations of [50] is convexified using a second-order conic relaxation of the expression used to calculate the squared magnitude of line currents, as described in [61]. According to the authors, this approach yields exact solutions given that certain conditions are met, including a radial system topology and upper bounds on bus voltage magnitudes. By leveraging convex characteristics, the proposed optimization framework—described in the following sections—balances accuracy and computational feasibility, making it suitable for real-world applications. The MINLP formulation for aggregate modeling of invisible DERs is presented next.

2.3. Problem Statement

Consider Figure 2.2, identifying the observable and unobservable sections of a representative distribution feeder comprised of NB buses. The diagram of Figure 2.2(a) represents the main trunk of the feeder, with each node representing unmetered laterals with loads and DERs. The system visibility is limited because only one meter is present at the substation end-node. Under normal operation with few DERs, the system topology and estimations can be used to infer the state of unobserved voltages and branch flow magnitudes. However, estimating the state of the distribution system becomes a challenge with the massive integration of invisible DERs.

Despite this challenge, it is possible to obtain NY equivalent representations of invisible DERs, as depicted in Figure 2.2(b), which leads to accurate power flow calculations. The measured power system quantities (e.g., voltage magnitudes and substation flows) are matched while the remaining unobserved variables are estimated considering the steady-state output of equivalent DERs. This is the basis of the MINLP formulation presented in this chapter.

Figure 2.2. (a) Radial distribution feeder with observed and unobserved sections, including invisible DERs; (b) equivalent system representation with aggregate DER models.



2.3.1. Objective function

The objective function minimizes the sum of weighted squared measurement residuals, encompassing estimates on measured voltage magnitudes. In instances where measurements for active and reactive power injection or branch flows exist, their respective squared residuals can be incorporated into (2.1) following the same structure.

$$\text{minimize } \|K^U \odot (U - \tilde{U})\|_F^2 \quad (2.1)$$

where $\odot$ represents the element-wise product of two matrices of the same dimensions. Subscript F denotes the Frobenius norm, computed as the sum of squares of all elements of the resulting matrix.

Matrices U and $\tilde{U}$ contain squared voltage magnitude estimates and measurements, respectively, referring to each metered bus with NDT data entries in historical data. Matrix K^U also has dimension $NDT \times NM$ and contains weights for calculating of the weighted Frobenius norm of the measurement residuals matrix. This is a symmetric and positive definite matrix. Larger weight values assign relative

importance to each dataset, ensuring that more accurate measurements are more influential on the results. This is particularly useful when a heterogeneous combination of voltage measurements exist—with some meters having higher accuracy class—and to allow the introduction of lower-fidelity data, such as pseudo-measurements. Note that the analyses presented in this chapter assume that all voltage magnitude measurements have the same weight value. This implies that the corresponding metering devices available in the targeted system have the same accuracy class.

2.3.2. Power flow and system operation constraints

Constraints (2.2)-(2.7) model the electrical behavior of a radial distribution system and ensure the solution is physically feasible. These equations correspond to the line-wise power flow model proposed by [50]. An important feature of the model is the absence of voltage angle, obtained by rearranging the quadratic equation referring to the voltage difference between two ends of a line (see Appendix A).

Eq. (2.2) is a product of applying Kirchoff's voltage law for the pi model of a line or transformer between buses i and j . Eqs. (2.3) and (2.4) account for the active and reactive power losses, respectively. Eqs. (2.5) and (2.6) deal with the node-wise active and reactive power balances, respectively, and were modified to explicit the power output of solar photovoltaic (PV) and wind (WT) DERs. Eq. (2.7) corresponds to the second-order conic relaxation for calculating the square of the magnitude of the current flowing in branch l and time t .

This relaxation yields a convex formulation and is adopted to enhance solver efficiency while ensuring optimality. As noted in Section 2.2.2., this relaxation is exact when addressing radial power systems with the line-wise power flow model, even in the conditions of reverse power flow [53], [61]. Note that these expressions are evaluated for every time interval in historical data. This notation was omitted here for clarity; however, all parameters and variables with subscript t are updated for each time interval.

$$U_{j,t} + 2 \left(P_{l,t}^S \cdot R_l + Q_{l,t}^S \cdot X_l - \frac{U_{i,t}}{2} \right) + J_{l,t} \cdot Z_l^2 = 0, \quad (i, j) \in l \quad (2.2)$$

$$P_{l,t}^F + P_{l,t}^S + R_l \cdot J_{l,t} = 0 \quad (2.3)$$

$$Q_{l,t}^F + Q_{l,t}^S + X_l \cdot J_{l,t} = 0 \quad (2.4)$$

$$[M] \begin{bmatrix} P_t^F \\ P_t^S \end{bmatrix} - U_t \cdot G = P_t^L - P_t^{PV} - P_t^{WT} \quad (2.5)$$

$$[M] \begin{bmatrix} Q_t^F \\ Q_t^S \end{bmatrix} + U_t \cdot B = Q_t^L \quad (2.6)$$

$$U_{j,t} \cdot J_{l,t} \geq P_{l,t}^{S^2} + Q_{l,t}^{S^2} \quad (2.7)$$

where U_j is the squared voltage magnitude at node j ; J_l is the squared magnitude of the current at line l ; Z_l , R_l , and X_l refer to the impedance, resistance, and reactance elements of a line, respectively; superscripts F and S indicate branch flows at the first

and second ends of a line. The susceptance and conductance shunt elements of a line are B and G , respectively.

Constraints (2.8) and (2.9) enforce operational limits on squared node voltage and branch current flow magnitudes, respectively. Traditional state estimation methods often do not impose these constraints, allowing limit violations. However, considering that the proposed method intends to determine aggregate DER models for accurate power flow calculations, these constraints ensure that equivalent models do not lead into unrealistic system operating conditions, improving the fidelity of the resulting model.

$$\underline{U}_k \leq U_{k,t} \leq \overline{U}_k \quad (2.8)$$

$$J_{l,t} \leq \overline{J}_l \quad (2.9)$$

2.3.3. DER allocation constraints

Eqs. (2.10)-(2.15) deal with the allocation of equivalent DERs. The active power injections from local solar and wind DERs are computed for every time interval using (2.10) and (2.11), respectively. The output is computed as the product of the maximum DER generating capacity allocated at bus k , and the corresponding functions direct solar irradiance and wind speed. Although these equations model the behavior of equivalent aggregate DER models, setting a limit for power generation—the results of variables with overbars—ensures that the actual generation constraints of existing DERs are enforced. Moreover, since the individual settings of each DER are unknown, voltage/power controls are not considered. However, reactive power constraints can be introduced to consider operation in constant power factor modes.

$$P_{k,t}^{PV} = \overline{P}_k^{PV} \cdot p_t^{PV} \quad (2.10)$$

$$P_{k,t}^{WT} = \overline{P}_k^{WT} \cdot p_t^{WT} \quad (2.11)$$

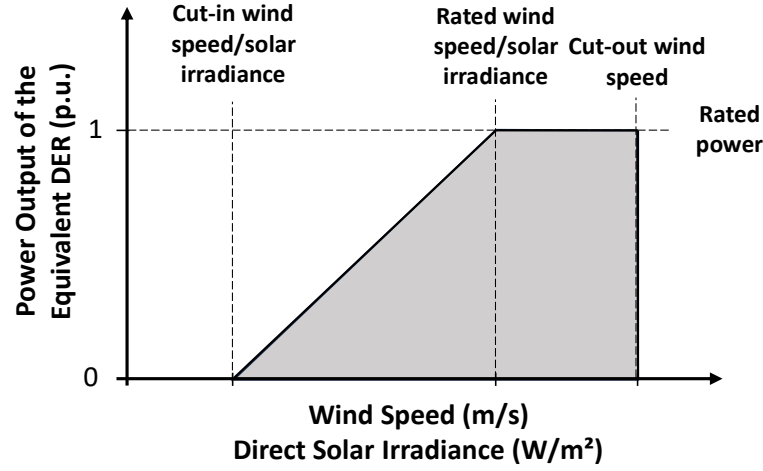
Piecewise linear functions can be derived from manufacturer data to model the power output of PV and W DERs as functions of direct solar irradiance and wind speed, respectively. The following expressions consider traditional cut-in, rated, and cut-out values for determining outputs, as depicted in Figure 2.3.

$$p_t^{PV} = \begin{cases} \max\{K^{PV} \cdot (h_t - \underline{h}), 0\}, & \text{if } h_t \leq h^r \\ 1, & \text{if } h_t \geq h^r \end{cases}$$

$$p_t^{WT} = \begin{cases} \max\{K^{WT} \cdot (w_t - \underline{w}), 0\}, & \text{if } w_t \leq w^r \\ 1, & \text{if } w^r \leq w_t < \overline{w} \\ 0, & \text{if } w_t \geq \overline{w} \end{cases}$$

where h and w refer to direct solar irradiance and wind speed values. The superscript r refers to rated values, while values with underbars and overbars refer to minimum and maximum, threshold values to determine non-zero output, respectively.

Figure 2.3. Power output of equivalent DERs as piecewise linear functions of weather inputs.



The total generating capacity per DER type at the feeder level is an input to the model. These values are defined as the parameters P_{total}^{PV} and P_{total}^{WT} for the total PV and WT capacity, respectively. Therefore, the sum of all aggregate DER models allocated by the proposed algorithm must equal these values, as expressed in (2.12) and (2.13). Notice that these summations are limited to $k = \{2, 3, \dots, NB\}$ to exclude the substation end-node as a candidate for equivalent DER allocation.

$$\sum_{k=2}^{NB} \bar{P}_k^{PV} = P_{total}^{PV} \quad (2.12)$$

$$\sum_{k=2}^{NB} \bar{P}_k^{WT} = P_{total}^{WT} \quad (2.13)$$

The maximum number of equivalent DERs that can be integrated into the system is controlled by the count of candidate buses, NY . This nonnegative integer parameter is introduced to constrain the binary decision variable $Y_k^{DER} \in \{0, 1\}$. Parameter NY limits the zero “norm” of $\mathbf{Y}^{DER}$ as per (2.14). This enables the peak generating capacity of a DER to be nonzero when this variable attains unity value for a specific bus k , as in (2.15). Here, C represents a sufficiently large number.

$$\|\mathbf{Y}^{DER}\|^0 \leq NY, \quad Y_k^{DER} \in \{0, 1\} \quad (2.14)$$

$$0 \leq \bar{P}_k^{PV}, \bar{P}_k^{WT} \leq C \cdot Y_k^{DER}, \quad C \gg 1 \quad (2.15)$$

Including a binary variable introduces complexity to the solution; however, this parameter is required since the proposed method intends to develop reduced equivalent representations of innumerable DERs using aggregate DER models that are adjustable depending on the targeted accuracy. If the control over the number of resulting DER models is not required, constraints (2.14) and (2.15) can be disregarded, converting the formulation into a linear problem with theoretically faster solution times.

However, this defeats the purpose of developing simpler yet accurate models for low-observable distribution systems with invisible DERs.

2.3.4. Load determination constraints

Next, constraints (2.16)-(2.19) deal with the customer load determination to distinguish their steady-state behavior from of the equivalent DERs. Constraints (2.16) and (2.17) update, for every time interval, the active and reactive power consumption for all load buses k , respectively. This considers the total aggregate consumption at the feeder level at a specific time instant, P_t^L and Q_t^L , and the share of the total load attributed to bus k , D_k^L .

$$P_{k,t}^L = D_k^L \cdot P_t^L \quad (2.16)$$

$$Q_{k,t}^L = D_k^L \cdot Q_t^L \quad (2.17)$$

Having a fixed value of D_k^L set to each bus assumes that the peak customer demand remains unchanged within the interval determined in historical data (e.g., hours to weeks of continuous system operation). This assumption simplifies many external (e.g., social and economic) factors outside of the scope of this research that may affect electricity consumption patterns and the rate of demand increase within a reduced time frame. Moreover, having local loads represented as shares of the total net load at the feeder level implies that similar consumption patterns are observed among each load node. Therefore, this assumption represents well distribution feeders constituted mostly by a single class of customers, e.g., residential, commercial, or industrial.

Since variable D_k^L corresponds to percentages, its summation over all the system load buses must equal unity value, as in (2.18). This ensures that equivalent load models approximate data from substation flow measurements. As per DER constraints, the decision whether the load consumption at bus k assumes a nonzero value is determined by a binary variable Y_k^L , which is defined in (2.19).

$$\sum_{k=2}^{NB} D_k^L = 1 \quad (2.18)$$

$$0 \leq \overline{D_k^L} \leq Y_k^L, \quad Y_k^L \in \{0,1\} \quad (2.19)$$

This approach assumes limited information on the actual locations and peak consumption of customer loads, as the only information used is a time series of net load measurements at the substation exit. The D_k^L values can be precomputed if accurate peak customer load data exists, removing the need for the binary variable for allocating loads.

2.4. Solution Method

The problem formulated in (2.1)-(2.19) represents a MINLP approach for developing aggregate DER models representing invisible DERs and estimating unobserved system states. This problem can be directly solved using a convex-MINLP

capable solver such as Bonmin [62]. This solver was selected based on its suitability for the problem type, prior use in the literature [63], availability—the solver is open-source, integrated into COIN-OR [64], a well-known collection of optimization tools—and its compatibility with different modeling frameworks. The formulation was implemented in Python [65], and Pyomo [66] was used to model the MINLP problem. The Bonmin solver was accessed through the AMPL Python API package [67].

Moreover, considering Bonmin's default internal threshold tolerance for convergence of 10^{-8} , the objective function is scaled by a factor of 10^6 to improve numerical stability. Scaling up the objective avoids arithmetic underflow problems and other precision limitations where rounding errors or numerical noise may dominate. This scaling factor was chosen based on a sensitivity analysis ranging from 10^3 to 10^9 . Results demonstrated that simulations with a weight of 10^6 tend to converge faster and provide similar solutions to scenarios with increased weight values. As the weight decreases (lower than 10^6), there is an increase in estimation errors, which is more significant than the decrease in convergence time considering the problem treated in this work. Considering the proposed formulation and Bonmin, multiplying by values larger than 10^9 affected convergence stability and, therefore, was not considered further.

2.4.1. Modeling assumptions

The simulation setup was designed based on the following assumptions: (1) the developed formulation supports the planning and operation of distribution systems under massive DER integration and must be compatible with existing advanced distribution management system platforms; (2) the aggregate solar and wind DER capacity at the feeder level is known, while specific details such as type, size, location, and power generation data are not available; (3) the operational topology of the target distribution system is radial; (4) all voltage magnitude reporting devices have the same accuracy class and therefore have the same weight in the objective function; and (5) measurement errors, when introduced, are independent and follow Gaussian distribution. Based on these assumptions, a summary of all inputs and outputs associated with the resulting MINLP formulation is provided in Table 2-2.

The maximum number of candidate buses is considered an optional input. By default, if a value for the parameter NY is not provided, the algorithm assumes that all buses (except the substation end-node) are candidates for equivalent DER allocation. Note that the feeder-level aggregated peak generating capacity of invisible DERs can be estimated using historical measurements at the substation exit, as reported in [31].

Table 2-2. Expected inputs and outputs of the proposed MINLP formulation for aggregate modeling of invisible DERs.

Inputs	Outputs
• System data: system topology, branch impedances, and bus voltage and branch current magnitude limits. • Operational data: feeder-level aggregated load (a time series) and generating power capacity of invisible DERs (a scalar value per DER type). • Measurement data: meter locations, if available, and time-series of the respective bus voltage magnitudes. • Weather data: time series of direct solar irradiance and wind speed for the same intervals as measurement and operational data. • Maximum candidate buses: the maximum number of buses eligible for equivalent DER allocation (optional).	• Equivalent DER allocation: definition of optimal locations and peak generating capacities of equivalent aggregate DER models representing innumerable invisible resources. • Reconstructed system states: time-series of observed and unobserved system variables considering measured and estimated data. This includes estimates on bus voltages magnitudes, power injections, and branch flows.

2.4.2. Test system data and preparation

In this chapter, the resulting formulation is employed to estimate system states and allocate equivalent DERs for a 10-bus, a 69-bus [68], and an 873-bus [69] test system. These were selected to test the proposed framework under different system scales and topological and operational characteristics. Note that detailed information on each test system is provided in Appendix B and the files used in simulation are available at [70].

The 10-bus test system consists of a 27.6 kV primary radial distribution feeder supplying 16 MW and 2 MVar of load. The total impedance of the conductors is $19.5 + j13.1$ Ohm (based on AWG 3/0 and 50-km length), equally split into nine branches. Two DER integration scenarios are devised. The first, accounts for a total aggregate generating capacity of 2.67 MW and 1.33 of PV and WT DERs, respectively. This results in a 25% DER penetration level, considering the ratio of the total DER generating capacity at the feeder level and the total peak aggregated load consumption. The second scenario foresees the massive integration of innumerable DERs, represented by a system with 100% DER penetration level, resulting in a total aggregate generating capacity of 10.67 MW and 5.33 MW for PV and WT DERs, respectively.

The resulting distribution of loads and DERs across the system is presented in Table 2-3. Note that bus 0 is omitted since it is assumed that there is no consumption and generation at the substation end-node. This information is solely used for

generating measurement data through power flow calculations. The line-wise power flow model of [50] was used for this purpose. Note that all values are in MW, except for Q^L , which is in MVar.

Table 2-3. Load and DER capacity per bus at the MV level for the 10-bus test feeder.

Bus	Pen	1	2	3	4	5	6	7	8	9
P^L		1.77	1.77	1.77	1.77	1.77	1.77	1.77	1.77	1.77
Q^L		0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.22
P^{PV}	25%	0.53		0.53		0.53		0.53		0.53
P^{WT}	25%	0.27		0.27		0.27		0.27		0.27
P^{PV}	100%	2.13		2.13		2.13		2.13		2.13
P^{WT}	100%	1.07		1.07		1.07		1.07		1.07

The IEEE 69-bus radial test system [68] has been extensively used in research, and was chosen to assess the performance of the proposed model when dealing with a larger radial distribution system with laterals. The system works at the nominal voltage of 12.66 kV with a total load of 3.802 MW and 2.694 MVar. It was modified to incorporate DERs placed on the nodes where the largest loads are located. The 25% DER penetration scenario has a total of 0.48 MW of both solar and wind DERs. DERs in the 100% penetration scenario remain connected to the same system nodes but were scaled to match the respective penetration values aggregated at the feeder level.

Lastly, the 873-bus radial distribution system of [69] was chosen to expand the application of the algorithm to the representation of a radial distribution system with multiple feeders connected to the same substation busbar. The system comprises seven feeders, supplying a total aggregated load of 33.6 MW + j19.8 MVar at the 27.6 kV level.

The system was modified to integrate 12.6 MW and 4.2 MW of solar and wind DERs, respectively (50% DER penetration scenario). DERs were placed in nodes with the highest loads up to 40% of the total number of nodes containing a DER. The individual generating capacities of invisible DERs were determined using the normal distribution. Each feeder has customer loads connected to every bus, except for the source node, with information summarized in Table 2-4. Each feeder was modeled and solved separately. Values are in MW, except for Q^L , which is in MVar.

These test systems and the respective DERs shared weather data from *renewables.ninja* [71], corresponding to the region of Toronto, Canada. This dataset consists of solar irradiance and wind speed values for one year of duration, divided into one-hour intervals (8760 time steps). The linearized DER power output calculation models considered cut-in and rated direct solar irradiance of 150 and 800 W/m², respectively. For wind DERs, cut-in, rated, and cut-out wind speeds are 2.8, 10, and 20 m/s, respectively.

Table 2-4. Load and DER capacity per feeder for the 873-bus distribution test system.

Feeder	F1	F2	F3	F4	F5	F6	F7	Total
NB	89	108	147	116	137	221	55	873
P^L	4.03	4.25	6.25	5.19	5.83	5.80	2.25	33.60
Q^L	2.16	2.53	3.48	3.21	3.51	3.63	1.28	19.80
P^{PV}	1.51	1.59	2.34	1.95	2.19	2.17	0.85	12.60
P^{WT}	0.50	0.53	0.78	0.65	0.74	0.72	0.28	4.20

2.4.3. Estimation performance metrics

The estimation performance is assessed with Mean Absolute Error (MAE) and Relative Error (RE). The Root-Mean-Square (RMSE), minimum absolute error (Min AE), and maximum absolute error (Max AE) values are used to gauge the dispersion result. The overall accuracy of estimations (ACC) is referred to as the complementary percentage value to the average relative error calculated using all time steps and system nodes or branches. For bus voltage magnitudes:

$$MAE_U = \frac{\sum_{t=1}^{NDT} \sum_{k=1}^{NB} |U_{k,t} - \tilde{U}_{k,t}|}{NDT \cdot NB} \quad [\text{p.u.}]$$

$$RE_{U_{k,t}} = \left| \frac{U_{k,t} - \tilde{U}_{k,t}}{\tilde{U}_{k,t}} \right| \quad [\%]$$

$$RMSE_U = \sqrt{\frac{\sum_{t=1}^{NDT} \sum_{k=1}^{NB} (U_{k,t} - \tilde{U}_{k,t})^2}{NDT \cdot NB}} \quad [\text{p.u.}]$$

$$\text{Min } AE_U = \min(|U - \tilde{U}|), \quad \text{Max } AE_U = \max(|U - \tilde{U}|) \quad [\text{p.u.}]$$

$$ACC_U = 1 - \frac{\sum_{t=1}^{NDT} \sum_{k=1}^{NB} RE_U}{NDT \cdot NB} \quad [\%]$$

For branch flows, replace voltage magnitude variables by P^F and NB by the number of lines in the system NT . The next sections present case studies along with estimation error results.

The choice over acceptable error thresholds for estimates depends on data availability, computational feasibility, and the specific application requirements within planning and operation of power distribution systems. For example, applications involving real-time control or safety-critical operations—such as voltage regulation, fault location and isolation, protection setting, and coordination—typically require higher accuracy estimations (e.g., <5%). Tighter error constraints are also expected under well-observed system conditions. On the other hand, there are applications where approximate solutions (e.g., <15%) still provide significant value. This is particularly valid in planning and trend analysis, especially considering the added benefit of integrating equivalent DER models representing invisible resources active in the system.

2.5. Results and Discussion

The first case studies (from Subsection 2.5.1 to Subsection 2.5.9) focus on the 10-bus system to showcase the performance of the resulting formulation under various conditions. Later sections are dedicated to larger test systems to expand the initial observations. The presence of voltage magnitude measurements is referred to as Meter Coverage Level (MCL). This may refer to an absolute value (the total number of metered buses in the system) or a percentage value corresponding to the share of the total number of buses that are being metered.

2.5.1. Ideal case (10-bus system)

Table 2-5 presents estimation results for when voltage magnitude measurements are available for the entire system (100% MCL). These measurements are assumed to be 100% accurate, and the algorithm is allowed to spread equivalent DERs across the entire network ($NY = 9$). Also, three months of historical data (2,160 one-hour intervals) are considered in simulations. The base voltage and base power for per unit calculations are 27.6 kV and 10 MVA, respectively.

Low absolute errors can be observed in both voltage magnitude and branch flow estimations for the 25% and 100% DER penetration scenarios. This highlights the capability of the proposed algorithm to replicate existing measurements and estimate unobserved branch power flows solely utilizing voltage magnitude measurements. These outcomes are dependent on model parameter configurations and target system characteristics, subjects explored in the subsequent sections.

Table 2-5. Estimation Performance: 10-bus system, $NY = 9$, 100% MCL

Metric	25% DER Penetration		100% DER Penetration	
	Voltage Magnitude Errors (p.u.)	Branch Active Power Flow Errors (p.u.)	Voltage Magnitude Errors (p.u.)	Branch Active Power Flow Errors (p.u.)
MAE	$1.25 \cdot 10^{-4}$	$3.22 \cdot 10^{-3}$	$7.09 \cdot 10^{-3}$	$1.38 \cdot 10^{-3}$
RMSE	$3.02 \cdot 10^{-4}$	$7.28 \cdot 10^{-3}$	$1.09 \cdot 10^{-2}$	$7.09 \cdot 10^{-3}$
Min AE	$1.96 \cdot 10^{-8}$	$1.17 \cdot 10^{-6}$	$2.45 \cdot 10^{-4}$	$5.71 \cdot 10^{-8}$
Max AE	$1.48 \cdot 10^{-3}$	$9.16 \cdot 10^{-2}$	$2.94 \cdot 10^{-2}$	$1.44 \cdot 10^{-1}$
ACC	99.99%	99.41%	99.99%	99.86%

2.5.2. Impact of the voltage measurement availability (10-bus system)

This subsection explores the sensitivity analysis concerning the quantity of voltage magnitude meters deployed in the system. Figure 2.4 and Figure 2.5 depict relative error values for voltage magnitude and branch active power flow estimations, respectively, across different MCLs and 25% and 100% DER penetration levels. The

number measurements ranges from two buses being monitored to the deployment of ten voltage meters across the entire system, and 3-month historical data.

For each number of metered buses, all possible combinations are computed for $NY = 1$ and $NY = 9$ by varying meter locations. It is assumed that the meter at the substation bus (bus 0) is always available. The results are presented as boxplots whose boxes mark the range of the central 50% of the data, with a central line marking the median value. Lines extend from each box to capture the interval of the remaining data, with '+' indicating outliers.

Figure 2.4. Relative voltage magnitude estimation errors for all metering scenarios combining the NY values of 1 and 9.

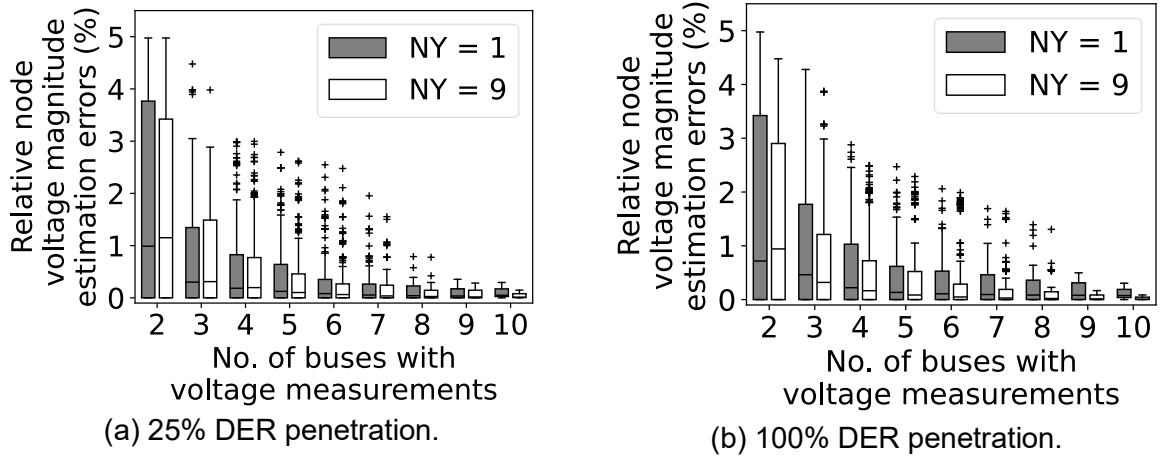
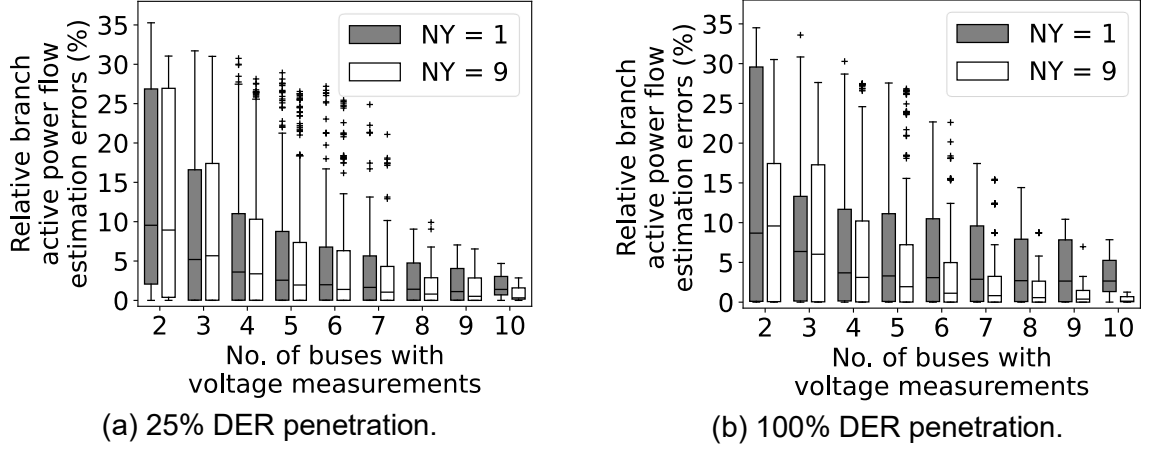


Figure 2.4(a) and Figure 2.4(b) show that the proposed formulation can estimate voltage magnitudes with errors up to 5% for both the 25% and 100% DER penetration levels and this error decreases with the number of metered buses. Having a larger value for NY enables the proposed algorithm to slightly improve estimation results. However, a more significant difference is observed for branch flow estimations, which is depicted in Figure 2.5.

Figure 2.5(a) and Figure 2.5(b) show that the proposed formulation can estimate unobserved branch flows with errors up to 35% for both the 25% and 100% DER penetration levels, respectively. The errors decrease as the number of voltage measurements increases. These results show that the utility should guarantee a reasonable coverage of voltage measurements to meet the target accuracy. For instance, if seven buses are monitored, the branch power flows can be estimated with errors up to 10% (neglecting outliers).

Figure 2.5. Boxplots of relative branch flow estimation errors for all metering scenarios combining the NY values of 1 and 9.



In all scenarios, having a larger value for NY improves estimation results as the proposed algorithm leverages the increased degrees of freedom for the equivalent DER allocation problem. For the case with 100% DER penetration, $NY = 1$ leads to deteriorated results, especially when seven or more voltage measurements are available. This happens due to the limitation of allocating a single (larger) equivalent DER, affecting the calculation of neighboring branch flows. The influence of parameter NY is explored next.

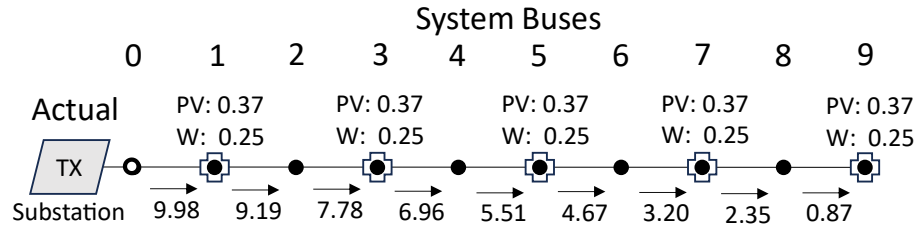
2.5.3. Impact of the choice of NY (10-bus system)

The parameter NY determines the maximum number of system nodes where equivalent DERs can be allocated. This parameter was introduced to provide control over the characteristics of the solution. It constitutes a trade-off relationship between solution sparsity (fewer aggregated DERs) and optimization performance (more degrees of freedom). A system representation with fewer equivalent DERs (a low NY value) can provide accurate power flow calculations while reducing the computational time required for planning studies with many scenarios. On the other hand, opting for higher NY values eases the solution of the formulated problem and may lead to a distribution of equivalent DERs that approximates better the invisible resources. Still, it increases the model's complexity, particularly in larger systems.

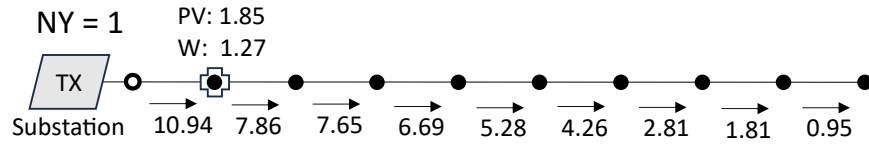
Considering the same conditions of previous tests, Figure 2.6 illustrates a particular example of how the parameter NY can affect the accuracy of estimated results. Values are in MW. Figure 2.6(a) depicts actual branch active power flows and the power outputs of solar and wind DERs. Figure 2.6(b) and Figure 2.6(c) show estimated results for the same time interval when $NY = 1$ and $NY = 9$, respectively. For this snapshot, the total aggregate load consumption is 11.47 MW + j1.42 MVar. The direct solar irradiance is 707 W/m² and the wind speed is 9.46 m/s, leading to a total aggregate power output of 1.85 MW and 1.27 MW for solar and wind DERs, respectively. All voltage measurements are available and considered 100% accurate.

Note that the idea is not to find the actual case in terms of DER and load allocations, but rather finding an alternative, simplified representation where power flows are the same.

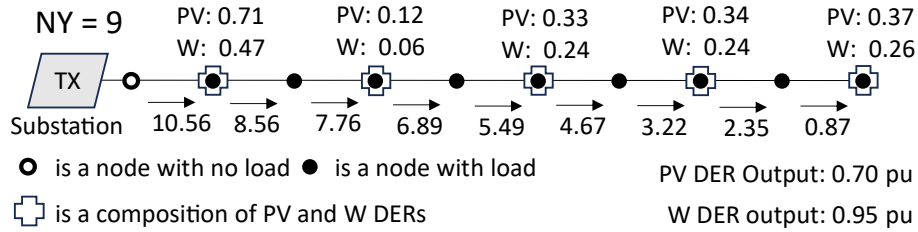
Figure 2.6. Actual and estimated branch active power flows and equivalent DER output for different NY .



(a) Actual results.



(b) Branch flow results when $NY = 1$.



(c) Branch flow results when $NY = 9$.

Considering this snapshot, the average relative error for estimated branch active power flows correspond to 4.41% and 3.20% when NY is 1 and 9, respectively, with maximum absolute errors of 1.33 MW and 0.63 MW. Considering the entire three months of estimated data and NY equal to 1, the minimum, mean, and maximum absolute errors for branch flow estimations are $8.24 \cdot 10^{-6}$, $1.39 \cdot 10^{-2}$, and $2.46 \cdot 10^{-1}$ p.u., respectively, while these values when NY is 9, considering the 25% DER penetration scenario, are given in Table 2-5. This shows that NY can be significant in scenarios with higher meter coverage levels. Figure 2.6(b) shows a single composition of solar and wind DERs on bus 1, while Figure 2.6(c) illustrates that the proposed formulation was able to precisely place equivalent DERs on the same buses where actual DERs are located, despite being allowed to include more ($NY = 9$). These results demonstrate its capability to approximate the spatial distribution of equivalent DERs when there is sufficient information available.

Figure 2.7 expands this analysis to include different meter coverage levels and values of NY . In this example, voltage measurements are introduced in the same order as buses are numbered (from the substation to the farthest node). In other words, when two meters are available, only buses 0 and 1 are monitored; buses 0, 1, and 2 for three meters, and so on. Values presented refer to the accuracy of the branch active power flow estimation, for different meter coverage levels and NY values. Note that, for all meter coverage levels, the accuracy levels tend to reach the same value when the parameter NY assumes a value larger than or equal to the actual number of DERs in the system ($NY \geq 5$). This shows that the algorithm converges close to the same solution given a large enough number of equivalent DERs is allowed to be placed via the value set to NY .

There is, on average, a 1.51% and a 0.34% increase for every new voltage meter introduced to the system and each increment of NY , respectively, for the 25% DER penetration scenario. Considering the 100% DER penetration level results, branch flow estimation accuracy increases on average 0.45% for every increment of NY , and 0.89% for each new voltage measurement available.

Figure 2.7. Branch active power flow estimation accuracy (ACC, %) considering the variation of NY and the number of metered buses.

NY=1	87.9	90.2	91.8	92.1	93.3	94.1	94.3	94.8	94.9
NY=2	87.9	90.5	92.7	93.8	95.9	96.9	97.0	97.1	97.5
NY=3	87.9	90.5	92.7	94.5	96.2	97.1	97.9	98.4	98.5
NY=4	87.9	90.5	92.7	94.7	96.2	97.5	98.4	98.7	98.8
NY=5	87.9	90.5	92.7	94.7	96.2	97.5	98.4	99.1	99.4
NY=6	87.9	90.5	92.7	94.7	96.2	97.5	98.4	99.1	99.4
NY=7	87.9	90.5	92.7	94.7	96.2	97.5	98.4	99.1	99.4
NY=8	87.9	90.5	92.7	94.7	96.2	97.5	98.4	99.1	99.4
NY=9	87.9	90.5	92.7	94.7	96.2	97.5	98.4	99.1	99.4
	2	3	4	5	6	7	8	9	10

No. of buses with voltage measurements

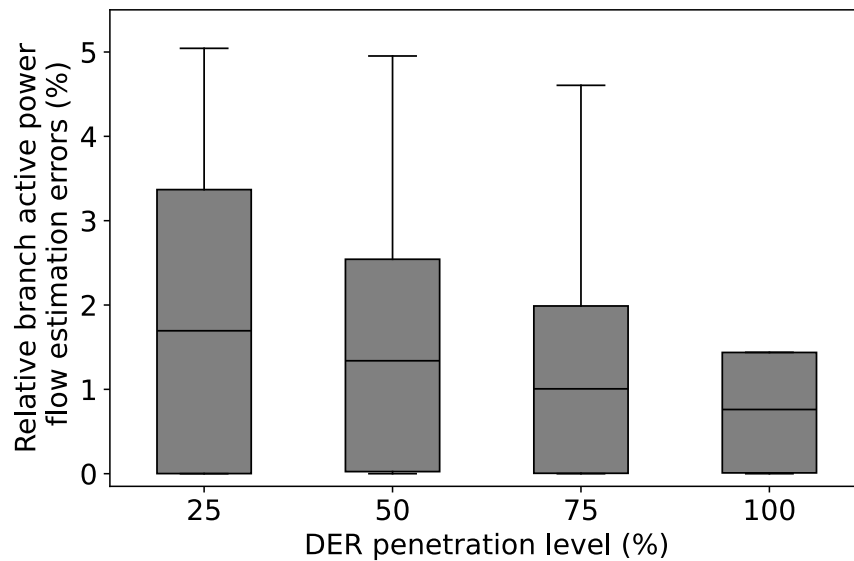
(a) 25% DER penetration.

NY=1	92.4	93.3	94.1	94.3	94.6	94.8	94.8	94.9	95.0
NY=2	92.4	93.8	95.4	95.9	97.2	97.3	97.5	97.7	98.0
NY=3	92.4	93.9	95.5	96.8	98.0	98.3	98.4	98.7	98.7
NY=4	92.4	94.5	96.2	97.6	98.1	98.8	98.9	99.1	99.3
NY=5	92.4	94.5	96.2	97.6	98.6	99.4	99.8	99.9	99.9
NY=6	92.4	94.5	96.2	97.6	98.6	99.4	99.8	99.9	99.9
NY=7	92.4	94.5	96.2	97.6	98.6	99.4	99.8	99.9	99.9
NY=8	92.4	94.5	96.2	97.6	98.6	99.4	99.8	99.9	99.9
NY=9	92.4	94.5	96.2	97.6	98.6	99.4	99.8	99.9	99.9
	2	3	4	5	6	7	8	9	10
	No. of buses with voltage measurements								
	(b) 100% DER penetration.								

2.5.4. Impact of DER penetration (10-bus system)

The branch flow estimation accuracy is also influenced by the total aggregated amount of DERs integrated into the system. Figure 2.8 illustrates that errors tend to decrease as DER penetration increases. DER penetration is taken as the ratio between the maximum aggregate DER generating and load capacities at the feeder level. These results are generated by varying NY from 1 to 9, assuming that all bus voltage magnitude measurements are available. The DER penetration levels of 25%, 50%, 75%, and 100% are obtained by scaling the DERs described in Table 2-3 to a total aggregate generating capacity of 4 MW, 8 MW, 12 MW, and 16 MW, respectively, in a feeder with the maximum aggregate consumption of 16 MW.

Figure 2.8. Branch flow estimation errors for different DER penetration levels.



2.5.5. Impact of the location of DERs (10-bus system)

This section explores how the spatial distribution of invisible DERs affects estimation accuracy and the allocation of equivalent DERs. While keeping the distribution of customer loads as per Table 2-3, solar and wind DERs are redistributed so the "center of generation" is located at the beginning (Case 1), middle (Case 2), and end (Case 3) of the feeder. The total installed capacities of PV and W DERs remain 2.67 MW and 1.33 MW, respectively (25% DER penetration scenario). *NY* is chosen to be 3. The system is assumed to have 100%-accurate voltage measurements at every node (100% MCL), and 3-month historical data are adopted to conduct the analysis.

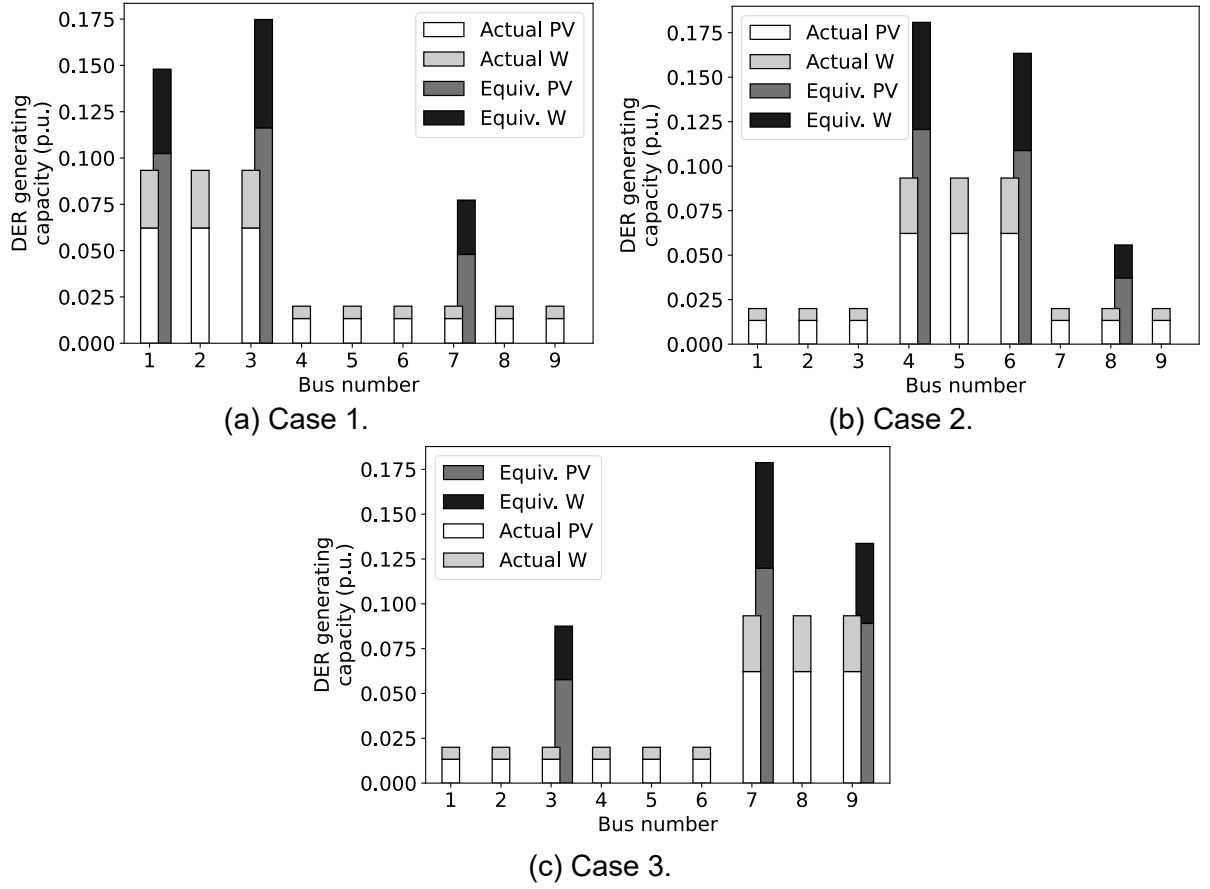
Table 2-6 presents the absolute error values for the estimation of branch active power flow in each of the three cases. The results demonstrate that the accuracy of estimation results remains insensitive to the specific locations of invisible DERs. Values are in p.u. and the base power is 10 MVA.

Table 2-6. Absolute error values for the estimation of branch flows for different spatial distribution of PV DERs

Metric	Case 1	Case 2	Case 3
25% Penetration			
MAE (p.u.)	$1.39 \cdot 10^{-3}$	$1.48 \cdot 10^{-3}$	$1.43 \cdot 10^{-3}$
Min AE (p.u.)	$2.33 \cdot 10^{-6}$	$2.01 \cdot 10^{-6}$	$2.04 \cdot 10^{-6}$
Max AE (p.u.)	$4.70 \cdot 10^{-3}$	$5.95 \cdot 10^{-3}$	$5.97 \cdot 10^{-3}$
100% Penetration			
MAE (p.u.)	$1.56 \cdot 10^{-3}$	$1.65 \cdot 10^{-3}$	$1.58 \cdot 10^{-3}$
Min AE (p.u.)	$7.48 \cdot 10^{-7}$	$8.25 \cdot 10^{-7}$	$1.26 \cdot 10^{-6}$
Max AE (p.u.)	$6.43 \cdot 10^{-3}$	$6.55 \cdot 10^{-3}$	$7.44 \cdot 10^{-3}$

Equivalent DER allocation results and actual DER positions for the three cases are depicted in Figure 2.9. These results demonstrate the capability of the proposed algorithm to approximate the spatial distribution of hidden DERs with a reduced number of equivalent representations. This unintended yet valuable outcome displays the proposed formulation's ability in discerning the steady-state behavior of equivalent DERs to be placed along the feeder.

Figure 2.9. Equivalent DER allocation for distinct locations of hidden DERs in the 25% DER penetration scenario.

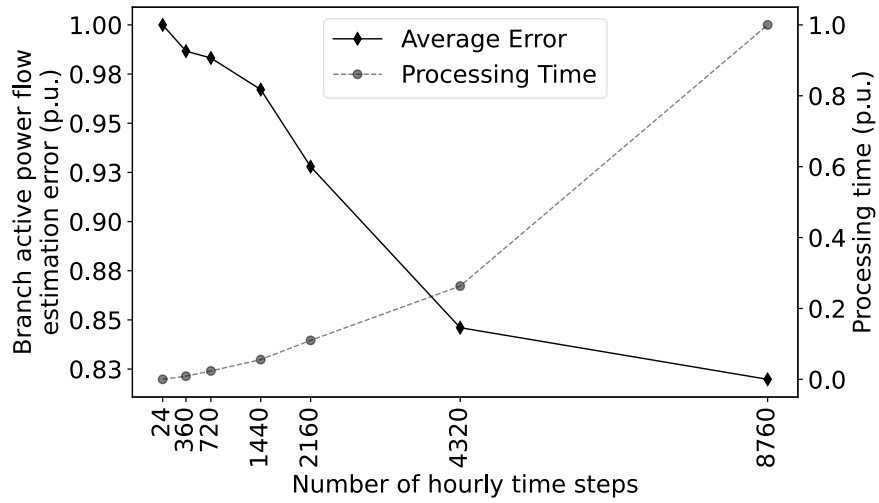


2.5.6. Impact of the length of historical data (10-bus system)

The length of historical data is another factor that influences the accuracy of estimation results. Notably, the largest absolute error values for unobserved branch flow estimations occur when a reduced horizon of historical data is considered. The results in Figure 2.10 refer to the MAE values for branch flow estimates, considering that the entire system is covered with voltage meters and parameter NY varies from 1 to 9. It starts with 24 time steps (representing a day) and ends with 8,760 time steps (representing a year), leading to a total decrease of nearly 17% in the magnitude of branch flow estimation errors.

Figure 2.10 also shows that this reduction comes at the expense of an increase in computational time, which presents a quadratic relation with the number of time steps considered. Note, however, that the average errors in the calculation of (observed) voltage magnitudes increase slightly with the addition of more snapshots. This is expected as error accumulates as the algorithm attempts to fit an increasing set of known values.

Figure 2.10. Average branch flow estimation and processing time values varying the length of historical data.



2.5.7. Impact of meter accuracy (10-bus system)

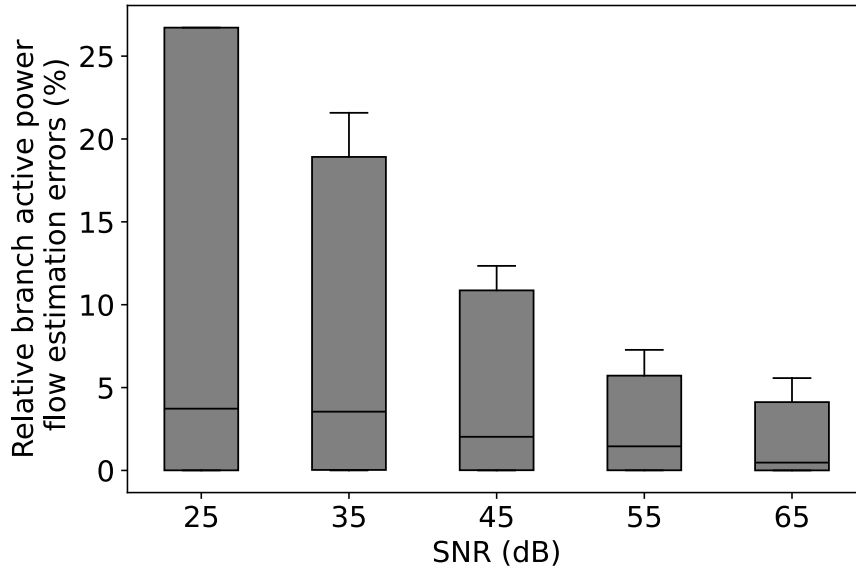
This subsection focuses on evaluating the impact of voltage measurement accuracy on the accuracy of branch flow estimations. Here, the measurement accuracy is expressed in terms of the Signal-to-Noise Ratio (SNR), defined as

$$SNR = 20 \cdot \log_{10} \left(\frac{1}{ACC_{meter}} \right) \quad [\text{dB}]$$

where, ACC_{meter} refers to meter accuracy, representing the maximum allowed percentage deviation from the true measured value.

Figure 2.11 depicts estimation results for different SNRs applied to voltage measurement data achieved by adding different levels of Gaussian noise to the input dataset. These results are generated by varying NY from 1 to 9, assuming that all bus voltage magnitude measurements are available with SNRs of 25, 35, 45, 55, and 65. Results show that when all voltage measurements present an SNR of 45, the highest branch flow estimation error reaches 13% for a time step with an instantaneous branch active power flow of 13.94 MW. These results were obtained for the 25% DER penetration scenario.

Figure 2.11. Branch flow estimation error for different SNRs.



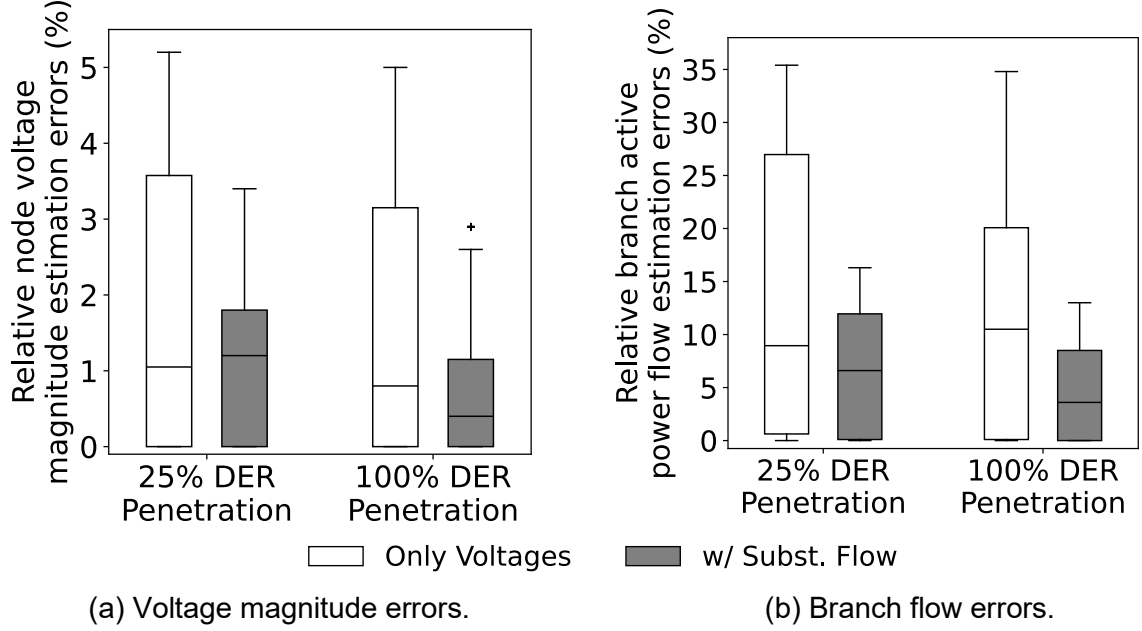
2.5.8. Impact of substation metering (10-bus system)

Up to this point, only voltage magnitude measurements were considered inputs to test the proposed method. However, it is a common practice for utilities to include various types of meters at distribution substations. A measurement typically available consists of the active and reactive power flows from the substation to the feeder. Therefore, substation flow data is introduced to the scenarios with the highest estimation errors presented in Figure 2.4 and Figure 2.5. These are the instances where only two voltage magnitude measurements exist. Results include all the possible combinations with different NY values, two voltage magnitude measurements (one always being the voltage at the substation-end bus), and power flows at the end of the substation. Figure 2.12 compares estimation errors for the 25% and 100% DER penetration scenarios.

As in previous studies, higher DER penetrations present finer results. The introduction of flow measurements reduced the maximum branch flow estimation errors from around 35% to a range between 12% and 17%. However, this effect is less pronounced for the estimation of unobserved voltage magnitudes (from around 5% to between 2.5% and 3.5%).

These results display the benefits of introducing substation flow measurements and the flexibility of the proposed method in considering other measured quantities besides bus voltage magnitudes. Note that this case study involves a system with limited visibility (only the substation flow and two voltage magnitude measurements are available) where traditional state estimation methods may falter. In contrast, the proposed method can still provide a solution and determine equivalent DER models.

Figure 2.12. Comparison of results considering branch flow measurements at the substation exit.



2.5.9. Analysis of solution robustness (10-bus system)

MINLP problems inherently have complex solution spaces due to the presence of both continuous and discrete variables. This complexity arises because changes in parameter values can lead to different solutions that satisfy the problem's requirements but vary in their numerical values. A sensitivity analysis is therefore conducted to assess the robustness of results in the presence of multiple solutions. The study considered a scenario with 25% DER penetration, using 720 power flow snapshots representing a month of historical data at 1-hour intervals, and assumed error-free measurements.

The lowest objective function value of $R_{opt} = 1577.37530515$ is achieved when $NY \geq 5$, being five, the actual number of DERs in the system. This results in an average squared voltage magnitude estimate error of $2.19 \cdot 10^{-7}$ per bus and per time step. Next, constraint (2.20) was added to the formulated problem to limit the calculations of the original objective function to the lowest error achieved, R_{opt} , while the objective function was modified to maximize (and later minimize) squared voltage magnitude measurement residuals—equivalent to maximizing or minimizing the squared RMSE error metric—as in (2.21).

$$\|K^U \odot (U - \tilde{U})\|_F^2 \leq R_{opt} \quad (2.20)$$

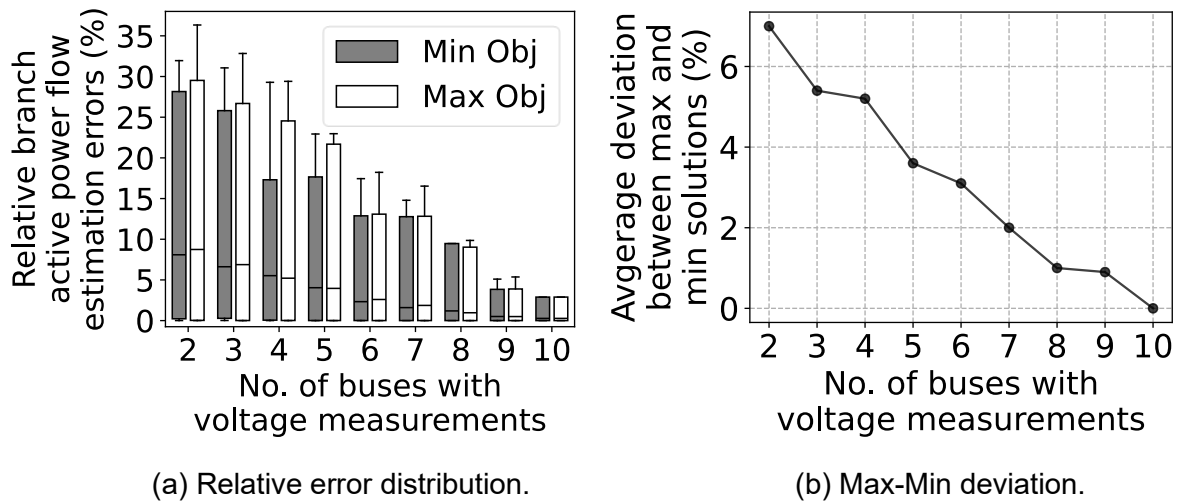
$$\text{maximize (or minimize)} \sum_{t=1}^{NDT} \sum_{k=1}^{NM} (U_{k,t} - \tilde{U}_{k,t})^2 \quad (2.21)$$

The relative errors of estimating unobserved branch flows considering different meter coverage levels for both the maximization and minimization objectives are

depicted in Figure 2.13(a). As in previous studies, estimates tend to get more accurate as system visibility increases. Despite the different objectives, results evidence a small range of values at which estimation errors may vary, thus indicating the robustness of the proposed approach despite the diversity of optimal solutions.

This analysis is complemented by Figure 2.13(b), which depicts the decreasing deviation between the maximum and minimum solutions (normalized by the maximum) as the number of voltage measurements increases. Percentage values account for the average deviation between max and min results considering all available time steps and $NY \geq 5$.

Figure 2.13. Relative branch flow estimation errors when maximizing and minimizing the squared RMSE metric.

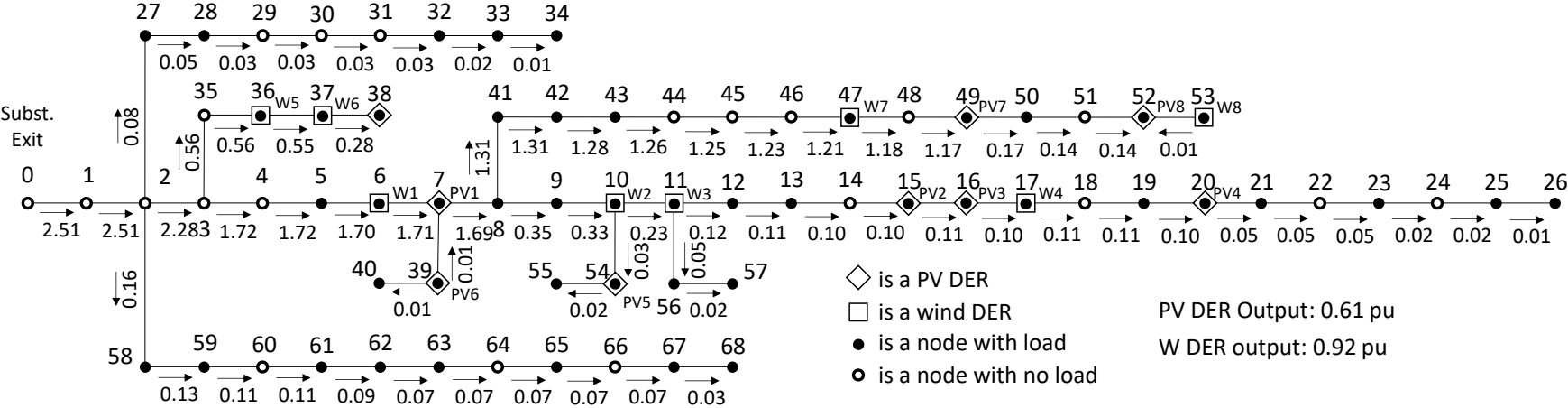


2.5.10. A larger distribution feeder with laterals (69-bus system)

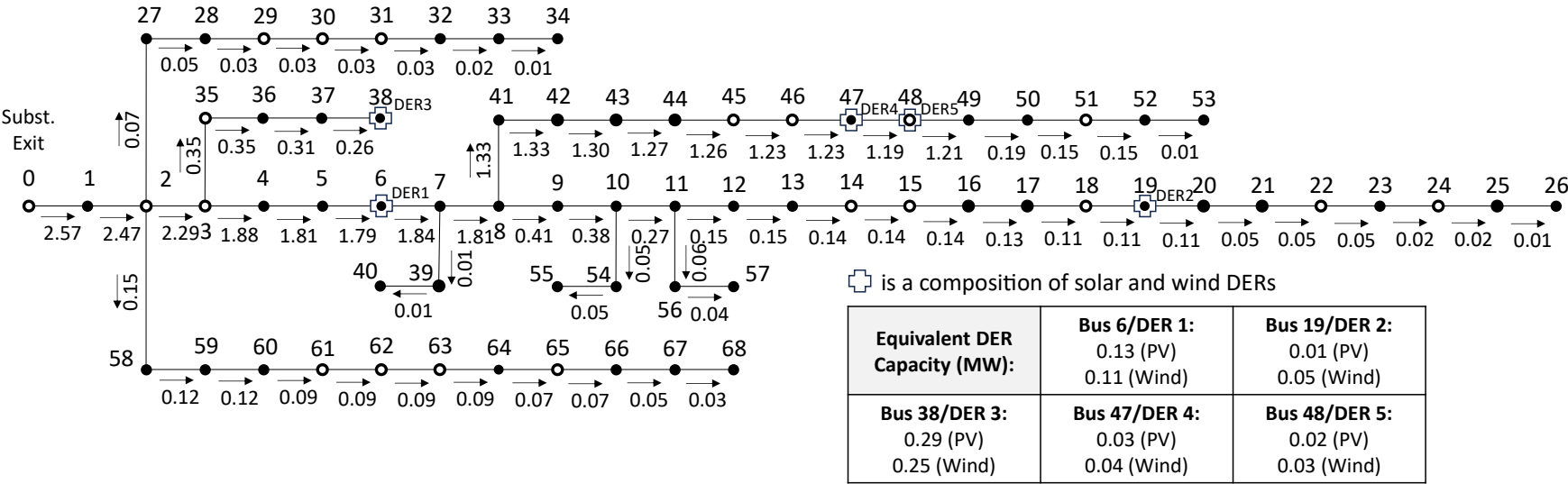
The IEEE 69-bus radial test system [68] was chosen to assess the performance of the proposed model when dealing with a larger radial distribution system with laterals. The system works at the nominal voltage of 12.66 kV with a total load of 3.802 MW and 2.694 MVar. It was modified to incorporate DERs placed on the nodes where the largest loads are located—see Figure 2.14(a).

Considering 720 one-hour time steps (a month of historical data), Table 2-7 presents the results for the application of the proposed formulation considering the ideal case with all the system nodes having 100%-accurate voltage measurements available. The parameter NY is set to be 5 to stress the capability of providing accurate estimations with a reduced number of equivalent DER representations.

Figure 2.14. Example of application for the modified IEEE 69-bus radial test system with a reduced number of equivalent DERs.



(a) Branch flow results for the original test system at $t = 257$.



(b) Branch flow results for the equivalent system representation at $t = 257$. NY is 5.

Table 2-7. Estimation Performance: 69-bus system, $NY = 5$, 100% MCL

Metric	25% DER Penetration		100% DER Penetration	
	Voltage Magnitude Errors (p.u.)	Branch Active Power Flow Errors (p.u.)	Voltage Magnitude Errors (p.u.)	Branch Active Power Flow Errors (p.u.)
MAE	$1.90 \cdot 10^{-4}$	$2.93 \cdot 10^{-3}$	$6.48 \cdot 10^{-3}$	$2.07 \cdot 10^{-3}$
RMSE	$5.14 \cdot 10^{-3}$	$1.03 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$	$3.53 \cdot 10^{-3}$
Min AE	$5.64 \cdot 10^{-8}$	$1.51 \cdot 10^{-6}$	$1.49 \cdot 10^{-5}$	$2.64 \cdot 10^{-6}$
Max AE	$5.57 \cdot 10^{-4}$	$5.58 \cdot 10^{-3}$	$3.77 \cdot 10^{-2}$	$1.61 \cdot 10^{-2}$
ACC	99.99%	98.31%	99.99%	99.04%

The 25% DER penetration scenario has a total of 0.48 MW of both solar and wind DERs. DERs in the 100% penetration scenario remain connected to the same system nodes but were scaled to match the respective penetration values aggregated at the feeder level. Low error values show that the proposed algorithm can replicate existing measurements and estimate branch power flows solely utilizing voltage magnitude measurements. Figure 2.14(b) depicts how the estimated active power branch flows are disposed of in the resulting system representation with five equivalent DERs being allocated. As in previous case studies, estimation results are affected by the number of measurements available and the adjustment of parameter NY .

This is depicted in Fig. 13, with results generated by varying the meter coverage level by 10% increments of the number of voltage measurements available, until the system is fully monitored. Meters are placed at random locations selected using uniform distribution and assuming that the measurements at bus 0 (the substation end-node) are always available and 100% accurate. On average, for every 10% increase in the number of voltage measurements available, an improvement of 3.7% in branch flow estimation accuracy is observed.

Figure 2.15. Branch active power flow estimation accuracy (ACC, %) varying the number of metered buses and the choice of NY .

$NY=5$	71.3	82.7	88.4	91.7	93.3	94.8	95.3	97.2	98.0	98.3
$NY=15$	71.8	83.4	89.3	92.5	94.3	96.3	97.5	98.7	99.0	99.2
	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
	Percentage of buses with voltage measurements									

2.5.11. A distribution system with multiple feeders (873-bus system)

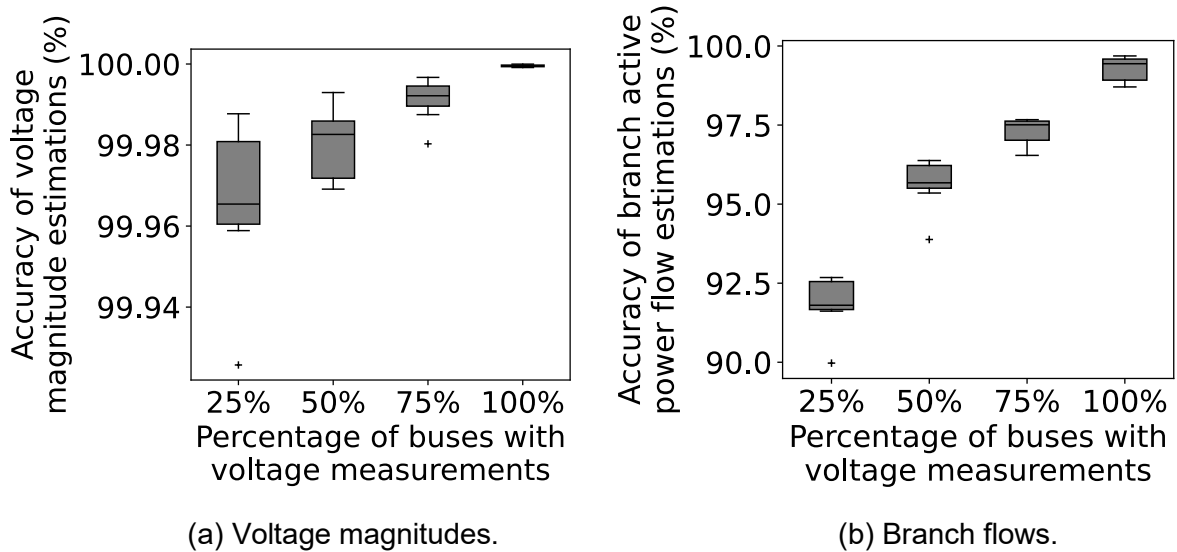
The following case study expands the application of the algorithm to the representation of a radial distribution system with multiple feeders connected to the same substation busbar. The 873-bus radial distribution system of [69] comprises

seven feeders, supplying a total aggregated load of 33.6 MW + j19.8 MVar at the 27.6 kV level. The system was modified to integrate 12.6 MW and 4.2 MW of solar and wind DERs, respectively (50% DER penetration scenario). Table 2-8 presents branch active power flow estimation results when voltage measurements are fully available. These results consider 48 power flow snapshots (two days of metered data points with 1-hour intervals). The parameter NY is 5 for each feeder, meaning there is a reduction from 350 nodes with actual DERs in the entire system to 35, with five aggregate DERs in each feeder.

Table 2-8. Branch Flow Estimation Performance: 873-bus system, $NY = 5$ (per feeder), 100% MCL

	MAE (p.u.)	RMSE (p.u.)	Min AE (p.u.)	Max AE (p.u.)	ACC (%)
Feeder 1	$4.31 \cdot 10^{-3}$	$6.98 \cdot 10^{-3}$	$3.13 \cdot 10^{-6}$	$3.63 \cdot 10^{-2}$	99.0
Feeder 2	$6.12 \cdot 10^{-3}$	$8.92 \cdot 10^{-3}$	$6.00 \cdot 10^{-6}$	$3.57 \cdot 10^{-2}$	98.7
Feeder 3	$2.18 \cdot 10^{-3}$	$3.24 \cdot 10^{-3}$	$2.64 \cdot 10^{-6}$	$1.26 \cdot 10^{-2}$	99.6
Feeder 4	$3.24 \cdot 10^{-3}$	$5.42 \cdot 10^{-3}$	$6.32 \cdot 10^{-6}$	$3.15 \cdot 10^{-2}$	99.4
Feeder 5	$2.69 \cdot 10^{-3}$	$4.20 \cdot 10^{-3}$	$2.18 \cdot 10^{-6}$	$2.44 \cdot 10^{-2}$	99.5
Feeder 6	$2.66 \cdot 10^{-3}$	$4.53 \cdot 10^{-3}$	$1.27 \cdot 10^{-6}$	$2.17 \cdot 10^{-2}$	99.5
Feeder 7	$2.95 \cdot 10^{-3}$	$6.39 \cdot 10^{-3}$	$4.02 \cdot 10^{-6}$	$4.93 \cdot 10^{-2}$	98.8

Figure 2.16. Estimation accuracy across the seven feeders considering different meter coverage levels.



Accurate ($\geq 90\%$) results evidence the algorithm's ability to estimate unobserved branch flows using only bus voltage magnitudes, even in a scenario with fewer DER representations. Next, Figure 2.16 presents the distribution among the existing feeders of the accuracy of branch flow and node voltage magnitude estimations for different voltage meter coverage levels. Meters were randomly placed in the system following uniform distribution, except for the substation end-node, which is always available. Results show that the algorithm can estimate unobserved branch active power flows with errors below 10%, even for the scenario with fewer node voltage magnitude measurements. *NY* is 5 in all cases. Measurements are considered 100% accurate.

2.6. Chapter Summary

This chapter presented a convex mixed-integer nonlinear problem formulation designed to develop aggregate DER models for equivalent representation of low-visibility radial distribution feeders with invisible DERs. The formulated problem aims at estimating branch power flows and bus voltage magnitudes while strategically allocating equivalent DERs to reproduce sparse system measurements and represent innumerable DERs hidden from the system operator. The resulting DER models can be integrated into traditional power flow solution methods to enable accurate calculations—which is the focus of Chapter 4 of this document.

A distinction of this method compared to conventional state estimation algorithms consists of lesser measurement data requirements to provide solutions in unobservable or poorly observable system conditions where these methods may falter. Case studies were realized to display the formulation's efficacy in estimating missing power system variables. These studies considered relevant factors, such as the number of available meters, adjustment of model parameters, meter accuracy, length of historical data, DER integration, and system scale.

In Chapter 3, different linearized versions of the convex-MINLP formulation are tested developed to enhance computational performance while maintaining numerical accuracy. Moreover, the modeling of invisible DERs is enhanced by introducing hybrid DER models (e.g., PV+BESS) to improve the representation of modern radial distribution systems with different DER configurations.

3. A MILP MODEL FOR ENHANCED AGGREGATE MODELING OF INVISIBLE HYBRID DERs²

3.1. Chapter Introduction

This chapter continues addressing the challenge of improving steady-state analysis for radial distribution systems with invisible DERs by presenting a Mixed-Integer Linear Programming (MILP) formulation to develop equivalent aggregate DER models representing numerous invisible resources. The linearized formulation is obtained by addressing the nonlinear and nonconvex nature of power flow equations, yielding a simplified yet accurate model with reduced solution times. The resulting DER models capture the diverse characteristics of invisible resources, now including hybrid DERs that combine the flexibility of BESS and PV units. With faster solution times, the method presented in this chapter enables the development of equivalent DER models for systems with more dynamic operating conditions, such as with many possible topology changes. In this sense, the method is suited to operational planning and tasks such as system reconfiguration, which require rapid solutions, while addressing invisible DERs.

As per Chapter 2, this chapter begins by providing minimal introductions to the key concepts serving as the basis for the modeling aspects of the research. In this sense, a review of linearization approaches applied to the modeling of distribution systems is provided. Next, the proposed MILP formulation is discussed, emphasizing changes in relation to the original MINLP optimization problem.

3.2. Linearization in Power Distribution Systems Modeling

The many different approaches for linearizing power flow equations can be categorized into two main groups, namely models based on assumptions over typical physical characteristics of radial power distribution systems and on mathematical approximations to existing nonlinear functions present in the formulation. The models in the first group often consider that voltage angles and magnitudes vary within relatively narrow boundaries. For example, in [72] and [73], the authors leverage small angle deviations across lines to simplify the calculation of bus voltage magnitudes. In [74] and [75], the authors assume that voltage magnitudes tend to be close to 1.0 p.u. to calculate branch power flows. Similarly, the authors in [76] assume negligible system losses, disregarding the calculation of the current magnitude for the traditional *Distflow* model for radial distribution systems [77]. However, although simplifying the model, assuming negligible losses may lead to larger errors in regions with higher current levels (e.g., the substation exit), as noted in [78].

The second group aims to determine mathematical expressions to approximate the power flow equations. These methods include Taylor series expansions, as employed in [79] for calculating branch power flows, and in [58] to approximate power

² This chapter's contents were published in part in the paper: P. N. Vasconcelos, F. C. L. Trindade, and B. Venkatesh, "Linearized Optimization for Enhanced Aggregate Modeling of Invisible Hybrid Distributed Energy Resources," *IET Generation, Transmission & Distribution*, vol. 19, no. 1, p. e70088, May 2025.

injections when modeling customer loads, both for node-injection power flow models. A more popular approach consists of piecewise linear approximations, which discretize nonlinear functions into linear segments, controlling accuracy and computational feasibility via the number of introduced segments. In [80] and [81], piecewise linear approximations define the calculation of squared active and reactive branch flows. In optimization applications, McCormick envelopes are often employed to calculate line currents of the line-wise power flow model, relaxing the nonlinear expression by introducing a set of linear constraints [76], [82].

Unlike methods that rely solely on simplifying assumptions or purely mathematical relaxations, the proposed MILP approach integrates both elements by approximating the nonlinear power flow equations using piecewise linearization and McCormick envelopes when addressing the equivalent behavior of aggregated invisible DERs. The formulation allocates (locates and sizes) aggregate DER models while estimating unobserved voltage magnitudes and branch flows.

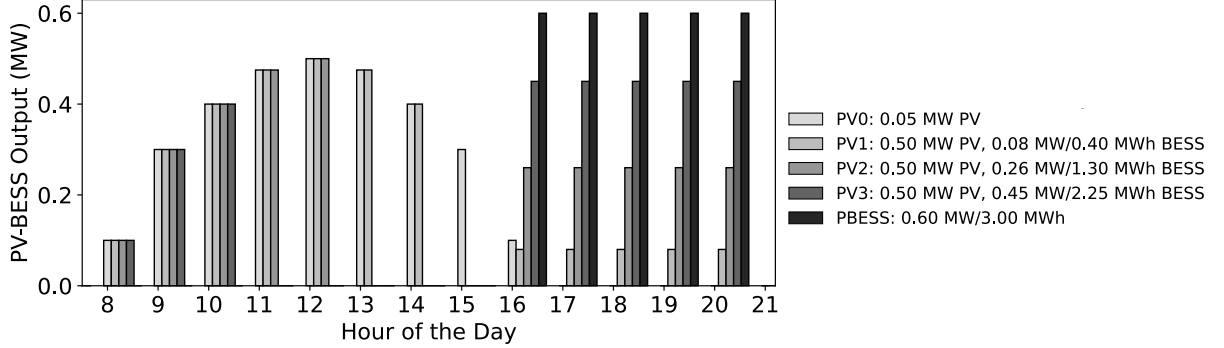
3.3. Problem Statement

3.3.1. Nature of hybrid PV-BESS DERs

Some of the power flow and DER allocation constraints presented in this section employ four different PV models (numbered from 0 to 3) to represent varying levels of energy storage capacity in hybrid PV-BESS DERs. These models capture different power generation profiles based on the time the battery charging begins. Since it is assumed that the BESS is recharged exclusively using locally generated solar power, the charging start times reflect different relative sizing of PV and BESS capacities. Earlier charging times correspond to hybrid DER configurations where the BESS has a larger capacity relative to PV generation, requiring a longer duration of available solar power to reach full charge. Conversely, later charging start times indicate scenarios where PV generation is larger, enabling shorter charging windows. BESS DERs discharge only during peak hours—between hours 16 and 21—when electricity rates are highest in a typical time of use (TOU) tariff strategy. Each combination is displayed in Figure 3.1.

Model PV0 corresponds to a scenario where only the PV is active, with no associated energy storage. The remaining models—PV1, PV2, and PV3—correspond to hybrid PV-BESS configurations where the BESS starts charging at hours 11, 13, and 15, respectively. Figure 3.1 also includes a case where only the output of the BESS is active, representing a combination where the production of the local PV is sufficient to charge the battery fully, and there is no export of excess PV power to the grid. PVs export their power output to the grid whenever the local storage resources are not charging and the solar irradiance level allows for a non-zero power output. Therefore, the inherent variability of these resources is accounted for in the model.

Figure 3.1. Comparison of the power output profiles of the different hybrid PV-BESS models. The BESS charges exclusively through PV output and discharge between hours 16 and 21.



3.3.2. Objective function

The objective function remains unchanged from the original formulation. It minimizes the squared Frobenius norm of the resulting matrix, computed as the sum of the squares of measurement residuals considering measured and estimated node voltage magnitudes. Objective function weights, defined in matrix $\mathbf{K}^U$ attribute relative importance to more accurate data.

$$\text{minimize } \|\mathbf{K}^U \odot (\mathbf{U} - \tilde{\mathbf{U}})\|_F^2 \quad (3.1)$$

3.3.3. Power flow constraints

The basis for the power flow constraints (3.2)-(3.6) also remains unchanged from the original formulation. Eq. (3.5) was modified to introduce the power output of BESS DERs and to refer to the different PV DER models considering the presence of local energy storage.

$$U_{j,t} + 2 \left(P_{l,t}^S \cdot R_l + Q_{l,t}^S \cdot X_l - \frac{U_{l,t}}{2} \right) + J_{l,t} \cdot Z_l^2 = 0, \quad (i, j) \in l \quad (3.2)$$

$$P_{l,t}^F + P_{l,t}^S + R_l \cdot J_{l,t} = 0 \quad (3.3)$$

$$Q_{l,t}^F + Q_{l,t}^S + X_l \cdot J_{l,t} = 0 \quad (3.4)$$

$$[\mathbf{M}] \begin{bmatrix} \mathbf{P}_t^F \\ \mathbf{P}_t^S \end{bmatrix} - \mathbf{U}_t \cdot \mathbf{G} = \mathbf{P}_t^L - \mathbf{P}_t^{PV,x} - \mathbf{P}_t^{WT} - \mathbf{P}_t^{BESS} \quad (3.5)$$

$$[\mathbf{M}] \begin{bmatrix} \mathbf{Q}_t^F \\ \mathbf{Q}_t^S \end{bmatrix} + \mathbf{U}_t \cdot \mathbf{B} = \mathbf{Q}_t^L \quad (3.6)$$

Originally, the calculation of the squared magnitude of the current flowing in line l at time t was being performed employing a second-order conic (SOC) relaxation as

$$U_{j,t} \cdot J_{l,t} \geq P_{l,t}^{S^2} + Q_{l,t}^{S^2}$$

This relation is linearized to benefit from the improved computational efficiency and scalability of MILP problems. The bilinear term on the left side is addressed using McCormick envelopes [83], while the right-hand side is approximated using piecewise

linear functions [87], resulting in (3.7). The choice of McCormick envelopes is motivated by their widespread adoption in optimization problems involving bilinear terms, as established in [85]. This approach preserves convexity, a desirable property for maintaining the quality of the solution when using standard optimization solvers [86]. Moreover, when considering a line-wise power flow model, this approach yields an improved approximation of the steady-state behavior of radial distribution systems compared to other linearization techniques, as discussed in the results section.

$$W_{j,l,t} \geq \sum_{k=1}^i \lambda_k^P \cdot p_k^2 + \sum_{k=1}^j \lambda_k^Q \cdot q_k^2 \quad (3.7)$$

where the domains of $P_{l,t}^{S^2}$ and $Q_{l,t}^{S^2}$ are divided into i and j breakpoints, respectively.

Each interval is represented by a weight λ_k^P or λ_k^Q and an associated squared term p_k^2 or q_k^2 . Variable $W_{m,l,t}$ is an auxiliary variable introduced to represent the linear approximation of the bilinear term, i.e., $W_{m,l,t} \approx U_{m,t} \cdot J_{l,t}$. Two sets of linear constraints impose its upper and lower bounds. Expressions (3.7a) and (3.7b) establish the lower limit, while (3.7c) and (3.7d) define the upper bounds. Parameters with lower and overbars refer to minimum and maximum reference values of the original squared node voltage magnitude and squared line current magnitude variables.

$$W_{j,l,t} \geq \underline{U} \cdot J_{l,t} + U_{j,t} \cdot \underline{J} - \underline{U} \cdot \underline{J} \quad (3.7a)$$

$$W_{j,l,t} \geq \overline{U} \cdot J_{l,t} + U_{j,t} \cdot \overline{J} - \overline{U} \cdot \overline{J} \quad (3.7b)$$

$$W_{j,l,t} \leq \overline{U} \cdot J_{l,t} + U_{j,t} \cdot \underline{J} - \overline{U} \cdot \underline{J} \quad (3.7c)$$

$$W_{j,l,t} \leq \underline{U} \cdot J_{l,t} + U_{j,t} \cdot \overline{J} - \underline{U} \cdot \overline{J} \quad (3.7d)$$

Constraints (3.8) and (3.9) enforce operational limits on squared bus voltage and line current magnitudes.

$$\underline{U}_k \leq U_{k,t} \leq \overline{U}_k \quad (3.8)$$

$$J_{l,t} \leq \overline{J}_l \quad (3.9)$$

3.3.4. DER allocation constraints

Constraints (3.10), (3.11), and (3.12) update, for every time interval, the power output of equivalent aggregate PV, WT, and BESS DERs using piecewise linear functions derived from manufacturer data. The output of BESS DERs is modeled according with a TOU tariff strategy. The BESS resources are allowed to discharge when electricity rates are highest, during the peak-hours between 4 PM and 9 PM, and this behavior is depicted in Figure 3.2. The TOU_t function in (3.12) changes its values from 0 to 1 depending on the time of the day.

As inferred in (3.10) and in the computation of p_t^{PV} , the values of x correspond to different hybrid DER models that consider both PV and BESS resources available locally. Parameter t_x consists of the time of the day when the associated energy storage model begins charging, consuming the PV output during this process.

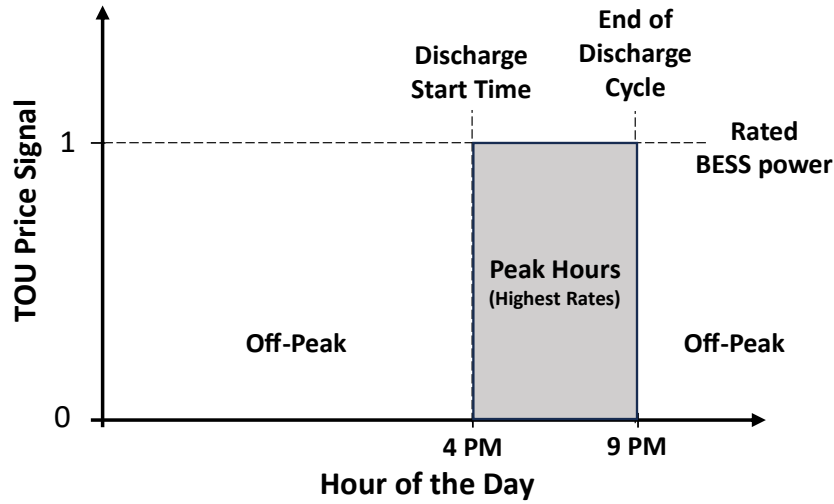
However, PVs generate power whenever the solar irradiance level exceeds predefined threshold values, ensuring that the inherent variability of renewable energy generation is accounted for in the model. BESS DERs export energy to the grid only during peak hours in a typical time-of-use tariff strategy. Variables with overbars represent the peak generating capacity attributed to each equivalent DER.

$$P_{m,t}^{PV} = \overline{P_m^{PV_x}} \cdot p_t^{PV_x}, \forall x \in \{0,1,2,3\} \quad (3.10)$$

$$P_{m,t}^{WT} = \overline{P_m^{WT}} \cdot p_t^{WT} \quad (3.11)$$

$$P_{m,t}^{BESS} = \overline{P_m^{BESS}} \cdot TOU_t \quad (3.12)$$

Figure 3.2. TOU pricing signal used to schedule BESS DERs.



Piecewise linear functions are derived from manufacturer data to model the power outputs of PV and WT DERs as functions of direct solar irradiance and wind speed, respectively. The following expressions consider traditional cut-in, rated, and cut-out parameter values for determining power outputs.

$$p_t^{PV_x} = \begin{cases} \max\{K^{PV} \cdot (h_t - \underline{h}), 0\}, & \text{if } h_t \leq h^r \wedge t < t_x \\ 1, & \text{if } h_t \geq h^r \wedge t < t_x, \\ 0, & \text{if } t \geq t_x \end{cases} \quad \forall x \in \{0,1,2,3\}$$

$$p_t^{WT} = \begin{cases} \max\{K^{WT} \cdot (w_t - \underline{w}), 0\}, & \text{if } w_t \leq w^r \\ 1, & \text{if } w^r \leq w_t < \overline{w} \\ 0, & \text{if } w_t \geq \overline{w} \end{cases}$$

Constraints (3.13), (3.14), and (3.15) limit the summation of the allocated generating capacity of equivalent DERs. Note that all PV models add up to a single value as the actual composition of PV-BESS DERs is unknown. Eq. (3.16) determines that the maximum capacity of each DER model may be higher than zero only for the nodes where the binary variable assumes the unity value. The maximum number of equivalent DER models is determined in (3.17) via the parameter NY .

$$\sum_{k=2}^{NB} \sum_{x=0}^3 \bar{P}_k^{PV_x} = P_{total}^{PV} \quad (3.13)$$

$$\sum_{k=2}^{NB} \bar{P}_k^{WT} = P_{total}^{WT} \quad (3.14)$$

$$\sum_{k=2}^{NB} \bar{P}_k^{BESS} = P_{total}^{BESS} \quad (3.15)$$

$$0 \leq \bar{P}_k^{PV_x}, \bar{P}_k^{WT}, \bar{P}_k^{BESS} \leq C \cdot Y_k^{DER}, \quad C \gg 1 \quad (3.16)$$

$$\sum_{k=2}^{NB} Y_k^{DER} \leq NY, \quad Y_k^{DER} \in \{0,1\} \quad (3.17)$$

3.3.5. Load determination constraints

The same constraints are employed to determine equivalent load models. Constraints (3.18) and (3.18) update, for every time interval, the active and reactive power consumption for all load buses. This is done by multiplying the total feeder-level consumption by the corresponding share of the total load at each bus. Constraint (3.20) ensures that the summation of all load factors equals the total feeder-level load. These are determined using a binary variable for all load buses in the system, as in (3.21).

$$P_{k,t}^L = D_k^L \cdot P_t^L \quad (3.18)$$

$$Q_{k,t}^L = D_k^L \cdot Q_t^L \quad (3.19)$$

$$\sum_{k=2}^{NB} D_k^L = 1 \quad (3.20)$$

$$0 \leq \bar{D}_k^L \leq Y_k^L, \quad Y_k^L \in \{0,1\} \quad (3.21)$$

3.4. Solution Method

The same implementation and solution method using Python and Pyomo was employed to test the resulting MILP formulation problem described in (3.1)-(3.21). Moreover, the same modeling assumptions are assumed in this chapter, related to data availability, DER information at the feeder level, system radial topology, and error distribution of existing measurement data.

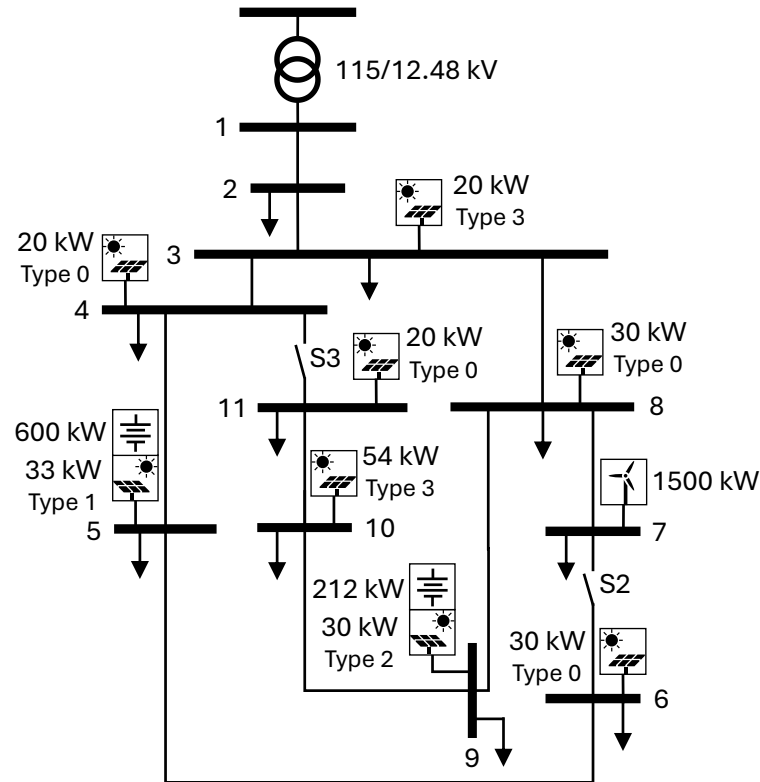
3.4.1. Test system data and preparation

In this chapter, the resulting formulation is employed to estimate system states and allocate equivalent DERs for an 11-bus [87] and a 240-bus [88] distribution test systems. These were selected to test the proposed framework under different system scales and topological and operational characteristics. Note that, as per the previous chapter, detailed information on each test system is provided in Appendix B and the files used in simulation are available at [89].

The 11-bus test feeder involves a typical North-American distribution feeder supplying 4.32 MW + j1.43 MVar of load at the 12.48 kV level. The total generating

capacity of PV DERs is 0.23 MW, while 1.50 MW is for WT, and 0.81 MW/4.06 MWh is for BESS. This corresponds to a 40% DER penetration level by taking the ratio between the total DER capacity and peak customer demand at the feeder level. The radial topology is obtained by leaving all tie-line switches and connections to neighboring feeders open. Figure 3.3 depicts the test system along with DER information. PV-BESS DER types were selected based on the presence of both resources in the original system. Note that this information is only used to generate system measurements through power flow calculations and is not disclosed to the proposed method, aligning with what would be available in practice, considering the presence of invisible DERs.

Figure 3.3. Diagram of the test system displaying DER information.



Although the 11-bus test feeder is a synthetic approximation of a real North-American MV distribution feeder, the data presented in [88] for the 240-bus test system corresponds to a real distribution grid located in the Midwest U.S. that belongs to a municipal utility. The 240-bus test system is a 10-MVA rating radial distribution system consisting of three separate feeders supplied by a 69/13.8 kV substation transformer. According to the authors, the system is a fully observable network with smart meters installed at all customer locations. The system has 240 nodes and 37 km of primary feeder conductors with overhead and underground sections.

A 50% penetration scenario was devised by distributing DERs across 107 load buses in the system (50% of the total number of load buses) over the three existing distribution feeders connected to the same substation bus-bar. In total, there are 95 PV (74 PV + 21 hybrid PV-BESS DERs) and 35 WT DERs. Details of each feeder are laid out in Table VIII.

Table 3-1. Load and DER capacity per feeder for the 240-bus distribution test system.

Feeder	NB	P^L (MW)	Q^L (MVar)	P^{PV} (MW)	P^{WT} (MW)	P^{BESS} (MW/MWh)
A	17	0.71	0.27	0.28	0.07	0.052/0.258
B	60	1.75	0.78	0.70	0.18	0.182/0.910
C	163	1.60	0.58	0.64	0.16	0.085/0.425
Total	240	4.06	1.63	1.62	0.41	0.319/1.595

The same weather dataset (See Appendix B) referring to the location of Toronto, Canada, was used to execute the simulations described next. The linearized DER power output calculation models considered cut-in and rated direct solar irradiance of 150 and 800 W/m², respectively. For wind DERs, cut-in, rated, and cut-out wind speeds are 2.8, 10, and 20 m/s, respectively.

This dataset was processed to get each month's average day (24-hour profile), reducing the total number of time steps from 8,760 to 288. This corresponds to the length of historical data used in simulations. The data was then utilized to generate measurements via power flow calculations using the line-wise power flow model of [50]. Consequently, all meter data in the results section correspond to values derived from this simulation. In practical applications, however, such data would be reported from field measurements.

3.4.2. Estimation performance metrics

To complement the assessment of the described method when providing estimations on unobserved voltage magnitudes and branch flows, the 95th percentile of absolute error (P95 AE) values are also computed. This metric provides the error value below which 95% of the observed deviations fall, capturing a typical upper bound of estimation errors. Note that the following subsections refer to two sets of results: the observed system, which consists of data reported by assumed existing measurements; and the entire network, where errors are calculated using power flow results as references. In the latter case, simulated data is used to extend evaluations and ensure a more comprehensive assessment of the performance of the proposed model.

3.5. Results and Discussion

The first case studies (from Subsection 3.5.1 to Subsection 3.5.6) focus on the 11-bus system to showcase the performance of the resulting formulation under various conditions. Subsection 3.5.6 presents a comparison between the performance of the

method described in this chapter with the original MINLP model from Chapter 2, and alternative linearized formulation versions. Subsections 3.5.7 is dedicated to the 240-bus test system.

3.5.1. Ideal case (11-bus system)

Table 3-2 presents the results for observed bus voltages and unobserved branch flows, assuming 100% MCL, meaning all system buses have voltage measurements. The solution can allocate the exact number of DERs in the system, i.e., $NY = 9$. This means there is no reduction in the number of DERs represented for this case. Historical data is comprised of 288 one-hour time steps with error-free measurements. The bases for per unit calculation are 12.47 kV and 10 MVA.

Results show that with voltage measurements amply available, the proposed method can reproduce the existing voltage information while accurately ($ACC > 95\%$) estimating unobserved branch flows. Taking the maximum historical branch active power flow in the system as reference (4.23 MW, at the substation exit), the mean and 95th percentile error values correspond to around 2.9% and 12.6%, respectively, indicating that high errors ($\geq 15\%$) are limited to a small portion of the branch flow estimates.

Table 3-2. Estimation Performance: 11-bus system, $NY = 9$, 100% MCL, 40% DER.

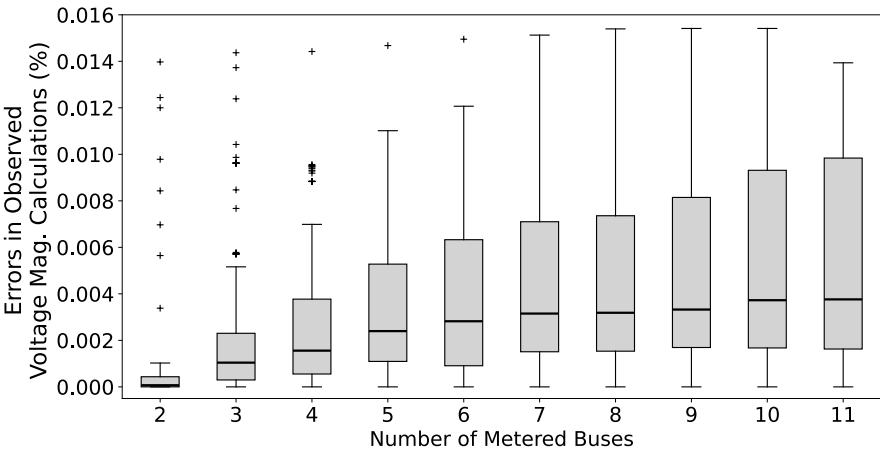
Error Metric	Observed Voltages	Unobserved Branch Flows
MAE (p.u.)	$1.92 \cdot 10^{-5}$	$1.22 \cdot 10^{-2}$
Max AE (p.u.)	$1.41 \cdot 10^{-4}$	$9.72 \cdot 10^{-2}$
P95 AE (p.u.)	$9.90 \cdot 10^{-5}$	$5.31 \cdot 10^{-2}$
ACC (%)	99.99	97.29

3.5.2. Impacts of the number of available measurements (11-bus system)

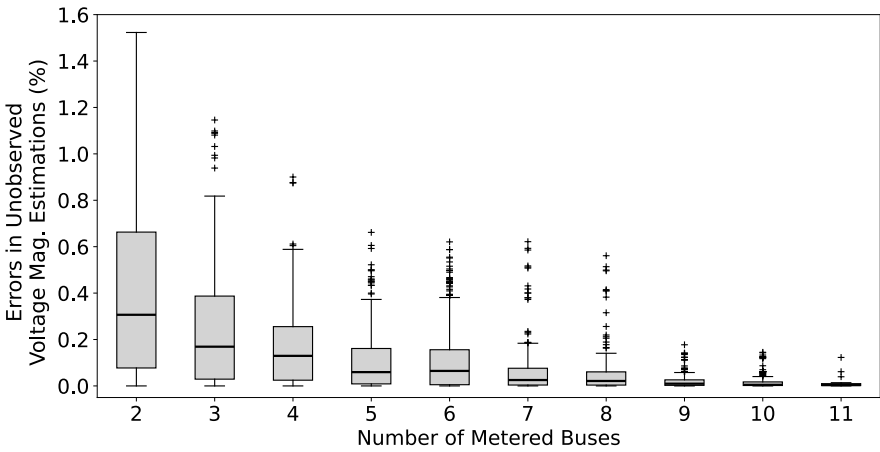
Notably, errors in estimating unobserved quantities reduce as more metered buses exist in the system. On the other hand, errors calculated for observed voltage magnitudes increase as each new measurement introduces variability to the input data. Figure 3.4 shows relative error results. Each boxplot consists of data from simulations considering a fixed number of metered buses—from two buses to the system fully covered with voltage magnitude measurements (100% MCL). Since multiple combinations exist considering the different possible meter locations, ten different scenarios were solved for each MCL by varying meter locations. The voltage at the substation exit (Bus 1) is always available, while the remaining metered buses

were selected using a uniform distribution. Parameter NY varies from 1 to 9 in all cases. The ‘+’ sign represents outliers.

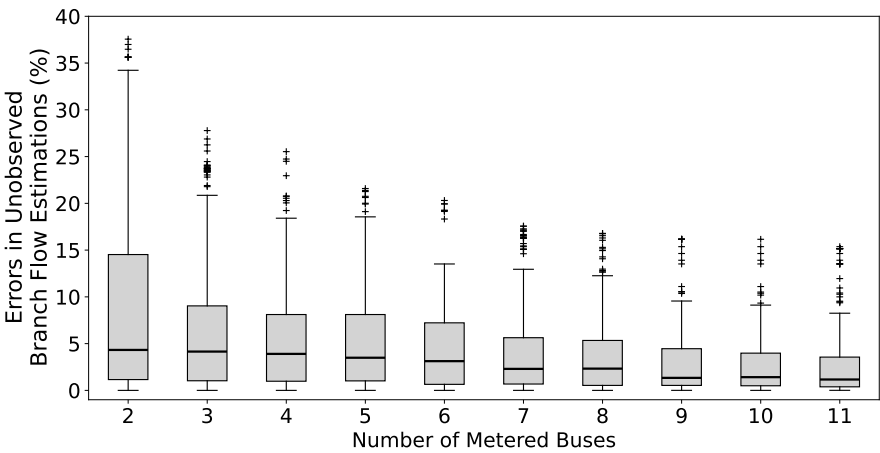
Figure 3.4. Relative estimation errors under different meter coverage levels. (a) Observed voltages. (b) Unobserved voltages. (c) Unobserved branch flows.



(a)



(b)



(c)

Unobserved voltage magnitude estimation errors remain below 1.6%, even in cases with few measurements available. High values ($\geq 15\%$ error) in branch flow estimations are limited to scenarios with fewer than five metered buses and outliers in results. The average error value stays below 5% in all scenarios, while the maximum error values remain below 10% when nine or more metered buses are present in the system (ignoring outliers).

3.5.3. Impacts of the resulting number of DER models (11-bus system)

The number of equivalent DER models, set by parameter NY , balances computational efficiency and accuracy. Fewer DER models speed up subsequent power flow calculations, while having a larger number of models representing many invisible DERs improves accuracy, yet increasing complexity, potentially burdensome as the system size and number of scenarios increase. Figure 3.5 shows the accuracy of unobservable branch flow estimations for varying MCLs and values of NY . Each value corresponds to the average accuracy of ten scenarios solved for each condition. As in the previous section, these variations consider different meter locations while assuming the voltage at the substation end-node available as input to the model.

Figure 3.5. Branch flow estimation accuracy (ACC, %) varying the meter coverage level and the maximum number of aggregate DERs (setting of parameter NY).

NY=1	83.2	87.5	89.1	90.9	93.2	93.8	94.8	94.8	95.8	95.8
NY=2	83.7	87.6	90.4	91.6	93.3	94.1	95.1	95.7	96.3	96.8
NY=3	84.8	89.7	91.6	93.3	94.0	94.3	95.4	95.7	96.3	96.8
NY=4	87.0	92.6	92.7	93.4	94.0	94.3	95.6	96.1	96.3	96.8
NY=5	87.2	93.3	93.7	94.0	94.6	95.8	95.8	96.2	96.5	96.9
NY=6	87.3	93.3	94.2	94.5	95.1	95.8	96.0	96.2	96.7	96.9
NY=7	87.7	93.5	94.5	94.8	95.7	95.8	96.1	96.2	96.8	96.9
NY=8	88.1	93.6	94.6	94.9	95.8	96.2	96.3	96.6	96.8	96.9
NY=9	90.1	94.7	94.7	95.9	96.3	96.6	96.8	96.8	96.9	97.3
	2	3	4	5	6	7	8	9	10	11
	Number of Metered Buses									

There is, on average, a 1.3% estimation accuracy increase for every increment of the number of equivalent DER models the proposed method is allowed to introduce—also, an average increase of 0.5% for each new voltage measurement introduced. However, note that the positive effect of NY in increasing estimation accuracy is more pronounced for the cases with fewer metered buses in the system.

3.5.4. Impacts of the DER penetration level (11-bus system)

The DER penetration level, computed as the ratio of the total DER generating capacity to the peak customer load at the feeder level, also impacts estimation

accuracy. Table 3-3 shows unobservable branch flow estimation results for the ideal scenario ($MCL = 100\%$, $NY = 9$) under varying DER penetration levels, achieved by increasing DER capacity from what was presented in Figure 3.3, while keeping customer loads unchanged. As DER penetration increases, the algorithm demonstrates consistent improvements in estimation accuracy. Results suggest that higher DER penetrations allow the algorithm to better differentiate DER steady-state behavior from customer loads in measurements. This enables more precise allocation of equivalent DERs, thus enhancing unobserved branch flow estimations.

Two statistical tests were employed to confirm that the proposed algorithm maintains stable performance across the different penetration levels. Groups with no statistically significant difference in means and variance are identified using both Welch-t and Levene tests and denoted by matching superscripts (e.g., 1, 2, 3). The former test [90] checks if two groups have different means, especially when the groups might have different variances and sample sizes. Moreover, the latter test [91] checks whether different groups have equal variances, regardless of their means.

Table 3-3. Estimation Performance: 11-bus system, $NY = 9$, 100% MCL, Varying DER Penetration

Error Metric (for Unobserved Branch Flows)	MAE (p.u.)	Max AE (p.u.)	P95 AE (p.u.)	ACC (%)
20% DER¹	$2.25 \cdot 10^{-2}$	$3.19 \cdot 10^{-1}$	$1.66 \cdot 10^{-1}$	95.00
40% DER¹	$1.22 \cdot 10^{-2}$	$9.72 \cdot 10^{-2}$	$5.31 \cdot 10^{-2}$	97.29
60% DER¹	$1.14 \cdot 10^{-2}$	$7.76 \cdot 10^{-2}$	$4.82 \cdot 10^{-2}$	97.47
80% DER²	$9.42 \cdot 10^{-3}$	$6.29 \cdot 10^{-2}$	$4.09 \cdot 10^{-2}$	97.91
100% DER²	$8.54 \cdot 10^{-3}$	$6.01 \cdot 10^{-2}$	$3.73 \cdot 10^{-2}$	98.10

^{1,2}Groups with no statistically significant difference in means and variances based on Welch-t and Levene tests (p-values <0.01).

3.5.5. Impacts of measurement and model parameter errors (11-bus system)

This section analyzes the impacts of four primary error sources on the estimation of unobserved branch flows: (1) noise in voltage magnitude data, (2) uncertainties in branch impedance parameters, (3) variations in the feeder-level aggregated capacity of invisible DERs, and (4) errors in net load measurements aggregated at the feeder level.

- **Voltage magnitude measurement errors**

Table 3-4 demonstrates the impact of varying noise levels in voltage magnitude data. Noise levels, ranging from 0.1% to 6%, were generated using a Gaussian distribution applied uniformly to all measurements. Percentage values correspond to the maximum allowable deviation from actual data. Performance remains stable up to 1.0% noise, but beyond 3.0% significant accuracy degradation occurs. This highlights a potential threshold where the proposed method struggles to maintain high estimation accuracy.

Table 3-4. Estimation Performance: 11-bus system, $NY = 9$, 100% MCL, 40% DER, Noise in Measurements

Error Metric (for Unobserved Branch Flows)	MAE (p.u.)	Max AE (p.u.)	P95 AE (p.u.)	ACC (%)
0.1% Error	$2.23 \cdot 10^{-2}$	$1.21 \cdot 10^{-1}$	$8.06 \cdot 10^{-2}$	96.54
0.2% Error	$2.53 \cdot 10^{-2}$	$1.31 \cdot 10^{-1}$	$8.94 \cdot 10^{-2}$	95.75
0.5% Error¹	$2.69 \cdot 10^{-2}$	$1.42 \cdot 10^{-1}$	$1.36 \cdot 10^{-1}$	95.24
1.0% Error¹	$2.77 \cdot 10^{-2}$	$1.52 \cdot 10^{-1}$	$1.43 \cdot 10^{-1}$	95.13
3.0% Error²	$3.22 \cdot 10^{-2}$	$1.80 \cdot 10^{-1}$	$1.68 \cdot 10^{-1}$	93.91
6.0% Error²	$3.39 \cdot 10^{-2}$	$1.80 \cdot 10^{-1}$	$1.71 \cdot 10^{-1}$	91.71

^{1,2}Groups with no statistically significant difference in means and variances.

- **Branch impedance errors**

Table 3-5 evaluates the effect of branch impedance inaccuracies on the estimation of unobserved branch flows. Errors were introduced as percentage deviations in resistance and reactance values, with error-free voltage magnitude measurements. Estimation accuracy remains consistent across tested impedance error levels. This demonstrates the robustness of the proposed method with respect to this error source.

Table 3-5. Estimation Performance: 11-bus system, $NY = 9$, 100% MCL, 40% DER, Branch Impedance Errors

Error Metric (for Unobserved Branch Flows)	MAE (p.u.)	Max AE (p.u.)	P95 AE (p.u.)	ACC (%)
1% Error¹	$1.22 \cdot 10^{-2}$	$9.57 \cdot 10^{-2}$	$5.22 \cdot 10^{-1}$	97.27
5% Error¹	$1.25 \cdot 10^{-2}$	$9.56 \cdot 10^{-2}$	$5.29 \cdot 10^{-2}$	97.17
15% Error¹	$1.25 \cdot 10^{-2}$	$9.68 \cdot 10^{-2}$	$5.23 \cdot 10^{-2}$	97.15

¹Group with no statistically significant difference in means and variances.

- **Total feeder-level DER capacity errors**

Table 3-6 shows the impact of aggregated DER capacity deviations on estimation accuracy. Errors imply an inaccurate value presented to the model as the total DER generating capacity at the feeder level, per fuel type. The same error was attributed to all DERs in each case, with error-free measurements and network parameters.

Table 3-6. Estimation Performance: 11-bus system, $NY = 9$, 100% MCL, 40% DER, Total DER Capacity Errors

Error Metric (for Unobserved Branch Flows)	MAE (p.u.)	Max AE (p.u.)	P95 AE (p.u.)	ACC (%)
-15% Error	$1.34 \cdot 10^{-2}$	$9.65 \cdot 10^{-2}$	$5.30 \cdot 10^{-2}$	96.78
-5% Error	$1.25 \cdot 10^{-2}$	$9.57 \cdot 10^{-2}$	$5.27 \cdot 10^{-2}$	97.15
-1% Error	$1.24 \cdot 10^{-2}$	$9.36 \cdot 10^{-2}$	$5.06 \cdot 10^{-2}$	97.22
+1% Error¹	$1.23 \cdot 10^{-2}$	$9.72 \cdot 10^{-2}$	$5.31 \cdot 10^{-2}$	97.23
+5% Error¹	$1.25 \cdot 10^{-2}$	$9.74 \cdot 10^{-2}$	$5.32 \cdot 10^{-2}$	97.16
+15% Error¹	$1.26 \cdot 10^{-2}$	$9.77 \cdot 10^{-2}$	$5.34 \cdot 10^{-2}$	97.11

¹Group with no statistically significant difference in means and variances.

Results show that variations in DER capacity have minimal impact on accuracy, with only slight variations ($\leq 1\%$) in overall accuracy values. Note that branch flow

errors are computed across all time steps in historical data. This includes periods when DER output is expected to be zero to insufficient solar irradiance or wind speeds. During these intervals the calculated power injection from equivalent DERs is also zero, leaving branch flow estimation unaffected, even if the total installed DER capacity is not perfectly estimated.

- **Errors in net load measurements aggregated at the feeder level**

Table 3-7 illustrates that while a 1% error in net load measurements has negligible effects on unobserved branch flow estimations, higher deviations—particularly at the 15% error level—result in a substantial decline in estimation accuracy. This drop is more pronounced than other error sources analyzed in this section.

Table 3-7. Estimation Performance: 11-bus system, $NY = 9$, 100% MCL, 40% DER, Net Load Measurement Errors

Error Metric (for Unobserved Branch Flows)	MAE (p.u.)	Max AE (p.u.)	P95 AE (p.u.)	ACC (%)
1% Error	$1.23 \cdot 10^{-2}$	$9.71 \cdot 10^{-2}$	$5.32 \cdot 10^{-2}$	97.25
5% Error	$1.33 \cdot 10^{-2}$	$9.62 \cdot 10^{-2}$	$5.51 \cdot 10^{-2}$	93.33
15% Error	$1.92 \cdot 10^{-2}$	$9.82 \cdot 10^{-2}$	$6.51 \cdot 10^{-2}$	84.56

This deterioration of results happens due to the proposed model's reliance on aggregated load consumption data to distribute the total load in the system across all load nodes and accurately determine the impacts of power injections from customer loads on measurements. Unlike voltage magnitude errors, which affect local measurement points, feeder-level net load errors propagate across multiple buses, influencing estimation results.

3.5.6. Quality of the linearized solution (11-bus system)

In this section, the performance of the linearized formulation is benchmarked against the MINLP model presented on Chapter 2 and two other linearized approaches. Note that the MINLP model was adapted to include the constraints associated with hybrid DER models. The first linearized benchmark model is obtained by assuming that the voltage profile is leveled at 1.0 p.u. when calculating the squared magnitude of the current flowing in the system's lines, eliminating the bilinear term multiplying line currents and node voltage magnitudes. Power losses are computed

normally using (3.3) and (3.4), with updated voltage magnitude values. This approach was addressed in [74] and [75].

The second linearized model depends on the negligibility of power losses. By neglecting active and reactive power losses, the line current variable can be eliminated from the model, which leads to the eradication of the SOC constraint and simplification of the remaining power flow equations. This approach was discussed in [76] and [78], with the authors in [78] stating its possible limitations, especially when addressing highly loaded and longer distribution feeders. Table 3-8 provides an overview of all models and their respective changes in the formulation implemented for testing.

Faster solution times than those of the proposed model are expected from benchmarked linear approaches due to a reduced number of constraints used to model the system's steady-state behavior through power flow equations. For the lossless MILP model, (3.7a)-(3.7d) are removed, while constraint (3.7) and line current variables are removed altogether from the formulation in the MILP 1.0 p.u. model.

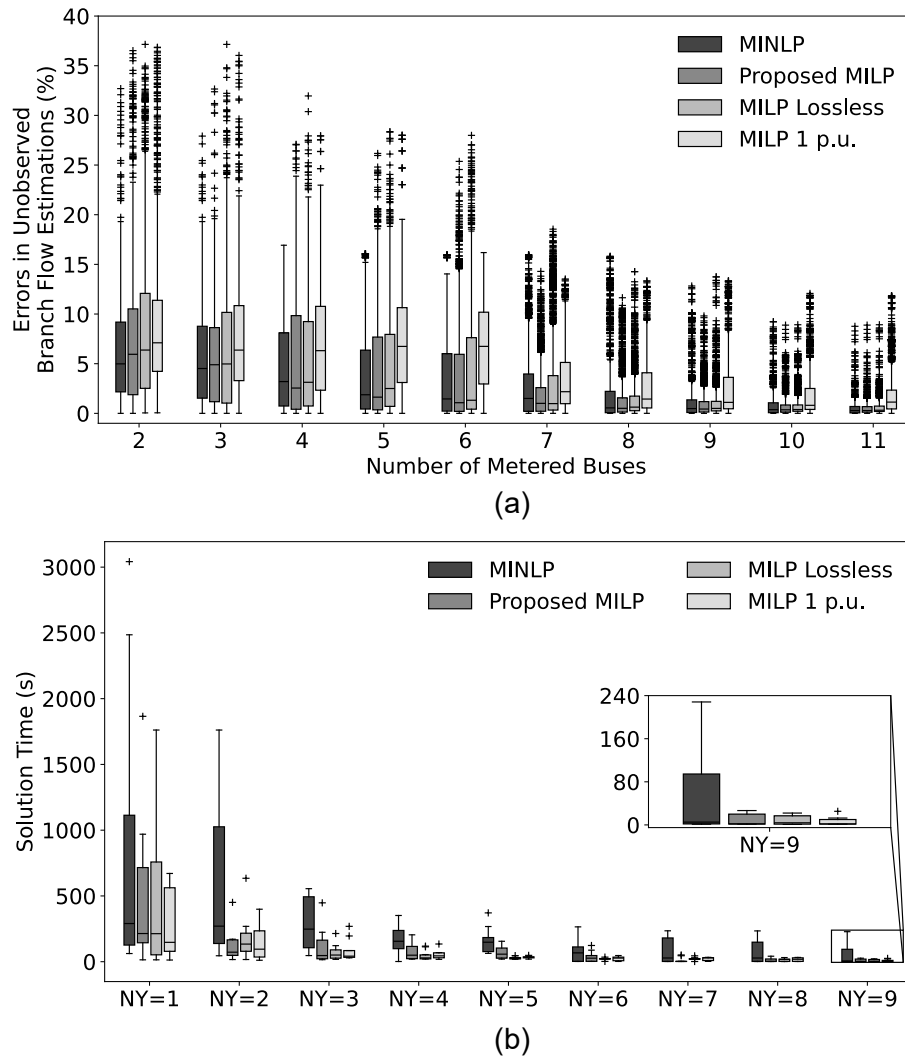
Table 3-8. Formulation Comparison of the Proposed Model and Linear and Nonlinear Benchmark Approaches

Model	Description	Constraint (3.7)
Proposed MILP	Formulation as described by (3.1)-(3.21).	$W_{j,l,t} \geq \sum_{k=1}^i \lambda_k^p \cdot p_k^2 + \sum_{k=1}^j \lambda_k^q \cdot q_k^2$
MINLP (Chapter 2)	Formulation as described by (2.1)-(2.19). SOC relaxation used to calculate line current magnitudes.	$U_{m,t} \cdot J_{l,t} \geq P_{l,t}^2 + Q_{l,t}^2$
MILP Lossless	Neglected losses. Line current variables are removed.	Removed from the formulation.
MILP 1.0 p.u.	Voltages are leveled at 1.0 p.u. to calculate line currents.	$J_{l,t} \geq \sum_{k=1}^i \lambda_k^p \cdot p_k^2 + \sum_{k=1}^j \lambda_k^q \cdot q_k^2$

Figure 3.6 presents boxplots of branch flow estimation errors for the linearized and nonlinear approaches. Voltage magnitude measurements are introduced sequentially as the buses are numbered, with the substation end-node (Bus 1) always available. The parameter NY varies from 1 (a single equivalent DER) to 9 (the actual number of DERs in the original system). There is an overall decreasing trend in estimation errors as the number of metered buses increases for all methods. For lower MCLs (with 2 to 6 metered buses), the median error values range from approximately 5% to 15%, with the MILP 1.0 p.u. method exhibiting higher errors and the largest spread.

For higher MCLs, the median errors decrease significantly for most methods, staying below 5% and with fewer extreme outliers. In general, the MINLP method from Chapter 2 shows lower errors across different MCLs compared to all approaches, with the proposed MILP formulation performing slightly better than the lossless MILP model. This shows that the proposed approach, linearized using McCormick envelopes, provides a reasonable trade-off between computational performance and accuracy compared to the linear variants.

Figure 3.6. Comparisons between the proposed method and other linearized and nonlinear approaches: (a) unobserved branch flow estimation errors, and (b) solution time.



This trade-off is further evidenced by Figure 3.6(b). The MINLP model has higher solution times compared to all MILP models, particularly for small values of NY —with outliers reaching beyond 3000 seconds. The inset for $NY = 9$ shows that

though solution times are much lower than for smaller NY , the MINLP model still presents a tenfold increase in solution time when compared to the MILP models.

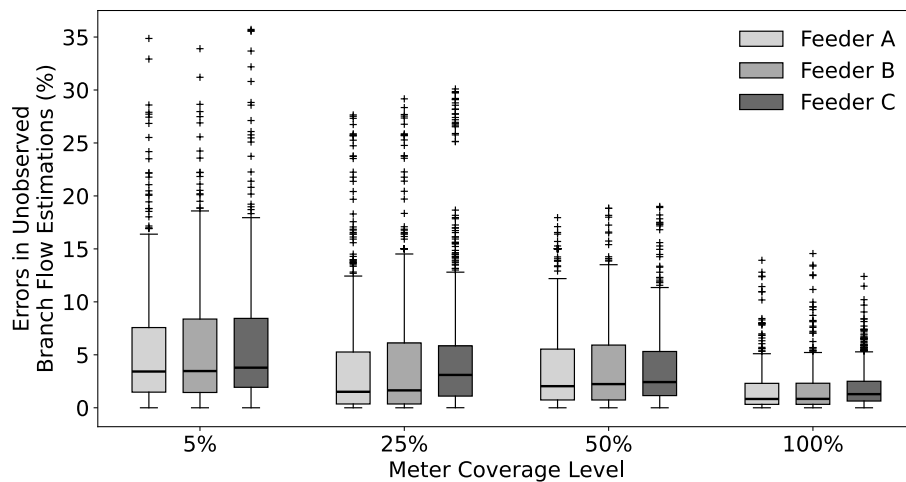
Note that higher solution times happen for lower NY values due to an increased combinatory complexity. For example, if $NY = 1$, the solver must choose a single bus to allocate equivalent DERs from a large set of buses, leading to a highly combinatorial decision problem. Also, since a binary decision variable controls this behavior, having fewer candidate buses for equivalent DER allocation introduces stronger discreteness and makes the problem harder to solve.

3.5.7. A larger distribution system with multiple feeders (240-bus system)

The next results consider the 240-bus distribution test system—a 10-MVA system consisting of three separate feeders supplied by a 69/13.8 kV substation transformer. A 50% penetration scenario was devised by distributing DERs across 107 load buses in the system (50% of the total number of load buses) over the three existing distribution feeders connected to the same substation bus-bar. Figure 3.7 shows branch flow estimation results for Feeders A, B, and C under varying MCLs ranging from 5% to 100%—from a system with sparse voltage magnitude measurements to a fully measured system.

Parameter NY is considered 5 for each feeder, resulting in 15 aggregate DER models in a system with originally 128 DERs. Accurate results display the performance of the proposed method independent of the system size in terms of the number of nodes. At 5% MCL, there is considerable variability in error, particularly for Feeder C, which has the largest number of nodes and a more distributed DER setup. At 100% MCL, all feeders exhibit low errors, reflecting the advantage of full network coverage with voltage magnitude measurements.

Figure 3.7. Branch flow estimation accuracy for all system feeders considering varying meter coverage levels. 50% DER penetration scenario.



Considering the modeling and solution of the proposed formulation for both test systems and the individual feeders on the 240-bus test system, the scalability of the proposed model can be discussed in terms of the computational time required to achieve a solution. Notably, the number of buses in the system directly influences the solution time. For instance, the 11-bus feeder with 288 hourly time steps solves in approximately four minutes ($NY = 1$), while the model requires forty-five minutes for the 163-bus feeder (Feeder C) within the 240-bus test system. While computational performance remains relevant, the method is valid for practical implementation for larger-scale systems, especially considering that it only requires solving once for a given system topology and load/generation level. Therefore, its solution time is not a major limitation for real-world applications.

3.6. Chapter Summary

This chapter continued addressing the challenge of developing aggregate DER models to represent innumerable invisible hybrid DERs in radial distribution systems. Case studies were devised to display the effectiveness of the linearized formulation in locating and sizing equivalent DER models and leveraging the allocated resources to estimate unobserved bus voltage magnitudes and branch flows. The presented analyses showed low average errors ($\leq 5\%$) for estimating unobserved branch flows using limited voltage magnitude data. Higher error values ($\geq 15\%$) were limited to a few cases outside the 95% percentile of estimates.

As per Chapter 2, the sensitivity of estimation results about different testing conditions was discussed. Subsection 3.5.6 compared the performance of the MILP-based approach, the MINLP method from Chapter 2, and two benchmark linearized formulations. The MINLP method from Chapter 2 generally shows lower errors across different meter coverage levels compared to all approaches, with the proposed MILP formulation performing slightly better than the lossless MILP model.

However, in terms of average solution times, the MINLP model has higher solution times than all its counterparts, particularly for small values of NY —with outliers reaching beyond 3000 seconds. The linearized formulation presented in this chapter, employing McCormick envelopes, showed a reasonable trade-off between computational performance and accuracy compared to the linear variants, with a tenfold decrease in solution time compared to the MINLP model.

4. INTEGRATION INTO AN INDUSTRY-STANDARD TOOL FOR DISTRIBUTION SYSTEMS MODELING AND ANALYSIS

4.1. Chapter Introduction

Distribution system decision-making seeks frequent power flow solutions to determine the system state and execute optimal reconfiguration to manage feeder loading. With the growing number of DERs connected to feeders, these algorithms require longer solution times, impeding operational decisions. Hence, utilities require methods to develop aggregate DER models that balance computational efficiency with numerical accuracy. As discussed in previous chapters, existing solutions to increase system and DER visibility rely on extensive measurement infrastructure, which becomes impractical at scale. Thus, there is a need for approaches that can be effective under low system observability, enabling accurate feeder and DER modeling.

Building upon the methodologies developed in previous chapters—where equivalent aggregate DER models were formulated using convex MINLP (Chapter 2) and MILP (Chapter 3) approaches—this chapter focuses on their integration into an industry-standard distribution system analysis tool, OpenDSS [44], which was chosen based on its widespread adoption in the industry and the research community. By implementing the developed equivalent models in OpenDSS, this chapter aims to validate their ability to accurately reproduce system behavior while reducing computational burden.

Previous chapters referred to the single-phase modeling of radial distribution systems, assuming balanced three-phase system conditions. Therefore, another outcome of the integration with OpenDSS is enabling the assessment of the impacts of aggregate DER models in unbalanced systems. Different scenarios are devised by varying the level of system unbalance caused by unbalanced load distributions across all phases. Tested scenarios consider the traditional node voltage and line current unbalance limits of 2% (per bus) and 20% (at the substation exit), respectively [20].

This chapter continues by detailing how previous results were integrated into OpenDSS. Next, simulations and results are discussed, along with an algorithm used to reduce system representation further by removing empty nodes and yield faster solution times with high numerical accuracy.

4.2. Integration with OpenDSS

To facilitate analysis and subsequent processing of results data, the outputs of the Python scripts implemented to solve the optimization problems referred to on Chapter 2 and Chapter 3 were saved on a spreadsheet file format, which are not directly readable by the OpenDSS engine. An auxiliary Python script was implemented to parse results—specifically, to gather system information, equivalent DER and load data, and time-series of load consumption and DER generation profiles. This approach generates a text file (.dss extension) compatible with OpenDSS, following extensive

documentation for model setting and data formatting of [92]. This procedure was facilitated by the DSS-Extensions platform [93].

Consequently, two main system files were generated for each test system used: an original file, with all loads and DERs explicitly modeled (assuming complete system and DER visibility), and a reduced system version with the outcomes of the proposed optimization models. For the analysis under unbalanced system conditions, different unbalanced load distributions were generated by keeping constant the total load over the three phases, but with different peak load values at phases A, B, and C. Voltage unbalance levels are calculated per system node using the root-mean-square voltage magnitudes at the fundamental frequency, as

$$\text{Voltage Unbalance} = 100 \cdot \frac{V^-}{V^+} \quad [\%]$$

where V^+ and V^- are the magnitudes of the positive and negative sequence voltages, respectively.

The current unbalance is calculated for the total feeder current at the substation exit. In normal operating conditions, this value is usually 10-20%.

$$\text{Current Unbalance} = 100 \cdot \frac{I^-}{I^+} \quad [\%]$$

where I^+ and I^- are the magnitudes of the positive and negative sequence of the current at the substation exit, respectively.

4.3. Comparisons Between Detailed and Reduced System Models

The first test cases involve the 240-bus test system presented in Chapter 3. Two testing conditions are considered: first, a comparison between two three-phase unbalanced system models is performed; second, a comparison between two unbalanced system models with balanced and unbalanced equivalent DERs.

The first analysis involves the detailed model with all loads and DERs represented assuming complete system visibility (Model I) and a reduced system version with three-phase balanced DERs modeled according to the outcomes of the proposed MILP model of Chapter 3 (Model II). The second analysis consists of a comparison between two modified versions of Model II to generate system unbalance by redistributing peak customer loads per system phase. Then, the different system models are created by introducing DERs as three-phase balance models (the outcome of the proposed algorithms) and as three-phase unbalanced models following the same unbalance of customer loads. The latter model assumption is valid once it is expected that DERs might be also installed at customer sites with single- and two-phase connections. Moreover, it is assumed that DERs are not allowed to increase the overall system unbalance, according to [20].

Comparisons between Model I and Model II for the 240-bus test system is done in Subsection 4.3.1. These comparisons are redone in Subsection 4.3.2 for the 10-bus

and the 873-bus systems of Chapter 2. Subsection 4.3.2 also introduces an algorithm to remove empty system nodes (i.e., nodes with no load or DER) from the equivalent system files. The algorithm combines lines leading to empty nodes, reducing the total number of nodes and lines in the files corresponding to model II, while yielding accurate voltage and branch flow computations for non-empty nodes.

4.3.1. 240-bus distribution test system

- **Three-phase balanced system analysis**

The first tests consist of comparing results between three versions of the 240-bus test system: a three-phase balanced model with all DERs and loads explicitly modelled assuming complete system visibility (Model I); the equivalent system representation with 15 aggregate DER models derived from the proposed model (Model II, with 5 DERs per feeder, $NY = 5$, $MCL = 100$, 40% DER penetration, error-free measurements); and a modified version of Model I (Model I-II)—a three-phase balanced system version considering that all DERs remain invisible to the system operation and therefore cannot be accurately integrated into the system model.

The latter version assumes that the total feeder-level DER generating capacity is equally distributed across the 15 system nodes with highest peak load consumptions. This is a reasonable approach, as large electricity consumers tend to procure DERs to economically offset the local consumption. Note that the modified system version (Model I-II) results in the same number of aggregate DERs considered in Model II.

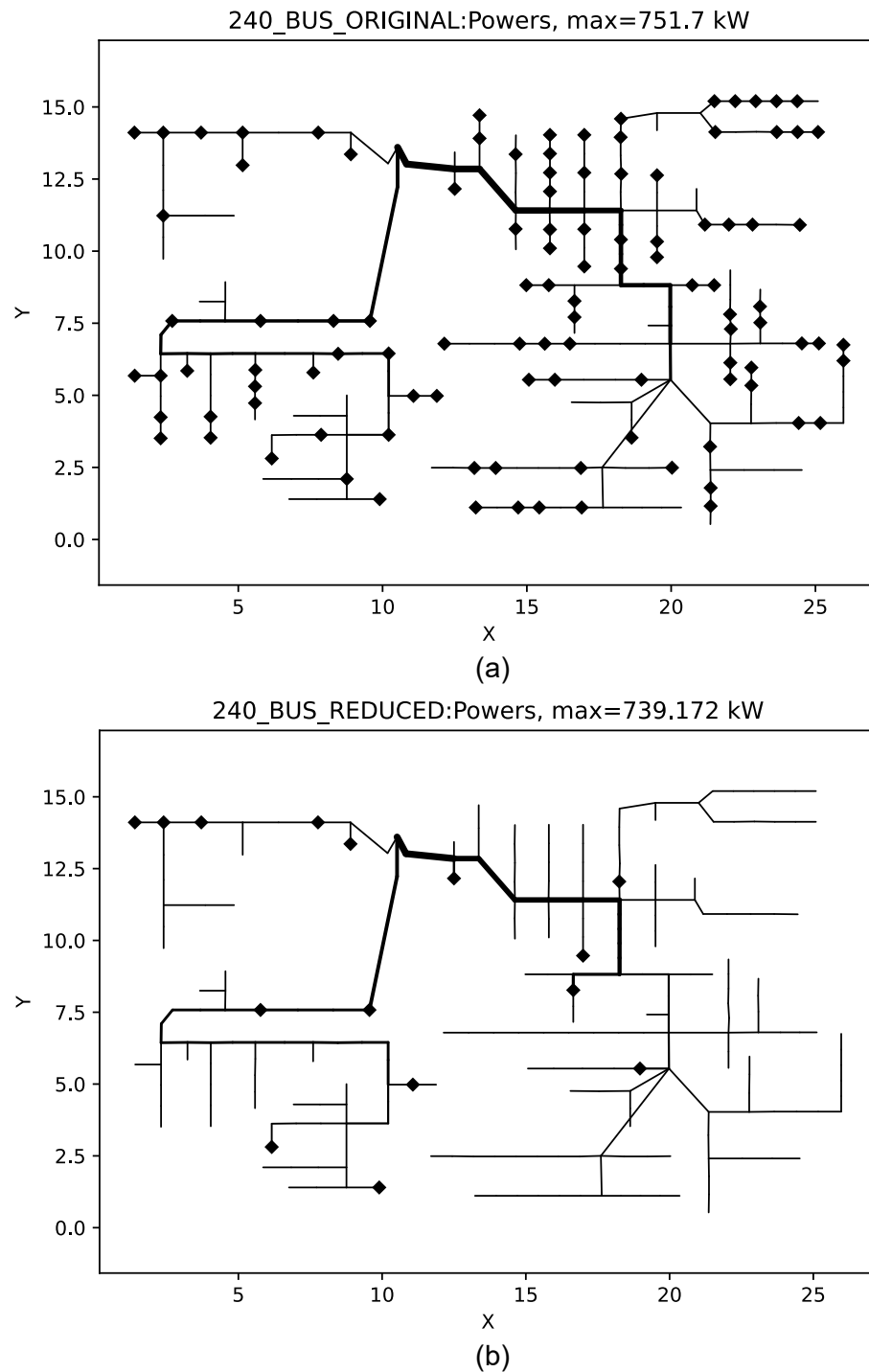
Figure 4.1 illustrates the layouts of the detailed and reduced versions of the test system. Line thicknesses in plots are scaled based on the maximum branch flow values, typically observed at the substation exit. The axis labels refer to the (X,Y) coordinates of system locations. Marked locations in the plots indicate points where DERs connect to the network.

The relative error between the maximum branch flow values of systems I and II is approximately 1.60%, indicating that the reduced system closely approximates the original in terms of branch flow calculations. Considering the deviations in power flow calculations between the original and the reduced system files in OpenDSS, the p.u. MAE, P95 AE, and Max AE values for voltages magnitudes are 0.0583, 0.0966, and 0.1438, respectively.

For Model I-II, the maximum branch flow reported is 755.63 kW. This, in relation to the full system representation of Model I, results in a 2.22% deviation. This result indicates that the proposed method preserves the characteristics of the original system better, with 1.60% deviation. Also, the MAE and Max AE values increase to 0.0758 and 0.1601, respectively, from 0.0583 and 0.1438. Moreover, if DER information is completely absent from system data (e.g., not including feeder-level totals for each DER type), DERs would not be modeled, resulting in a maximum branch flow of 935.69

kW. This shows a 26.56% deviation from the complete system representation of Model I and would lead to highly inaccurate assessments of the operating conditions.

Figure 4.1. System plots for the 240-bus case with DER locations marked. (a) Detailed system with 128 DERs. (b) Reduced version with 15 equivalent DERs.



- **Unbalanced three-phase power flow compatibility check**

In the unbalanced three-phase version of the 240-bus system (Model III), the difference between the most loaded phase (Phase A) from the least loaded phase (Phase B) is 13.88%, with the maximum bus voltage unbalance of 0.09% and current unbalance at the substation exit of 6.99%, respectively. Phase current unbalance levels ranging from 6.99% to 20% were generated by redistributing loads over the three phases while keeping constant the total peak customer load summation per node. Voltage unbalances are calculated for all system buses, while the current value refers to the substation exit. Table 4-1 summarizes unbalance results (i.e., maximum node voltage unbalance and substation current unbalance) for all the tested scenarios. The table also shows comparison results between the system representations with aggregate DER models being introduced as balance and unbalanced three-phase models. Low deviations ($\leq 5\%$) evidence that the representation of equivalent DERs as three-phase balanced model is adequate, even under highly unbalanced system conditions.

Table 4-1. Branch Flow Deviations Under Different System Unbalance Levels

Calculated Current Unbalance (at the substation exit)	6.99%	10%	15%	20%
Maximum Node Voltage Unbalance	0.09%	0.12%	0.16%	0.21%
Average Branch Flow Deviation (between feeder models with balanced and unbalanced aggregate DER models)	1.16%	1.74%	2.00%	3.07%

4.3.2. 10-bus and 873-bus test systems

Table 4-2 summarizes power flow results obtained for the 10-bus and the 873-bus test systems, comparing the performance of the original (Model I) and reduced (Model II) system files. The reduced system with aggregated DERs shows low average voltage magnitude errors. Similarly, average feeder power errors remain low, though maximum errors reach higher values. However, the 95th percentile errors for feeder powers are 9.26% and 8.93% for the 10-bus and 873-bus systems, respectively, indicating that higher errors are limited to a small portion of the system branches.

Table 4-2. Power flow results obtained using OpenDSS. Balanced three phase models for the 10- and 873-bus test systems.

System	Total Demand	DERs (#, MW) – Original System	DERs (#, MW) – Reduced System	MAE (Max AE) – Voltages	MAE (Max AE) – Branch Flows
10-bus	16.0 MW	9 DERs (4 MW)	3 DERs (4 MW)	0.71% (2.32%)	2.98% (26.11%)
873-bus	33.6 MW	350 DERs (16.8 MW)	35 DERs (16.8 MW)	0.13% (1.95%)	2.01% (31.72%)

- **Algorithm for further system reduction**

Empty nodes (i.e., a node with no load or DER) might exist in the reduced system versions with aggregated DER models. Therefore, Algorithm 1 was implemented to combine lines leading to empty nodes, resulting in a further reduced system model in terms of the total number of nodes and lines—which, in turn, affect average power flow calculation times.

Algorithm 1: System Reduction Procedure for Simplifying OpenDSS Models

Input: OpenDSS (.dss) file containing system information.

Parse the input file to extract lists of active loads and DERs.

Parse the input file to extract line connections between nodes.

S_i is the set of lines where bus i is the sending node.

repeat for all lines $l_1: i \rightarrow j$, $l_1 = 1, 2, \dots, NT$

if j is an empty node with $D_j^L, \bar{P}_j^{PV}, \bar{P}_j^{WT}, \bar{P}_j^{BESS} = 0$ **do**

if $S_i = \emptyset$ **do**

remove line l_1 and node n from system data.

else do

for all $i \in S_i$ **do**

create line $l_3 = l_1 + l_2$, $l_3: i \rightarrow k$.

remove line $l_2: j \rightarrow k$ from list of lines.

end

end

end

until all lines are processed or removed.

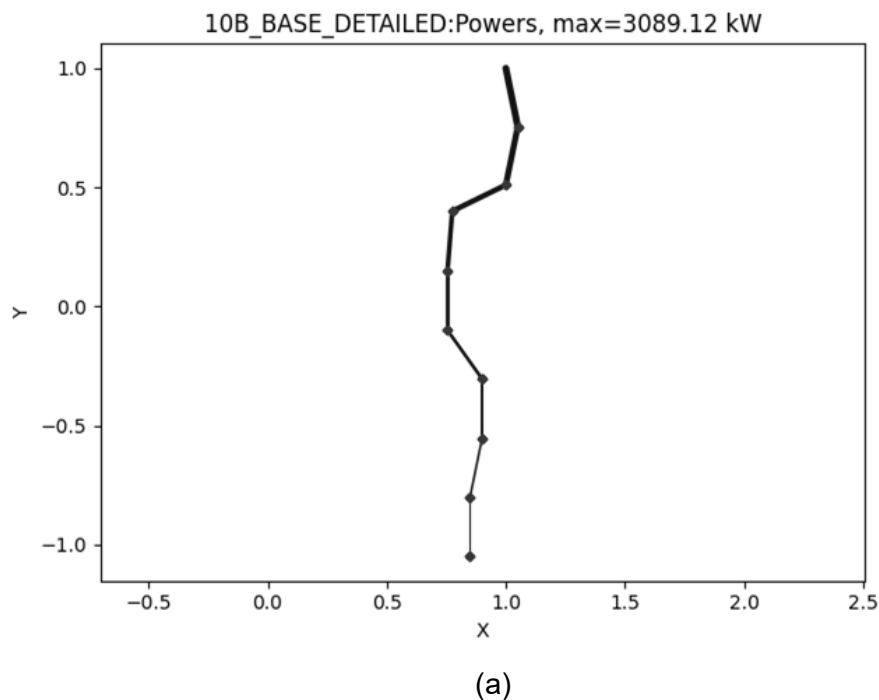
Figure 4.2 and Figure 4.3 illustrate the original and reduced 10-bus and 873-bus system cases. For the 10-bus system, the reduction consolidates the DERs from 9 units into 3, while the 873-bus system reduces the 350 original DERs to 35 aggregated units. DER locations are marked in each plot. The X and Y axes correspond to fictitious geographical coordinates of system node data used to generate the plots. Line thickness corresponds to the branch power flow magnitude at each segment, scaled to the maximum value registered at the top of the figure in kW, typically at the substation end-node.

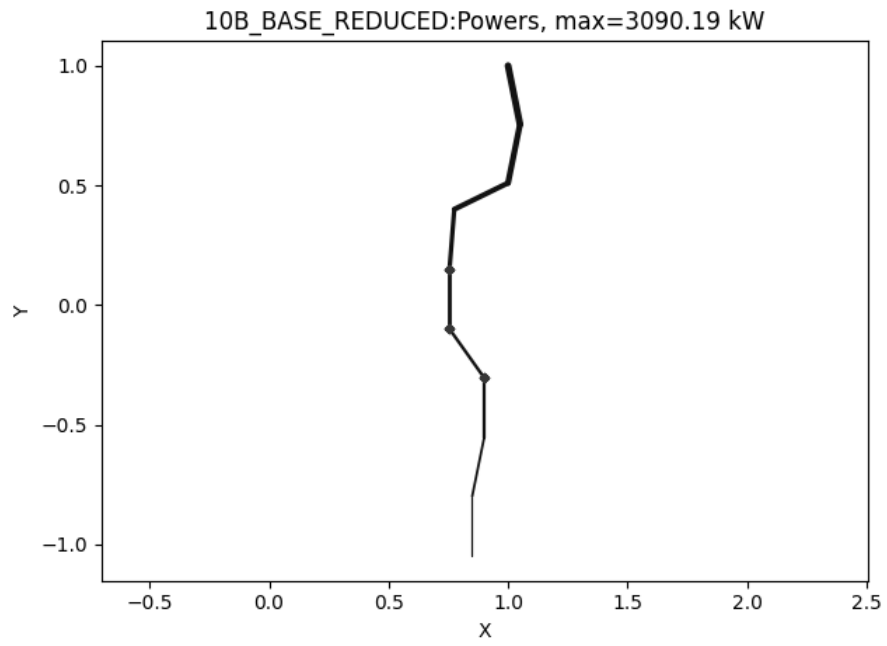
Finally, Table 4-3 compares the solution times required for 10,000 power flow calculations in OpenDSS using both test systems' original and reduced system versions. For the 10-bus system, there was a 29% time reduction, while there was a 40% reduction for the 873-bus system. This evidences the computational advantages of the reduced models, especially for larger systems when solving many power flow scenarios.

Table 4-3. Solution time after 10,000 power flow calculations in OpenDSS.

System	Solution Time	
	Original System	Reduced System
10-bus	0.21 s	0.15 s (↓29%)
873-bus	9.40 s	5.63 s (↓40%)

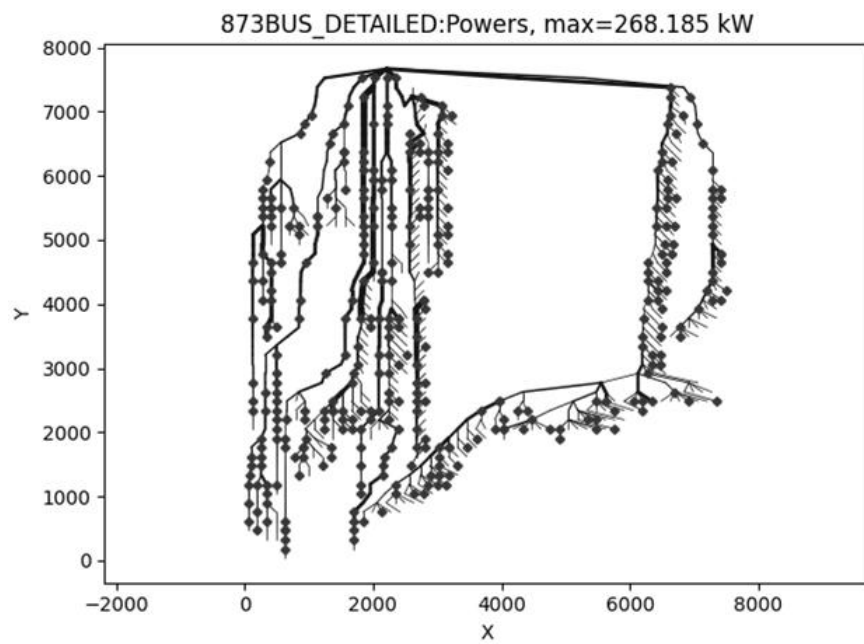
Figure 4.2. 10-bus test feeder. (a) Original system. (b) Reduced system.



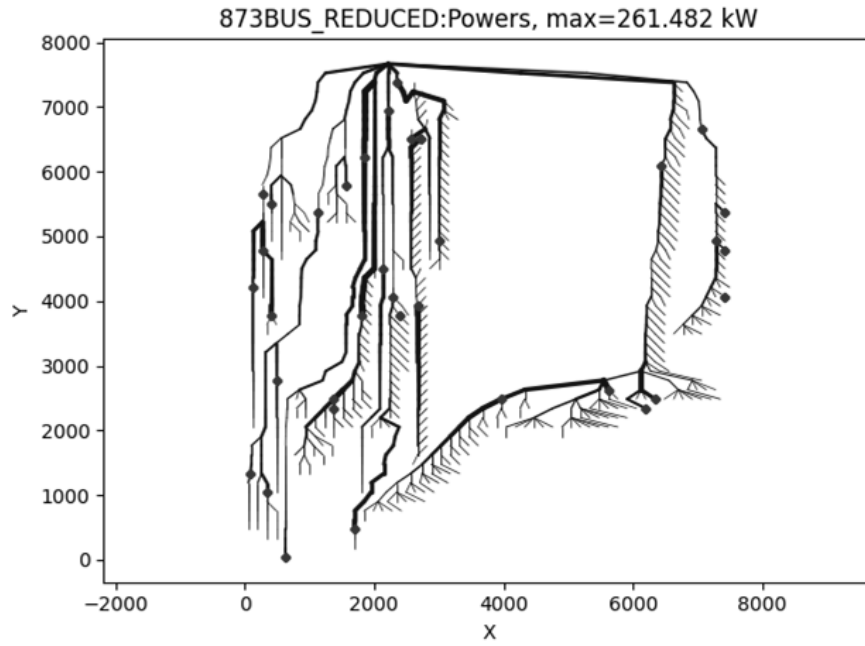


(b)

Figure 4.3. 873-bus test system. (a) Original system. (b) Reduced system.



(a)



(b)

4.4. Chapter Summary

This chapter presented an alternative approach to validate the dual outcomes of the MIP formulations presented in Chapter 2 and Chapter 3—the development of aggregate DER models and estimation of unobserved system states for radial distribution systems. This was done by processing the resulting system information (e.g., location and size of aggregate DER models, peak customer load per node, system topology) and parsing it into an OpenDSS-compatible data structure.

By introducing equivalent feeder models into an industry-standard tool for system modeling and analysis, this approach allows accurate and scalable modeling of DER-rich distribution feeders. Aggregate DER models reduce the computational burden when addressing distribution feeders with innumerable DERs while preserving high numerical accuracy.

Simulation results on 10-bus and 873-bus test systems show branch flow estimation accuracy levels above 95%, even under limited observability conditions. The average computational time required to calculate 10,000 power flow solutions was reduced by around 29% and 40% for the 10-bus and the 873-bus test systems, respectively. Furthermore, the three-phase power flow compatibility check performed for the 240-bus systems evidenced that introducing aggregate DER models into three-phase feeder models in OpenDSS yield accurate ($\leq 5\%$ deviation) results, even under highly unbalanced system conditions.

5. CONCLUSION

5.1. Chapter-Wise Summary

The research presented in this thesis addressed the challenge of improving steady-state analysis in radial distribution systems containing numerous invisible DERs. Optimization-based methods aimed to develop equivalent aggregate DER models representing innumerable invisible resources and to leverage resulting DER models to enhance power flow calculations and enable estimation of unobserved system states without relying on dense measurement infrastructure.

Chapter 2 presented a convex Mixed-Integer Nonlinear Programming (MINLP) formulation to aggregate invisible DERs and estimate system states such as branch power flows and node voltage magnitudes. The formulation enabled accurate system analysis while minimizing measurement requirements and addressing the limitations of traditional state estimation methods, especially when considering the traditional low-observability of power distribution systems. Average branch flow estimation results below 10% were observed using voltage magnitude measurements, even under low meter coverage levels.

Chapter 3 extended the work by introducing linearized versions of the convex MINLP formulation. The MILP-based approach provided an efficient trade-off between computational speed and numerical accuracy, achieving an average tenfold reduction in solution time compared to the MINLP formulation despite having an increased number of constraints due to the introduction of McCormick envelopes and piece-wise linear approximations of the original formulation. Comparisons with benchmark linearized versions commonly employed when addressing radial distribution systems using the line-wise power flow model were performed, demonstrating that the proposed formulation achieves more numerically accurate results at comparable solution times despite having more constraints than the reference linear approaches. Additionally, hybrid DER models (PV-BESS DERs) were introduced to better represent the diversity of modern distribution systems.

Chapter 4 presented an approach to validate the developed aggregate DER models by integrating them into OpenDSS, an industry-standard simulation tool for distribution system analysis. Simulations demonstrated high numerical accuracy (over 95%) when modeling distribution feeders with aggregate DER models and reduced computational burden (up to 40% reduction of average power flow calculation times) across various network sizes and system conditions, including unbalanced, three-phase systems.

5.2. Thesis Contributions

Based on the thesis objectives laid out in Section 1 and the research outcomes presented in the subsequent chapters, this research offers contributions to the field of distribution system modeling. The main takeaways can be summarized as follows. This research:

- Devised MIP-based optimization formulations (MINLP and MILP) to develop aggregate DER models for radial distribution networks, addressing the challenge of low DER observability in these systems.
- The resulting formulations leverage system data that might be readily available to system operators (e.g., substation flow measurements, voltage magnitude measurements from existing metering devices at the bulk feeder level, weather information).
- Introduced hybrid DER models (e.g., PV-BESS) to better capture DER configurations' diverse and evolving nature in modern distribution systems.
- Proposed a framework for estimating unobserved system states, such as branch flows and node voltages, without relying on dense measurement infrastructure.
- Achieved high numerical accuracy (over 90%) when estimating branch flows using node voltage magnitude measurements, even under limited system visibility conditions.
- Demonstrated how the developed aggregate DER models could be integrated into industry-standard simulation tools like OpenDSS, facilitating practical implementation.
- Achieved high numerical accuracy (over 95%) when using aggregate DER models in three-phase distribution system models under balanced and unbalanced conditions. Also, achieved substantial reductions in computational time (up to 40%) for power flow calculations across various test systems.

Note that the proposed methods are particularly suited for cases where innumerable invisible DERs exist, making it impractical for utilities to collect detailed data for each site. If DER penetration is extremely low, no significant change in utilities' practices is needed.

5.3. Directions for Future Research

While the approaches developed in this thesis offer advances in the scalable modeling of DER-rich distribution systems, several areas remain for future research can be pursued. Examples include:

- **Enhanced System Modeling:** Future work could extend the formulated problems to include additional system equipment and operational constraints for more detailed simulations. Examples include introducing voltage regulation equipment, transformer tapping, different load models, and dynamic topological changes.
- **Advanced Hybrid DER Models:** Further research could focus on refining the hybrid DER models to include additional configurations, such as those incorporating demand response and improved representation of battery energy storage DERs, considering the diverse typical behaviors of DER owners based on various socioeconomic factors.

- **Develop Application Cases:** By leveraging aggregate DER models and the system state estimation outcomes of this research, future work could focus on developing application cases for distribution utilities. Examples could involve optimizing meter placement strategies, enhancing feeder reconfiguration and load transfer procedures, and improving existing power flow, optimal power flow, and state estimation workflows.
- **Adaptation to Meshed Networks:** Adapting the aggregate DER modeling techniques to handle more complex network topologies, such as weakly meshed or fully meshed distribution systems, would expand the applicability of the methods to a broader range of power distribution cases.
- **Field Validation and Utility Collaboration:** Collaborating with utilities to validate these models in real-world systems would provide valuable insights and help fine-tune these methods for practical, large-scale deployment.

6. REFERENCES

- [1] J. R. Aguero and A. Khodaei, "Grid modernization, DER integration & utility business models – trends & challenges," *IEEE Power and Energy Mag.*, vol. 16, no. 2, pp. 112-121, 2018.
- [2] U. G. Onu, A. C. Z. Souza, B. D. Bonatto, "Drivers of microgrid projects in developed and developing economies," *Util. Policy*, vol. 80, p. 101487, 2023.
- [3] H. A. Ibrahim, et al., "Sustainability of power generation for developing economies: A systematic review of power sources mix," *Energy Strat. Rev.*, vol. 47, p. 101085, 2023.
- [4] M. Wei, C. A. McMillan, and S. de la Rue du Can, "Electrification of industry: potential, challenges, and outlook," *Curr. Sustainable Renewable Rep.*, vol. 6, pp. 140-148, 2019.
- [5] Y. Zhang, et al. "Mitigating power grid impact from proactive data center workload shifts: a coordinated scheduling strategy integrating synergistic traffic-data-power networks," *Appl. Energy*, vol. 377, no. 1, p. 124697, 2025.
- [6] B. Uzum, Y. Yoldas, S. Baheci, and A. Onem, "Comprehensive review of transmission system operators-distribution system operators collaboration for flexible grid operations," *Electric Power Syst. Res.*, vol. 227, p. 109976, 2024.
- [7] S. Ruester, S. Schwenen, C. Batlle, and I. Pérez-Arriaga, "From distribution to smart distribution systems: Rethinking the regulation of electricity DSOs," *Util. Policy*, vol. 31, pp. 229-237, 2014.
- [8] H. Farias de Barros, M. -C. Alvarez-Herault, B. Raison and Q. T. Tran, "Optimal AC/DC Distribution Systems Expansion Planning From DSO's Perspective Considering Topological Constraints," *IEEE Trans. Power Deliv.*, vol. 38, no. 5, pp. 3417-3428, 2023.
- [9] P. Srivastava, et al. "Voltage regulation in distribution grids: a survey", *Annu. Rev. Control.*, vol. 55, pp. 165-181, 2023.
- [10] J. A. Guzmán-Henao, R. I. Bolaños, O. D. Montoya, L. F. Grisales-Noreña, and H. R. Chamorro, "On integrating and operating distributed energy resources in distribution networks: a review of current solution methods, challenges, and opportunities," *IEEE Access*, vol. 12, pp. 55111-55133, 2024.
- [11] B. P. Cancian, J. C. G. Andrade, W. Freitas, "Technical-economic evaluation of EV fast charging station with distributed energy resources," *J. Control Autom. Elec. Syst.*, vol. 33, pp. 1724-1738, 2022.
- [12] D. Pudjianto, C. Ramsay, and G. Strbac, "Virtual power plant and system integration of distributed energy resources," *IET Ren. Power Gen.*, vol. 1, no. 11, pp.10-16, 2007.
- [13] A. Srivastava, et al., "Distribution system behind-the-meter DERs: estimation, uncertainty quantification, and control," *IEEE Trans. Power Syst.*, vol. 40, no. 1, pp. 1060-1077, 2024.
- [14] A. Baviskar, K. Das, M. Koivisto and A. D. Hansen, "Multi-voltage level active distribution network with large share of weatherdependent generation," *IEEE Trans. Power Syst.*, vol. 37, no. 6, pp. 4874-4884, 2022.
- [15] Y. Li, G. Hao, Y. Liu, Y. Yu, Z. Ni, and Y. Zhao, "Many-objective distribution network reconfiguration via deep reinforcement," *IEEE Trans. Power Deliv.*, vol. 37, no. 3, pp. 2230-2244, 2022.
- [16] S. Li, W. Wu and Y. Lin, "Robust data-driven and fully distributed Volt/VAR control for active distribution networks with multiple virtual power plants," *IEEE Trans. Smart Grid*, vol. 13, no. 4, pp. 2627-2638, 2022.

- [17] C. Qin, S. Bajagain, S. Pannala, A. K. Srivastava, and A. Dubey, "Enhanced situational awareness for DER-rich distribution systems under pre- and post-outage," *IEEE Trans. Power Deliv.*, vol. 38, no. 1, pp. 677-691, 2023.
- [18] M. M. Hosseini, L. Rodriguez-Garcia, and M. Parvania, "Hierarchical combination of deep reinforcement learning and quadratic programming for distribution system restoration," *IEEE Trans. Sustain. Energy*, vol. 14, no. 2, pp. 1088-1098, 2023.
- [19] J. Peppanen, M. Hernandez, J. Deboever, M. Rylander and M. J. Reno, "Distribution load-modeling – Survey of the industry state, current practices and future needs," in *2021 North American Power Symposium (NAPS)*, College Station, TX, USA, 2021.
- [20] Hydro One, "Distributed generation technical interconnection requirements at voltages 50kV and below," DT-10-015 R3, 2013 [Online]. Available: https://hydroone.com/businessservices_/generators_/Documents/Distributed%20Generation%20Technical%20Interconnection%20Requirements.pdf, accessed Nov. 12, 2024.
- [21] Brazil, "Law no. 14.300, January 6, 2022". [In Portuguese]. Available: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2022/lei/l14300.htm, accessed Nov. 12, 2024.
- [22] N. H. Leite, C. P. G. Lascano, H. G. V. Moraes, L. C. P. da Silva, "Impact of the Net-Metering Policies on Solar Photovoltaic Investments for Residential Scale: A Case Study in Brazil," *J. Renew. Energy*, pp. 120788, 2024.
- [23] Energy Research Office, "Data Panel of Micro and Mini Distributed Generation," EPE [In Portuguese]. Available: <https://dashboard.epe.gov.br/apps/pdgd/>, accessed Feb. 01, 2024.
- [24] E. C. Kara, C. M. Roberts, M. Tabone, L. Alvarez, D. S. Callaway and E. M. Stewart, "Disaggregating solar generation from feeder-level measurements," *Sustain. Energy, Grids Netw.*, vol. 13, pp. 112-121, 2018.
- [25] W. Li, M. Yi, M. Wang, Y. Wang, D. Shi and Z. Wang, "Real-time energy disaggregation at substations with behind-the-meter solar generation," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 2023-2034, 2021.
- [26] A. F. Jaramillo, et al., "Effective identification of distributed energy resources using smart meter net-demand data," *IET Smart Grid*, vol. 5, pp. 120-135, 2022.
- [27] X. He, L. Chu, R. C. Qiu, Q. Ai, Z. Ling and J. Zhang, "Invisible units detection and estimation based on random matrix theory," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 1846-1855, 2020.
- [28] N. Mahdavi, D. Weeraddana and Y. Guo, "Probabilistic estimation of PV generation at customer and distribution feeder levels using net demand data," *IEEE Trans. Smart Grid*, vol. 14, no. 3, pp. 1974-1984, 2023.
- [29] F. Bu, R. Cheng and Z. Wang, "A two-layer approach for estimating behind-the-meter PV generation using smart meter data," *IEEE Trans. Power Syst.*, vol. 38, no. 1, pp. 885-896, 2023.
- [30] M. Sun, C. Feng, and J. Zhang, "Factoring behind-the-meter solar into load forecasting: case studies under extreme weather," in *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT-NA)*, Washington, USA, 2020, pp. 1-5.
- [31] A. R. Nikzad, A. A. Mohamed, B. Venkatesh, and J. Penarandra, "Estimating aggregate capacity of connected DERs and forecasting feeder power flow with limited data availability," *IEEE Open Access J. Power Energy*, vol. 11, pp. 266-279, 2024.

- [32] C. Gómez-Peces, S. Grijalva, M. J. Reno and L. Blakely, "Estimation of PV location based on voltage sensitivities in distribution systems with discrete voltage regulation equipment," in *2021 IEEE Madrid PowerTech*, Madrid, Spain, 2021, pp. 1-5.
- [33] M. Markovic, A. Sajadi, A. Florita, R. Cruickshank and B. M. Hodge, "Voltage estimation in low-voltage distribution grids with distributed energy resources," *IEEE Trans. Sustain. Energy*, vol. 12, no. 3, pp. 1640-1650, 2021.
- [34] P. L. Donti, Y. Liu, A. J. Schmidtt, A. Bernstein, R. Yang and Y. Zhang, "Matrix completion for low-observability voltage estimation," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2520-2530, 2020.
- [35] A. Akrami and H. Mohsenian-Rad, "Joint sparse estimation of sensitivity distribution factors and power flows in low-observable power distribution systems," in *2023 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT-NA)*, Washington, USA, 2023, pp. 1-5.
- [36] K. Parlaktuna, E. Dursun, M. C. Cinalioglu and M. Gol, "Improved situational awareness for power distribution systems based on WLAV estimator and direct load flow solution," in *2022 57th International Universities Power Engineering Conference (UPEC)*, Istanbul, Turkey, 2022.
- [37] Y. John, Venkataramanan, A. M. Annaswamy, A. Dubey and A. K. Srivastava, "Improved observability in distribution grids using correlational measurements," *IEEE Access*, vol. 10, pp. 27320-27329, 2022.
- [38] B. A. Acuña Acurio, D. E. Barragán, J. C. Amezcua, M. J. Rider, and L. C. P. Silva, "Design and implementation of a machine learning state estimation model for unobservable microgrids," *IEEE Access*, vol. 10, pp. 123387-123398, 2022.
- [39] T. Su, J. Zhao, Y. Pei, and F. Ding, "Probabilistic physics-informed graph convolutional network for active distribution system voltage prediction," *IEEE Trans. Power Syst.*, vol. 38, no. 6, pp. 5969-5972, 2023.
- [40] J. Zhang, J. Zhao, J. Yang, and J. Zhao, "Deep multi-fidelity Bayesian data fusion for probabilistic distribution system voltage estimation with high penetration of PVs," *IEEE Trans. Power Syst.*, vol. 39, no. 2, pp. 3661-3672, 2024.
- [41] H. Li, Y. Weng, V. Vittal, and E. Blasch, "Distribution grid topology and parameter estimation using deep-shallow neural network with physical consistency," *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 655-666, 2024.
- [42] S. Grafenhorst, K. Förderer, and V. Hagenmeyer, "Distribution grid monitoring based on feature propagation using smart plugs," *Energy Inform.*, vol. 7, no. 116, 2024.
- [43] S. Bhela, V. Kekatos and S. Veeramachaneni, "Enhancing observability in distribution grids using smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5953-5951, 2018.
- [44] Electric Power Research Institute, OpenDSS, EPRI [Online]. Available: <https://www.epri.com/pages/sa/opensdss>, accessed Nov. 12, 2024.
- [45] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, "A Survey on State Estimation Techniques and Challenges in Smart Distribution Systems," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2312-2322, 2019.
- [46] G. Cheng, Y. Lin, A. Abur, A. Gómez-Expósito, and W. Wu, "A Survey of Power System State Estimation Using Multiple Data Sources: PMUs, SCADA, AMI, and Beyond," *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 1129-1151, 2024.

- [47] M. Hernandez, J. Peppanen, J. Deboever, and M. Rylander, "Enhanced Load Modeling: survey results on industry load modeling practices," *Electric Power Research Institute (EPRI), Palo Alto, CA, Technical Update 3002020036*, 2020.
- [48] N. Alguacil, H. Herrero and C. Solares, "Observability Analysis and Restoration for Distribution Grids: A Robust Greedy Numerical Algorithm," *IEEE Trans. Smart Grid*, vol. 14, no. 6, pp. 4542-4552, 2023.
- [49] W. Wei, J. Wang, N. Li and S. Mei, "Optimal Power Flow of Radial Networks and Its Variations: A Sequential Convex Optimization Approach," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2974-2987, 2017.
- [50] Aldik and B. Venkatesh, "Reactive power planning using convex line-wise power balance equations for radial distribution systems," *IET Gener. Transm. Distrib.*, vol. 14, no. 12, pp. 2399-2406, 2020.
- [51] R. G. Cespedes, "New method for the analysis of distribution networks," *IEEE Trans. Power Deliv.*, vol. 5, no. 1, pp. 391-396, 1990.
- [52] M. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. Power Deliv.*, vol. 4, no. 1, pp. 735-743, 1989.
- [53] S. Huang, Q. Wu, J. Wang and H. Zhao, "A Sufficient Condition on Convex Relaxation of AC Optimal Power Flow in Distribution Networks," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1359-1368, 2017.
- [54] A. A. Mohamed and B. Venkatesh, "Line-wise power flow and voltage collapse," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3768-3778, 2018.
- [55] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press: New York, USA, 2009.
- [56] A. J. Monticelli, *Load Flow in Electricity Networks [Fluxo de Carga em Redes de Energia Elétrica]*, Edgard Blucher: São Paulo, 1983.
- [57] J. A. Taylor, *Convex Optimization of Power Systems*, Cambridge University Press: Cambridge, UK, 2015.
- [58] J. S. Giraldo, P. P. Vergara, J. C. López, P. H. Nguyen, and N. G. Paterakis, "A Linear AC-OPF Formulation for Unbalanced Distribution Networks," *IEEE Trans. Ind. Appl.*, vol. 57, no. 5, pp. 4462-4472, 2021.
- [59] J. S. Giraldo, J. A. Castrillon, J. C. López, M. J. Rider, and C. A. Castro, "Microgrids Energy Management Using Robust Convex Programming," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4520-4530, 2019.
- [60] D. Han, J. Jian, and L. Yang, "Outer Approximation and Outer-Inner Approximation Approaches for Unit Commitment Problem," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 505-513, 2014.
- [61] M. Farivar and S. H. Low, "Branch flow model: relaxations and convexification – Part I," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2554-2564, 2013.
- [62] P. Bonami, et al. "An algorithm framework for convex mixed integer nonlinear programs," *Discrete Optim.*, vol. 5, no. 2, pp. 186-204, 2008.
- [63] J. Kronqvist, D. E. Bernal, A. Lundell and I. E. Grossmann, "A review and comparison of solvers for convex MINLP," *Optim. Eng.*, vol. 20, pp. 397-255, 2019.
- [64] P. Bonami, et al., "Basic open-source nonlinear mixed integer programming," *COIN-OR* [Online], 2005. Available: <https://www.coin-or.org/Bonmin/>, accessed Feb. 12, 2025.
- [65] Python Software Foundation, "Python 3.13.3 Documentation," [Online]. Available: <https://docs.python.org/3.13/>, accessed Feb. 24, 2025.
- [66] M. L. Bynum, et al. *Pyomo—Optimization modeling in Python*. 3rd ed., vol. 67, Berlin/Heidelberg, Germany: Springer, 2021.

- [67] AMPL Optimization, "AMPL Python API," [Online]. Available: <https://amplpy.ampl.com/>, accessed Nov. 12, 2024.
- [68] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Deliv.*, vol. 4, no. 1, pp. 725-734, 1989.
- [69] R. Kavaseri and C. Ababei, *REDS: Repository of Distribution Systems*. <https://www.dejazzer.com/reds.html> (accessed September 30, 2024).
- [70] P. Vasconcelos, "10-bus, 69-bus, and 873-bus radial distribution feeder datasets for power flow calculation," [Online]. Available: <https://doi.org/10.21227/5n30-5402>, accessed Feb. 24, 2025.
- [71] S. Pfenninger and I. Staffel, "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, pp. 1251-1265, 2016.
- [72] H. Ahmadi, J. R. Martí, and A. von Meier, "A linear power flow formulation for three-phase distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 5012-5021, 2016.
- [73] M. Markovic and B.-M. Hodge, "Parameterized linear power flow for high fidelity voltage solutions in distribution systems," *IEEE Trans. Power Syst.*, vol. 38, no. 5, pp. 4391-4403, 2023.
- [74] K. Liu, C. Wang, W. Wang, Y. Chen, and H. Wu, "Linear power flow calculation of distribution networks with distributed generation," *IEEE Access*, vol. 7, pp. 44686-44695, 2019.
- [75] J. Yang, N. Zhang, C. Kang, and Q. Xia, "A state-independent linear power flow model with accurate estimation of voltage magnitude," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3607-3617, 2017.
- [76] M. S. Javadi, C. S. Gouveia, L. M. Carvalho, and R. Silva, "Optimal power flow solution for distribution networks using quadratically constrained programming and McCormick relaxation technique," in *IEEE Industrial and Commercial Power Systems Europe (I&CPS 2021)*, Bari, Italy, 2021, pp. 1-6.
- [77] M. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. Power Deliv.*, vol. 4, no. 1, pp. 735-743, 1989.
- [78] B. Alizadeh, M. Sheibani, S. M. Hashemi, and A. Marini, "On the accuracy of linear DistFlow method: a comparison survey," in *9th International Conference on Technology and Energy Management (ICTEM 2024)*, Besharh, Iran, 2024, pp. 1-5.
- [79] B. Li, C. Wan, Y. Li, Y. Jiang, and P. Yu, "Generalized linear-constrained optimal power flow for distribution networks," *IET Gen. Transm. Distrib.*, vol. 17, no. 6, pp. 1298-1309, 2023.
- [80] Z. Yang, H. Zhong, Q. Xia, A. Bose, and C. Kang, "Optimal power flow based on successive linear approximation of power flow equations," *IET Gen. Transm. Distrib.*, vol. 10, no. 14, pp. 3654-3662, 2016.
- [81] R. Golzanari, S. Hasanzadeh, E. Heydarian-Forushani, and I. Kamwa, "Coordinated active and reactive power management for enhancing PV hosting capacity in distribution networks," *IET Ren. Power Gen.*, 2023.
- [82] Y. Zhou, et al., "Convex optimal power flow based on power injection-based equations and its applications in bipolar DC distribution network," *Electr. Power Syst. Res.*, vol. 230, p. 110271, 2024.
- [83] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part I-convex underestimating problems," *Math Program.*, vol. 10, no. 1, pp. 147-175, 1976.

- [84] B. Hamann and J.-L. Chen, "Data point selection for piecewise linear curve approximation," *Comput. Aided Geom. Des.*, vol. 11, no. 3, pp. 289-301, 1994.
- [85] H. Hijazi, C. Coffrin, and P. V. Hentenryck, "Convex quadratic relaxations for mixed-integer nonlinear programs in power systems," *Math. Prog. Comp.*, vol. 9, pp. 321-376, 2017.
- [86] S. Boyd and L. Vandenberghe, *Convex Optimization*, New York: Cambridge University Press, 2009.
- [87] Kai Strunz, et al., "Benchmark systems for network integration of renewable and distributed energy resources," Task Force C6.04.02, Technical Report 575, 2014.
- [88] F. Bu, Y. Yuan, Z. Wang, K. Dehghanpour, and A. Kimber, "A time-series distribution test system based on real utility data," in *2019 North American Power Symposium (NAPS)*, Wichita, KS, USA, 2019, pp. 1-6.
- [89] P. Vasconcelos, "11-bus and 240-bus radial distribution feeder datasets for power flow calculation," [Online]. Available: <https://doi.org/10.21227/pgh4-zm86>, accessed Feb. 24, 2025.
- [90] B. L. Welch, "The generalization of 'Student's' problem when several different population variances are involved," *Biometrika*, vol. 34, no. 1, pp. 28–35, 1947.
- [91] H. Levene, "Robust tests for equality of variances." I. Olkin (Ed.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, pp. 278–292. Stanford University Press, 1960.
- [92] Electric Power Research Institute, "Introduction to OpenDSS", EPRI [Online]. Available: <https://opendss.epri.com/IntroductiontoOpenDSS.html>, accessed Nov. 12, 2024.
- [93] D. Krishnamurthy and P. Meira, "DSS-Extensions," [Online]. Available: <https://dss-extensions.org/index.html>, last accessed Mar. 18, 2025.

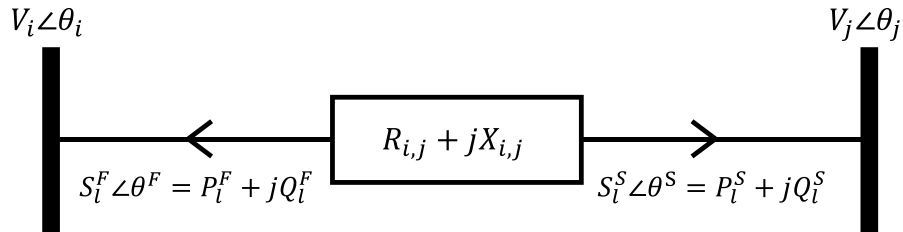
APPENDIX A—DEDUCTION OF THE LINE-WISE POWER FLOW EQUATIONS

In this section, the set of line-wise power balance equations for radial power systems is presented. This deduction provides evidence that this model does not rely on simplifying assumptions of the steady-state behavior of power systems and, therefore, yields accurate results that can be directly compared to other power flow models. If voltage angles are required, they can be calculated as a subsequent step after the convergence of the preferred solution method.

Consider the series impedance element of the pi model of a distribution line or transformer between buses i and j of Figure A.1. Subscripts F and S refer to the power flows in the first and second ends of the line, respectively. Note that the shunt admittance elements are omitted here for clarity but are considered as power injection sources in (A.14) and (A.15).

The variables to be determined in this model are U , P^S , P^F , Q^S , and Q^F . The total number of equations (and therefore, variables) is $NLB + 4 \cdot NT$, with NLB and NT being the number of load buses and number of system branches, respectively. NB refers to the total number of buses in the system.

Figure A.1. Representation of the l -th line segment connecting buses i and j .



Using the voltage difference between the elements of line l , connecting buses i and j (Kirchoff's Voltage Law), yields:

$$V_i \angle \theta_i - I_l \cdot Z_l - V_j \angle \theta_j = 0 \quad (\text{A.1})$$

Rearranging and applying $S_l^S \angle \theta_S = V_j \angle \theta_j \cdot I_j^*$:

$$V_i \angle \theta_i - V_j \angle \theta_j = \left(\frac{S_l^S \angle \theta_S}{V_j \angle \theta_j} \right)^* \cdot Z_l \angle \theta_Z \quad (\text{A.2})$$

Multiplying both sides by the conjugate of $V_j \angle \theta_j$:

$$V_i \angle \theta_i \cdot (V_j \angle \theta_j)^* - V_j^2 = (S_l^S \angle \theta_S)^* \cdot Z_l \angle \theta_Z \quad (\text{A.3})$$

Realizing the multiplication and converting into rectangular representation:

$$V_i \cdot V_j \cdot [\cos(\theta_i - \theta_j) + j \sin(\theta_i - \theta_j)] - V_j^2 = (P_l^S - jQ_l^S) \cdot (R_l + jX_l) \quad (\text{A.4})$$

Taking the real and imaginary parts of (A.4) result in (A.5) and (A.6), respectively:

$$V_i \cdot V_j \cdot \cos(\theta_i - \theta_j) = P_l^S R_l + Q_l^S X_l + V_j^2 \quad (\text{A.5})$$

$$V_i \cdot V_j \cdot \sin(\theta_i - \theta_j) = P_l^S X_l - Q_l^S R_l \quad (\text{A.6})$$

Since $\cos^2 + \sin^2 = 1$, we can add the squares of (A.5) and (A.6) and rearrange to eliminate the resulting elements with opposing signals:

$$V_j^4 + 2V_j^2 \left(P_l^S R_l + Q_l^S X_l - \frac{V_i^2}{2} \right) + (R_l^2 + X_l^2) \cdot (P_l^{S^2} + Q_l^{S^2}) = 0 \quad (\text{A.7})$$

By making a substitution of the squared voltage magnitudes as $U = V^2$:

$$U_j^2 + 2U_j \left(P_l^S R_l + Q_l^S X_l - \frac{U_i}{2} \right) + (R_l^2 + X_l^2) \cdot (P_l^{S^2} + Q_l^{S^2}) = 0 \quad (\text{A.8})$$

Finally, dividing (A.8) by U_j and replacing $(P_l^{S^2} + Q_l^{S^2}) \cdot U_j^{-1}$ by J_l —with J_l representing the square of the magnitude of the current flowing in line l :

$$U_j + 2 \left(P_l^S R_l + Q_l^S X_l - \frac{U_i}{2} \right) + (R_l^2 + X_l^2) \cdot J_l = 0 \quad (\text{A.9})$$

$$J_l = \frac{P_l^{S^2} + Q_l^{S^2}}{U_j} \quad (\text{A.10})$$

Accounting for the power loss over the line impedance and the incoming power flows at the first and second ends of line l , we have:

$$S^F \angle \theta_F + S^S \angle \theta_S + Z_l \angle \theta_Z \cdot I_l^2 = 0 \quad (\text{A.11})$$

By substituting the squared line current magnitude of (A.10) and taking the real and imaginary parts of (A.11), we introduce two expressions that account for the line active and reactive losses:

$$P_l^F + P_l^S + R_l \cdot J_l = 0 \quad (\text{A.12})$$

$$Q_l^F + Q_l^S + X_l \cdot J_l = 0 \quad (\text{A.13})$$

Finally, node-wise power balance equations for active and reactive power are introduced as summations of all incoming branch flows for a specific node, as well as accounting for the local consumption, generation, and the contribution of shunt admittance elements.

$$\sum_{(k,l) \in \Omega F_i} P_{k,i}^S + \sum_{(i,j) \in \Omega S_i} P_{i,j}^F - U_i \cdot G_i = P_i^{Load} - P_i^{Gen} \quad (\text{A.14})$$

$$\sum_{(k,l) \in \Omega F_i} Q_{k,i}^S + \sum_{(i,j) \in \Omega S_i} Q_{i,j}^F + U_i \cdot B_i = Q_i^{Load} - Q_i^{Gen} \quad (\text{A.15})$$

where ΩF_i and ΩS_i refer to the sets of lines that have bus i in their first and second ends, respectively. These expressions can be rewritten by constructing a matrix $\mathbf{M}$ that

is similar to the concept of a bus incidence matrix as follows. For each line l in a system with NT lines:

$$\begin{aligned} [\mathbf{MF}]_{i,l} &= [\mathbf{M}]_{i,l} = 1 && \text{If bus } i \text{ is the first bus for line } l \\ [\mathbf{MS}]_{j,l} &= [\mathbf{M}]_{j,NT+1} = 1 && \text{If bus } j \text{ is the second bus for line } l \\ [\mathbf{MF}]_{i,l} &= [\mathbf{M}]_{i,l} = 1 && \text{otherwise} \end{aligned}$$

Note that the dimension of $\mathbf{M}$ is $NB \times 2 \cdot NT$ and that it can be divided into two smaller matrices as in (A.16).

$$[\mathbf{M}] = [[\mathbf{MF}][\mathbf{MS}]] \quad (\text{A.16})$$

Using $\mathbf{M}$, we can rewrite the node-wise active and reactive power balance equations in matrix form, resulting in (A.17) and (A.18), respectively.

$$[\mathbf{M}] \begin{bmatrix} \mathbf{P}^F \\ \mathbf{P}^S \end{bmatrix} - \mathbf{U} \cdot \mathbf{G} = \mathbf{P}^{Load} - \mathbf{P}^{Gen} \quad (\text{A.17})$$

$$[\mathbf{M}] \begin{bmatrix} \mathbf{Q}^F \\ \mathbf{Q}^S \end{bmatrix} + \mathbf{U} \cdot \mathbf{B} = \mathbf{Q}^{Load} - \mathbf{Q}^{Gen} \quad (\text{A.18})$$

In summary, the line-wise model for representing radial distribution systems consists of solving the set of equations constituted by (A.9), (A.12), (A.13), (A.17), and (A.18) to determine U , P^S , P^F , Q^S , and Q^F . This model was first proposed in [50], with the authors presenting the construction of the Jacobian Matrix for solving the model using the traditional Newton-Raphson method.

APPENDIX B—TEST SYSTEM DATA USED IN SIMULATIONS

This appendix lists the sources to the original data for all test systems used in simulations, as well as provides access to the actual spreadsheets used in this work. Changes to original data consisted of introducing DERs into feeder models, as well as generating bus coordinates for plotting in OpenDSS.

- **10-Bus Radial Distribution Test Feeder:**

Original Source: This test system was developed during this research.

Data Access: <https://doi.org/10.21227/5n30-5402>

- **11-Bus Cigré Radial Distribution Feeder:**

Original Source: Kai Strunz, et al., “Benchmark systems for network integration of renewable and distributed energy resources,” Task Force C6.04.02, Technical Report 575, 2014.

Data Access: <https://doi.org/10.21227/pgh4-zm86>

- **IEEE 69-Bus Radial Distribution Test System:**

Original Source: M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," IEEE Trans. Power Deliv., vol. 4, no. 1, pp. 725-734, 1989.

Data Access: <https://doi.org/10.21227/5n30-5402>

- **Iowa State University’s 240-Bus Radial Distribution Test System:**

Original Source: F. Bu, Y. Yuan, Z. Wang, K. Dehghanpour, and A. Kimber, “A time-series distribution test system based on real utility data,” in *2019 North American Power Symposium (NAPS)*, Wichita, KS, USA, 2019, pp. 1-6.

Data Access: <https://doi.org/10.21227/pgh4-zm86>

- **North Dakota State University’s 873-Bus Radial Distribution Feeder:**

Original Source: R. Kavaseri and C. Ababei, *REDS: Repository of Distribution Systems*. <https://www.dejazzer.com/reds.html> (accessed September 30, 2024).

Data Access: <https://doi.org/10.21227/5n30-5402>

- **Weather Data**

The weather data with one-year wind speed and direct solar irradiance values with 1-hour intervals (8760 data points) is available alongside the abovementioned system files. The 24-hour monthly average values used in Chapter 3 are provided in Table B-1.

Table B-1. Summary of the average monthly 24-hour weather data profile used in simulations.

Time	H	M	Wspd (m/s)	Irrad (W/m ²)	Time	H	M	Wspd (m/s)	Irrad (W/m ²)	Time	H	M	Wspd (m/s)	Irrad (W/m ²)
0	0	Jan	7.9	0.0	96	0	May	7.5	0.0	192	0	Sep	9.1	0.0
1	1	Jan	7.9	0.0	97	1	May	7.5	0.0	193	1	Sep	9.0	0.0
2	2	Jan	7.9	0.0	98	2	May	7.5	0.0	194	2	Sep	8.8	0.0
3	3	Jan	8.0	0.0	99	3	May	7.6	0.0	195	3	Sep	8.6	0.0
4	4	Jan	8.0	0.0	100	4	May	7.6	0.0	196	4	Sep	8.5	0.0
5	5	Jan	8.0	0.0	101	5	May	7.0	0.0	197	5	Sep	8.1	0.0
6	6	Jan	7.9	0.0	102	6	May	6.8	13.3	198	6	Sep	7.6	2.6
7	7	Jan	7.9	0.7	103	7	May	7.4	97.1	199	7	Sep	7.6	47.9
8	8	Jan	8.0	84.7	104	8	May	7.7	243.6	200	8	Sep	8.0	152.4
9	9	Jan	8.1	217.1	105	9	May	7.9	376.9	201	9	Sep	8.3	270.8
10	10	Jan	8.1	372.0	106	10	May	7.9	470.6	202	10	Sep	8.4	395.3
11	11	Jan	8.1	454.8	107	11	May	8.0	507.9	203	11	Sep	8.6	474.6
12	12	Jan	8.1	469.0	108	12	May	8.0	503.4	204	12	Sep	8.8	503.0
13	13	Jan	8.1	435.0	109	13	May	8.0	463.9	205	13	Sep	8.9	475.4
14	14	Jan	8.0	334.6	110	14	May	8.1	384.3	206	14	Sep	8.9	362.9
15	15	Jan	8.1	192.8	111	15	May	8.1	269.5	207	15	Sep	8.8	256.0
16	16	Jan	8.2	50.7	112	16	May	8.0	150.2	208	16	Sep	8.6	133.2
17	17	Jan	8.2	0.1	113	17	May	7.8	56.1	209	17	Sep	8.7	40.6
18	18	Jan	8.0	0.0	114	18	May	8.0	3.5	210	18	Sep	9.0	0.7
19	19	Jan	7.7	0.0	115	19	May	8.4	0.0	211	19	Sep	9.1	0.0
20	20	Jan	7.4	0.0	116	20	May	8.4	0.0	212	20	Sep	9.2	0.0
21	21	Jan	7.5	0.0	117	21	May	8.1	0.0	213	21	Sep	9.2	0.0
22	22	Jan	7.6	0.0	118	22	May	8.0	0.0	214	22	Sep	9.1	0.0
23	23	Jan	7.8	0.0	119	23	May	8.0	0.0	215	23	Sep	9.0	0.0
24	0	Feb	6.5	0.0	120	0	Jun	7.3	0.0	216	0	Oct	8.8	0.0
25	1	Feb	6.5	0.0	121	1	Jun	7.4	0.0	217	1	Oct	8.8	0.0
26	2	Feb	6.6	0.0	122	2	Jun	7.3	0.0	218	2	Oct	8.8	0.0
27	3	Feb	6.8	0.0	123	3	Jun	7.0	0.0	219	3	Oct	8.8	0.0
28	4	Feb	7.0	0.0	124	4	Jun	6.6	0.0	220	4	Oct	8.8	0.0
29	5	Feb	7.0	0.0	125	5	Jun	5.7	0.0	221	5	Oct	8.8	0.0
30	6	Feb	7.1	0.0	126	6	Jun	5.6	15.5	222	6	Oct	8.7	0.0
31	7	Feb	7.2	9.1	127	7	Jun	6.0	98.2	223	7	Oct	8.4	27.4
32	8	Feb	7.2	71.9	128	8	Jun	6.4	190.8	224	8	Oct	8.7	157.1
33	9	Feb	7.4	177.8	129	9	Jun	6.5	295.6	225	9	Oct	9.3	347.2
34	10	Feb	7.6	284.4	130	10	Jun	6.5	354.2	226	10	Oct	9.7	465.5
35	11	Feb	7.8	352.1	131	11	Jun	6.5	380.6	227	11	Oct	9.9	505.4
36	12	Feb	7.9	350.9	132	12	Jun	6.6	389.0	228	12	Oct	9.9	506.0
37	13	Feb	7.9	306.1	133	13	Jun	6.5	377.9	229	13	Oct	9.8	479.5
38	14	Feb	7.7	229.9	134	14	Jun	6.3	315.0	230	14	Oct	9.5	394.4
39	15	Feb	7.4	137.6	135	15	Jun	6.1	235.7	231	15	Oct	9.2	264.1
40	16	Feb	7.3	50.5	136	16	Jun	5.9	147.8	232	16	Oct	9.2	130.1
41	17	Feb	7.4	4.5	137	17	Jun	5.8	64.9	233	17	Oct	9.4	20.0
42	18	Feb	7.5	0.0	138	18	Jun	6.0	8.2	234	18	Oct	9.5	0.0
43	19	Feb	7.3	0.0	139	19	Jun	6.3	0.0	235	19	Oct	9.5	0.0
44	20	Feb	7.2	0.0	140	20	Jun	6.5	0.0	236	20	Oct	9.3	0.0
45	21	Feb	7.1	0.0	141	21	Jun	6.6	0.0	237	21	Oct	9.0	0.0
46	22	Feb	6.9	0.0	142	22	Jun	6.8	0.0	238	22	Oct	8.8	0.0

47	23	Feb	6.7	0.0	143	23	Jun	7.0	0.0	239	23	Oct	8.6	0.0
48	0	Mar	8.0	0.0	144	0	Jul	7.2	0.0	240	0	Nov	8.2	0.0
49	1	Mar	7.9	0.0	145	1	Jul	7.3	0.0	241	1	Nov	8.1	0.0
50	2	Mar	7.7	0.0	146	2	Jul	7.2	0.0	242	2	Nov	8.0	0.0
51	3	Mar	7.7	0.0	147	3	Jul	7.0	0.0	243	3	Nov	7.9	0.0
52	4	Mar	7.8	0.0	148	4	Jul	6.8	0.0	244	4	Nov	7.8	0.0
53	5	Mar	8.0	0.0	149	5	Jul	6.2	0.0	245	5	Nov	7.8	0.0
54	6	Mar	7.8	1.6	150	6	Jul	5.6	10.9	246	6	Nov	8.0	0.0
55	7	Mar	7.4	55.4	151	7	Jul	5.9	103.5	247	7	Nov	8.0	1.9
56	8	Mar	7.4	193.2	152	8	Jul	6.2	222.6	248	8	Nov	7.9	64.8
57	9	Mar	7.6	342.9	153	9	Jul	6.2	327.4	249	9	Nov	8.0	211.3
58	10	Mar	7.8	449.2	154	10	Jul	6.3	423.6	250	10	Nov	8.3	331.8
59	11	Mar	7.9	511.2	155	11	Jul	6.2	475.1	251	11	Nov	8.6	403.0
60	12	Mar	8.0	499.2	156	12	Jul	6.1	492.1	252	12	Nov	8.6	432.2
61	13	Mar	8.2	448.4	157	13	Jul	6.0	456.9	253	13	Nov	8.4	383.1
62	14	Mar	8.3	378.2	158	14	Jul	5.9	391.5	254	14	Nov	8.2	301.0
63	15	Mar	8.2	265.6	159	15	Jul	5.8	283.2	255	15	Nov	8.1	195.6
64	16	Mar	8.0	144.5	160	16	Jul	5.7	171.5	256	16	Nov	8.2	80.8
65	17	Mar	8.3	41.3	161	17	Jul	5.6	76.8	257	17	Nov	8.2	0.6
66	18	Mar	8.7	1.0	162	18	Jul	5.9	10.8	258	18	Nov	8.2	0.0
67	19	Mar	8.7	0.0	163	19	Jul	6.3	0.0	259	19	Nov	8.3	0.0
68	20	Mar	8.5	0.0	164	20	Jul	6.6	0.0	260	20	Nov	8.4	0.0
69	21	Mar	8.3	0.0	165	21	Jul	6.8	0.0	261	21	Nov	8.4	0.0
70	22	Mar	8.2	0.0	166	22	Jul	6.8	0.0	262	22	Nov	8.4	0.0
71	23	Mar	8.0	0.0	167	23	Jul	6.9	0.0	263	23	Nov	8.4	0.0
72	0	Apr	8.7	0.0	168	0	Aug	7.2	0.0	264	0	Dec	7.3	0.0
73	1	Apr	8.6	0.0	169	1	Aug	7.2	0.0	265	1	Dec	7.4	0.0
74	2	Apr	8.5	0.0	170	2	Aug	7.0	0.0	266	2	Dec	7.4	0.0
75	3	Apr	8.3	0.0	171	3	Aug	6.8	0.0	267	3	Dec	7.3	0.0
76	4	Apr	8.1	0.0	172	4	Aug	6.6	0.0	268	4	Dec	7.2	0.0
77	5	Apr	7.9	0.0	173	5	Aug	6.0	0.0	269	5	Dec	7.1	0.0
78	6	Apr	7.5	4.8	174	6	Aug	5.4	5.7	270	6	Dec	7.0	0.0
79	7	Apr	7.7	53.9	175	7	Aug	5.6	88.2	271	7	Dec	7.0	0.0
80	8	Apr	8.2	153.6	176	8	Aug	6.1	219.5	272	8	Dec	6.9	33.4
81	9	Apr	8.4	274.5	177	9	Aug	6.4	369.5	273	9	Dec	6.7	128.9
82	10	Apr	8.6	368.9	178	10	Aug	6.7	509.4	274	10	Dec	6.6	252.4
83	11	Apr	8.5	429.2	179	11	Aug	6.8	589.4	275	11	Dec	6.7	355.5
84	12	Apr	8.3	454.6	180	12	Aug	6.8	620.6	276	12	Dec	6.9	399.4
85	13	Apr	8.3	422.0	181	13	Aug	6.9	565.6	277	13	Dec	7.0	386.1
86	14	Apr	8.2	363.0	182	14	Aug	7.0	464.5	278	14	Dec	7.2	314.7
87	15	Apr	8.1	263.1	183	15	Aug	6.9	322.7	279	15	Dec	7.7	189.2
88	16	Apr	7.8	151.6	184	16	Aug	6.6	179.3	280	16	Dec	8.0	44.7
89	17	Apr	7.9	57.1	185	17	Aug	6.6	73.0	281	17	Dec	8.2	0.0
90	18	Apr	8.2	1.7	186	18	Aug	7.0	4.4	282	18	Dec	8.2	0.0
91	19	Apr	8.4	0.0	187	19	Aug	7.2	0.0	283	19	Dec	8.0	0.0
92	20	Apr	8.3	0.0	188	20	Aug	7.4	0.0	284	20	Dec	7.7	0.0
93	21	Apr	8.3	0.0	189	21	Aug	7.5	0.0	285	21	Dec	7.6	0.0
94	22	Apr	8.4	0.0	190	22	Aug	7.4	0.0	286	22	Dec	7.4	0.0
95	23	Apr	8.6	0.0	191	23	Aug	7.4	0.0	287	23	Dec	7.4	0.0

APPENDIX C—MAIN PUBLICATIONS RELATED TO THIS RESEARCH

- *P. N. Vasconcelos, F. C. L. Trindade, B. Venkatesh, W. Freitas, A. C. Zambroni de Souza, and G. N. Taranto, “A Mixed-Integer Nonlinear Model to Support the Operation of Distribution Systems with Hidden DERs,” IEEE Transactions on Power Delivery, vol. 40, no. 1, pp. 484-496, Feb. 2025.*

This paper presents the original convex MINLP formulation for aggregate DER modeling and estimating unobserved system states in radial distribution systems. This is an outcome of a research project in partnership with a distribution utility company in Ontario, Canada. (Chapter 2)

- *P. N. Vasconcelos, F. C. L. Trindade, and B. Venkatesh, “Linearized Optimization for Enhanced Aggregate Modeling of Invisible Hybrid Distributed Energy Resources,” IET Generation, Transmission & Distribution, vol. 19, no. 1, p. e70088, May 2025.*

This paper explores linearized formulations to enhance the solution of the aggregate DER allocation and system estimation problems. The resulting formulation enables the modeling of hybrid DERs (e.g., PV+BESS), provides estimates with high numerical accuracy, and results that are twice as fast as the original formulation. The integration of results with OpenDSS, an industry-standard power systems modeling and analysis tool, is also discussed. (Chapter 3)

The content of Chapter 4 was, in part, organized as a conference paper and accepted for presentation at the 2025 CIGRE International Symposium. The conference will be held in Montréal, Canada, from Sep. 29, 2025, to Oct. 3, 2025.

Other publications were realized during this degree (from 2020 to 2025) but are not directly related to this work. The main (unrelated) research outputs involve topics in engineering education (two journal papers), multidisciplinary dynamic system models (a conference paper and a book chapter), and other optimization applications into power systems planning (two conference papers). A complete list of publications is available at:

<https://scholar.google.ca/citations?user=YwiXb9sAAAAJ&hl=en&oi=ao>.