Data-enabled approaches for planning and operation of smart and sustainable power distribution system

by

Reza Nematirad

B.Sc., Amirkabir University of Technology, 2018

M.Sc., Amirkabir University of Technology, 2021

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Mike Wiegers Department of Electrical and Computer Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2026

Abstract

The increasing penetration of renewable energy and digital monitoring technologies is transforming the power distribution grid into a more complex, data-rich environment. While these advancements promise improved sustainability and efficiency, they also introduce new challenges in planning, operation, and real-time decision-making. This dissertation presents a unified data-driven framework that addresses two fundamental challenges in modern power distribution systems: optimal integration of photovoltaic (PV) and battery storage systems, and enhanced situational awareness through advanced data analytics.

In the first part of this dissertation, a statistical and optimization-based methodology is developed for determining the optimal sizing of PV and battery storage systems in rural power networks. The proposed approach captures the inherent uncertainties of solar irradiance and electricity demand through Monte Carlo simulation and time-series clustering. By evaluating generated scenarios, the framework identifies PV–battery configurations that satisfy a specified peak-demand reduction requirement while maximizing financial benefit. A probabilistic formulation based on Gamma distributions and Kolmogorov–Smirnov goodness-of-fit testing is introduced to estimate the 95th-percentile peak-demand threshold, enabling statistical evaluation of system sizing under uncertainty. The resulting methodology provides utilities and planners with practical guidelines for cost-effective PV–battery deployment.

The second part of the dissertation focuses on improving situational awareness in distribution networks through electricity demand forecasting and distribution system state estimation (DSSE) using advanced metering infrastructure (AMI) data. A deep learning forecasting model, Seasonal Periodic Decomposition Network (SPDNet), is developed to capture the seasonal, periodic, and short-term variations in residential electricity demand. The proposed

model combines seasonal-trend decomposition with periodic decomposition to capture temporal dependencies at different scales and generate forecast-based pseudo-measurements for state estimation. To address the limited availability of near-real-time AMI measurements, Bayesian Weighted Least Squares (BWLS) and Variational Bayesian Matrix Completion (VBMC) are employed to estimate the voltage states of the distribution network. The resulting estimates are combined using calibrated Maximum Ratio Combining (MRC), which accounts for the bias and uncertainty associated with each estimator. The results demonstrate that the calibrated fusion approach improves voltage-estimation accuracy compared with the individual estimators and conventional MRC. Furthermore, a comparison between individual and transformer-aggregated load forecasting shows that reducing the number of forecasting profiles from 646 individual loads to 58 aggregated groups provides nearly identical BWLS state-estimation accuracy while reducing the number of forecasting tasks.

Collectively, the proposed frameworks contribute to the planning and operation of modern power distribution systems by combining statistical modeling, optimization, machine learning, and state estimation. The results demonstrate the effectiveness of probabilistic PV–battery sizing for peak-demand reduction and economic benefit, as well as the potential of forecasting-aided and calibrated state estimation to improve situational awareness under limited measurement availability.

Data-enabled approaches for planning and operation of smart and sustainable power distribution system

by

Reza Nematirad

B.Sc., Amirkabir University of Technology, 2018

M.Sc., Amirkabir University of Technology, 2021

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Mike Wiegers Department of Electrical and Computer Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2026

Approved by:

Major Professor
Dr. Anil Pahwa

Copyright

© Reza Nematirad 2026.

Abstract

The increasing penetration of renewable energy and digital monitoring technologies is transforming the power distribution grid into a more complex, data-rich environment. While these advancements promise improved sustainability and efficiency, they also introduce new challenges in planning, operation, and real-time decision-making. This dissertation presents a unified data-driven framework that addresses two fundamental challenges in modern power distribution systems: optimal integration of photovoltaic (PV) and battery storage systems, and enhanced situational awareness through advanced data analytics.

In the first part of this dissertation, a statistical and optimization-based methodology is developed for determining the optimal sizing of PV and battery storage systems in rural power networks. The proposed approach captures the inherent uncertainties of solar irradiance and electricity demand through Monte Carlo simulation and time-series clustering. By evaluating generated scenarios, the framework identifies PV–battery configurations that satisfy a specified peak-demand reduction requirement while maximizing financial benefit. A probabilistic formulation based on Gamma distributions and Kolmogorov–Smirnov goodness-of-fit testing is introduced to estimate the 95th-percentile peak-demand threshold, enabling statistical evaluation of system sizing under uncertainty. The resulting methodology provides utilities and planners with practical guidelines for cost-effective PV–battery deployment.

The second part of the dissertation focuses on improving situational awareness in distribution networks through electricity demand forecasting and distribution system state estimation (DSSE) using advanced metering infrastructure (AMI) data. A deep learning forecasting model, Seasonal Periodic Decomposition Network (SPDNet), is developed to capture the seasonal, periodic, and short-term variations in residential electricity demand. The proposed

model combines seasonal-trend decomposition with periodic decomposition to capture temporal dependencies at different scales and generate forecast-based pseudo-measurements for state estimation. To address the limited availability of near-real-time AMI measurements, Bayesian Weighted Least Squares (BWLS) and Variational Bayesian Matrix Completion (VBMC) are employed to estimate the voltage states of the distribution network. The resulting estimates are combined using calibrated Maximum Ratio Combining (MRC), which accounts for the bias and uncertainty associated with each estimator. The results demonstrate that the calibrated fusion approach improves voltage-estimation accuracy compared with the individual estimators and conventional MRC. Furthermore, a comparison between individual and transformer-aggregated load forecasting shows that reducing the number of forecasting profiles from 646 individual loads to 58 aggregated groups provides nearly identical BWLS state-estimation accuracy while reducing the number of forecasting tasks.

Collectively, the proposed frameworks contribute to the planning and operation of modern power distribution systems by combining statistical modeling, optimization, machine learning, and state estimation. The results demonstrate the effectiveness of probabilistic PV–battery sizing for peak-demand reduction and economic benefit, as well as the potential of forecasting-aided and calibrated state estimation to improve situational awareness under limited measurement availability.

Table of Contents

List of Figures xi

List of Tables xvi

List of Nomenclature xviii

Acknowledgements xx

Dedication xxi

Preface xxii

1 Introduction 1

1.1 Research Questions 4

1.2 Dissertation Organization 8

2 Literature Review 10

2.1 Building Smart Communities to Advance Solar Energy in Rural America . . 10

2.1.1 Single-Objective Optimization Methods 11

2.1.2 Multi-Objective Optimization Methods 11

2.1.3 Stochastic Optimization Methods 12

2.1.4 Robust Optimization Methods 13

2.2 Leveraging Smart Meter Data for Enhanced Situational Awareness of Distribution Systems 13

2.2.1 Distribution System State Estimation 13

2.2.2 Electricity Demand Forecasting 18

3 Background . . . 23
3.1 Distribution System State Estimation . . . 23
3.2 Distribution System State Estimation Techniques . . . 23
3.2.1 Variational Bayesian Matrix Completion . . . 23
3.2.2 Bayesian Weighted Least Squares . . . 31

4 Optimal Sizing of PV–Battery Systems . . . 39
4.1 Methodology . . . 40
4.2 Proposed Approach 1 . . . 45
4.3 Proposed Approach 2 . . . 49
4.4 Benefit Analysis . . . 53
4.5 Modified Monte Carlo Simulation . . . 58
4.6 Results . . . 66
4.7 Approach 1 . . . 72
4.7.1 Financial Analysis . . . 76
4.7.2 Modified Monte Carlo Simulation . . . 81
4.8 Approach 2 . . . 84
4.8.1 Benefit Analysis . . . 87
4.8.2 Modified Monte Carlo Simulation . . . 89
4.9 Summary . . . 93

5 Electricity demand forecasting . . . 97
5.1 Proposed Forecasting Module . . . 97
5.1.1 SPDNet Methodology . . . 102
5.1.2 Results . . . 112
5.1.3 Uncertainty Quantification . . . 119
5.1.4 Summary . . . 123

6 Calibrated Fusion for Distribution System State Estimation . . . 125

6.1 Estimation Framework . . . 126
6.1.1 Bayesian Weighted Least Squares . . . 127
6.1.2 Bayesian Matrix Completion . . . 128
6.2 Estimator Fusion . . . 130
6.2.1 Classical MRC . . . 130
6.2.2 Calibrated MRC . . . 131
6.3 Case Study . . . 134
6.3.1 Feeder and Measurements . . . 134
6.3.2 Scenario 1: Aggregated Loads . . . 135
6.3.3 Scenario 2: Individual Loads . . . 136
6.4 Results . . . 137
6.4.1 Scenario 1: Aggregated Loads . . . 137
6.4.2 Scenario 2: Individual Loads . . . 157
6.5 Summary . . . 161

7 Conclusion and Future Work . . . 163
7.1 Conclusions . . . 163
7.2 Future Directions . . . 168

Bibliography . . . 171

List of Figures

4.1 Daily load, PV generation, and modified load 41

4.2 Modified load with flattening line 43

4.3 Updated peak with undersized battery 45

4.4 Organizational flowchart of the proposed statistical peak-constrained PV–battery sizing approach (method 1). 50

4.5 Organizational flowchart of the proposed elbow-based PV–battery sizing approach (method 2). 53

4.6 Process of determining the conditional probabilities. 67

4.7 Organizational flowchart to generate load and solar irradiance scenarios based on their interdependencies. 68

4.8 Histogram of the daily peak demand hours for 2019 70

4.9 Histogram of the required daily battery sizes that effectively flattened the load profiles for PV = 2000 kW. 71

4.10 Fitted Gamma PDF for updated daily peaks with PV = 2000 kW and battery = 3000 kWh. The vertical marker shows the 95th-percentile peak (2041.996 kW). 74

4.11 Evolution of Gamma shape and scale parameters for PV = 2000 kW as the battery size increases. The recalibrated point that enforces $F_\Gamma(2000\text{ kW}) = 0.95$ falls closest to the curve corresponding to a 4000 kWh battery. 75

4.12 Comparative financial impacts of PV–battery installations, showing energy, peak costs, and equipment costs for varying system sizes. 77

4.13 Present-worth benefit over 20 years for all PV–battery combinations that satisfied the 95% peak-flattening requirement. The 2000 kW + 4000 kWh case produces the maximum benefit of USD 812,648. 78
4.14 Sensitivity analysis of the benefits with variation in the economic parameters. The black triangles indicate the maximum benefit for each scenario. 80
4.15 Elbow curve of the K-means clustering for January load profiles. 82
4.16 Results of the January load-demand clustering using self-organizing maps. . 82
4.17 Histogram and fitted normal distribution for 500 Monte Carlo runs of 20-year benefit for the optimal PV–battery system (PV = 2000 kW, battery = 4000 kWh). 83
4.18 Fitted Gamma PDF for updated daily peaks with PV = 2000 kW and battery = 4000 kWh. The dashed line marks the 95th percentile (1,985 kW). 85
4.19 Ninety-fifth percentile (95%) peak-demand thresholds versus battery size for PV = 2000 kW. A clear diminishing-return region appears beyond ∼4.6 MWh. 86
4.20 Elbow identification for PV = 2000 kW using the maximum-distance method. The maximum distance occurs at a battery size of 4,600 kWh, with a corresponding 95% peak-demand threshold of 1,971 kW. 87
4.21 Present-worth benefit over 20 years for elbow-based (Approach 2) PV–battery combinations. The maximum benefit ($ 820,373) occurs at PV = 1,200 kW and battery = 3,200 kWh. 88
4.22 Gap statistic and Silhouette coefficient for different numbers of clusters for July load-demand profiles. 90
4.23 Hierarchical clustering results for July load-demand profiles, identifying three main clusters. 91
4.24 Histogram of 500 Monte Carlo 20-year benefit samples for the selected PV–battery system (PV = 1,200 kW, battery = 3,200 kWh), with fitted Normal distribution. 92

5.1 Example of one-week residential electricity load demand. 98
5.2 Illustrative 24-hour-ahead forecasts from different model families. 101
5.3 Systematic workflow of the proposed SPDNet forecasting module. 102
5.4 Overview of the Periodical Decomposition Module (PDM) Framework to reshape the 1D input tensor into k distinct 2D tensors. 106
5.5 Overall architecture of Transformer encoder processing to capture global dependencies. 111
5.6 Forecasted and actual electricity demand for load 1 at sequence length $S = 96$ and prediction horizon $P = 96$. Legends are shown only in the first subplot for clarity. 115
5.7 Average training time (seconds) per epoch for different models. All experiments are trained with an input length of 720, batch size of 32, and prediction horizons of {1,4,24,48,96,192,384,672,1344,2688}. 118
5.8 Comparison of actual and forecasted aggregated electricity demand using SPDNet, RNN, LSTM, GRU, ConvLSTM, Transformer, Informer, and PatchTST over a 24-hour prediction horizon at 15-minute resolution. 120
5.9 Day-ahead forecasting errors and data-driven uncertainty bands for two representative aggregated load groups. 122

6.1 Single-line diagram of the distribution feeder used in the case study. 134
6.2 Phase-A voltage magnitude and angle estimated by VBMC at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated 95% uncertainty interval. 139
6.3 Phase-B voltage magnitude and angle estimated by VBMC at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated 95% uncertainty interval. 140

6.4 Phase-C voltage magnitude and angle estimated by VBMC at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated 95% uncertainty interval. 141
6.5 Phase-A voltage magnitude and angle estimated by BWLS at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated uncertainty interval. 142
6.6 Phase-B voltage magnitude and angle estimated by BWLS at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated uncertainty interval. 143
6.7 Phase-C voltage magnitude and angle estimated by BWLS at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated uncertainty interval. 144
6.8 Phase-A voltage magnitudes estimated by BWLS, VBMC, and standard MRC at 13:00 on July 8, 2023. 146
6.9 Phase-B voltage magnitudes estimated by BWLS, VBMC, and standard MRC at 13:00 on July 8, 2023. 147
6.10 Phase-C voltage magnitudes estimated by BWLS, VBMC, and standard MRC at 13:00 on July 8, 2023. 148
6.11 Phase-A voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 13:00 on July 8, 2023. 150
6.12 Phase-B voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 13:00 on July 8, 2023. 151
6.13 Phase-C voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 13:00 on July 8, 2023. 152
6.14 Phase-A voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 22:00 on July 8, 2023. 153

6.15 Phase-B voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 22:00 on July 8, 2023. 154
6.16 Phase-C voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 22:00 on July 8, 2023. 155
6.17 Phase-A voltage-magnitude MAE over a 24-hour period for BWLS, VBMC, and bias-aware MRC in the aggregated-load case. 158
6.18 Voltage-magnitude MAE of calibrated MRC across 353 secondary-side bus-phase nodes, averaged over 96 timestamps on July 8, 2023. 160
6.19 Distribution of secondary-side bus-phase nodes across four voltage-magnitude MAE ranges using calibrated MRC, averaged over 96 timestamps on July 8, 2023. 161

List of Tables

4.1 Technical and economic parameters used in the rural distribution utility case study. 69

4.2 KS statistic test results for the fitted PDFs for PV = 2000 kW and various battery sizes. 73

4.3 Optimal battery sizes for PV–battery systems to flatten 95% of the daily peak demands to 2000 kW. 76

4.4 Economic comparison before and after PV and PV–battery installation (present-worth values over 20 years). 79

4.5 Financial benefits for optimal PV–battery combinations obtained from Approach 2. 88

4.6 Detailed economic comparison for the best Approach 2 configuration (PV = 1,200 kW, battery = 3,200 kWh). All values are present worth over 20 years. 89

5.1 Advantages and disadvantages of electricity demand forecasting models. . . . 100

5.2 Forecasting results for load 1 with normalized values, using sequence lengths $S = \{96, 384, 720\}$ and prediction horizons $P = \{1, 4, 24, 48, 96\}$. 116

5.3 Performance metrics (MSE and MAE) for SPDNet and LSTM models across various prediction horizons for different loads. 117

5.4 SPDNet forecasting accuracy for the 58 aggregated transformer-level load groups, measured using mean squared error (MSE) and mean absolute error (MAE). 119

6.1 Phase-A voltage-magnitude estimation accuracy for the feeder snapshot at 16:00 on July 8, 2023. 156

6.2 Calibrated MRC performance under different AMI observability and measurement-noise levels, in the feeder case study. 156
6.3 Overall BWLS state estimation accuracy at common primary-side buses using aggregated and individual load forecasts. 159
6.4 Phase-wise BWLS state estimation accuracy at common primary-side buses using aggregated and individual load forecasts. 159

Nomenclature

AC Alternating current

AMI Advanced metering infrastructure

BESS Battery energy storage system

BiLSTM Bidirectional long short-term memory

BWLS Bayesian weighted least squares

CDF Cumulative distribution function

ConvLSTM Convolutional long short-term memory

DC Direct current

DLinear Decomposition linear model

DSSE Distribution system state estimation

DTW Dynamic time warping

FFT Fast Fourier transform

FFN Feed-forward network

GRU Gated recurrent unit

ITC Investment tax credit

KS Kolmogorov–Smirnov

LSTM Long short-term memory

MAE Mean absolute error

MAPE Mean absolute percentage error

MRC Maximum ratio combining

MSE Mean squared error

O&M Operation and maintenance

PDC Peak-demand charge

PDF Probability density function

PDM Periodical decomposition module

PMU Phasor measurement unit

PV Photovoltaic

RMSE Root mean squared error

RNN Recurrent neural network

SCADA Supervisory control and data acquisition

SOM Self-organizing map

SPDNet Seasonal Periodic Decomposition Network

STDM Seasonal-trend decomposition module

TIIC Total initial investment cost

VBMC Variational Bayesian matrix completion

WLS Weighted least squares

Acknowledgments

This dissertation would not have been possible without the invaluable guidance, support, and encouragement of my advisor, Dr. Anil Pahwa. He has been far more than an academic advisor to me—he has been a mentor, role model, and source of inspiration throughout my doctoral journey. His vision, patience, and wisdom have deeply shaped my professional and personal growth. I am profoundly grateful for his constant encouragement, thoughtful feedback, and for always standing by me through every challenge and success.

I would also like to express my sincere appreciation to my co-advisor, Dr. Bala Natarajan, for his insightful guidance, constructive criticism, and continuous support throughout this research. His expertise in signal processing and data analytics greatly enriched the depth and quality of this work.

I extend my gratitude to the members of my doctoral committee for their time, valuable feedback, and thoughtful suggestions that strengthened this dissertation. Their diverse perspectives and expertise helped refine both the technical and practical aspects of my research.

Finally, I owe my deepest gratitude to my family, whose love and sacrifices made this journey possible. My parents have always believed in me, encouraged my curiosity, and provided the foundation for everything I have achieved. Their unwavering support, patience, and faith gave me the strength to persevere through every challenge. To my loved ones who stood by me during difficult times, your encouragement and understanding have been my greatest motivation.

Dedication

This dissertation is dedicated to my beloved parents, whose unconditional love, endless sacrifices, and unwavering support have been the foundation of my achievements. Their encouragement, patience, and belief in me have given me the strength to pursue my dreams, even through the most challenging times.

I also dedicate this work to my family, whose love and support have been a constant source of motivation and inspiration throughout this journey.

This accomplishment is a reflection of their sacrifices, faith, and encouragement, and I dedicate it to them with my deepest love and gratitude.

Preface

This dissertation, entitled "Data-enabled Approaches for Planning and Operation of Smart and Sustainable Power Distribution System," is submitted for the degree of Doctor of Philosophy in the Mike Wiegers Department of Electrical and Computer Engineering at Kansas State University. The research presented herein was carried out under the guidance of my major advisor, Dr. Anil Pahwa, and my co-advisor, Dr. Balasubramaniam Natarajan. To the best of my knowledge, the work presented in this dissertation is original, except where appropriate acknowledgments and references have been provided.

The research presented in this dissertation focuses on the optimal planning of photovoltaic–battery systems and the development of data-driven methodologies for electricity demand forecasting and distribution system state estimation. Portions of this research have been published in peer-reviewed journals and conference proceedings or disseminated as research preprints. The publications and related scholarly works are listed below.

1. R. Nematirad, A. Pahwa, B. Natarajan, and H. Wu, "Optimal sizing of photovoltaic-battery system for peak demand reduction using statistical models," Frontiers in Energy Research, vol. 11, article 1297356, 2023.

2. R. Nematirad, A. Pahwa, and B. Natarajan, "A Novel Statistical Framework for Optimal Sizing of Grid-Connected Photovoltaic-Battery Systems for Peak Demand Reduction to Flatten Daily Load Profiles," Solar, vol. 4, no. 1, pp. 179–208, 2024.

3. R. Nematirad and A. Pahwa, "Solar radiation forecasting using artificial neural networks considering feature selection," in 2022 Kansas Power and Energy Conference (KPEC), pp. 1–4, 2022.

4. R. Nematirad, A. Pahwa, and B. Natarajan, "Times2D: Multi-Period Decomposition and Derivative Mapping for General Time Series Forecasting," in Proceedings of the AAAI Conference on Artificial Intelligence, 2025.

5. R. Nematirad, A. Pahwa, and B. Natarajan, "Are Data Embeddings effective in time series forecasting?," Transactions on Machine Learning Research, 2025.

6. R. Nematirad, A. Pahwa, and B. Natarajan, "SPDNet: Seasonal-Periodic Decomposition Network for Advanced Residential Demand Forecasting," arXiv preprint, arXiv:2503.22485, 2025.

7. R. Nematirad, A. Pahwa, and B. Natarajan, "Calibrated Maximum Ratio Combining for Fusion-Based Distribution System State Estimation," in 2026 IEEE Power and Energy Society General Meeting, 2026.

Chapter 1

Introduction

The increasing penetration of renewable energy resources and monitoring technologies is transforming the power distribution grid into a more complex, data-rich, and dynamic environment. While these advancements promise greater sustainability, efficiency, and resilience, they also introduce unprecedented challenges in system planning, operation, and real-time decision-making. Modern distribution networks must now balance variable renewable generation with fluctuating demand, manage large volumes of heterogeneous data from advanced sensors, and maintain reliability under uncertain conditions. Addressing these intertwined challenges requires a holistic, data-driven perspective that combines physical modeling, statistical analysis, and machine learning. This dissertation presents a unified framework that integrates these perspectives to tackle two fundamental issues in the modernization of power distribution systems: the optimal integration of photovoltaic (PV) and battery energy storage systems for sustainable energy planning, and the enhancement of situational awareness through advanced data fusion and intelligent analytics. Together, these two research thrusts aim to advance both the long-term sustainability and the real-time intelligence of future electric distribution networks.

The shift towards renewable energy is a response to growing environmental concerns and the finite supply of non-renewable resources. Among the available options, photovoltaic (PV) systems have emerged as one of the most promising technologies due to their ability to

generate electricity directly from abundant solar energy without emitting greenhouse gases. According to projections from the U.S. Department of Energy's Solar Energy Technologies Office and the National Renewable Energy Laboratory, solar energy could provide up to forty percent of the nation's electricity by 2035 [1]. Despite this remarkable potential, PV systems continue to face several technical and economic limitations that hinder their large-scale adoption.

For a photovoltaic system to be economically viable, it must achieve measurable reductions in both energy consumption and peak demand charges. In practice, however, some PV installations primarily reduce daytime energy consumption and have minimal influence on system peak demand, which can occurs during the early morning or late evening hours when irradiance levels are negligible. This temporal mismatch between solar generation and peak load periods results in only partial financial benefits for utilities. To mitigate this limitation, battery energy storage systems (BESS) have emerged as complementary technologies capable of storing surplus solar energy during periods of low demand and discharging it during peak demand hours. When properly sized and controlled, a PV–BESS configuration can flatten the daily load profile, reduce peak demand, and enhance both technical performance and financial returns. However, battery integration introduces additional complexity—such as installation cost, degradation, and round-trip energy losses—that must be carefully considered in system design.

The optimal design of PV–BESS systems is thus a multi-dimensional challenge that involves balancing performance, reliability, and cost. Conventional sizing techniques often rely on deterministic optimization frameworks that treat solar irradiance and load demand as fixed quantities. In reality, these parameters are inherently uncertain and vary significantly due to weather fluctuations, seasonal changes, and consumer behavior. Neglecting these uncertainties can lead to either undersized systems that fail to meet demand or oversized systems that are economically inefficient. To overcome this limitation, this dissertation develops a statistical optimization framework that explicitly accounts for the stochastic nature of both solar generation and electricity demand. The framework integrates statistical modeling, Monte Carlo simulation, and time-series clustering to capture variability across daily

and seasonal scales. By fitting probability distributions to updated daily peak demands and evaluating the 95th-percentile peak-demand thresholds, the framework identifies PV and battery capacities that satisfy the specified peak-demand reduction requirement while maximizing financial benefit.

In addition to sustainable energy planning, improving situational awareness is another important challenge in modern power distribution systems. Distribution utilities require accurate information about system operating conditions to ensure reliable and efficient operation. Distribution system state estimation plays an important role in this process by estimating bus-voltage magnitudes and phase angles using available measurements. Advanced metering infrastructure provides detailed information about customer electricity consumption and offers opportunities to improve system monitoring. However, unlike conventional real-time monitoring devices, AMI measurements are not always available in near real time due to communication limitations, delayed data collection, and missing measurements. Consequently, distribution networks often operate under low-observability conditions, making accurate state estimation challenging.

To address these limitations, electricity demand forecasting can be used to generate pseudo-measurements for unavailable load data. Machine learning techniques have demonstrated the ability to capture complex temporal patterns in electricity demand, including seasonal variations, periodic behavior, and short-term fluctuations. However, forecasting errors introduce uncertainty into the generated pseudo-measurements, which can affect state estimation accuracy. Furthermore, relying on a single state estimation technique may lead to suboptimal results, particularly when the available measurements are limited or contain uncertainties. Combining complementary estimation methods provides an opportunity to exploit different sources of information and improve the accuracy of estimated system states.

Motivated by these challenges, the second part of this dissertation develops a data-driven framework for enhancing situational awareness using AMI data. First, a deep learning forecasting model, Seasonal Periodic Decomposition Network, is developed to forecast residential electricity demand and quantify forecasting uncertainty. The resulting forecasts are used

as pseudo-measurements in Bayesian Weighted Least Squares, while Variational Bayesian Matrix Completion provides complementary estimates by exploiting spatial and temporal correlations in incomplete AMI data. The estimates obtained from these two methods are then combined using calibrated Maximum Ratio Combining, which accounts for the bias and uncertainty associated with each estimator.

1.1 Research Questions

To address the challenges outlined in this dissertation, the following research questions are investigated:

Question 1: How can photovoltaic (PV) and battery energy storage systems (BESS) be optimally sized to achieve maximum peak-demand reduction and economic benefit under uncertain solar irradiance and load conditions?

Question 2: What statistical and computational methods can be employed to account for the stochastic variability of solar generation and electricity demand when determining the optimal PV–BESS configuration?

Question 3: How can advanced metering infrastructure (AMI) data be effectively utilized to enhance situational awareness in low-observable distribution networks?

Question 4: How can machine learning models trained on historical AMI data be used to forecast active and reactive power demand and generate pseudo-measurements for use in distribution system state estimation?

Question 5: How can Variational Bayesian Matrix Completion (VBMC) and Bayesian Weighted Least Squares (BWLS) methods be integrated to improve the accuracy and robustness of state estimation under limited or missing AMI measurements?

Question 6: How can the potential bias in the estimates obtained from Variational Bayesian Matrix Completion (VBMC) and Bayesian Weighted Least Squares (BWLS) be characterized, modeled, and mitigated within the fusion process so that the final combined estimate remains unbiased and statistically consistent?

These research questions collectively address the dual objectives of this dissertation: (1)

optimizing the planning of photovoltaic and battery storage systems to improve the long-term sustainability of distribution networks, and (2) enhancing the operational intelligence of these networks through AMI-based situational awareness and data-driven state estimation.

To achieve these objectives, this research develops two integrated frameworks corresponding to the two thrusts. The first focuses on photovoltaic–battery system optimization by incorporating the statistical characteristics of solar generation and load demand into a statistical model for optimal sizing. The second focuses on improving system observability using AMI data through load forecasting, statistical estimation, and fusion of limited measurements. The estimates obtained from Variational Bayesian Matrix Completion (VBMC) and Bayesian Weighted Least Squares (BWLS) are combined using calibrated Maximum Ratio Combining (MRC), which accounts for the bias and uncertainty associated with each estimator to improve voltage-state estimation accuracy.

Contributions

To address the research questions, this dissertation develops a complete data-driven methodology for (i) optimal sizing of photovoltaic–battery (PV–battery) systems with explicit consideration of peak demand reduction, statistical uncertainty, and long-term economic benefit, and (ii) enhancing situational awareness of distribution systems using AMI data, forecasting, and calibrated fusion of multiple state estimators. The major contributions are summarized below.

- **Formulation of peak-demand-aware PV–battery sizing.** This work treats peak demand reduction as a primary design objective. The PV–battery system is required to flatten daily peaks up to a specified threshold with a prescribed probability (95%), which aligns the sizing problem with the way rural and municipal utilities are billed for both energy and demand.

- **Statistical methodology for updated peak modeling.** For a given PV size, modified daily load profiles are generated by subtracting PV generation from the original

load, the daily needed battery size is determined using an area-balancing procedure that includes roundtrip efficiency and utilization factor, and updated daily peaks are fitted with suitable probability distributions to obtain 95th-percentile peak-demand thresholds for every PV–battery pair.

- **Elbow-based selection of battery size.** After computing the 95th-percentile peak-demand threshold for a range of battery capacities corresponding to one PV size, an elbow-point rule on the "95th-percentile peak-demand threshold vs. battery size" curve is introduced to select the battery size that gives the largest useful reduction in peak without oversizing.

- **Realistic scenario generation using time-series clustering and Monte Carlo simulation.** Daily solar irradiance and load data are clustered on a monthly basis and the resulting conditional probabilities are used inside Monte Carlo simulation to generate realistic yearly scenarios that preserve the dependence between irradiance and load. This step validates the selected PV–battery sizes under practical operating conditions.

- **Benefit-driven selection among feasible PV–battery pairs.** Several PV–battery combinations that satisfy the 95% technical requirement are evaluated using a 20-year economic/benefit analysis that includes installation cost, battery replacement, operation and maintenance, and energy and demand charge savings. The final PV–battery size is selected as the one that meets the technical requirement and provides the highest financial benefit to the utility.

- **Application to a real rural distribution system.** Using three years of hourly load and solar irradiance data collected from a small town in Kansas, USA, the proposed methodology identifies PV–battery combinations that can flatten most of the daily peaks and provide positive financial benefits, demonstrating the applicability of the framework to rural distribution systems.

- **AMI-based data preprocessing and load aggregation.** A data-cleaning and

aggregation procedure is developed for AMI measurements to remove long missing segments, fill short gaps, and group loads at the transformer level so that the data can be used directly in forecasting and state estimation.

- **SPDNet forecasting model.** A deep learning forecasting model, Seasonal Periodic Decomposition Network (SPDNet), is developed to capture the seasonal, periodic, and short-term variations in residential electricity demand. The model combines seasonal-trend decomposition and periodic decomposition to capture temporal dependencies at different scales and generate forecast-based pseudo-measurements for distribution system state estimation.

- **Power flow analysis on a mixed-phase distribution feeder.** The AMI-based framework is tested on a mixed-phase distribution grid in OpenDSS to obtain the network voltages and power flows corresponding to the forecasted and completed AMI data, providing the system model and reference values needed for state estimation.

- **State estimation using VBMC.** Variational Bayesian Matrix Completion (VBMC) is used to reconstruct missing or unpolled AMI measurements by exploiting spatial and temporal correlations in the AMI dataset, producing completed measurements with associated uncertainty.

- **State estimation using BWLS.** Bayesian Weighted Least Squares (BWLS) is used to estimate system states by using the forecasted active and reactive power obtained from the AMI forecasting models as pseudo-measurements together with the network model.

- **Estimator fusion using MRC.** The estimates obtained from VBMC and BWLS are combined using Maximum Ratio Combining (MRC) to exploit the complementary information provided by statistical completion and physics-based estimation.

- **Calibration to account for estimator bias.** A calibrated MRC approach is developed to account for the bias and uncertainty associated with the individual estimators.

The proposed method incorporates bias correction and residual mean-squared-error-based weighting to improve voltage-state estimation accuracy.

- **Evaluation of load aggregation for state estimation.** The effect of load aggregation on BWLS state estimation is investigated by comparing individual and transformer-aggregated load forecasts. The results demonstrate that reducing the number of forecasting profiles from 646 individual loads to 58 aggregated groups provides nearly identical voltage-estimation accuracy while reducing the number of forecasting tasks.

1.2 Dissertation Organization

This dissertation is organized to follow the two main research thrusts introduced in Chapter 1: optimal planning of photovoltaic–battery systems for peak-demand reduction, and AMI-based situational awareness and state estimation for low-observable distribution networks. Chapter 1 presents the motivation, research questions, and main contributions of the dissertation. Chapter 2 reviews related work on photovoltaic and battery system sizing, statistical and probabilistic methods for modeling solar irradiance and electricity demand, and existing approaches for AMI-based monitoring, electricity demand forecasting, and distribution system state estimation.

Chapter 3 provides the theoretical background for distribution system state estimation, with particular emphasis on Variational Bayesian Matrix Completion (VBMC) and Bayesian Weighted Least Squares (BWLS).

Chapter 4 develops the statistical optimization framework for optimal sizing of photovoltaic–battery systems. It presents the two proposed sizing approaches, probabilistic modeling of updated daily peak demands, the 95th-percentile peak-demand threshold, and the economic benefit analysis. The chapter also introduces time-series clustering and modified Monte Carlo simulation to evaluate the selected PV–battery configurations under uncertain operating conditions.

Chapter 5 focuses on electricity demand forecasting using AMI data. It describes the preprocessing and aggregation of individual customer load profiles and develops the Seasonal Periodic Decomposition Network (SPDNet) for forecasting electricity demand. The proposed model is evaluated using individual and transformer-aggregated load profiles, and the forecasting errors are analyzed to quantify the uncertainty associated with forecast-based pseudo-measurements.

Chapter 6 presents the distribution system state estimation framework using limited AMI measurements and forecast-based pseudo-measurements. It describes the application of BWLS and VBMC, followed by estimator fusion using conventional and calibrated Maximum Ratio Combining (MRC). The proposed methods are evaluated on a mixed-phase distribution feeder, and the effects of measurement observability, measurement noise, and load aggregation on voltage-state estimation accuracy are investigated.

Finally, Chapter 7 summarizes the main findings and contributions of the dissertation and discusses possible future research directions.

Chapter 2

Literature Review

2.1 Building Smart Communities to Advance Solar Energy in Rural America

The transition toward renewable energy has been motivated by growing environmental concerns and the finite nature of conventional energy resources [2]. Among the available renewable technologies, photovoltaic (PV) systems have received particular attention because they can generate electricity from an abundant and environmentally friendly source of energy [3]. Recent studies have reported significant advances in PV technology, including monolithic perovskite/silicon tandem cells with certified efficiencies close to 29% [4], [5], and improved module designs that can operate reliably in harsh conditions [6]. Despite these advances, PV systems alone do not necessarily satisfy the technical and economic requirements of distribution utilities, especially when utilities pay for both total energy and peak demand. In many service territories, the daily or seasonal peak occurs when PV generation is negligible, such as winter mornings or spring and fall evenings, and therefore PV installations do not contribute to peak-demand reduction on those days.

Because peak-demand charges can be significant, battery energy storage systems (BESS) have been proposed as complementary technologies to shift energy from off-peak to peak hours, improving the potential financial benefit of PV installations. However, storage intro-

duces additional costs, roundtrip losses, and degradation effects, so determining the optimal size of a PV–battery system that can meet a specified peak with high probability while remaining economically viable has become a central topic in the literature [7]. Existing works on PV–battery optimal sizing can be grouped into four main categories.

2.1.1 Single-Objective Optimization Methods

Single-objective optimization focuses on optimizing one performance index of the PV–battery system, such as total cost, energy savings, or reliability [8]. Studies of residential electricity systems have examined the potential for reducing energy costs through changes in electricity use and the integration of distributed energy resources. Economic assessment of residential PV systems with lithium-ion batteries has shown that their financial performance depends on storage costs, electricity tariffs, and the selected capacities [9]. The temporal resolution of the load and PV profiles also affects self-consumption and the resulting system sizes [10]. Other studies have developed lifetime cost-based methodologies to determine the optimal PV–battery size over the project horizon using realistic test cases [11]. These single-objective approaches are straightforward and computationally tractable, but their narrow focus can overlook other important operational needs of distribution systems, such as voltage impacts or grid interaction.

2.1.2 Multi-Objective Optimization Methods

Multi-objective optimization addresses the limitation of single-objective methods by considering several goals at the same time, such as cost, energy efficiency, voltage profile, losses, and environmental impact [12]. A multi-objective model based on particle swarm optimization was proposed to minimize power losses and improve voltage and loadability when PV and battery systems are integrated into the distribution network [13]. Another multi-objective framework for grid-connected PV–battery systems used machine learning techniques to minimize purchased energy, maximize battery state of charge, and reduce emissions [14]. Scenario-based multi-objective optimization has been applied to rural PV–battery sys-

tems and has shown that coordinated planning can reduce fluctuations in power exchanged with the grid by more than 80% across different weather conditions [15]. In the residential context, optimization strategies for home energy management that include PV and storage have also been proposed to integrate sustainable energy sources while maintaining acceptable grid performance [16]. Although multi-objective methods better reflect practical operation, they still commonly assume that solar irradiance and load demand are known or can be accurately forecasted and, therefore, may not fully represent real operating conditions in rural systems.

2.1.3 Stochastic Optimization Methods

Stochastic optimization has been introduced to explicitly account for the variability and uncertainty of solar generation and electricity demand [17]. In this approach, multiple scenarios are generated to represent different realizations of irradiance and load, and the PV–battery sizing problem is solved so that the system performs well on average or minimizes expected cost across all scenarios. This allows the resulting design to be more resilient to changes in weather patterns or customer behavior [18]. A stochastic model for PV–battery sizing showed that PV panel cost and efficiency have a strong impact on the optimal solution when uncertainty is included [19]. Another study used stochastic optimization to determine the size of battery storage integrated with PV systems while minimizing battery cost and grid energy imports, and it concluded that combining financial and technical objectives is necessary to obtain economically feasible designs [20]. Integrated stochastic frameworks have also shown that tariffs and unit costs are major drivers in determining optimal PV–battery sizes [21]. The main drawback of stochastic methods is that they often require a large number of scenarios and, in many cases, fitting historical data into probability distribution functions, which can increase computational complexity and may be difficult when the available data are limited or nonstationary [22].

2.1.4 Robust Optimization Methods

Robust optimization addresses uncertainty by defining bounds or uncertainty sets for the key parameters instead of relying on their probability distributions [23]. This is useful when there is not enough data to estimate distributions or when the distributions do not pass statistical tests. Robust optimization seeks PV–battery sizes that remain feasible for all values within the specified uncertainty sets, which improves consistency and resilience against parameter variations [24]. A two-stage robust optimization model was proposed for PV systems with battery units where polyhedral uncertainty sets were used to model PV and load variability [25]. Robust optimization for grid-connected PV–battery systems has also shown that there is a trade-off between minimizing the mean levelized cost of electricity and minimizing its standard deviation, so that more conservative solutions can be obtained when higher reliability is needed [26]. The limitation of robust methods is that, if the uncertainty bounds are chosen too large, the resulting solution can be overly conservative and lead to unnecessary oversizing of the battery, which may not be acceptable for rural or municipal utilities with limited investment capacity.

However, many of these studies still do not explicitly model the probability of meeting a specified peak or link the optimal sizing step to the actual billing structure of rural utilities. The methodology developed in this dissertation is motivated by these gaps and combines statistical modeling of updated daily peaks, scenario generation, and benefit-based selection to support solar energy deployment in rural communities.

2.2 Leveraging Smart Meter Data for Enhanced Situational Awareness of Distribution Systems

2.2.1 Distribution System State Estimation

Modern distribution systems are increasingly equipped with advanced metering infrastructure (AMI), which provides high-resolution customer-level measurements that can be ex-

ploited to improve system visibility and operational awareness. However, AMI measurements are not available from all customers at all times because of communication bandwidth limits, polling policies, and missing data. As a result, the distribution system state estimation (DSSE) problem remains low-observable, and the main question in the literature has been how to reconstruct or estimate the unobserved quantities from a limited set of real measurements. Existing works can be grouped according to the way they handle this unobservability: (1) methods that enforce observability by adding pseudo-measurements, (2) sparsity-based methods that exploit structure in the data instead of creating many pseudo-measurements, (3) low-rank matrix/tensor completion methods that impute missing values using spatial and temporal correlations, and (4) methods that explicitly integrate measurements arriving at multiple time scales and from irregular sampling processes.

Classical DSSE with Pseudo-Measurements

Weighted least squares (WLS) has been the traditional approach for DSSE and is still the reference point for most later work [27]. The WLS estimator is widely used because it is computationally efficient when the measurement set is adequate and mostly free of gross errors [28]. In distribution networks, however, real-time measurements are scarce, so historical-load-based pseudo-measurements are routinely added to the measurement set to enforce observability before WLS is run [29]. This WLS plus pseudo-measurement approach has two main limitations: (i) it depends on historical data or predictive models to construct pseudo-measurements, and (ii) its estimation accuracy can degrade when the pseudo-measurements do not reflect the current operating condition [30].

A second and equally important limitation of classical WLS is its non-robustness to bad data. Measurements telemetered to the control center may contain random, communication, or systematic errors; even a single outlier can distort the WLS solution because the quadratic cost over-weights large residuals [31]. For this reason, traditional WLS-based DSSE is almost always accompanied by bad-data processing based on statistical tests, most commonly the chi-square test for detection and the largest normalized residual test for identification and

removal of the corrupted measurement [32]. This sequential elimination can improve estimation accuracy, but it also reduces measurement redundancy and may increase computation time, especially when several bad measurements are present at once.

To improve robustness, several variants of WLS have been proposed. Iterative reweighted least squares (RWLS) dynamically adjusts the weights of suspicious measurements during the iterations, which increases robustness without running separate bad-data tests [33]. Non-quadratic estimators [34], weighted least absolute value (WLAV) [35], and high–breakdown-point estimators such as least median of squares [36] and least trimmed squares [37] have also been investigated to handle multiple or leverage-type outliers. These methods can tolerate a much higher level of contamination than plain WLS — sometimes up to half of the redundant measurements — but they do so at the cost of higher computation and, in some cases, more difficult convergence properties [38]. Mixed-integer and heuristic formulations have likewise been explored to deal with measurement, topology, and parameter errors, although their computational cost can increase as the system size grows.

From the perspective of this dissertation, the key takeaway from the classical DSSE literature is that the whole WLS–plus–pseudo-measurement pipeline was designed for a data regime that assumes (i) enough redundancy to run bad-data tests and (ii) reasonably synchronized, reasonably accurate measurements. In modern distribution systems with AMI-based measurements that are sparse, delayed, or missing, these assumptions no longer hold. This motivates the shift toward sparsity-based DSSE, matrix/tensor completion, and learning-based multi–time-scale aggregation methods described in the following subsections.

Sparsity-Based DSSE

To avoid generating a large number of pseudo-measurements, sparsity-based DSSE methods use the fact that distribution-system states or their transformed versions are often sparse or smooth over the network. Compressive sensing (CS) was one of the earliest approaches in this group [39]. It exploits spatio-temporal correlation between loads and distributed generation to reconstruct real and reactive power measurements and then estimate voltage

magnitude and angle using far fewer measurements than WLS would need [40]. This idea was later extended from single-phase to three-phase distribution systems and combined with dynamic compression for spatial and spatio-temporal data [41]. Recursive or online variants have also been proposed to track the system as new measurements arrive [42], and CS has even been used to assist in topology identification when only limited smart grid data are available. All of these methods reduce the dependence on pseudo-measurements by using measurement structure, but they still have to deal with corrupted data and model errors.

Low-Rank Matrix and Tensor Completion

A related line of work observed that AMI and feeder measurements, when stacked over time and nodes, often form low-rank matrices or tensors [43]. This led to the use of matrix completion and tensor completion for DSSE in low-observable conditions [44]. In these methods, missing elements in the measurement matrix are imputed by finding a low-rank approximation that is consistent with the available data. The low rank here comes from spatial correlation between neighboring nodes and from temporal correlation of loads. Once the missing entries are filled, power-flow constraints can be imposed to obtain the final state estimates. Matrix completion with noise-resilient power-flow constraints has been shown to provide accurate estimates even when a large portion of measurements is missing [45], and robust variants have been proposed to handle bad data explicitly [46]. Work on data-loss handling for PMU-based systems using matrix completion and CS further confirmed that completion-based methods are well suited for distribution grids where data dropouts are common [47]. Tensor completion extends the same idea to fully exploit spatio-temporal structure and has been demonstrated to work in very low-observable systems [48]. Compared with pure CS, these methods rely more on low-rank structure than on sparsity in a transform domain, but the goal is the same: estimate the unobserved states without creating a large number of pseudo-measurements.

Integrating Multi-Time-Scale and Irregular AMI Measurements

Even when AMI is available, the measurements are often reported at slower rates (minutes or hours) than other field devices, and some values are delayed or missing. Early approaches tried to reconcile such multi-rate data by simple linear interpolation or extrapolation before running WLS [49], but that approach does not really use the network-wide spatio-temporal relationships and is limited to only two time scales. Later work adjusted measurement variances according to credibility to handle non-synchronized data [43], or used extended Kalman filters to cope with irregular sampling and communication delays [50]. These approaches improve the classical WLS-based DSSE but still require a moderate level of measurement redundancy and can become computationally heavy when applied at scale.

Because AMI data is often unevenly sampled and arrives from different customers at different times, recent work has turned to probabilistic and learning-based models that can naturally handle irregular time series. Bayesian estimators that model the probabilistic dependence of measurements on system states have been proposed [51], and DSSE methods that use predicted measurements together with the subset of measurements available at time t have been developed, but they typically assume fairly high redundancy and do not fully address missing-measurement scenarios [52]. Dynamic DSSE approaches based on Kalman filtering and load-evolution models also exist, but they can be computationally burdensome because of repeated matrix inversions and they are often demonstrated only for balanced or fully observable systems [53].

Neural-network-based models have also been explored to process AMI measurements that are sparse, delayed, or sampled at different rates [54]. Early studies used recurrent architectures such as vanilla RNN, LSTM, and GRU to learn load and voltage patterns from historical smart-meter data and then use the learned model to fill in missing entries or forecast the next few time steps [55]. These models work well when the data are regularly sampled, but their performance degrades when the AMI stream has long gaps or irregular reporting intervals. To reduce the impact of irregular sampling, later work introduced continuous-time or event-driven variants in which the hidden state is propagated between observations and

updated only when new data arrive, allowing the model to retain the actual time stamps of AMI measurements [56]. Other neural formulations incorporated stochastic components or noise-driven latent states to better capture sudden changes in consumption and to generate realistic trajectories for unobserved customers [57]. What all of these NN-based approaches have in common is that they learn the temporal structure directly from data and therefore do not require manually designed interpolation rules. However, most of them still provide point estimates of the missing AMI values and must be combined with a separate uncertainty treatment if the filled-in data are to be used inside a probabilistic DSSE.

Overall, the literature shows a clear evolution from WLS with pseudo-measurements, through sparsity-based DSSE that exploits network structure, to low-rank matrix/tensor completion that imputes missing AMI data, and finally to probabilistic and learning-based models that can integrate AMI measurements arriving at multiple time scales. The common objective across these categories is to improve situational awareness using limited, delayed, or partially missing smart meter data, and to do so without requiring full observability of the distribution network.

2.2.2 Electricity Demand Forecasting

Since the estimation frameworks above assume the availability of high-quality pseudo-measurements derived from AMI data, it is necessary to examine residential and distribution-level load forecasting techniques that can generate them. Residential electricity demand, however, shows more complex temporal behavior than feeder- or substation-level demand. It typically exhibits multiple seasonality, periodicity, and sharp short-term fluctuations [58]. Multiple seasonality refers to patterns that repeat on daily, weekly, and monthly bases, while periodicity refers to cycles that recur at fixed intervals [59]. There are also within-period correlations (load at 8:00 depends on load at 7:45 of the same day) and between-period correlations (load at 8:00 today depends on load at 8:00 on previous days). In addition, abrupt changes often occur because of household activities or weather-driven heating and cooling [60]. For this reason, many utilities forecast aggregated loads (customers behind one transformer or

feeder) because aggregation smooths out these high-frequency variations and improves forecastability [61]. Yet, as smart communities emerge, and as AMI-based DSSE may benefit from customer-level pseudo-measurements, there is increasing interest in forecasting individual residential loads, which remain more variable and harder to predict than aggregated loads [62].

Traditional statistical models such as ARIMA and Exponential Smoothing have been widely used for short-term load forecasting [63]. ARIMA treats future values as linear combinations of past observations and errors [64], while Exponential Smoothing handles trend and seasonality efficiently [65]. These models work well for linear, stationary, and smoothly seasonal series, but they struggle with the nonlinear, non-stationary, and spike-like behavior present in residential AMI time series [66]. To address nonlinearities, Multilayer Perceptrons (MLPs) have been employed [67]. MLPs can learn complex input–output mappings, but standard feedforward MLPs do not maintain a recurrent hidden state and typically capture temporal dependencies through the supplied input window [68].

Because of these limitations, more recent work has focused on deep learning models that explicitly capture temporal dependence, nonlinear interactions, and multi-periodic structures in electricity demand [69]. These models leverage hierarchical feature learning and adaptive optimization to learn patterns directly from AMI data, which is suitable for the large datasets available in smart distribution grids.

Recurrent Neural Networks

Recurrent Neural Networks (RNNs) and their variants are among the earliest deep models used for load forecasting because they maintain a hidden state that carries information from previous time steps and thus can model sequential dependence [70]. In [71], deep RNNs were applied to New York ISO data and outperformed a deep feedforward network for both 24-hour and 168-hour horizons. A GRU-based day-ahead multivariate forecasting model in [72] showed better performance than ARIMA when forecasting large-scale grid data. LSTM-based univariate forecasting in [73] achieved better accuracy than SVM and decision trees. A

bidirectional LSTM (BiLSTM) for building load in [74] improved accuracy by about 23.97% over a standard LSTM. A 3D-LSTM that converts time series into 3D continuous frames was proposed in [75] and demonstrated superior performance even against Transformer models. A parallel CNN–LSTM architecture was reported in [76] to capture both local and long-term patterns. Despite these advances, RNN-type models still face vanishing or exploding gradients for very long sequences and may not capture very short, high-frequency fluctuations without help from convolutional front-ends.

Multilayer Perceptrons

MLP-based models remain attractive when the main goal is fast and robust forecasting with a moderate number of exogenous variables. In [77], an MLP was designed for net-load forecasting in low-voltage networks using temperature and solar irradiance; it outperformed LSTM and random forest in accuracy and computation. A feedforward MLP for sub-hourly forecasting that incorporates electric vehicle and air-conditioning loads as inputs was proposed in [78] and outperformed XGBoost and Prophet. N-BEATS, which is essentially a stack of MLP blocks, was tested for smart-grid short-term forecasting in [79] and showed better performance than ARIMA, SVM, RF, CNN, and BiLSTM. In [80], N-HiTS and a Transformer were both used for power demand forecasting; N-HiTS achieved comparable accuracy while being more computationally efficient for long input windows. Unlike recurrent models, these MLP-type models do not maintain a recurrent hidden state. Their forecasting performance therefore depends on how temporal information is represented in the input features and the model architecture.

Convolutional Neural Networks

Convolutional Neural Networks (CNNs) have been adapted to time series to extract local temporal features through sliding convolutional filters [81]. A pure CNN for daily demand prediction was introduced in [82] and outperformed SVR, MLP, and RNN across several substations. A CNN–LSTM hybrid was proposed in [83], where CNN layers detect local

patterns and the LSTM decoder tracks long-term dependencies. A CNN–BiLSTM model for residential demand was introduced in [84], demonstrating that combining convolution and bidirectional recurrence is beneficial for high-frequency residential AMI data. Temporal Convolutional Networks (TCNs) with causal and dilated convolutions were applied in [85] and outperformed LSTM and GRU. A more recent hybrid that combines CNN, Transformer, and LSTM was presented in [86]; it first maps 1D time series into 2D feature maps to improve feature extraction and then uses attention to model long-range dependencies. These CNN-based and hybrid approaches are particularly suitable for AMI data because they can capture the sharp load spikes that arise from individual appliance usage, which classical statistical models miss [87–89].

Transformers

Transformer-based models have been introduced to load forecasting to better capture both local and global dependencies without relying on sequential processing [90]. A Transformer with sparse attention for household load forecasting was proposed in [91], using historical load and weather as exogenous inputs, and achieved better accuracy and computational efficiency than RNN and LSTM models. The Informer architecture [90], enhanced with time2vec embedding and 1D convolution, was used in [92] for day-ahead forecasting and showed improved accuracy over LSTM and standard Transformer when temperature and dew point were included. A Crossformer model with hierarchical encoder–decoder and variational mode decomposition was reported in [93], based on [94], to capture cross-dimensional dependencies in multivariate load data. PatchTST, which combines a patch-based Transformer with TCN, was applied in [95] for multi-load forecasting in smart buildings and outperformed LSTM, GRU, RNN, Transformer, and Informer both in accuracy and efficiency. Transformer-style models can involve substantial computational cost, particularly for long input sequences [96], although their attention-based structures can capture long-range temporal relationships.

Overall, the literature shows a progression from classical statistical forecasting, through MLPs and sequence models, to hybrid CNN–RNN and attention-based architectures. For the

purposes of this dissertation, these models are relevant not only as stand-alone forecasting tools but as building blocks for generating high-quality pseudo-measurements that can be fed into state estimation models when AMI data are missing or delayed.

Chapter 3

Background

3.1 Distribution System State Estimation

In this work, two complementary distribution system state estimation (DSSE) techniques are used to deal with the limited observability that arises when only a subset of advanced metering infrastructure (AMI) meters is polled at a given time. The first technique is a Variational Bayesian Matrix Completion (VBMC) method that treats the available AMI and feeder measurements as an incomplete data matrix and reconstructs the missing entries by exploiting their spatial and temporal correlations [46]. The second technique is a Bayesian Weighted Least Squares (BWLS) state estimator that uses network model information together with pseudo-measurements obtained from the load forecasting module to estimate the system states [27]. Using both methods allows us to combine data-driven reconstruction of missing measurements with physics-based estimation of bus voltages.

3.2 Distribution System State Estimation Techniques

3.2.1 Variational Bayesian Matrix Completion

In AMI-based distribution systems, the fundamental difficulty is that the system operator rarely receives a complete, synchronized matrix of measurements. Different customers are

polled at different times, some values are delayed or corrupted, and communication limits prevent real-time acquisition from every node. One way to view this situation is to arrange all available AMI data into a rectangular matrix where the rows correspond to spatial locations (customers, service transformers, or buses) and the columns correspond to time indices. The result is a matrix that is only partially observed. Variational Bayesian Matrix Completion (VBMC) provides a probabilistic framework for reconstructing the missing entries of this matrix while at the same time quantifying the uncertainty of the reconstruction. A linearized distribution power-flow model is subsequently used to relate the reconstructed power injections to voltage estimates.

Problem formulation

Let $Y \in \mathbb{R}^{m \times n}$ denote the partially observed measurement matrix. Each row $i \in \{1, \ldots, m\}$ represents a node or load point in the distribution system, and each column $j \in \{1, \ldots, n\}$ represents a time instant (e.g., 15-minute or hourly AMI interval). Let

$$P \in \{0, 1\}^{m \times n}$$

be the binary observation mask (distinct from the active-power variable) that indicates which entries are observed:

$$P_{ij} = \begin{cases} 1, & \text{if the AMI system reported } Y_{ij} \text{ at time } j, \\ 0, & \text{otherwise.} \end{cases}$$

The goal is to estimate a complete matrix $X \in \mathbb{R}^{m \times n}$ such that

$$X_{ij} \approx Y_{ij} \quad \text{whenever } P_{ij} = 1,$$

and such that the unobserved entries X_{ij} for $P_{ij} = 0$ are inferred from the structure present in the observed entries.

The central modeling assumption of matrix completion is that the complete matrix X is (approximately) low rank. That is, there exist latent factors

$$A \in \mathbb{R}^{m \times r}, \qquad B \in \mathbb{R}^{n \times r}, \qquad r \ll \min(m, n),$$

such that

$$X = AB^{\top}. \tag{3.1}$$

Intuitively, each row i is represented by a low–dimensional latent vector $\mathbf{a}_i^{\top}$ (the i-th row of A), and each time instant j is represented by another low–dimensional latent vector $\mathbf{b}_j^{\top}$ (the j-th row of B). Their inner product $\mathbf{a}_i^{\top}\mathbf{b}_j$ reconstructs the missing measurement at position (i, j).

To connect the latent factors to the observations, VBMC assumes a Gaussian likelihood on the *observed* entries:

$$Y_{ij} \mid \mathbf{a}_i, \mathbf{b}_j, \beta \sim \mathcal{N}\big(\mathbf{a}_i^{\top}\mathbf{b}_j,\, \beta^{-1}\big), \qquad \text{for all } (i, j) \text{ with } P_{ij} = 1, \tag{3.2}$$

where β is the observation-noise precision (inverse variance). Unobserved entries do not contribute to the likelihood.

Hierarchical Bayesian priors

The above likelihood alone would lead to a maximum-likelihood matrix factorization that can overfit and requires the rank r to be specified. The Bayesian formulation addresses this limitation by placing Gaussian priors on the latent factors and by *tying* those priors through a set of automatic–relevance–determination (ARD) hyperparameters.

For every row factor $\mathbf{a}_i \in \mathbb{R}^r$ and every column factor $\mathbf{b}_j \in \mathbb{R}^r$, we assume

$$\mathbf{a}_i \mid \Gamma \sim \mathcal{N}\big(\mathbf{0},\, \Gamma^{-1}\big), \quad i = 1, \ldots, m, \tag{3.3}$$

$$\mathbf{b}_j \mid \Gamma \sim \mathcal{N}\big(\mathbf{0},\, \Gamma^{-1}\big), \quad j = 1, \ldots, n, \tag{3.4}$$

where

$$\Gamma = \mathrm{diag}(\gamma_1, \gamma_2, \ldots, \gamma_r)$$

is a diagonal precision matrix shared by *all* rows and columns. Each γ_k controls the relevance of the k-th latent dimension across the whole matrix. Large γ_k shrinks the k-th components of both $\mathbf{a}_i$ and $\mathbf{b}_j$ toward zero, which effectively removes that dimension from the factorization.

To complete the hierarchy, each γ_k is assigned a Gamma prior, using the shape–rate parameterization,

$$\gamma_k \sim \mathrm{Gamma}(a_{\gamma_0}, b_{\gamma_0}), \qquad k = 1, \ldots, r, \tag{3.5}$$

and the noise precision β can also be given a Gamma prior,

$$\beta \sim \mathrm{Gamma}(a_{\beta_0}, b_{\beta_0}). \tag{3.6}$$

This hierarchical model is sometimes called a *Bayesian probabilistic matrix factorization with ARD* and it is particularly suitable for AMI data because it allows the algorithm to start with a comparatively large rank and to prune unnecessary directions automatically.

Variational approximation

The exact posterior

$$p(A, B, \Gamma, \beta \mid Y, P)$$

is not tractable because of the coupling between all unknowns in the likelihood (3.2). VBMC therefore adopts a mean–field variational approximation of the form

$$q(A, B, \Gamma, \beta) = q(A)\, q(B)\, q(\Gamma)\, q(\beta),$$

and seeks the factorized distribution q that minimizes

$$\mathrm{KL}\big(q \,\|\, p(A, B, \Gamma, \beta \mid Y, P)\big),$$

or, equivalently, maximizes the evidence lower bound (ELBO). Under the Gaussian–Gamma structure of (3.3)–(3.6), all the variational factors admit closed-form updates, so the algorithm proceeds by cycling through updates for $q(A)$, $q(B)$, $q(\Gamma)$, and $q(\beta)$.

Because the likelihood factorizes over matrix entries, and because the priors over the rows of A and B are independent, the variational posteriors over individual rows are also Gaussian:

$$q(\mathbf{a}_i) = \mathcal{N}\big(\boldsymbol{\mu}_{a_i}, \Sigma_{A,i}\big), \qquad q(\mathbf{b}_j) = \mathcal{N}\big(\boldsymbol{\mu}_{b_j}, \Sigma_{B,j}\big),$$

with means and covariances derived below.

Update for the row factors

Fix the current variational distributions for B, Γ, and β. Consider a single row i of the matrix. Let

$$\mathcal{O}_i = \{j \mid P_{ij} = 1\}$$

be the set of time indices for which an observation exists at node i. The contribution of row i to the ELBO depends on

$$\{Y_{ij} : j \in \mathcal{O}_i\}, \quad \{\boldsymbol{\mu}_{b_j}, \Sigma_{B,j} : j \in \mathcal{O}_i\}, \quad \mathbb{E}[\Gamma], \quad \mathbb{E}[\beta].$$

Because the likelihood is Gaussian and the prior is Gaussian, the resulting variational posterior over $\mathbf{a}_i$ is also Gaussian with precision (inverse covariance)

$$\Sigma_{A,i}^{-1} = \mathbb{E}[\beta] \sum_{j \in \mathcal{O}_i} \big(\boldsymbol{\mu}_{b_j}\boldsymbol{\mu}_{b_j}^\top + \Sigma_{B,j}\big) + \mathbb{E}[\Gamma], \tag{3.7}$$

and mean

$$\boldsymbol{\mu}_{a_i} = \mathbb{E}[\beta]\, \Sigma_{A,i} \sum_{j \in \mathcal{O}_i} Y_{ij}\boldsymbol{\mu}_{b_j}. \tag{3.8}$$

Intuitively, $\Sigma_{A,i}^{-1}$ aggregates the second moments of the column factors corresponding to the observed entries in row i, and $\boldsymbol{\mu}_{a_i}$ is a weighted least–squares fit of those factors to the

observed AMI values in row i, regularized by Γ.

Update for the column factors

By symmetry, for a fixed column j, let

$$\mathcal{O}^j = \{i \mid P_{ij} = 1\}.$$

The variational posterior over $\mathbf{b}_j$ is Gaussian with precision

$$\Sigma_{B,j}^{-1} = \mathbb{E}[\beta] \sum_{i \in \mathcal{O}^j} \left(\boldsymbol{\mu}_{a_i} \boldsymbol{\mu}_{a_i}^\top + \Sigma_{A,i}\right) + \mathbb{E}[\Gamma], \tag{3.9}$$

and mean

$$\boldsymbol{\mu}_{b_j} = \beta \, \Sigma_{B,j} \sum_{i \in \mathcal{O}^j} Y_{ij} \boldsymbol{\mu}_{a_i}. \tag{3.10}$$

Thus the algorithm alternates between (3.7)–(3.8) for all rows and (3.9)–(3.10) for all columns until convergence.

Update for ARD precisions

After all $\{\boldsymbol{\mu}_{a_i}, \Sigma_{A,i}\}$ and $\{\boldsymbol{\mu}_{b_j}, \Sigma_{B,j}\}$ have been updated, the expected squared amplitude of each latent dimension k can be computed as

$$S_k = \sum_{i=1}^{m} \left(\mu_{a_i,k}^2 + [\Sigma_{A,i}]_{kk}\right) + \sum_{j=1}^{n} \left(\mu_{b_j,k}^2 + [\Sigma_{B,j}]_{kk}\right),$$

where $\mu_{a_i,k}$ denotes the k-th element of $\boldsymbol{\mu}_{a_i}$. Combining this with the Gamma prior (3.5), the variational update for γ_k takes the standard form

$$\mathbb{E}[\gamma_k] = \frac{a_{\gamma_0} + (m+n)/2}{b_{\gamma_0} + S_k/2}, \qquad k = 1, \ldots, r. \tag{3.11}$$

If S_k is small, $\mathbb{E}[\gamma_k]$ becomes large, and the k-th latent dimension is effectively pruned. Periodically, very large expected precisions can be used to remove the corresponding columns from A and B, which implements automatic rank determination.

Update for noise precision

The expected reconstruction error over the observed entries is

$$E = \sum_{(i,j):P_{ij}=1} \mathbb{E}\big[(Y_{ij} - \mathbf{a}_i^\top \mathbf{b}_j)^2\big].$$

Because the product of two independent Gaussian vectors has a closed-form second moment, each term can be written as

$$\mathbb{E}\big[(Y_{ij} - \mathbf{a}_i^\top \mathbf{b}_j)^2\big] = (Y_{ij} - \boldsymbol{\mu}_{a_i}^\top \boldsymbol{\mu}_{b_j})^2 + \boldsymbol{\mu}_{b_j}^\top \Sigma_{A,i} \boldsymbol{\mu}_{b_j} + \boldsymbol{\mu}_{a_i}^\top \Sigma_{B,j} \boldsymbol{\mu}_{a_i} + \mathrm{tr}(\Sigma_{A,i}\Sigma_{B,j}).$$

Combining this error with the Gamma prior (3.6), the update for β becomes

$$\mathbb{E}[\beta] = \frac{a_{\beta_0} + |\{(i,j) : P_{ij} = 1\}|/2}{b_{\beta_0} + E/2}. \tag{3.12}$$

If β is not to be estimated, it can be fixed to a user–specified value.

Posterior variance of reconstructed entries

For later fusion with a physics–based estimator, it is useful to retain not only the posterior mean

$$\hat{X}_{ij} = \boldsymbol{\mu}_{a_i}^\top \boldsymbol{\mu}_{b_j}$$

but also the posterior variance

$$\mathrm{var}(X_{ij}) = \boldsymbol{\mu}_{b_j}^\top \Sigma_{A,i} \boldsymbol{\mu}_{b_j} + \boldsymbol{\mu}_{a_i}^\top \Sigma_{B,j} \boldsymbol{\mu}_{a_i} + \mathrm{tr}(\Sigma_{A,i}\Sigma_{B,j}). \tag{3.13}$$

This variance quantifies the uncertainty associated with the reconstructed entry under the assumed low-rank model. The uncertainty of the resulting voltage estimates and their use in fusion are addressed in Chapter 6.

Embedding linearized power-flow relations

The formulation above treats the data matrix as a collection of measurements indexed by location and time. For distribution system state estimation, reconstructed power injections can be supplied to a linearized distribution power-flow model to obtain corresponding voltage estimates. At a selected time instant, let N_{load} denote the number of non-slack phase-nodes. The active and reactive load powers reconstructed from the relevant measurements are assembled into

$$\mathbf{pq} = \begin{bmatrix} -\mathbf{P}_{\text{load}} \\ -\mathbf{Q}_{\text{load}} \end{bmatrix} \in \mathbb{R}^{2N_{\text{load}}}.$$

The negative sign represents the load demand convention when the power-flow formulation uses positive power injections. Let M_Y be a complex voltage-sensitivity matrix, K_Y be a real voltage-magnitude-sensitivity matrix, and $\mathbf{w}$ be the no-load complex-voltage vector. With dimensions $M_Y \in \mathbb{C}^{N_{\text{load}} \times 2N_{\text{load}}}$ and $K_Y \in \mathbb{R}^{N_{\text{load}} \times 2N_{\text{load}}}$, the linearized relations are

$$\mathbf{V}_{\text{load}} = M_Y \, \mathbf{pq} + \mathbf{w}, \tag{3.14}$$

$$|\mathbf{V}|_{\text{load}} \approx K_Y \, \mathbf{pq} + |\mathbf{w}|. \tag{3.15}$$

The slack-bus voltages are fixed to the reference vector $\mathbf{v}_0$. These relations allow voltage estimates to be obtained from the completed active- and reactive-power data. If the power-flow relations are imposed during the matrix-completion iterations, that projection is an additional model-based step and is not part of the unconstrained variational updates derived above. The measurement matrix remains indexed by location and time; power and voltage variables must therefore be stored in explicitly defined measurement matrices or feature blocks rather than interpreted as the first two time columns.

Convergence and practical notes

The VBMC iterations continue until the change in the reconstructed matrix is below a small tolerance, typically measured as

$$\frac{\|X^{(t)} - X^{(t-1)}\|_F^2}{mn} < \varepsilon,$$

with ε on the order of 10^{-6}–10^{-8}, or until a maximum number of iterations is reached. Because of the ARD mechanism, the effective rank usually decreases during the iterations, which reduces computation and prevents overfitting. The outputs available from the VBMC formulation are:

- the completed measurement matrix X, together with voltage estimates obtained through the network model when applicable;
- the factor covariances $\{\Sigma_{A,i}\}$, $\{\Sigma_{B,j}\}$;
- the per–entry variances $\mathrm{var}(X_{ij})$ from (3.13);
- the learned noise precision β.

In the overall AMI-based situational awareness framework, VBMC plays the role of a data-driven, uncertainty-aware estimator. It reconstructs missing AMI values using low-rank structure and provides entrywise uncertainty estimates; a linearized network model can then relate reconstructed power injections to voltage estimates. Because it relies mainly on correlations in the data, VBMC may still retain small biases under very sparse sampling or under mismatched base voltages.

3.2.2 Bayesian Weighted Least Squares

To complement VBMC, this dissertation employs a second, physics-based estimator that explicitly enforces the nonlinear AC power-flow equations, incorporates the actual network

admittance matrix, and allows different confidence levels for measured and forecasted quantities. This estimator is a Bayesian extension of the classical weighted least squares (WLS) state estimator and is referred to here as Bayesian weighted least squares.

The main idea is straightforward: construct a measurement vector that combines (i) the available measurements from AMI or field devices and (ii) pseudo-measurements of active and reactive power obtained from the load-forecasting module. Each of these blocks is assigned a variance that reflects its reliability: actual AMI readings receive small variances, forecasted or imputed values receive larger variances. The BWLS estimator then solves a damped nonlinear least-squares problem to find the voltage phasors that best fit all of these heterogeneous observations. The measurement Jacobian also provides an approximate state-estimation covariance under a local linearization, which can be considered in subsequent estimator fusion.

Network and state parameterization

Consider a distribution network with n phase-nodes (i.e., each phase of a multi-phase bus is treated as a separate node), with complex bus admittance matrix

$$Y_{\text{bus}} = G + jB \in \mathbb{C}^{n\times n}.$$

Let $\mathcal{S} \subset \{1,\dots,n\}$ denote the set of slack (substation) nodes, typically one per phase at the head of the feeder, and let $\mathcal{N} = \{1,\dots,n\} \setminus \mathcal{S}$ denote the set of non-slack nodes. The complex bus voltage is written as

$$V_i = |V_i|e^{j\delta_i}, \qquad i = 1,\dots,n,$$

and we collect the unknown *states* into a real vector

$$\mathbf{x} = \begin{bmatrix} \boldsymbol{\delta}_{\mathcal{N}} \\ \mathbf{v}_{\mathcal{N}} \end{bmatrix} \in \mathbb{R}^{2|\mathcal{N}|}, \tag{3.16}$$

where $\boldsymbol{\delta}_{\mathcal{N}}$ is the vector of voltage angles at the non-slack nodes and $\mathbf{v}_{\mathcal{N}}$ is the vector of voltage magnitudes at the non-slack nodes. Voltages at the slack nodes are fixed to their nominal phase values (e.g., $1\angle 0°$, $1\angle -120°$, $1\angle 120°$ for a three-phase source) and are not estimated.

Measurement model

Let us define a stacked real-valued measurement vector $\mathbf{z} \in \mathbb{R}^m$. Depending on what is available at run time, $\mathbf{z}$ may contain:

- measured voltage magnitudes at a subset of nodes,
- measured real and reactive power injections at a subset of nodes,
- forecast-based (pseudo-) real and reactive power injections at the remaining nodes,
- optional measured real and imaginary parts of the complex voltage at a few nodes to stabilize the phase in multi-phase feeders.

We group these as

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_{|V|} \\ \mathbf{z}_P \\ \mathbf{z}_Q \\ \mathbf{z}_{\Re V} \\ \mathbf{z}_{\Im V} \end{bmatrix},$$

where any of the blocks may be empty. For each block there is a corresponding nonlinear measurement function evaluated at the current state:

$$\mathbf{h}(\mathbf{x}) = \begin{bmatrix} \mathbf{h}_{|V|}(\mathbf{x}) \\ \mathbf{h}_P(\mathbf{x}) \\ \mathbf{h}_Q(\mathbf{x}) \\ \mathbf{h}_{\Re V}(\mathbf{x}) \\ \mathbf{h}_{\Im V}(\mathbf{x}) \end{bmatrix}.$$

The standard AC power-injection equations provide $\mathbf{h}_P$ and $\mathbf{h}_Q$. For any bus i,

$$P_i(\mathbf{x}) = |V_i| \sum_{k=1}^{n} |V_k| \big(G_{ik} \cos(\delta_i - \delta_k) + B_{ik} \sin(\delta_i - \delta_k) \big), \tag{3.17}$$

$$Q_i(\mathbf{x}) = |V_i| \sum_{k=1}^{n} |V_k| \big(G_{ik} \sin(\delta_i - \delta_k) - B_{ik} \cos(\delta_i - \delta_k) \big). \tag{3.18}$$

The voltage-magnitude measurement function is simply

$$h_{|V|,i}(\mathbf{x}) = |V_i|,$$

while the real and imaginary parts are

$$h_{\Re V,i}(\mathbf{x}) = |V_i| \cos \delta_i, \qquad h_{\Im V,i}(\mathbf{x}) = |V_i| \sin \delta_i.$$

These latter two are particularly helpful in multi-phase feeders where each phase is physically distinct (A/B/C) and a phase-aware initialization is needed to avoid convergence to the wrong branch of the angle.

Observed versus pseudo-measurements

Let us denote by $\Phi \in \{0,1\}^{n \times 2}$ a binary availability matrix that tells, for each node, whether a real active-power measurement (P) and/or a real reactive-power measurement (Q) is present. Let

$$\hat{P}_i^{\text{meas}}, \quad \hat{Q}_i^{\text{meas}}$$

be those real measurements, in per unit, when available. The forecasting module, or an associated reactive-power estimation procedure, provides

$$\hat{P}_i^{\text{fore}}, \quad \hat{Q}_i^{\text{fore}}$$

for all nodes. The assembled injection data used in BWLS is then

$$\tilde{P}_i = \begin{cases} \hat{P}_i^{\text{meas}}, & \text{if } \Phi_{i,1} = 1, \\ \hat{P}_i^{\text{fore}}, & \text{if } \Phi_{i,1} = 0, \end{cases} \tag{3.19}$$

$$\tilde{Q}_i = \begin{cases} \hat{Q}_i^{\text{meas}}, & \text{if } \Phi_{i,2} = 1, \\ \hat{Q}_i^{\text{fore}}, & \text{if } \Phi_{i,2} = 0. \end{cases} \tag{3.20}$$

Thus every non-slack node has an associated active and reactive injection for the state estimator, but some are "hard" (measured) and some are "soft" (forecasted). This is the point where the load-forecasting methodology developed in Chapter 5 feeds into the DSSE framework: high-quality short-term forecasts produce pseudo-measurements with smaller error variances. In BWLS, the weights are inversely proportional to these variances, so more reliable measurements provide higher contribution to the state estimation.

Measurement-noise and prior modeling

We now assign an error model to each measurement type. Let

$$\begin{aligned} \sigma_{|V|} &: \text{standard deviation of measured voltage magnitudes (typically 0.005 p.u.),} \\ \sigma_P^{\text{meas}} &: \text{standard deviation of measured active power (e.g., 0.02 p.u.),} \\ \sigma_Q^{\text{meas}} &: \text{standard deviation of measured reactive power (e.g., 0.02 p.u.),} \\ \sigma_P^{\text{pseudo}} &: \text{standard deviation of forecasted active power (larger, e.g., 0.10–0.30 p.u.),} \\ \sigma_Q^{\text{pseudo}} &: \text{standard deviation of forecasted reactive power,} \\ \sigma_{\Re V}, \sigma_{\Im V} &: \text{small anchor variances for real/imag parts, e.g., } 10^{-3}. \end{aligned}$$

Form the diagonal covariance (or its inverse, the weight matrix) as

$$R = \operatorname{diag}\big(\sigma_{|V|}^2 \mathbf{1},\, \mathbf{r}_P^2,\, \mathbf{r}_Q^2,\, \sigma_{\Re V}^2 \mathbf{1},\, \sigma_{\Im V}^2 \mathbf{1}\big), \qquad W = R^{-1},$$

where $\mathbf{r}_P$ is assembled entrywise as

$$[r_P]_i = \begin{cases} \sigma_P^{\text{meas}}, & \text{if node } i \text{ has real } P \text{ measurement,} \\ \sigma_P^{\text{pseudo}}, & \text{otherwise,} \end{cases}$$

and similarly for $\mathbf{r}_Q$. This construction provides uncertainty-based weighting of measured and forecasted quantities. In the absence of an explicit prior distribution on the state vector, the objective below is a weighted least-squares formulation rather than a full Bayesian posterior estimator. Measured values are assigned variances according to their assumed accuracy, while forecasted pseudo-measurements are assigned variances reflecting their forecasting uncertainty. Any artificial voltage anchors must also be justified because overly small variances can bias the estimated states.

Nonlinear weighted least squares

With $\mathbf{z}$, $\mathbf{h}(\mathbf{x})$, and W defined, the BWLS problem is

$$\min_{\mathbf{x}} \; J(\mathbf{x}) = \frac{1}{2}\big(\mathbf{z} - \mathbf{h}(\mathbf{x})\big)^\top W \big(\mathbf{z} - \mathbf{h}(\mathbf{x})\big). \tag{3.21}$$

Because the AC power-flow relations (3.17)–(3.18) are nonlinear, (3.21) is solved iteratively. Let $\mathbf{x}^{(k)}$ be the current estimate. Define the residual

$$\mathbf{r}^{(k)} = \mathbf{z} - \mathbf{h}\big(\mathbf{x}^{(k)}\big),$$

and the Jacobian

$$H^{(k)} = \left.\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^{(k)}}.$$

The Gauss–Newton (or damped Levenberg–Marquardt) update is then

$$\big((H^{(k)})^\top W H^{(k)} + \lambda I\big)\Delta\mathbf{x} = (H^{(k)})^\top W \mathbf{r}^{(k)}, \tag{3.22}$$

where $\lambda \geq 0$ is a damping factor that is decreased when the step is successful and increased when the trial step does not reduce the cost. The state is updated as

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta\mathbf{x}.$$

This procedure is repeated until $\|\Delta\mathbf{x}\|_\infty$ and $\|(H^\top W\mathbf{r})\|_\infty$ are below small tolerances.

The structure of H reflects the measurement blocks:

- rows corresponding to $|V|$ are zero everywhere except the column of the corresponding voltage magnitude;
- rows corresponding to P and Q are the standard partial derivatives of (3.17)–(3.18) with respect to angles and magnitudes, including cross-couplings through the admittance matrix;
- rows corresponding to $\Re V$ and $\Im V$ contain simple derivatives $-|V_i|\sin\delta_i$, $\cos\delta_i$, etc.

Because the network is multi-phase, the slack buses must be locked to their phase-consistent angles at every iteration (e.g., $0, -120°, 120°$). The state vector (3.16) contains only non-slack buses, so this locking is done in the "full" voltage representation before evaluating $\mathbf{h}(\cdot)$.

Posterior covariance and uncertainty propagation

One advantage of solving (3.21) in a Gauss–Newton form is that the normal matrix

$$G = H^\top W H$$

is available at the solution. Under a local linearization and correctly specified measurement covariance, the inverse of G provides an approximation to the state-estimation error covariance:

$$\Sigma_{\mathbf{x}} \approx G^{-1}, \tag{3.23}$$

where the approximation assumes that R contains measurement-error variances in the appropriate units and that the local Jacobian has full column rank. If only relative measurement weights are specified, a residual-based scale estimate may be required; it should not be applied again when the absolute variances are already included in R. The diagonal entries of $\Sigma_{\mathbf{x}}$ then give the marginal variances of voltage angles and magnitudes at each non-slack node:

$$\sigma^2_{\delta,i} = [\Sigma_{\mathbf{x}}]_{i,i}, \qquad \sigma^2_{|V|,i} = [\Sigma_{\mathbf{x}}]_{|\mathcal{N}|+i,\,|\mathcal{N}|+i}.$$

Under an approximately Gaussian estimation-error model, these variances can be used to construct one-standard-deviation bands and approximate 95% intervals, for example,

$$\mathrm{CI}_{95}(|V_i|) = |V_i| \pm 1.96\,\sigma_{|V|,i}, \qquad \mathrm{CI}_{95}(\delta_i) = \delta_i \pm 1.96\,\sigma_{\delta,i}.$$

These uncertainty estimates are useful when forecast-based pseudo-measurements are used, although their validity depends on the measurement-error model and the local linearization.

Conceptually, BWLS can be seen as the "physics-first" counterpart to VBMC:

- VBMC imputes missing AMI entries using low-rank and spatio-temporal structure, and delivers per-entry variances (*data-driven, soft constraints*).
- BWLS enforces the AC power-flow equations with heterogeneous measurement variances, and provides approximate per-bus voltage covariances (*model-driven estimation*).

Both produce an estimate of voltages and powers, but each has its own bias/variance characteristics. BWLS can be biased when many injections are pseudo-measured and their forecast errors have temporal structure not captured in the covariance. VBMC can be biased when the low-rank assumption is stretched by abrupt changes. In Chapter 6, these two estimators are combined through calibrated Maximum Ratio Combining (MRC). The fusion uses calibration residuals to account for estimator bias and uncertainty when combining corresponding voltage-state estimates.

Chapter 4

Optimal Sizing of PV–Battery Systems

The shift toward photovoltaic (PV) generation in rural and small municipal utilities is motivated by the need to reduce both energy purchases (kWh) and peak demand charges (kW). As discussed in Chapter 2, some PV installations mainly reduce daytime energy consumption but often fail to reduce the system peak because the latter may occur in the early morning or late evening when solar irradiance is negligible. This temporal mismatch limits the financial benefit of PV for utilities that are billed on both energy and demand. Battery energy storage systems (BESS) can mitigate this limitation by shifting surplus solar generation from off-peak hours to peak hours, thereby flattening the daily load profile. However, adding storage also introduces additional cost, round-trip losses, and degradation, so the PV–battery sizes must be selected carefully.

The objective of this chapter is to develop a peak-demand-aware, statistically grounded sizing methodology for grid-connected PV–battery systems that

1. explicitly accounts for the stochastic nature of daily load and solar irradiance;
2. identifies, for each PV size, the battery capacity that reduces the *updated* daily peaks up to a specified probability level (95%); and

3. selects the final PV–battery pair based on long-term economic benefit to the utility.

The methodology builds on actual AMI load data and measured irradiance from a rural community and is intended to be transparent enough for utility engineers to implement.

4.1 Methodology

PV–Battery System Component Model

This study considered a grid-connected system composed of silicon-based PV panels, inverters, transformers, and battery energy storage. The PV array converts incident solar irradiance into DC electrical power. For a PV system with a rated size of X kW at standard test conditions, the simplified DC output power at hour h is modeled as

$$\mathrm{PV}(h) = \frac{X}{1000} \times I(h), \tag{4.1}$$

where

- $\mathrm{PV}(h)$ is the DC output power of the PV system at hour h (kW),
- $I(h)$ is the solar irradiance at hour h (W/m^2),
- 1000 W/m^2 is the reference (standard test condition) irradiance.

Equation (4.1) assumes that the PV array produces its rated power X kW when the irradiance is 1000 W/m^2; at other irradiance levels, the model assumes linear scaling and does not explicitly represent temperature-dependent changes in module efficiency.

The DC power is then converted to AC power through the inverter. The inverter performance is represented by a constant efficiency factor η_{inv}, so that the AC output power at hour h is

$$P_{\mathrm{inv}}(h) = \mathrm{PV}(h)\,\eta_{\mathrm{inv}}, \tag{4.2}$$

where $P_{\mathrm{inv}}(h)$ is the AC power delivered by the inverter (kW), and $\eta_{\mathrm{inv}} \in (0, 1]$ denotes the inverter efficiency.

Required Daily Battery Size

The primary task of the battery in the PV–battery system is to store excess energy generated by the PV panels during peak sunlight hours and provide energy during periods of high demand. To develop the proposed methodology, the first step was to calculate the daily load demand after PV installation, referred to as modified daily load profiles. Initially, a range of PV sizes, denoted from X_1 to X_n kW, was considered. For a PV system of X kW, the output AC power is given by Equation (4.2). We then generated the modified daily load profiles by subtracting PV generation from the original load profile for each day. For example, Figure 4.1 illustrates the original load, PV generation, and the modified load, derived by subtracting PV generation from the original load for a PV size of 2000 kW on a selected day in December 2020.

After calculating the modified load profiles for PV sizes ranging from X_1 to X_n for all days, the next step involved determining the required daily battery sizes. These are sizes that can effectively flatten the modified daily load profiles. This calculation assumes a daily battery

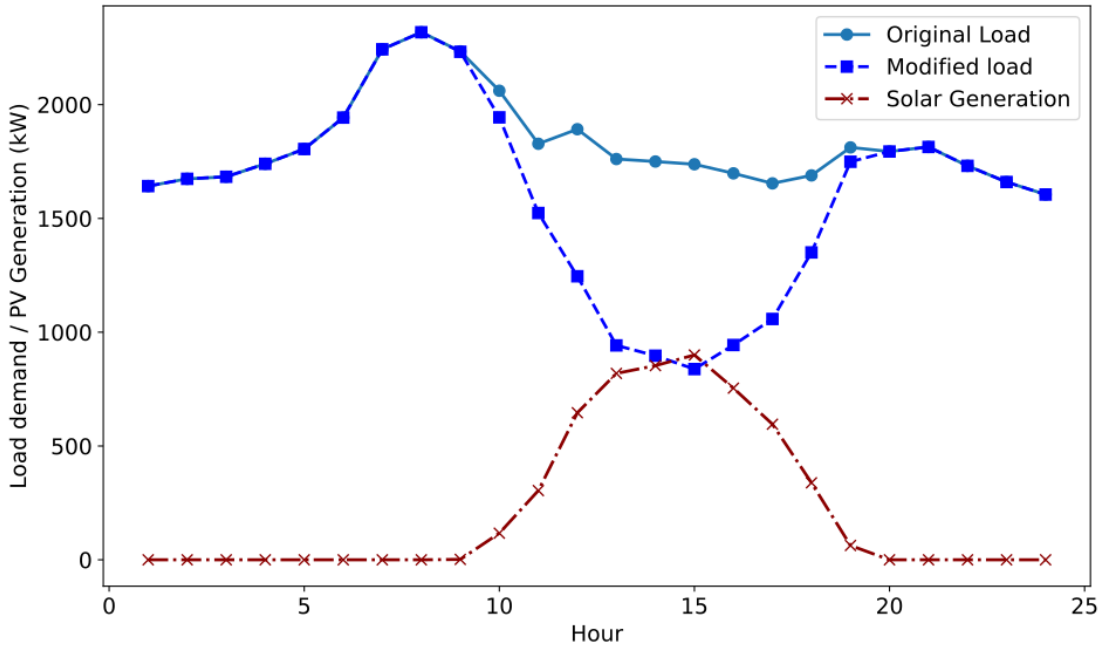


Figure 4.1: *Comparison of daily electrical load with and without PV solar generation for a 2000 kW PV system on a specific day in December 2020.*

cycle, which means the battery fully charges and discharges within a day. To determine the required daily battery size, a horizontal line was drawn across the modified daily load profile so that the area above the line was equal to the roundtrip efficiency multiplied by the area below the line:

$$\text{Area above the line} = \text{Area below the line} \times \text{Battery roundtrip efficiency}. \tag{4.3}$$

The lower area, corresponding to energy drawn for battery charging, was adjusted by the battery charging efficiency and the battery utilization factor, leading to a needed battery size of

$$\text{Needed battery size} = \frac{\text{Area below the line} \times \text{Battery efficiency}}{\text{Utilization factor}}. \tag{4.4}$$

This method ensures that the battery charges when the modified load is below this line and discharges when it is above, while keeping the peak load at the line value. For instance, Figure 4.2 illustrates the established horizontal line on the modified load profile for the day depicted in Figure 4.1, with a PV size of 2000 kW. On this particular day, the horizontal line was drawn such that the area above it equals 1994 kWh, which is the product of the area under the line (2210 kWh) and the roundtrip efficiency (0.9025). Using Equation (4.4), we determined that a battery size of 3000 kWh was needed to effectively flatten this particular modified load profile.

It is important to note that the inclusion of roundtrip efficiency in the battery sizing algorithm realistically simulates battery losses. Hence, while battery installation aims to reduce peak demands, operational losses may necessitate increased energy purchases from the grid.

For a PV system of size X kW, we calculated the required daily battery sizes that effectively flattened the modified daily load profiles for a duration of three years. However, it is not feasible to select different battery sizes for a single system based on the needs of different days. While choosing the largest battery size could flatten all daily load profiles, this approach is economically inefficient. Consequently, the objective is to identify an optimal

battery size that can adequately flatten the load curve on most days, while accepting the risk of inadequate flattening on a few days. As the PV–battery system is connected to the grid, on these specific days, the demand shortage can be met by purchasing power from the grid.

Updated Peaks

Following the established methodology, we first calculated the required daily battery sizes to flatten the modified load profiles for a given PV size of X kW for each day. The subsequent step included calculating the updated daily peak demands after the integration of a battery size of Y kWh into a PV system of size X kW. To achieve this, we selected battery sizes ranging from Y_1 to Y_m as potential sizes for a PV system of size X. For a specific day, if a selected battery size of Y kWh is greater than or equal to the needed battery size, then this Y can effectively flatten the modified load profile for that day. In these instances, the updated peaks for such days are equivalent to the values of the horizontal lines identified in the previous step. For example, as demonstrated in Figure 4.2, the needed battery size for the depicted modified load demand was 3000 kWh for a PV size of 2000 kW. Figure 4.2 shows the horizontal line at 1642 kW. If the selected battery capacity of Y kWh is greater than or equal to 3000 kWh, the peak demand for that day remains at 1642 kW after integrating the

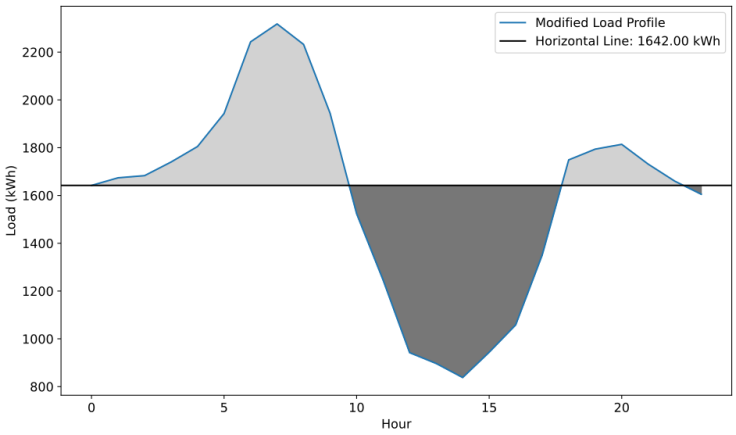


Figure 4.2: *Modified load profile and the adjusted horizontal line to flatten the load profile for a 2000 kW PV system on a specific day in December 2020. Light and dark areas show the battery discharging and charging period, respectively.*

battery size Y kWh and PV size of 2000 kW.

Conversely, if the chosen battery size Y kWh is smaller than the needed battery size for a specific day, it will not be sufficient to flatten the load profile for that day. In such cases, to determine the updated peak demands, we drew two new horizontal lines. The first line was positioned so that the area above it corresponded to the capacity of the selected battery Y kWh. That means the area above the first line equals the product of the selected battery capacity Y kWh, the battery roundtrip efficiency, and the battery utilization factor, all divided by the battery efficiency. This area represents the discharging period of the battery. The first horizontal line therefore represents the updated daily peak demand following the integration of a battery with a capacity of Y kWh into a PV system of size X kW. The second line was positioned such that the area between it and the modified load profile equals the battery charging area values divided by the battery roundtrip efficiency, denoting the battery charging area. For example, Figure 4.3 demonstrates the application of the proposed methodology on the day represented in Figure 4.1, with a PV system size of 2000 kW. With a smaller selected battery size, such as 2000 kWh, we drew two new horizontal lines. The first line was positioned so that the area above it reflects the discharging capacity of the battery. This was calculated as the product of the selected battery capacity (2000 kWh), the roundtrip efficiency, and the utilization factor, all divided by the battery efficiency. This first line, set at 1782 kW, indicates the updated peak demand for that day after integrating the battery size 2000 kWh and PV size 2000 kW systems. The area above the first line, corresponding to 1330 kWh, represents the energy discharged from the battery. Moreover, the second line was drawn such that the area under it equals 1330 kWh divided by a roundtrip efficiency of 0.9025, resulting in 1473.68 kWh. This area signifies the energy charged into the battery.

Following this methodology, for each PV system of size X kW paired with a battery capacity of Y kWh, we calculated the updated peak demands across all data days. Initially, we selected a starting PV size of X_1 and a range of battery capacities from Y_1 to Y_m, computing the updated daily peak demands for every PV and battery combination. This procedure was iteratively conducted for additional PV sizes up to X_n. Afterward, a scaled

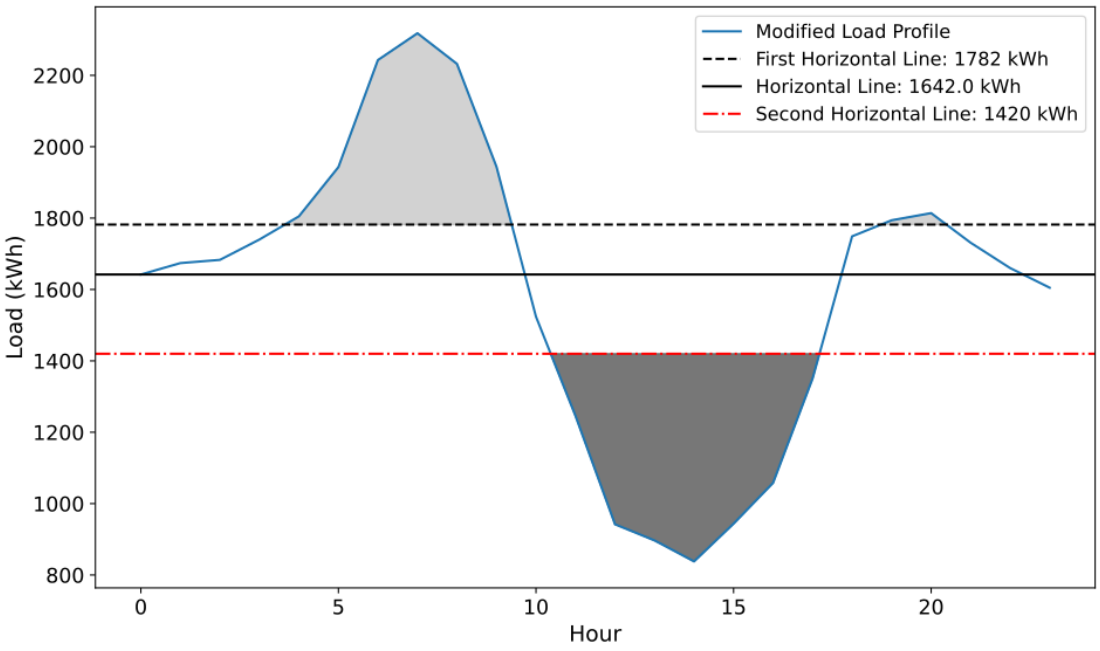


Figure 4.3: *Illustration of the required battery size and updated peak demand calculation for a 2000 kW PV system on a specific day in December 2020. Light and dark areas show the battery discharging and charging period, respectively.*

histogram was generated for the updated daily peak demand associated with each PV and battery combination. Consequently, for PV sizes ranging from X_1 to X_n and battery capacities from Y_1 to Y_m, a series of $n \times m$ scaled histograms were produced. To determine the optimal PV–battery system, these histograms provide the basis for the proposed statistical methodology.

4.2 Proposed Approach 1

Optimal PV–Battery Sizes

The proposed statistical methodology seeks to configure a PV–battery system capable of flattening, with high probability, the majority of daily load profiles up to a prescribed demand threshold. In this work, the target was set such that approximately 95% of the updated daily peaks lie at or below a specified limit, while the remaining 5% of the days are allowed to exceed this limit. This design choice introduces a controlled and clearly quantified level

of operational risk: on a small fraction of high-demand or low-solar days, the local PV–battery system may not be sufficient, and the utility is assumed to procure the deficit from the upstream supplier. Adopting a 5% exceedance probability is consistent with standard engineering risk practices and is supported by the distributional behavior of the updated peaks obtained from simulation [97].

This framework differs from deterministic sizing—which would enforce peak shaving on every single day and thus lead to an over-sized and uneconomical battery—by explicitly acknowledging the stochastic nature of load and solar irradiance. The goal is therefore *probabilistic* peak reduction: identify the PV size and the associated battery size such that 95% of the days can be flattened to or below the desired threshold, while the remaining days are served from the grid.

Statistical Analysis of Updated Peaks

Once the updated daily peak demands are computed for every combination of PV size X_i, $i = 1, \ldots, n$, and battery size Y_j, $j = 1, \ldots, m$, the next step is to model the empirical distribution of these peaks and extract the 95th-percentile value. This was accomplished by fitting well-known probability density functions (PDFs) to the scaled histograms of the updated peaks.

In this study, three PDFs were considered because their shapes matched well with the empirical histograms of daily peaks in PV–battery scenarios: the Gamma distribution, the Log-normal distribution, and the Beta distribution. Each of these distributions is widely used for modeling positive, right-skewed, or bounded nonnegative variables such as load and peak demand.

Gamma distribution. The Gamma distribution is useful for modeling strictly positive and right-skewed data. Its PDF is given by [98]

$$f(x) = \frac{1}{\beta\,\Gamma(\alpha)} \left(\frac{x-\mu}{\beta}\right)^{\alpha-1} \exp\left(-\frac{x-\mu}{\beta}\right), \tag{4.5}$$

where $\alpha > 0$ is the shape parameter, $\beta > 0$ is the scale parameter, μ is a location parameter, and $\Gamma(\alpha)$ is the Gamma function,

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt. \tag{4.6}$$

The Gamma distribution was often found to provide the closest match to the simulated updated peaks.

Beta distribution. When the updated peaks are normalized to lie in the interval $[0, 1]$, the Beta distribution becomes a suitable candidate. Its PDF is [99]

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1 - x)^{\beta-1}, \tag{4.7}$$

where $\alpha > 0$ and $\beta > 0$ control the shape of the distribution.

Log-normal distribution. For positively skewed data that span several magnitudes, the Log-normal distribution is also appropriate. Its PDF is [99]

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln(x) - \mu}{\sigma}\right)^2\right), \tag{4.8}$$

where μ and σ are the mean and standard deviation of the underlying normal random variable $\ln(x)$.

Goodness-of-fit. To decide which of the three candidate PDFs best represents the empirical updated-peak data for a given (X_i, Y_j), the Kolmogorov–Smirnov (KS) test was employed [100]. The KS statistic measures the maximum absolute difference between the empirical cumulative distribution function (CDF) of the simulated peaks and the theoretical CDF of the candidate PDF. The PDF with the smallest KS statistic and the largest p-value (typically $p > 0.05$) was selected as the best fit for that PV–battery combination [101]. This procedure was repeated for every pair (X_i, Y_j), yielding a collection of fitted distributions

over the entire sizing grid.

Step-by-step procedure

After fitting the PDFs, the optimal sizes are identified through the following steps.

1. **Selection of PV and battery sizes.** Start with a PV size X_1 and, for that PV size, compute the updated daily peak demand for all candidate battery sizes $Y_1, Y_2, \ldots, Y_m$. This produces m sets of updated peaks corresponding to the same PV size but different battery capacities.

2. **Histogram creation and PDF fitting.** For each (X_1, Y_j), build a scaled histogram of the updated daily peaks obtained from all historical (or simulated) days. Fit the Gamma, Log-normal, and Beta distributions to this histogram and, using the KS statistic, select the best-fitting PDF. Store the fitted parameters for that (X_1, Y_j) pair.

3. **Determining the 0.95 threshold.** For the selected PDF corresponding to (X_1, Y_j), compute the 95th-percentile threshold as

$$F^{-1}(0.95), \tag{4.9}$$

   where $F^{-1}(\cdot)$ is the inverse CDF of the fitted distribution. This value represents the peak level that 95% of the days will not exceed when the system operates with PV size X_1 and battery size Y_j. Repeating this for all $j = 1, \ldots, m$ gives a set of 95% thresholds for the same PV size.

4. **Optimal sizing criterion.** Suppose the utility specifies a technical peak limit of T kW (for example, the contracted demand level). If, for some (X_1, Y_j), the 95th-percentile value in (4.9) is equal to or less than T, then Y_j is an admissible battery size for PV size X_1. If none of the discrete battery sizes leads to a 95th-percentile equal to T, the parameters of the fitted PDF can be slightly adjusted so that

$$F(T; \text{parameters}) = 0.95, \tag{4.10}$$

i.e., the CDF evaluated at T equals 0.95. Solving (4.10) provides a refined set of PDF parameters associated with the desired threshold. Interpolation between the battery sizes Y_j and Y_{j+1} that bracket this refined solution can then be used to infer the battery size that would achieve exactly 95% compliance at threshold T.

The previous four steps are repeated for all PV sizes $X_1, X_2, \ldots, X_n$. The result is an $n \times m$ grid of candidate PV–battery pairs, from which the combinations satisfying the fitted 95% peak-demand criterion at T kW are identified. This matrix constitutes the *technically* acceptable solution set. At this stage, multiple PV–battery combinations will typically satisfy the 95% peak-flattening requirement. Therefore, a final selection must also consider economic criteria, which will be addressed in the subsequent sections. The organizational flowchart of the simulation procedure in this study is shown in Figure 4.4.

4.3 Proposed Approach 2

The first sizing method (Approach 1) treated the problem as "given a target peak T, find the PV–battery pair whose 95th-percentile updated peak does not exceed T." That approach is convenient when the utility already knows the desired demand ceiling. In some planning studies, however, the utility does *not* start from a fixed peak target; instead, it wants to see how the achievable 95% peak level *evolves* as the battery becomes larger for a given PV size, and then pick the smallest battery at which the benefit starts to saturate. Approach 2 provides exactly this alternative: it uses the same updated-daily-peak data as Approach 1, but it looks for an "elbow" in the curve of (battery size) $\mapsto$ (95% peak) for each PV size. In other words, Approach 1 is target-driven, while Approach 2 is shape-driven.

Statistical Modeling of Updated Daily Peaks

With prior knowledge of daily energy and peak load, and after constructing the modified daily load profiles as in the previous subsection, we first fix a PV size X (kW). For this fixed

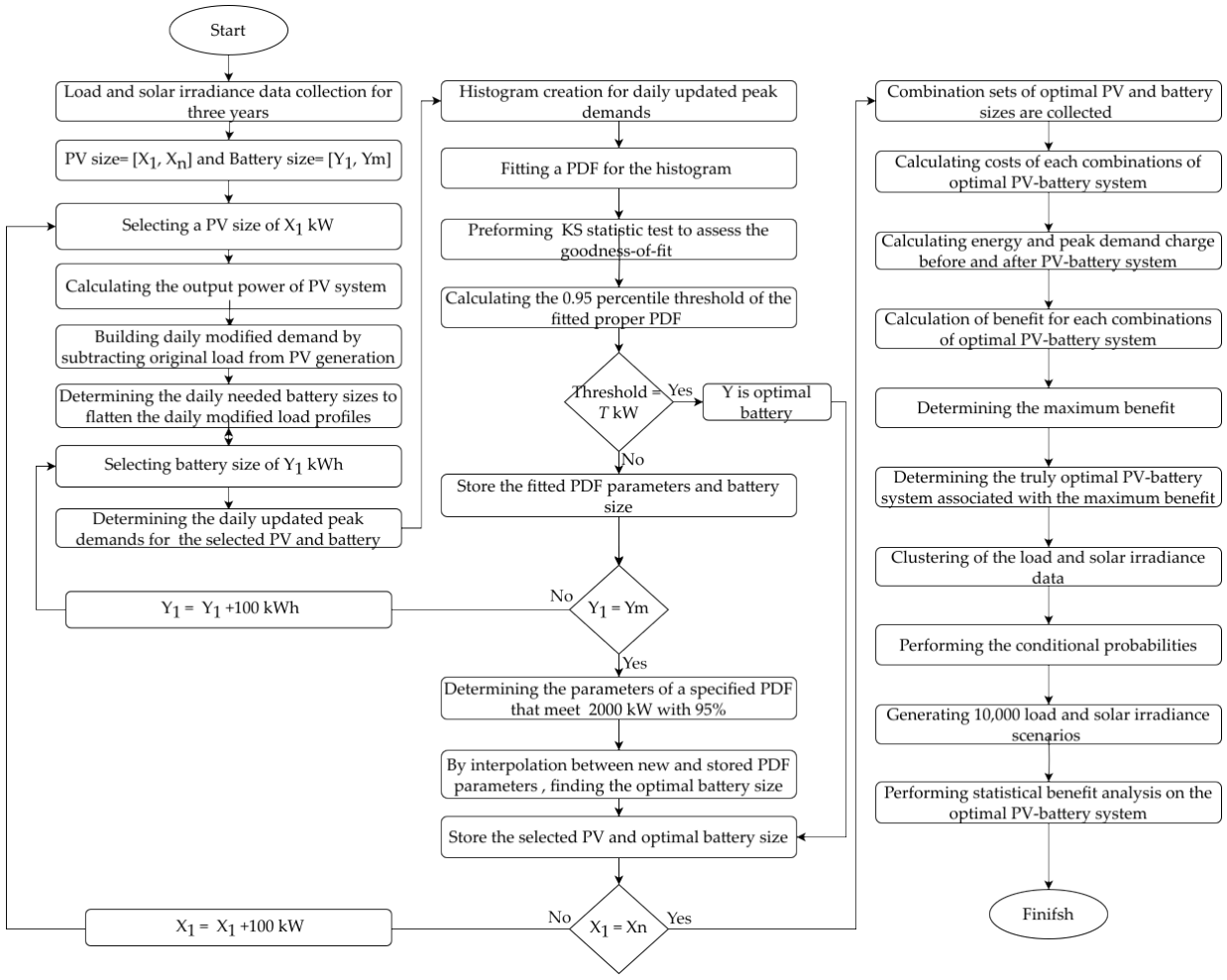


Figure 4.4: *Organizational flowchart of the proposed statistical peak-constrained PV–battery sizing approach (method 1).*

PV size, we define a range of candidate battery sizes

$$Y_1, Y_2, \ldots, Y_n,$$

and, for each Y_j, we compute the *updated* daily peak loads over all available days using the charging/discharging logic described earlier. This yields, for every battery size Y_j, a sample of daily peaks. From each such sample we can draw a scaled histogram and then fit an appropriate probability distribution function (PDF).

Consistent with the literature, we consider two flexible families whose shapes match positive, right-skewed peak data: the Gamma distribution and the Lognormal distribution.

The PDF of the Gamma distribution is written as [102]:

$$f(x) = \frac{\left(\frac{x-\mu}{\beta}\right)^{\alpha-1} \exp\left(-\frac{x-\mu}{\beta}\right)}{\beta\,\Gamma(\alpha)}, \qquad x \geq \mu,\ \alpha > 0,\ \beta > 0, \tag{4.11}$$

where μ, α, and β are the location, shape, and scale parameters, respectively, and $\Gamma(\alpha)$ is the Gamma function

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt. \tag{4.12}$$

The PDF of the Lognormal distribution, which is also suitable for strictly positive, positively skewed data such as updated daily peaks, is given by

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln(x) - \mu}{\sigma}\right)^2\right), \qquad x > 0, \tag{4.13}$$

where μ is the mean of the associated normal variable and σ is its standard deviation.

To evaluate how well each fitted PDF reproduces the empirical distribution of updated daily peaks, we employ the Kolmogorov–Smirnov (KS) goodness-of-fit test. The KS test compares the empirical cumulative distribution function (CDF) of the data with the CDF of the fitted PDF; a lower KS statistic and a higher p-value (typically $p > 0.05$) indicate a better fit. For every PV size X and every battery size Y_j, we retain the PDF (Gamma or Lognormal) that achieves the better KS performance.

Once the best-fitting PDF has been selected, we determine the peak-demand threshold associated with 95% probability. For a given PV size X, battery size Y, and fitted CDF $F(\cdot)$, we compute

$$F^{-1}(0.95) = T_{0.95}(X, Y), \tag{4.14}$$

which is the value of peak demand that the PV–battery system can meet with 95% probability. Repeating (4.14) for all battery sizes $Y_1, \ldots, Y_n$ associated with the fixed PV size X produces a sequence

$$T_{0.95}(X, Y_1),\ T_{0.95}(X, Y_2),\ \ldots,\ T_{0.95}(X, Y_n),$$

that monotonically (or nearly monotonically) decreases as the battery size grows. This sequence is the input to the elbow-based selection described next.

Elbow-Based Optimal PV–Battery Sizing

The central observation behind this heuristic is that increasing battery capacity initially provides substantial peak-demand reduction, while further increases beyond a certain point provide only marginal improvement. The elbow method is used to identify this transition in the relationship between battery capacity and the 95th-percentile peak-demand threshold.

For a fixed PV size X, the selected battery capacity $Y^\star$ corresponds to the elbow of the curve

$$Y \longmapsto T_{0.95}(X, Y),$$

where $T_{0.95}(X, Y)$ is the 95th-percentile peak-demand threshold obtained from (4.14).

To identify the elbow numerically, a straight line is drawn between the first and last points of the threshold curve over the evaluated battery-capacity range. The perpendicular distance between each point on the curve and this line is calculated. The battery capacity corresponding to the maximum distance is selected as $Y^\star$. This maximum-distance criterion identifies the bend in the curve, where the substantial peak reduction obtained from increasing battery capacity transitions to relatively small additional reductions.

By repeating this procedure for all PV sizes $X_1, X_2, \ldots, X_n$, a set of candidate PV–battery pairs is obtained:

$$\{(X_1, Y_1^\star),\ (X_2, Y_2^\star),\ \ldots,\ (X_n, Y_n^\star)\}.$$

Each pair has its own corresponding 95th-percentile peak-demand threshold. Unlike Approach 1, these pairs are not required to satisfy a common predefined peak-demand limit. Instead, the battery capacity for each PV size is selected based on diminishing returns in peak reduction.

The resulting candidate pairs are then evaluated using the economic analysis described in the subsequent sections. The final PV–battery configuration is selected as the pair with the highest financial benefit among these candidates. The organizational flowchart of the

simulation procedure is shown in Figure 4.5.

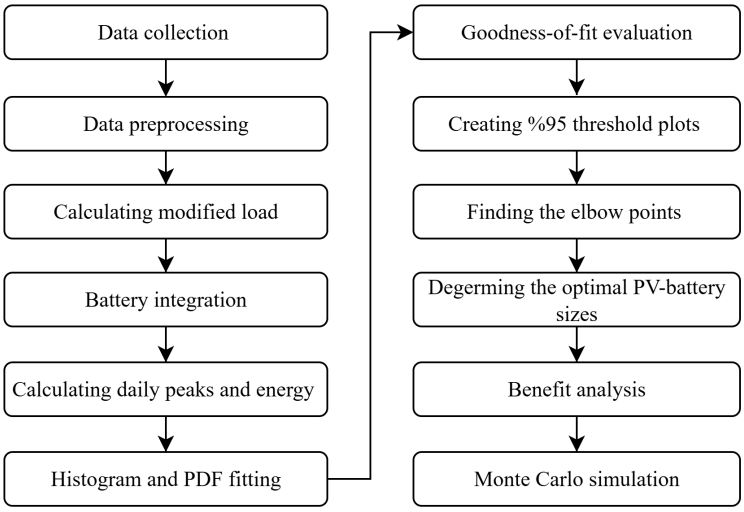


Figure 4.5: *Organizational flowchart of the proposed elbow-based PV–battery sizing approach (method 2).*

4.4 Benefit Analysis

To determine the truly optimal PV–battery system for installation, it is essential to assess each technically feasible PV–battery combination from an economic perspective. In this section, every pair (X, Y) obtained from the sizing procedures is evaluated in terms of upfront investment, recurring costs, and the present worth of energy and demand-charge savings. The goal is to identify the configuration that maximizes the net economic benefit over the project lifetime.

Initial Investment Cost

The total initial investment cost is composed of PV-side costs, battery costs (including future replacement, brought to present worth), and balance-of-system costs. For a PV size of X kW and a battery size of Y kWh, the following components are considered.

PV installation cost

$$C_{\text{pv}}^{\text{ini}} = X \cdot C_{\text{pv}} \tag{4.15}$$

where $C_{\text{pv}}^{\text{ini}}$ is the total PV installation cost (USD) and C_{pv} is the capital cost of PV panels (USD/kW).

Inverter cost

$$C_{\text{inv}}^{\text{ini}} = X \cdot \eta_{\text{inv}} \cdot K \cdot C_{\text{inv}} \tag{4.16}$$

where $C_{\text{inv}}^{\text{ini}}$ is the inverter installation cost (USD), C_{inv} is the unit inverter cost (USD/kW), η_{inv} is the inverter efficiency, and K is the oversizing factor of the inverter.

Labor cost

$$C_{\text{labor}}^{\text{ini}} = X \cdot C_{\text{labor}} \tag{4.17}$$

where $C_{\text{labor}}^{\text{ini}}$ is the labor cost (USD) and C_{labor} is the unit labor cost (USD/kW).

Equipment cost

$$C_{\text{eq}}^{\text{ini}} = X \cdot C_{\text{eq}} \tag{4.18}$$

where $C_{\text{eq}}^{\text{ini}}$ is the initial cost of auxiliary equipment (USD) and C_{eq} is the unit equipment cost (USD/kW).

Overhead cost.

$$C_{\text{over}}^{\text{ini}} = X \cdot C_{\text{over}} \tag{4.19}$$

where $C_{\text{over}}^{\text{ini}}$ is the total overhead (USD) and C_{over} is the unit overhead cost (USD/kW).

Battery cost. Because the battery lifetime is typically shorter than the project lifetime, at least one replacement is anticipated. The initial battery cost is

$$C_{\text{bat}}^{1,\text{ini}} = Y \cdot C_{\text{bat}}^{1} \tag{4.20}$$

where $C_{\text{bat}}^{1,\text{ini}}$ is the cost of the first battery (USD) and C_{bat}^{1} is the unit cost of the first battery (USD/kWh).

Assuming future battery prices will decline, the replacement battery cost is modeled as

$$C_{\text{bat}}^{2,\text{ini}} = Y \cdot C_{\text{bat}}^{2} \tag{4.21}$$

where $C_{\text{bat}}^{2,\text{ini}}$ is the nominal cost of the replacement battery (USD) and C_{bat}^{2} is its (reduced) unit cost (USD/kWh). This replacement occurs in year M, so its present worth is

$$C_{\text{bat}}^{2,\text{present}} = C_{\text{bat}}^{2,\text{ini}} \cdot \frac{1}{(1+i)^{M}} \tag{4.22}$$

where i is the discount rate. The total battery-related initial cost, accounting for the present worth of the replacement, is

$$C_{\text{bat}}^{\text{ini}} = C_{\text{bat}}^{1,\text{ini}} + C_{\text{bat}}^{2,\text{present}}. \tag{4.23}$$

Transformer and interconnection cost. Let $C_{\text{trans}}^{\text{ini}}$ denote the cost of the required transformer and interconnection hardware.

Total initial investment cost. Collecting all terms, the total initial investment cost (TIIC) is

$$\text{TIIC} = C_{\text{pv}}^{\text{ini}} + C_{\text{inv}}^{\text{ini}} + C_{\text{labor}}^{\text{ini}} + C_{\text{eq}}^{\text{ini}} + C_{\text{over}}^{\text{ini}} + C_{\text{trans}}^{\text{ini}} + C_{\text{bat}}^{\text{ini}}. \tag{4.24}$$

Operation, Maintenance, and Insurance Costs. Once the PV–battery system is in service, annual operation, maintenance, and insurance (O&M) costs must be included. These cover monitoring, periodic maintenance, cleaning, insurance, and minor replacements. We model the annual O&M cost as

$$C_{\text{O\&M}}^{\text{annual}} = X \cdot C_{\text{O\&M}}, \tag{4.25}$$

where $C_{\mathrm{O\&M}}$ is the unit annual O&M cost (USD/kW). Over an N-year project life and discount rate i, the present worth of these annual costs is

$$C_{\mathrm{O\&M}}^{\mathrm{present}} = C_{\mathrm{O\&M}}^{\mathrm{annual}} \left[\frac{1-(1+i)^{-N}}{i} \right]. \tag{4.26}$$

Peak-Demand and Energy Costs

For most small utilities, two streams of operating cost are dominant: the monthly (or 12-month) peak-demand charge and the energy charge.

Peak-demand charge. Let $P_m^{\max}$ be the monthly peak demand in month m (kW), and let C_{peak} be the demand-charge rate (USD/kW). Then the annual peak-demand charge in year n is

$$\mathrm{PDC}_n = \sum_{m=1}^{12} P_m^{\max} \cdot C_{\mathrm{peak}}. \tag{4.27}$$

The present value of peak-demand payments *before* PV–battery installation is

$$\mathrm{PDC}_{\mathrm{before}}^{\mathrm{present}} = \sum_{n=1}^{N} \left(\mathrm{PDC}_n \cdot \frac{1}{(1+i)^n} \right). \tag{4.28}$$

After PV–battery installation, the monthly peaks are reduced to $P_m'^{\max}$, so the annual charge becomes

$$\mathrm{PDC}'_n = \sum_{m=1}^{12} P_m'^{\max} \cdot C_{\mathrm{peak}}, \tag{4.29}$$

and its present value is

$$\mathrm{PDC}_{\mathrm{after}}^{\mathrm{present}} = \sum_{n=1}^{N} \left(\mathrm{PDC}'_n \cdot \frac{1}{(1+i)^n} \right). \tag{4.30}$$

Energy cost. Let P_h be the hourly energy consumption (kWh) and C_{energy} be the energy rate (USD/kWh). Then the annual energy cost in year n is

$$\text{EC}_n = \sum_{h=1}^{H} P_h \cdot C_{\text{energy}}, \tag{4.31}$$

where H is the total number of hours in the year. Its present value before PV–battery installation is

$$\text{EC}_{\text{before}}^{\text{present}} = \sum_{n=1}^{N} \left(\text{EC}_n \cdot \frac{1}{(1+i)^n} \right). \tag{4.32}$$

After PV generation and peak shaving, the annual energy cost becomes EC'_n, and its present value is

$$\text{EC}_{\text{after}}^{\text{present}} = \sum_{n=1}^{N} \left(\text{EC}'_n \cdot \frac{1}{(1+i)^n} \right). \tag{4.33}$$

Economic Benefit

The economic benefit over the project lifetime is obtained by comparing the present-worth cost stream without the PV–battery system and the present-worth cost stream with the PV–battery system.

Total cost before installation.

$$\text{TC}_{\text{before}} = \text{EC}_{\text{before}}^{\text{present}} + \text{PDC}_{\text{before}}^{\text{present}}. \tag{4.34}$$

Total cost after installation.

$$\text{TC}_{\text{after}} = \text{TIIC} + C_{\text{O\&M}}^{\text{present}} + \text{PDC}_{\text{after}}^{\text{present}} + \text{EC}_{\text{after}}^{\text{present}}. \tag{4.35}$$

Net economic benefit.

$$\text{Benefit} = \text{TC}_{\text{before}} - \text{TC}_{\text{after}}. \tag{4.36}$$

A positive value of (4.36) indicates that, over the lifetime of the project, the PV–battery system yields sufficient savings in energy and demand charges to offset the initial and recurring costs. In this dissertation, this calculation is carried out for every PV–battery pair that satisfied the technical 95% peak-flattening criterion in the previous sections. The pair (or small set of pairs) with the largest positive benefit is then identified as the economically optimal solution. Finally, to ensure robustness to variability in load and solar irradiance, these economically optimal solutions are further stress-tested using the Monte Carlo scenario framework described later.

4.5 Modified Monte Carlo Simulation

Monte Carlo simulation is a statistical technique that utilizes random sampling and probabilistic modeling for scenario analysis, decision-making, and predictive modeling. It operates by running a large number of simulations with random variables, thereby generating a wide range of possible outcomes and their probabilities [103]. However, in this study, the typical approach to Monte Carlo simulation requires modification due to the simultaneous consideration of two parameters: daily load and solar irradiance profiles. The goal is to generate multiple scenarios that accurately reflect the interdependence of solar irradiance and demand profiles throughout the year. A key challenge in this methodology is the logical pairing of load and solar irradiance profiles to reflect realistic conditions. For instance, it is not methodologically sound to associate a high-demand profile, typical of summer conditions, with a solar irradiance profile of winter. Such contradictions lead to skewed results and impair the reliability of the simulation. To address this, we integrated a clustering technique to categorize similar load demands and solar irradiance profiles into distinct groups and, subsequently, to generate realistic paired scenarios.

Time-series clustering involves grouping similar temporal data patterns into distinct clusters. To cluster the time-series data of load demands and solar irradiance, we categorized three years of data on a monthly basis. Subsequently, within each month, daily time series for load and solar irradiance were clustered separately. Different clustering techniques were

employed in the two approaches to examine the temporal patterns and variability of the daily profiles using different similarity and grouping strategies. The first approach employed K-means and self-organizing maps (SOM), while the second approach used K-medoids with dynamic time warping (DTW) and hierarchical clustering. Despite these differences, both approaches followed the same modified Monte Carlo simulation framework for generating paired load-demand and solar-irradiance scenarios. The clustering procedure involved the following steps [104]:

1. **Data preprocessing.** The first step involved comprehensive data preparation. This included cleaning the data to remove any inconsistencies or errors, addressing outliers, and ensuring that all data were correctly formatted. Subsequently, we normalized the data values to fall between 0 and 1. This standardization is crucial for comparability analysis. Then, we grouped the data monthly, aggregating the three years of data for further analysis.

2. **Similarity measures.** The objective of time-series clustering is to categorize time-series datasets into clusters in which datasets within each cluster exhibit maximum similarity among themselves and minimal similarity with datasets in other clusters. A similarity measure is therefore crucial in quantifying the degree of resemblance between two time series. In this study, we employed dynamic time warping (DTW), a technique that has demonstrated significant efficacy in assessing similarity, particularly in the energy-management sector [105]. DTW compares each point of one time series with multiple points of another, finding the best alignment by minimizing the cumulative distance between these matched points. By allowing such flexibility in the alignment, DTW effectively captures the inherent patterns and shapes within the time-series data, even when these occur at different rates or phases.

3. **Clustering algorithms.** The next step was to employ appropriate time-series clustering algorithms. Time-series clustering is a complex process, and validation of the clustering results is challenging. For this purpose, different clustering techniques were

employed in the two approaches to examine the temporal patterns and variability of daily load-demand and solar-irradiance profiles.

In Approach 1, K-means and self-organizing maps (SOM) were employed. K-means is a partitioning clustering algorithm that has shown effective performance in various power-system clustering applications [106]. It groups similar daily profiles around cluster centroids and has relatively low computational complexity. However, K-means requires the number of clusters to be specified in advance and can be sensitive to outliers and the initialization of centroids.

On the other hand, SOM is an unsupervised neural network that produces a structured, topologically ordered representation of the data during its training process [107]. SOM visualizes high-dimensional data in a low-dimensional map and preserves the topological relationships among the input profiles. This capability facilitates the identification of patterns and trends within complex time-series datasets [108]. However, SOM requires careful selection of the appropriate map size and learning parameters.

In Approach 2, K-medoids with dynamic time warping and hierarchical clustering were employed. K-medoids is a partitioning clustering algorithm similar to K-means, but it selects actual data points as the representatives of clusters rather than calculating their means. This characteristic makes K-medoids suitable for non-Euclidean similarity measures such as DTW and generally less sensitive to outliers than K-means. DTW measures the similarity between two daily profiles by allowing temporal alignment between their observations, making it useful for identifying similar load-demand and solar-irradiance patterns that may occur at slightly different times. Hierarchical clustering, on the other hand, groups daily profiles based on their pairwise similarities and organizes them into a hierarchical structure. Unlike partitioning methods, hierarchical clustering does not require the number of clusters to be specified before constructing the hierarchy. The resulting dendrogram provides a representation of the relationships among daily profiles and allows the selection of an appropriate number of clusters.

The different clustering techniques were considered to examine the daily profiles using complementary grouping strategies. While K-means and SOM identify clusters through centroid-based partitioning and topology-preserving learning, K-medoids and hierarchical clustering use representative observed profiles and pairwise similarities. Despite these differences, the resulting clusters were used in the same modified Monte Carlo simulation framework for generating paired load-demand and solar-irradiance scenarios.

K-means procedure: K-means clustering aims to partition the data into K clusters, in which each data point belongs to the cluster with the nearest centroid, serving as a prototype of the cluster. The K-means algorithm operates through the following iterative steps [109]:

1. Initialization: Randomly select K daily profiles as the initial centroids of the clusters.

2. Assignment step: Assign each daily profile (load or solar) to the nearest centroid based on the selected distance measure.

3. Update step: Recalculate the centroid of each cluster based on the profiles assigned to that cluster.

4. Convergence: Repeat the assignment and update steps until the positions of the centroids stabilize or the cluster assignments no longer change significantly.

5. Optimal number of clusters: Determining the optimal number of clusters is a critical aspect. We employed the Elbow method to identify this number. To apply the Elbow method, we first executed the K-means algorithm over a range of K values, then computed the within-cluster sum of squares (WCSS) for each K, and finally plotted these WCSS values against the number of clusters. By observing the WCSS curve, we identified the point where the rate of decrease in WCSS significantly slowed down. The K value at this elbow point was considered the optimal number of clusters, as it indicates a trade-off between increasing the number of clusters and reducing WCSS [110].

SOM procedure: SOM is a type of artificial neural network that is trained using unsupervised learning to produce a low-dimensional (typically 2D) map [111]. The methodology for applying SOM in this study is as follows:

1. Initialization: Initialize the SOM neural network with weight vectors, usually selected randomly within the range of the input data.

2. Competitive learning: For each daily profile, SOM identifies the best-matching unit (BMU) by finding the neuron whose weight vector is closest to the input profile.

3. Weight adjustment: The weights of the BMU and its neighboring neurons are adjusted to become more similar to the input profile, with the adjustment magnitude decreasing over time and with the distance from the BMU.

4. Iterative process: This competitive learning and weight-adjustment cycle is repeated over many epochs, allowing the SOM to evolve and form a map that reflects the intrinsic structure of the data.

5. Cluster visualization: The final SOM map groups similar daily profiles close to one another, enabling visual or algorithmic extraction of clusters.

K-medoids with DTW procedure: K-medoids is a partitioning clustering algorithm that groups similar daily profiles around representative observations, called medoids. Unlike K-means, the medoid of each cluster is selected from the original dataset. In this approach, DTW was used to calculate the similarity between daily load-demand and solar-irradiance profiles. The clustering procedure consists of the following steps [112]:

1. Initialization: Randomly select K daily profiles as the initial medoids of the clusters.

2. Distance calculation: Calculate the DTW distance between each daily profile and the selected medoids. DTW identifies an optimal alignment between two time series by minimizing the cumulative distance between their matched observations.

3. Assignment step: Assign each daily profile to the cluster whose medoid has the minimum DTW distance.

4. Medoid update: For each cluster, select the daily profile that minimizes the sum of DTW distances to all other profiles within that cluster. The selected profile becomes the new medoid.

5. Convergence: Repeat the assignment and medoid-update steps until the medoids no longer change or the clustering objective reaches a stable value.

6. Number of clusters: Evaluate different values of K to determine an appropriate number of clusters based on the similarity of daily profiles within each cluster and their separation from other clusters.

Hierarchical clustering procedure: Hierarchical clustering organizes daily profiles into groups based on their pairwise similarities. In the agglomerative approach, each daily profile initially forms a separate cluster, and the most similar clusters are progressively merged. The methodology consists of the following steps [113]:

1. Initialization: Consider each daily load-demand or solar-irradiance profile as an individual cluster.

2. Distance calculation: Calculate the pairwise distances between daily profiles using the selected similarity measure.

3. Cluster merging: Identify the two most similar clusters and merge them into a single cluster according to the selected linkage criterion.

4. Distance update: Recalculate the distances between the newly formed cluster and the remaining clusters based on the linkage criterion.

5. Iterative process: Repeat the merging and distance-update steps until all daily profiles are grouped into a single hierarchical structure.

6. Cluster selection: Examine the resulting dendrogram and select an appropriate level at which to divide the hierarchy into distinct clusters of daily profiles.

Finally, the clustering results were examined separately for each approach. In Approach 1, the results obtained from K-means and SOM were compared to identify consistent groupings of daily load-demand and solar-irradiance profiles. In Approach 2, K-medoids with DTW and hierarchical clustering were used to examine the grouping of daily profiles based on their temporal similarities. The resulting monthly clusters from each approach were then integrated into the same modified Monte Carlo simulation framework to generate realistic paired load-demand and solar-irradiance scenarios.

Modified Monte Carlo Simulation

After clustering the daily load and solar irradiance for each month, we implemented a common modified Monte Carlo simulation framework for both approaches to generate realistic scenarios. Although different clustering algorithms were employed, the subsequent scenario-generation procedure remained the same, using the resulting monthly clusters and their conditional probabilities. Suppose that, for a given month, we have

$$\{I_1, I_2, \ldots, I_m\} \quad \text{solar-irradiance clusters, and} \quad \{L_1, L_2, \ldots, L_n\} \quad \text{load-demand clusters.}$$

The steps for a typical month were as follows:

1. **Assign probabilities to solar-irradiance clusters.** For each solar-irradiance cluster I_i, $i = 1, 2, \ldots, m$, we calculated its probability as

$$P(I_i) = \frac{\text{number of daily profiles in } I_i}{\text{total number of daily profiles in the month}}. \tag{4.37}$$

2. **Establish probability intervals.** This was carried out by sequentially adding the probabilities of the clusters. For the first cluster I_1, its probability interval is

$$\text{PI}(I_1) = [0,\ P(I_1)]. \tag{4.38}$$

For the second cluster I_2, the interval is defined as

$$\mathrm{PI}(I_2) = \big(P(I_1),\, P(I_1) + P(I_2)\big], \tag{4.39}$$

and this continues for each cluster I_m, where

$$\mathrm{PI}(I_m) = \left(\sum_{k=1}^{m-1} P(I_k),\, \sum_{k=1}^{m} P(I_k)\right]. \tag{4.40}$$

3. **Random cluster selection for solar irradiance.** A random number $R \in [0, 1]$ was selected from a uniform distribution, and the solar-irradiance cluster I_i for which $R \in \mathrm{PI}(I_i)$ was selected as the irradiance type for that simulated day.

4. **Identify the corresponding historical days.** After selecting irradiance cluster I_i, we determined the specific historical days that belong to that irradiance cluster. These days form the admissible set for load–solar pairing.

5. **Match the days with load clusters.** For each identified day in the irradiance cluster I_i, we found its associated load-demand cluster from the set $\{L_1, \ldots, L_n\}$. This step allows us to build the joint frequency of observing load cluster L_j together with irradiance cluster I_i.

6. **Calculate the conditional probability for load clusters.** After selecting the solar-irradiance cluster I_i, the probability of each load-demand cluster conditioned on the selection of I_i was calculated as

$$P(L_j \mid I_i) = \frac{P(L_j \cap I_i)}{P(I_i)}, \tag{4.41}$$

where $P(L_j \cap I_i)$ is the joint empirical probability of days that were simultaneously in load cluster L_j and irradiance cluster I_i.

7. **Construct conditional probability intervals.** Similar to Steps 2 and 3, we constructed intervals for the conditional probabilities $\{P(L_j \mid I_i)\}_{j=1}^{n}$. A second random

number was drawn to select one of the load clusters according to these conditional probabilities.

8. **Random selection of a daily pair.** From the selected irradiance cluster I_i and the selected load cluster L_j, a specific historical day was randomly chosen. If multiple profile pairs were available within the selected clusters, one pair was selected uniformly at random so that each admissible day had an equal chance of being chosen.

9. **Scenario generation over the year.** By repeating Steps 1–8 for every day of the year, and then repeating the entire yearly generation multiple times, we obtained a large set of realistic daily load–solar scenarios. These scenarios preserved the interdependencies between solar irradiance and load-demand patterns observed in the historical data.

Fig. 4.6 illustrates the process of determining the conditional probabilities for a given month. The overall procedure for generating daily load-demand and solar-irradiance scenarios for both approaches is presented in Fig. 4.7. Although different clustering algorithms were employed, the same modified Monte Carlo simulation procedure was applied to both approaches. By considering the historical interdependencies between solar irradiance and load-demand profiles, this methodology generates realistic paired scenarios for the PV–battery sizing analysis.

4.6 Results

For this study, actual hourly solar irradiance and load-demand data were collected from a rural distribution utility in Kansas, USA, over a three-year period from January 2019 through December 2021. These data provided the basis for (i) constructing monthly clusters of daily load and irradiance profiles, (ii) generating realistic paired scenarios via the modified Monte Carlo simulation described in Section 4.8.2, and (iii) evaluating the technical and economic performance of candidate PV–battery configurations.

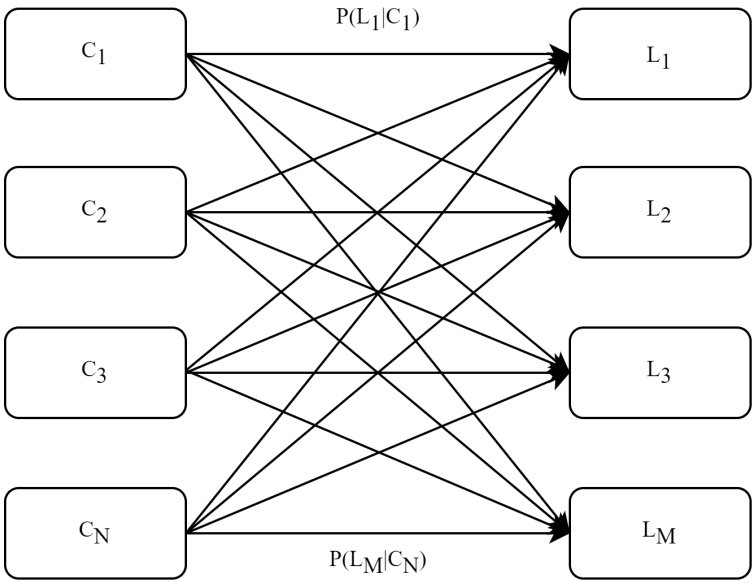


Figure 4.6: *Process of determining the conditional probabilities.*

For Approach 1, it was assumed that the distribution utility seeks to flatten the daily net-load profile (i.e., load after PV) up to a predefined demand threshold of 2000 kW. This target value is consistent with the observed feeder peak levels during the study period and represents a realistic planning limit for a small municipal system. As discussed in the sizing section for Approach 1, the statistical procedure identifies, for each PV nameplate size, the smallest battery energy capacity that can keep 95% of the updated daily peaks at or below this 2000 kW threshold, while allowing the remaining 5% of days to exceed the threshold.

For Approach 2, instead of using a predefined demand threshold, the 95th-percentile peak-demand threshold is calculated for different PV–battery configurations. The battery size is then selected based on the elbow point of the resulting threshold curve, which represents a balance between peak-demand reduction and battery capacity.

The economic analysis was performed assuming a 30% investment tax credit (ITC) for the initial PV–battery installation. Consistent with the assumptions of this study, the ITC was applied only to the first battery purchase and not to the replacement battery that is assumed to be installed later in the project life. A project lifetime of 20 years and a real discount rate of 8% were used to compute the present worth of all annual costs and savings.

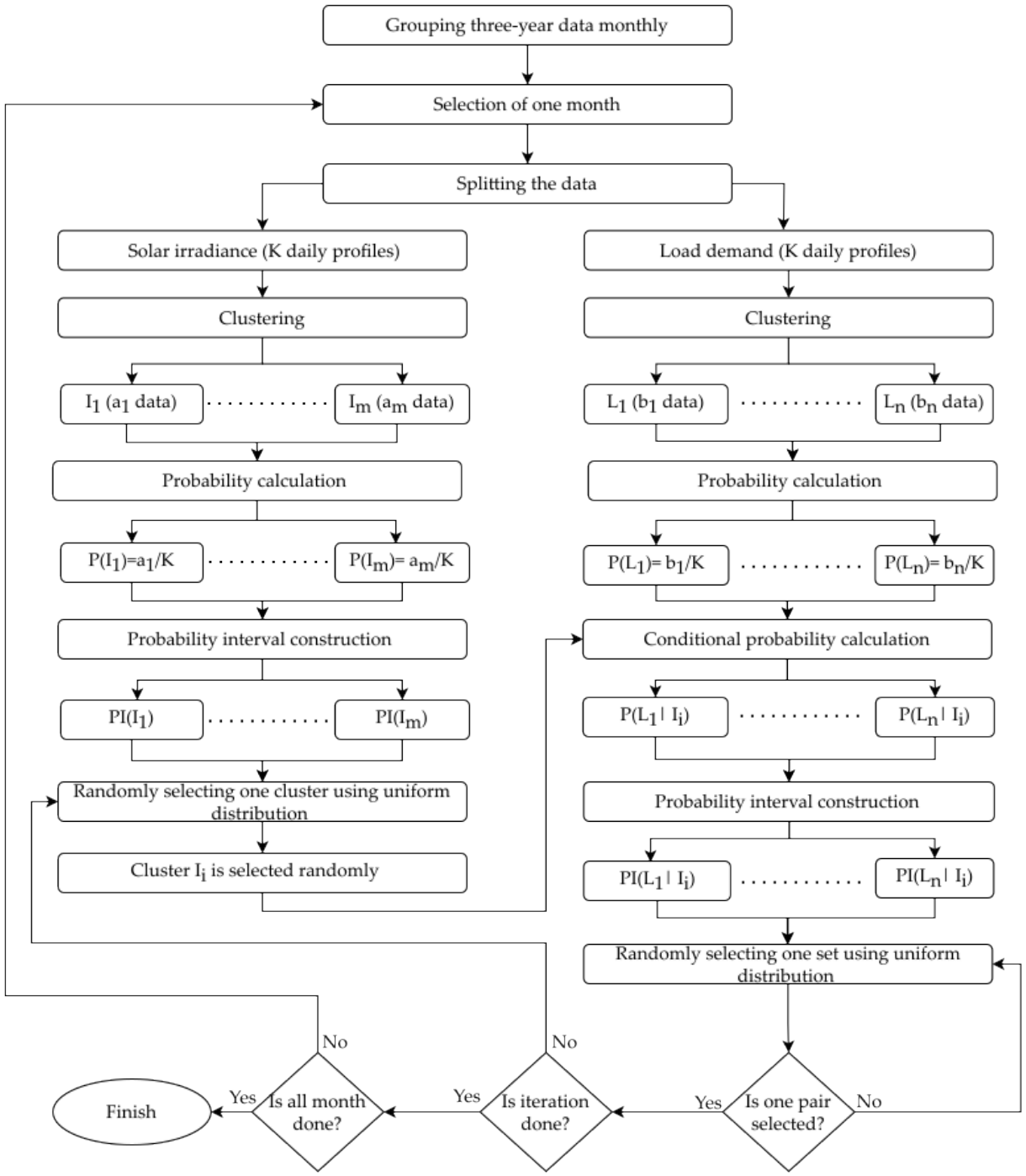


Figure 4.7: *Organizational flowchart to generate load and solar irradiance scenarios based on their interdependencies.*

Table 4.1 summarizes the technical and economic parameters used in the case study.

A complementary analysis was conducted to examine *when* daily peak demands occur and whether those hours coincide with high solar availability. To do this, daily peak hours

Table 4.1: *Technical and economic parameters used in the rural distribution utility case study.*

Parameter	Value
PV module cost	0.35 USD/W
Inverter cost	0.04 USD/W
Equipment / BOS cost	0.18 USD/W
Overhead cost	0.10 USD/W
Labor cost	0.10 USD/W
O&M cost	15 USD/kW-year
Transformer cost	150,000 USD
Inverter oversizing coefficient	1.2
Energy cost (purchase)	0.025 USD/kWh
Demand (power) charge	22 USD/kW
Federal tax credit (initial PV–battery)	30%
Initial battery cost	150 USD/kWh
Replacement battery cost	100 USD/kWh
Project lifetime	20 years
Discount rate	0.08
Battery round-trip efficiency	0.9025
Battery efficiency (discharge)	0.95
Battery utilization factor	0.7
Peak-flattening target	2000 kW

were extracted for every day of the three-year dataset and then grouped by month. The resulting distribution of peak hours is shown in Fig 4.8. In the winter months (December–February), the majority of daily peaks occurred in the early morning hours, when heating loads are high and solar irradiance is negligible. In the summer months (June–August), daily peaks shifted to the early evening window (approximately 18:00–20:00), driven largely by cooling loads and customer behavior. In both seasonal regimes, the dominant peak hours do *not* align with the hours of highest irradiance (typically around solar noon). This confirms the central motivation of this chapter: PV alone cannot reliably reduce the system peak in this service territory because the peak is temporally decoupled from solar production.

These empirical patterns justify the introduction of battery storage. By charging from surplus PV (midday in summer) or from the grid in off-peak hours (nighttime in winter), the battery can be dispatched during the empirically dominant peak windows identified in Figure 4.8.

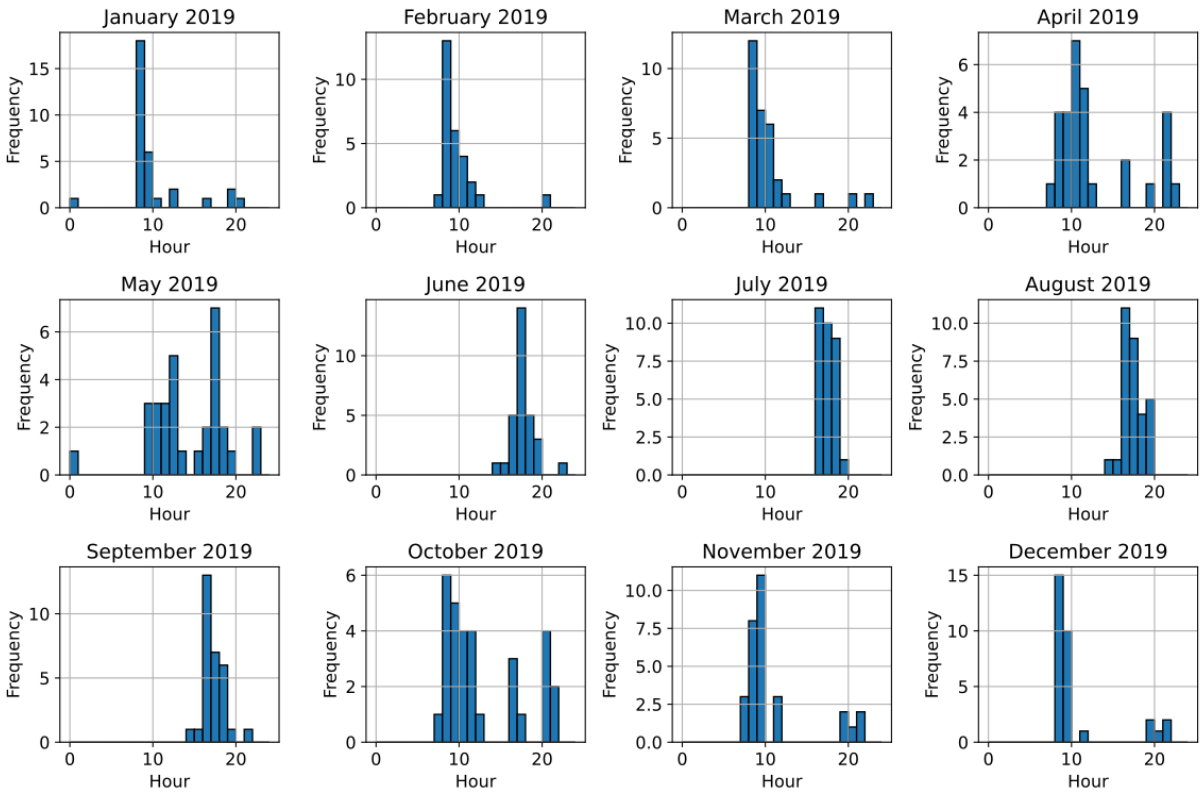


Figure 4.8: *Histogram of the daily peak demand hours for 2019*

Therefore, we considered PV system sizes ranging from $X_1 = 200$ kW up to $X_n = 10{,}000$ kW, and for each PV size we paired battery storage capacities from $Y_1 = 200$ kWh to $Y_m = 10{,}000$ kWh in increments of 100 kWh. For every candidate PV size, the modified daily load profile was obtained by subtracting the corresponding PV generation from the original hourly load profile for that day. On top of these modified profiles, the daily battery energy required to *flatten* (i.e., to hold the load at the balancing line introduced earlier) was computed day by day.

As an example, Fig 4.9 shows the resulting histogram for a 2,000 kW PV system. Each bar represents one day in the three-year dataset, and its height indicates the battery energy (kWh) that would have been needed on that day to flatten the PV-modified load. The histogram is clearly multi-peaked and fairly wide, which reflects substantial variability in both solar availability and load shape across the year. Some days require only a small amount of storage because midday PV aligns reasonably well with the load, while other days

require several megawatt-hours of storage because the peak occurs when PV output is low.

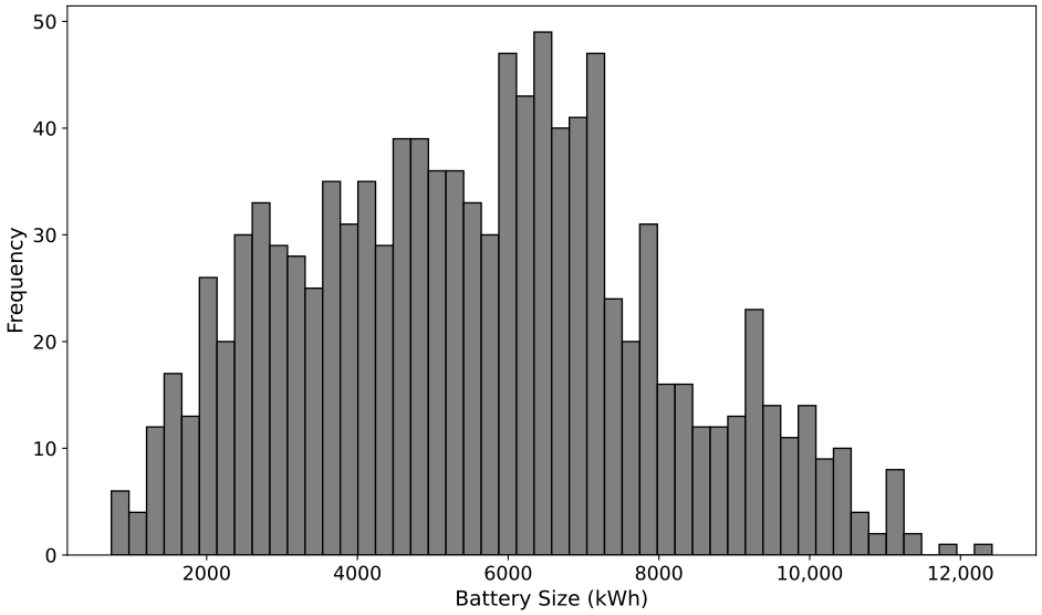


Figure 4.9: *Histogram of the required daily battery sizes that effectively flattened the load profiles for PV = 2000 kW.*

This spread of daily requirements makes it difficult to choose a single optimal battery size for the 2 MW PV case. A large battery (for example, 12,000 kWh) could cover nearly all days and achieve very good flattening performance, but it would be economically unattractive. A much smaller battery would cost less but would fail to flatten the load on the winter and cloudy days that populate the right tail of the histogram. The two sizing approaches presented in the next subsections are designed precisely to navigate this trade-off by targeting a high (95%) success probability rather than the single worst day. In the remainder of this chapter, the two proposed sizing approaches are applied to the dataset collected from a small town in Kansas, USA:

- **Approach 1** (PDF-based sizing): for every PV–battery pair, histograms of updated daily peaks are constructed and fitted with candidate PDFs (Gamma, Lognormal, Beta). The 95th-percentile value is extracted and compared with the target of 2000 kW. Pairs that satisfy the technical requirement are then sent to the benefit model.

- **Approach 2** (elbow-based sizing): for every PV size, the curve of "95% peak threshold

vs. battery size" is constructed. The battery size at the elbow point—where additional storage yields only marginal peak reduction—is selected as the technically meaningful size, and then evaluated economically.

4.7 Approach 1

Building upon the previous results of the required daily battery sizes for modified load-profile flattening, this section examines optimal sizing for PV–battery systems using a statistical (PDF-based) framework. For each PV and battery size combination, we first calculated the *updated* daily peak demands after PV–battery integration. Then, for every such PV–battery pair, we constructed scaled histograms of the updated daily peaks. These histograms capture, over the three-year dataset, how often a given updated peak value occurs when that specific PV size and that specific battery size are deployed.

Next, three probability density functions (PDFs) were fitted to every histogram: the Gamma distribution, the Log-normal distribution, and the Beta distribution. The intent was to determine which theoretical distribution best represents the empirical distribution of updated daily peaks so that percentiles (in particular, the 95th percentile) can be read off in a consistent manner. To assess the goodness-of-fit, the Kolmogorov–Smirnov (KS) statistic test was performed on all fitted PDFs.

The results revealed that the Gamma distribution fitted the updated daily peak demands most appropriately. Table 4.2 presents the KS-test results comparing the fits of the Gamma, Log-normal, and Beta distributions to the updated daily peak demands for a PV size of 2000 kW across various battery sizes. The KS statistics and p-values indicate that the Gamma distribution consistently offered a better fit for the data compared to the Log-normal and Beta distributions. This is evident from the generally lower KS statistic values and higher p-values for the Gamma distribution across all battery sizes. For instance, at a battery size of 2000 kWh, the Gamma distribution had a KS statistic of 0.031794 and a p-value of 0.65124, which are considerably more favorable than those for the Log-normal and Beta distributions, suggesting a more accurate and reliable fit.

Table 4.2: *KS statistic test results for the fitted PDFs for PV = 2000 kW and various battery sizes.*

Battery (kWh)	Log-normal		Gamma		Beta	
	KS_Statistic	p-Value	KS_Statistic	p-Value	KS_Statistic	p-Value
2000	0.041089	0.0480	0.031794	0.65124	0.080	0.005355
2500	0.035370	0.12579	0.029497	0.55618	0.110	0.027578
3000	0.040364	0.15470	0.026193	0.56416	0.090	0.005044
3500	0.026193	0.43233	0.023331	0.42345	0.120	0.042890
4000	0.062428	0.03700	0.027589	0.65164	0.160	0.014550
4500	0.039343	0.06544	0.034610	0.32136	0.130	0.004353
5000	0.046593	0.01660	0.030778	0.32103	0.140	0.000539
5500	0.051666	0.00554	0.062428	0.27564	0.150	0.000127
6000	0.054413	0.00292	0.034076	0.24565	0.250	5.58×10^{-5}

After identifying the Gamma distribution as the best-fitted PDF for the updated daily peak demands, the next step involved determining a 0.95 (i.e., 95%) threshold for every fitted Gamma PDF. For each fitted Gamma PDF, we calculated the value corresponding to the 95th percentile, effectively establishing a threshold that represents the maximum peak level that the PV–battery system is expected to flatten on 95% of the days.

For example, Figure 4.10 presents the fitted Gamma PDF for the updated daily peaks with a PV size of 2000 kW and a battery size of 3000 kWh. The Gamma distribution parameters were $\alpha = 26.805$ and $\beta = 57.668$, and the corresponding 95th-percentile threshold was 2041.996 kW. This means there is only a 5% chance of exceeding this peak demand when the system operates with PV = 2000 kW and battery = 3000 kWh.

For each combination of PV and battery, unique Gamma PDF parameters—including shape, scale, and the implied 95th-percentile threshold—were obtained in this manner. However, the utility's desired peak-demand threshold in this study was *exactly* 2000 kW with a 95% probability, indicating that the system should not exceed a 2000 kW peak demand more than 5% of the time. In many fitted cases, the 95th-percentile value from the empirical Gamma fit was *near* but not *exactly* 2000 kW (for instance, 2041.996 kW in the example above). Therefore, an additional recalibration step was introduced to enforce the desired technical requirement.

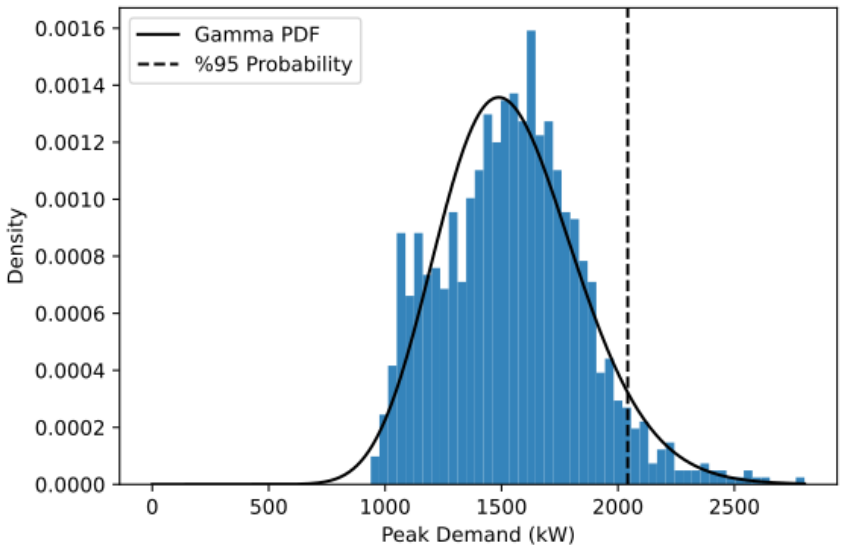


Figure 4.10: *Fitted Gamma PDF for updated daily peaks with PV = 2000 kW and battery = 3000 kWh. The vertical marker shows the 95th-percentile peak (2041.996 kW).*

Specifically, we recalibrated a Gamma distribution to meet this criterion by solving

$$F_\Gamma(T = 2000 \text{ kW}; \alpha, \beta) = 0.95, \tag{4.42}$$

where $F_\Gamma(\cdot)$ is the CDF of the Gamma distribution, and α and β are the shape and scale parameters, respectively. Since this constraint alone cannot uniquely determine both parameters, the previously fitted Gamma PDFs were used as reference points. For each fitted distribution, one parameter was held fixed while the other was determined numerically to satisfy (4.42). This procedure was repeated by alternately fixing the shape and scale parameters. The resulting parameter pairs were compared with the original fitted parameters using the sum of squared relative differences, and the pair with the minimum difference was selected.

For instance, with a PV size of 2000 kW, the recalibrated shape and scale parameters were found to be $\alpha = 25.99$ and $\beta = 57.92$, respectively. It is noteworthy that these recalibrated Gamma parameters aligned closely with the actual data-derived parameters ($\alpha = 26.805$, $\beta = 57.668$). This proximity suggests that the recalibrated parameters and

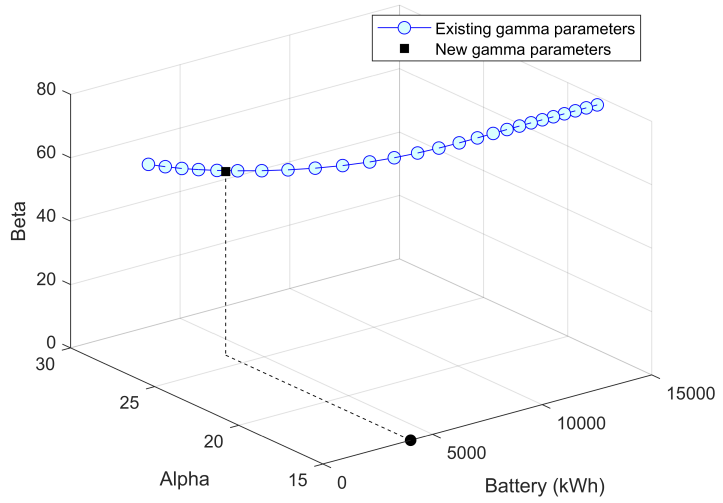


Figure 4.11: *Evolution of Gamma shape and scale parameters for PV = 2000 kW as the battery size increases. The recalibrated point that enforces* $F_\Gamma(2000\ kW) = 0.95$ *falls closest to the curve corresponding to a 4000 kWh battery.*

associated Gamma distribution are strongly representative of the actual daily peak-demand data, while now being exactly aligned with the utility's target of 2000 kW at the 95% probability level.

Figure 4.11 illustrates the shape and scale parameters of the fitted Gamma PDFs for a PV size of 2000 kW and various battery sizes. By aligning the newly recalibrated parameters on this curve, the corresponding optimal battery size was determined to be approximately 4000 kWh. In other words, for PV = 2000 kW, a battery of about 4000 kWh meets the estimated 95% peak-flattening criterion under the fitted model.

Accordingly, the proposed methodology performed well in determining the appropriate battery size to meet specific demand-flattening targets. By applying this method to the full range of PV sizes (from 200 kW to 10,000 kW), the optimal battery size for each PV configuration can be determined. Table 4.3 presents the optimal battery sizes for PV–battery systems that flatten 95% of daily peak demands to 2000 kW. An important observation from that table is that, for PV sizes between 500 kW and 1000 kW, optimal battery sizes were not defined—this implies that, within the study's tested battery range (up to 10,000 kWh), these PV systems did not admit a battery solution that could meet the 2000 kW-at-95%

criterion. As the PV system size increased to 1200 kW, a substantial battery capacity of about 9200 kWh was required, indicating a significant need for storage to effectively flatten peak demands for that relatively small PV size. As the PV system size continued to increase, there was a gradual *decrease* in the required battery size. This trend suggests that larger PV systems improve mid-day net-load relief and thus require relatively smaller batteries to achieve the same 95% peak-flattening objective.

This method uses three years of hourly data to obtain (i) empirical histograms of updated daily peaks, (ii) fitted Gamma PDFs for those histograms, and (iii) battery sizes for which the fitted probability of satisfying the utility's 2000 kW peak-demand target is at least 95%. These technically admissible PV–battery pairs are then passed to the benefit (economic) analysis to identify which configuration should be installed in practice.

4.7.1 Financial Analysis

The proposed statistical methodology produced a family of PV–battery configurations that can flatten 95% of the daily peak loads up to the technical target of 2000 kW. While satisfying this technical requirement is necessary, it is not sufficient for adoption by a small municipal utility; each technically feasible PV–battery pair must also be evaluated from an economic perspective over the full project horizon. This section presents the financial analysis framework and the results obtained for all optimal pairs.

To ensure a consistent economic basis, the three years of measured data (2019–2021) from a small town in Kansas, USA, were first grouped on a *monthly* basis. That is, the January load and irradiance data from 2019, 2020, and 2021 were combined to form a representative January; the same was done for February, and so forth through December. For every month, the average monthly peak demand charge and the average monthly energy cost were

Table 4.3: *Optimal battery sizes for PV–battery systems to flatten 95% of the daily peak demands to 2000 kW.*

PV (kW)	500	1000	1200	1500	2000	2500	3000	3500	4000
Battery (kWh)	—	—	9200	4400	4000	3800	3600	3400	3300

computed. These 12 representative months were then concatenated to form one synthetic year, and this year was projected over the 20-year project lifetime using the discounting procedure in Section 4.4 (i.e., present-worth of annual O&M, present-worth of demand and energy charges before and after PV–battery installation, and inclusion of transformer and replacement-battery costs).

Fig 4.12 summarizes the main economic components for all PV–battery combinations that satisfied the 95% peak-flattening criterion. As the PV nameplate size increases, the annual *energy* cost to the utility decreases, because larger PV systems offset more daytime energy. By contrast, the *battery* has only a limited effect on energy cost: it shifts energy in time but does not generate it. This is why, across the feasible set, the curves of energy cost are mainly separated by PV size, not by battery size.

A second observation from Figure 4.12 is that the monthly/annual *peak demand charge* shows only minor variation among the technically acceptable PV–battery configurations. This is expected: all of these configurations were constructed to flatten 95% of daily peaks to the same target (2000 kW), so their impact on the billed demand is nearly identical. Small differences are attributable to (i) the fact that the 5% tail of days is still supplied by the grid and (ii) small fitting errors introduced when the updated daily peaks were mapped to Gamma PDFs.

Equipment and capital costs, on the other hand, increase as PV size increases. Even though the optimal battery size tends to *decrease* for larger PV systems (see Table 4.3), the

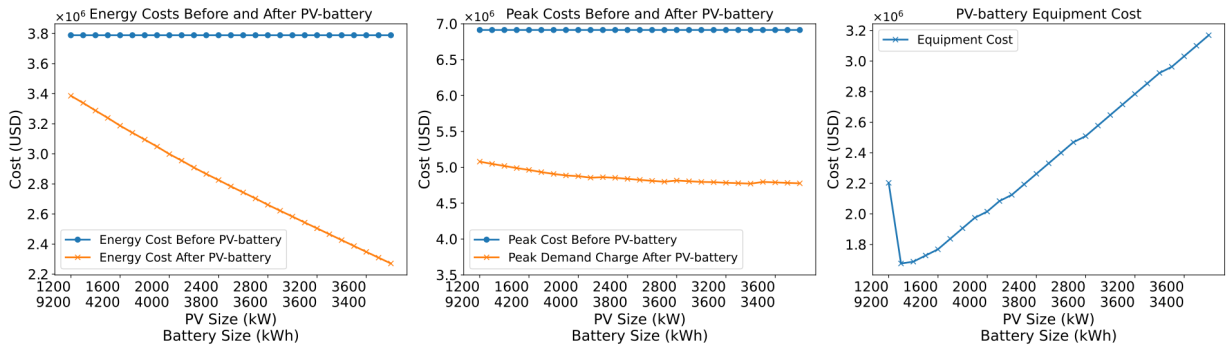


Figure 4.12: *Comparative financial impacts of PV–battery installations, showing energy, peak costs, and equipment costs for varying system sizes.*

increase in PV capacity dominates the mild battery reduction, so the total equipment cost rises with PV. One notable outlier is the case with 1200 kW PV and 9200 kWh battery: this point appears high on the cost curve because the battery is unusually large for that PV size. When the battery size drops in subsequent PV sizes, the equipment cost also drops.

To identify the single most attractive configuration, the present-worth benefit was computed for every PV–battery pair. The PV–battery size that yielded the *maximum* benefit was selected as the economically preferred solution. Figure 4.13 shows the 20-year benefit for all optimal pairs. A PV system of 2000 kW combined with a 4000 kWh battery delivered the highest benefit, equal to USD 812,648 over 20 years. This point is the intersection where (i) PV is large enough to capture significant energy savings, (ii) the battery is just large enough to meet the 95% peak-flattening requirement, and (iii) the combined capital cost is still moderate.

Table 4.4 reports the main economic quantities for the winning configuration. Two conclusions stand out. First, in the base year the total *peak-demand* cost of the feeder is nearly twice the energy cost, so any technology that credibly reduces billed demand has

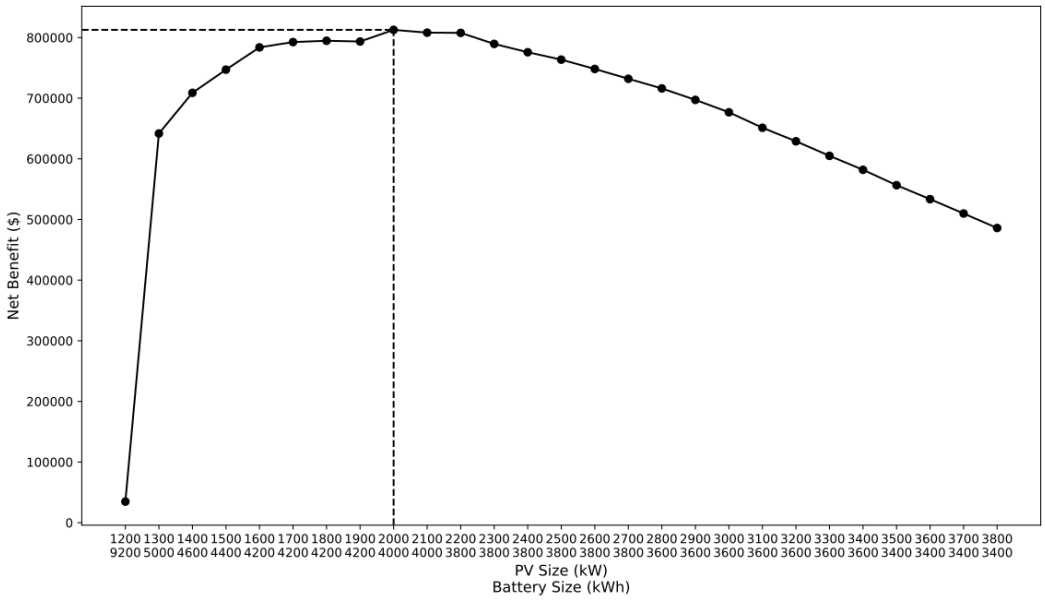


Figure 4.13: *Present-worth benefit over 20 years for all PV–battery combinations that satisfied the 95% peak-flattening requirement. The 2000 kW + 4000 kWh case produces the maximum benefit of USD 812,648.*

Table 4.4: *Economic comparison before and after PV and PV–battery installation (present-worth values over 20 years).*

	Before PV–Battery	PV Only	After PV–Battery
Equipment cost (USD)	0	1,638,688	2,015,246
Energy cost (USD)	3,788,907	3,036,927	3,023,569
Peak demand charge (USD)	6,913,926	6,472,805	4,901,679
Benefit (USD)	0	−445,587	812,648

leverage over the utility's bill. Second, installing *only* PV reduces the energy cost by about 19.8% but reduces peak-demand cost by only about 6%; the high up-front PV cost and yearly O&M make the PV-only option unattractive in net-present-value terms. In contrast, PV *plus* battery reduces peak demand by up to 29%, and even though the battery introduces about USD 25,000 of extra energy cost over the project life (due to round-trip losses), the demand-charge savings more than compensate, yielding the observed USD 812,648 benefit.

Sensitivity Analysis

The results above were obtained under a baseline set of economic assumptions: energy rate = USD 0.025/kWh, peak-demand charge = USD 22/kW, PV capital cost = USD 0.35/W, initial battery cost = USD 150/kWh, labor cost = USD 0.10/W, and tax credit = 30%. Because these parameters can vary across utilities and over time, a multi-parameter sensitivity analysis was carried out to test the robustness of the selected PV–battery sizes.

For each parameter, values above and below the baseline were tested while holding the other parameters fixed. The resulting benefits are shown in Fig 4.14. Each subplot corresponds to one parameter (energy rate, demand rate, PV cost, battery cost, labor cost, tax credit), and within each subplot the black triangle marks the maximum benefit for that scenario.

Three parameters showed a *direct* relationship with benefit: the peak-demand rate, the energy rate, and the tax-credit rate. As any of these increased, the 20-year benefit increased as well. This is intuitive: higher demand or energy rates magnify the dollar value of the technical reductions achieved by the PV–battery system; higher tax credits reduce the effective

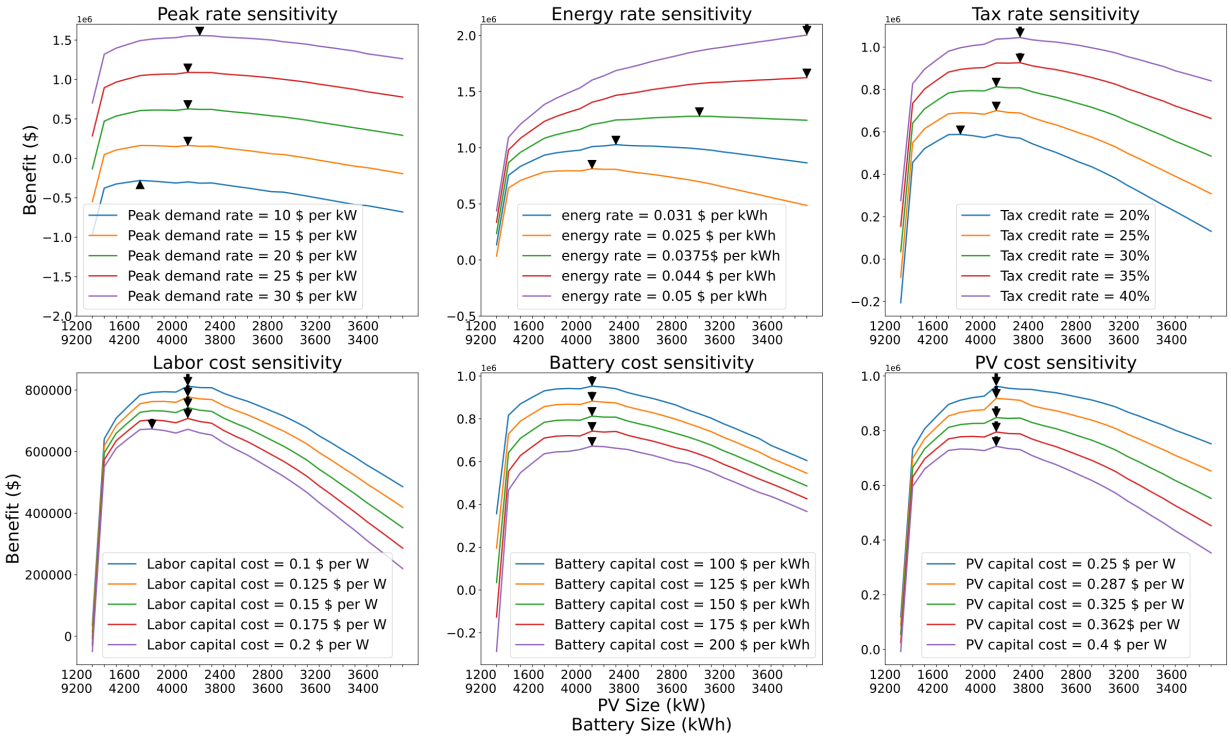


Figure 4.14: *Sensitivity analysis of the benefits with variation in the economic parameters. The black triangles indicate the maximum benefit for each scenario.*

up-front burden. Conversely, the capital costs of labor, PV modules, and batteries showed an *inverse* relationship with benefit: as these costs increased from baseline, the benefit curves shifted downward, but they did so fairly uniformly across the candidate PV–battery sizes. Because the downward shift was uniform, the *identity* of the optimal PV–battery pair did not change in most cases; only its benefit value changed. This is an important robustness property of the proposed statistical sizing method.

A noteworthy finding is that reducing the peak-demand rate to around USD 10/kW made all PV–battery options uneconomical (benefits became negative), even though technically they continued to flatten 95% of daily peaks. In contrast, increasing the *energy* rate above USD 0.0375/kWh not only increased the benefit but also made larger PV systems relatively more attractive, because PV directly offsets energy while batteries do not. Similarly, higher tax credits made larger PV–battery sizes feasible because the up-front cost penalty was mitigated.

Overall, the sensitivity study shows that:

- parameters that scale *operational savings* (energy rate, demand rate) and *capital relief* (tax credit) have the strongest impact on the benefit value and can, in some cases, shift the preferred PV size upward;
- parameters that scale *capital outlay* (PV cost, battery cost, labor cost) mainly shift the benefit curve up or down but do not dramatically change which PV–battery pair is best, because all technically feasible pairs were designed to flatten 95% of peaks to the same level.

Finally, because all of these economic results are still derived from historical 2019–2021 load and irradiance patterns, we validated the chosen PV–battery configuration under the modified Monte Carlo scenarios.

4.7.2 Modified Monte Carlo Simulation

In addition, to account for the inherent variability in load demand and solar irradiance, a modified Monte Carlo simulation was used to test the robustness of the optimal PV–battery configuration identified in the previous subsections (PV = 2000 kW, battery = 4000 kWh).

To enable realistic scenario generation, the three years of hourly data (2019–2021) were first grouped by month and clustered separately for load and for solar irradiance. For example, January contained 93 daily profiles (31 days $\times$ 3 years). After normalization to the $[0, 1]$ range, K-means clustering was applied with different numbers of clusters, and the within-cluster sum of squares (WCSS) was computed to form elbow curves. Figure 4.15 shows the elbow curve for the January load profiles. The reduction in WCSS was noticeable up to about five clusters; beyond that point, the improvement became marginal, so five clusters were retained for January. Repeating this process for the remaining months yielded comparable results for both load and irradiance.

Because cluster validation for time-series data is nontrivial, self-organizing maps (SOM) were also trained on the same monthly datasets to confirm the K-means outcome. The SOM results for January are shown in Figure 4.16. The SOM likewise identified five dominant groups of daily load shapes, consistent with the K-means elbow analysis. A few daily profiles

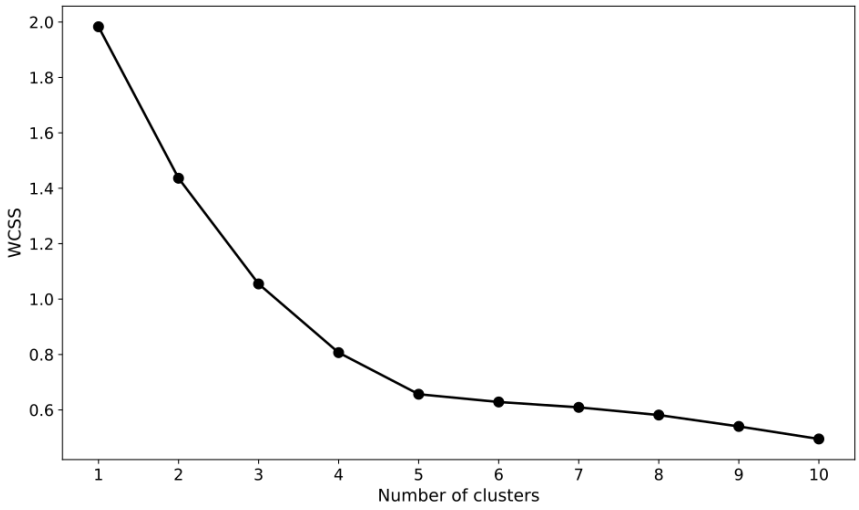


Figure 4.15: *Elbow curve of the K-means clustering for January load profiles.*

within a cluster exhibited minor hour-to-hour deviations (e.g., around hours 11–13 or 21–23), but the SOM did not split them into separate clusters, indicating that the overall daily pattern, rather than small local differences, dominates the grouping. This confirmed that the monthly clustering structure was stable enough to be used in scenario generation.

After monthly clusters for both load and irradiance were available, the modified Monte Carlo simulation generated 10,000 daily load–irradiance scenarios per month. These daily

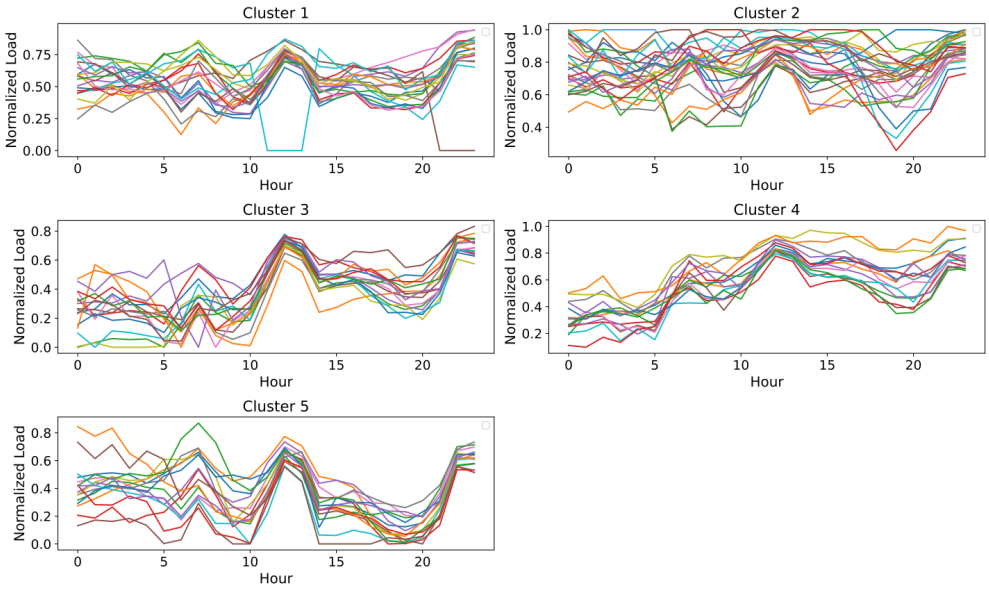


Figure 4.16: *Results of the January load-demand clustering using self-organizing maps.*

scenarios were then concatenated into 10,000 synthetic annual datasets. For the 20-year benefit analysis, 20 distinct daily pairs (one per year) were drawn without replacement from the synthesized dataset, and the full economic model (demand charge, energy charge, O&M, and discounted costs) was evaluated. This procedure was repeated 500 times, producing 500 independent 20-year benefit values for the (2000 kW, 4000 kWh) system.

The resulting benefit values are summarized in Figure 4.17, which shows the histogram of the 500 runs together with a fitted normal distribution. The average (mean) 20-year benefit was \$810,364, which is very close to the deterministic value reported earlier ($\approx$ \$812,000), indicating that the optimal system remains profitable under a wide range of plausible load–solar realizations. The standard deviation of the benefit samples was \$7,834.234, which implies a relatively tight spread around the mean. The first standard-deviation band, [802,529.82, 818,198.28], covered about 68% of all runs, and the second band, [794,695.59, 826,032.49], covered about 95% of the runs. Thus, even when interannual variability in both demand and irradiance is explicitly introduced, the vast majority of realizations produce a positive and practically similar benefit.

These results confirm that the PV–battery configuration selected in the previous section

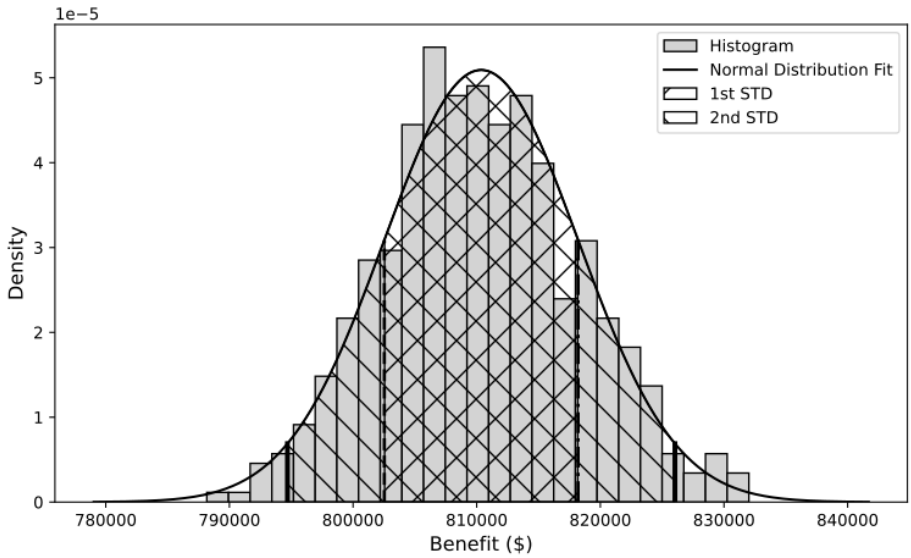


Figure 4.17: *Histogram and fitted normal distribution for 500 Monte Carlo runs of 20-year benefit for the optimal PV–battery system (PV = 2000 kW, battery = 4000 kWh).*

is not only optimal under the historical 2019–2021 dataset, but it also remains economically attractive when the same system is exposed to thousands of synthetic daily combinations that preserve the seasonal dependence between load and solar irradiance. At the same time, the analysis highlights the scope of the proposed framework: it is best suited for utilities (or large customers) that (i) pay a nontrivial demand charge, and (ii) are grid-connected so that the upper 5% of the daily peaks that are intentionally left unflattened can still be supplied by the upstream provider. Because all key economic inputs (energy rate, demand charge, tax credit, PV and battery unit costs) appear explicitly in the benefit model, the same validation procedure can be performed for other service territories by changing these inputs and re-running the Monte Carlo stage.

4.8 Approach 2

The second sizing strategy relies on the same underlying dataset and updated-peak construction as Approach 1, but it selects the battery size based on the shape of the 95% peak-threshold curve rather than using a predefined threshold of 2000 kW. The motivation is empirical: for a fixed PV size, increasing the battery capacity generally lowers the 95% peak threshold, but with *diminishing returns*. At some point, adding another 200–300 kWh of storage produces only a negligible improvement in the 95% threshold. Approach 2 identifies this point of diminishing returns as an *elbow* on the curve.

As discussed, the daily peak occurred in the early morning, when solar irradiance was zero. On such a day, PV installation alone cannot contribute to peak reduction, but a battery can discharge at the beginning of the day to shave the peak and then recharge around midday when PV becomes available. This same temporal misalignment appeared repeatedly in the three-year dataset, explaining why the 95% peak-demand thresholds do not decrease to zero even with relatively large battery capacities.

Using the three-year hourly dataset, modified load profiles were generated for PV nameplate capacities ranging from 200 kW to 10,000 kW in 100 kW steps. For every PV size, battery capacities from 1,000 kWh to 10,000 kWh were paired in 100 kWh steps, and for

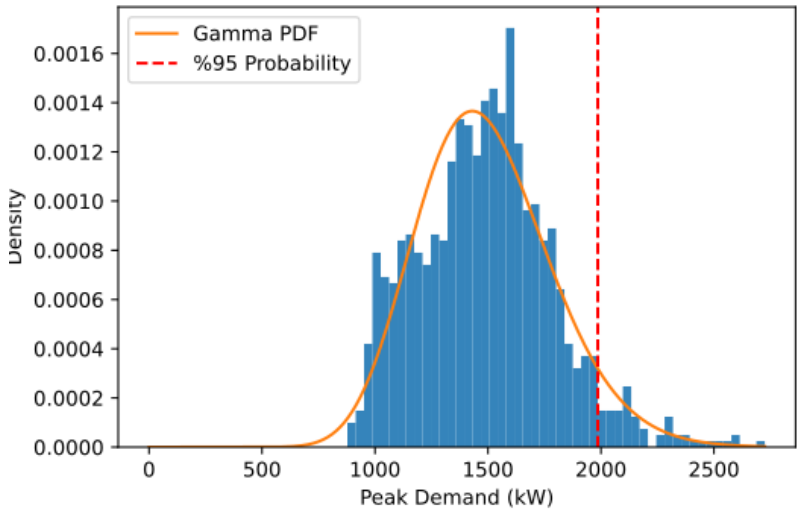


Figure 4.18: *Fitted Gamma PDF for updated daily peaks with PV = 2000 kW and battery = 4000 kWh. The dashed line marks the 95th percentile (1,985 kW).*

each PV–battery pair, the updated daily peak was computed for every day in the three-year period. As in Approach 1, scaled histograms of these updated daily peaks were formed, and Gamma and Lognormal PDFs were fitted. The goodness-of-fit was evaluated using the Kolmogorov–Smirnov (KS) test. The results indicated that the Gamma distribution provided a more consistent fit to the data, with p-values generally above the significance level of 0.05, while several Lognormal fits had p-values below 0.05. For example, for a PV size of 2000 kW and battery sizes of 2000 kWh and 3000 kWh, the Lognormal p-values were 0.0053 and 0.027, respectively. Therefore, the fitted Gamma PDFs were used to calculate the 95% peak-demand thresholds for different PV–battery combinations.

Fig. 4.18 shows the 95% peak threshold for a battery size of 4000 kWh and a PV size of 2000 kW using the corresponding fitted Gamma PDF. The peak-demand threshold with a 95% probability for the given case is 1,985 kW. This means that, with 95% probability, the updated daily peak demand is expected to be at or below 1,985 kW.

When these 95% thresholds are plotted versus battery size for a *fixed* PV size, a characteristic diminishing-return behavior appears. Fig. 4.19 shows this curve for the 2,000 kW PV case. The left part of the curve (up to about 4.4–4.8 MWh) is steep: each 100 kWh increase in

battery size yields a noticeable reduction in the 95% threshold. Beyond that point, the curve becomes much flatter, indicating that additional storage provides only small improvements in peak reduction.

Therefore, to determine the battery size for a PV size of 2000 kW, the elbow method is applied to the 95th-percentile peak-demand threshold curve shown in Fig. 4.19. Before the elbow point, increasing battery capacity provides substantial peak-demand reduction, whereas further increases beyond this point provide relatively small additional reductions. To identify the elbow numerically, a straight line is drawn between the first and last points of the curve, and the perpendicular distance from each point to this line is calculated. Figure 4.20 illustrates the resulting distances. The maximum distance occurs at a battery capacity of 4,600 kWh, identifying the elbow point for the 2000 kW PV system, with a corresponding 95th-percentile peak-demand threshold of 1,971 kW.

By repeating this procedure for all PV sizes between 200 kW and 10 MW, a set of PV–battery pairs was obtained, with each pair corresponding to the elbow point of its 95% peak-threshold curve. However, to determine the best PV and battery size, an economic

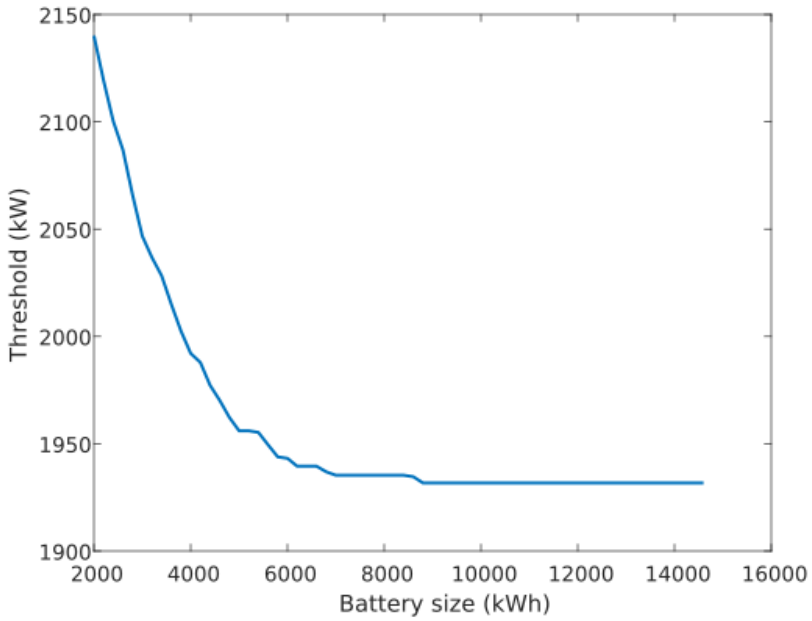


Figure 4.19: *Ninety-fifth percentile (95%) peak-demand thresholds versus battery size for PV = 2000 kW. A clear diminishing-return region appears beyond ∼4.6 MWh.*

benefit analysis must be conducted for all the selected combinations. The final PV–battery system size is determined by the combination that provides the highest overall economic benefit over the 20-year project lifetime.

4.8.1 Benefit Analysis

The benefit analysis was conducted to evaluate the financial profitability of the PV and battery size combinations obtained from the elbow-based sizing procedure over a 20-year planning horizon. Hourly load-demand and solar-irradiance data from 2019–2021 were first grouped on a monthly basis. For each month, energy charges and peak-demand charges were computed separately for all three years and then averaged to obtain representative monthly values. These monthly values were then annualized and discounted over 20 years to obtain the present-worth benefit for every PV–battery pair.

Table 4.5 reports the 20-year financial benefit for all optimal (elbow-based) PV–battery combinations. Figure 4.21 plots the same information; however, because of plotting limitations, the battery size corresponding to each PV size is not explicitly labeled in the figure.

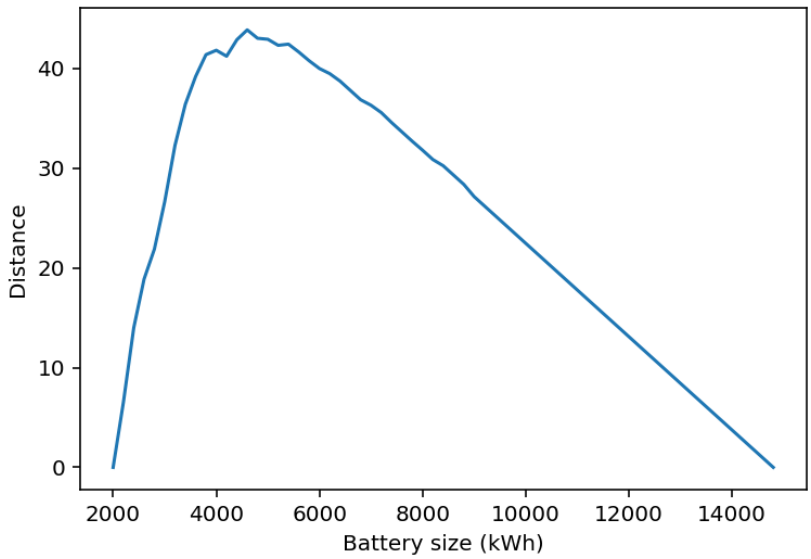


Figure 4.20: *Elbow identification for PV = 2000 kW using the maximum-distance method. The maximum distance occurs at a battery size of 4,600 kWh, with a corresponding 95% peak-demand threshold of 1,971 kW.*

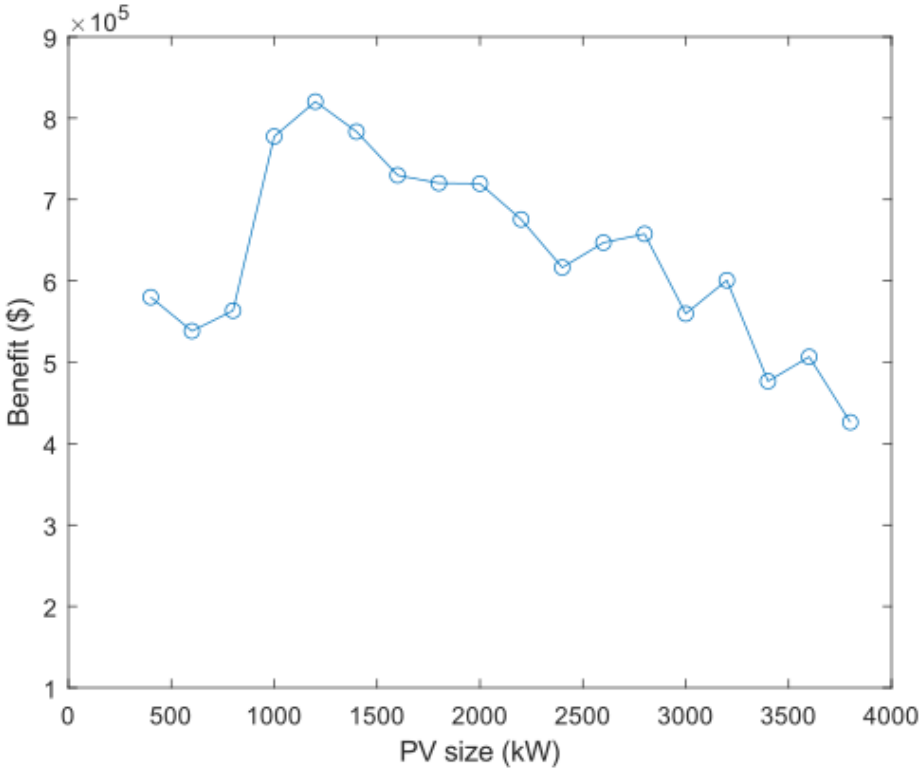


Figure 4.21: *Present-worth benefit over 20 years for elbow-based (Approach 2) PV–battery combinations. The maximum benefit ($ 820,373) occurs at PV = 1,200 kW and battery = 3,200 kWh.*

The best-performing configuration among all combinations was a PV size of 1,200 kW with a battery size of 3,200 kWh, which achieved a 95% peak threshold of 2,078 kW and a present-worth financial benefit of $820,373 over 20 years. Table 4.6 provides a detailed comparison of energy cost, peak-demand cost, and equipment cost for this optimal system, relative to the no PV–battery and PV-only cases.

Table 4.5: *Financial benefits for optimal PV–battery combinations obtained from Approach 2.*

PV (kW)	400	600	800	1000	1200	1400	1600	1800	2000
Battery (kWh)	6200	5500	5000	4800	3200	4200	4400	4600	4600
Benefit ($)	579,970	538,390	563,450	777,510	**820,373**	783,450	719,810	535,090	719,560
PV (kW)	2200	2400	2600	2800	3000	3200	3400	3600	3800
Battery (kWh)	5000	5400	5000	4600	4400	4200	4800	4400	4400
Benefit ($)	775,570	616,460	647,040	657,760	559,640	600,780	476,670	506,680	426,250

These results highlight several important observations. First, the original peak-demand cost (about $6.9 million, present worth) was almost twice the original energy cost (about $3.79 million), confirming that peak-demand charges dominate the utility bill and therefore must be directly targeted. Second, installing PV alone reduced energy cost by about 12.9% and peak-demand cost by only 5.58%, but the present-worth installation cost of $1,016,013 made the PV-only alternative uneconomic, with a net benefit of −$190,366. Third, adding the battery further reduced peak-demand charges to approximately $5.17 million, representing a 25.28% reduction from the baseline. Although battery charging and discharging increased energy costs by approximately $27,600 compared to the PV-only case, the demand-charge savings were sufficient to produce a positive overall benefit of $820,373. This confirms that, in this service territory and under the assumed tariffs, PV becomes economically attractive only when coupled with storage. Finally, to confirm that this benefit is not specific to the 2019–2021 weather–load combination, the optimal PV–battery system was later tested under clustered, Monte-Carlo–generated scenarios.

4.8.2 Modified Monte Carlo Simulation

To evaluate the robustness of the selected PV–battery configuration under variations in load demand and solar irradiance, the modified Monte Carlo simulation was applied using the methodology described in the previous section. For this approach, K-medoids with DTW and hierarchical clustering were employed to identify similar daily load-demand and solar-irradiance profiles.

Table 4.6: *Detailed economic comparison for the best Approach 2 configuration (PV = 1,200 kW, battery = 3,200 kWh). All values are present worth over 20 years.*

	Without PV–battery	**With PV only**	**With PV–battery**
Energy cost ($)	3,788,907	3,337,719	3,365,366
Peak cost ($)	6,913,926	6,539,467	5,166,301
Equipment cost ($)	0	1,016,013	1,350,792
Benefit ($)	–	−190,366	**820,373**

Time-Series Clustering

Three years of hourly load-demand and solar-irradiance data (2019–2021) were grouped by month and normalized before clustering. K-medoids with DTW was first applied to identify the appropriate number of clusters for each month. The gap statistic and Silhouette coefficient were used to evaluate the clustering results.

Fig. 4.22 presents the gap statistic and Silhouette coefficient for July load-demand profiles. Both metrics reach their maximum values at three clusters, indicating that three groups provide an appropriate representation of the daily load-demand patterns for July.

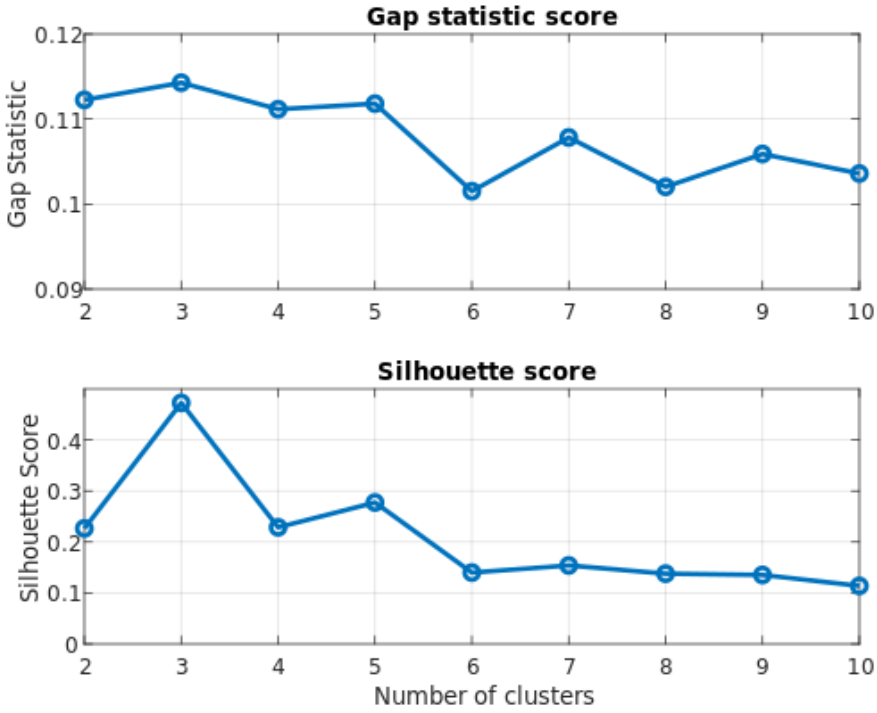


Figure 4.22: *Gap statistic and Silhouette coefficient for different numbers of clusters for July load-demand profiles.*

To further evaluate the clustering results, hierarchical clustering was applied using the same DTW distance matrix. Fig. 4.23 presents the resulting dendrogram for July load-demand profiles. By cutting the dendrogram at the level indicated by the red line, three main clusters were identified, consistent with the K-medoids results. The same clustering procedure was applied to the remaining months for both load demand and solar irradiance.

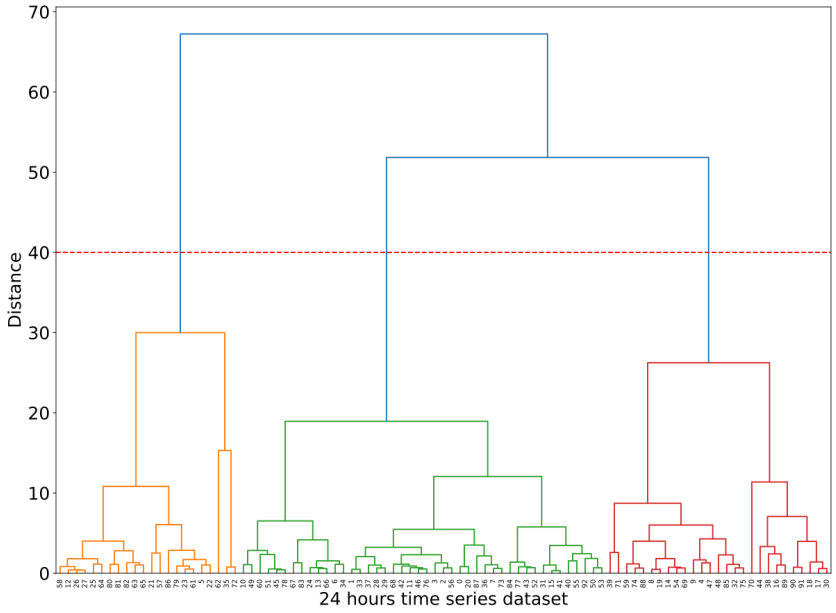


Figure 4.23: *Hierarchical clustering results for July load-demand profiles, identifying three main clusters.*

The resulting monthly clusters were subsequently used in the modified Monte Carlo simulation to generate paired daily load-demand and solar-irradiance scenarios while preserving their historical relationships.

Using the resulting monthly clusters, 10,000 synthetic annual scenarios were generated following the modified Monte Carlo simulation methodology. The selected PV–battery configuration from the elbow-based sizing approach, consisting of 1,200 kW PV and 3,200 kWh battery capacity with a 95% peak threshold of 2,078 kW, was evaluated under these scenarios. For the 20-year economic analysis, 20 annual scenarios were selected without replacement for each Monte Carlo run. This process was repeated 500 times, resulting in 500 samples of the 20-year economic benefit.

Fig. 4.24 presents the histogram of the 500 simulated benefit values along with the fitted Normal distribution. The mean 20-year benefit was \$822,440, with a standard deviation of \$7,554. The coefficient of variation was approximately 0.0092, indicating relatively low variability in the estimated economic benefit. The KS statistic for the fitted Normal distri-

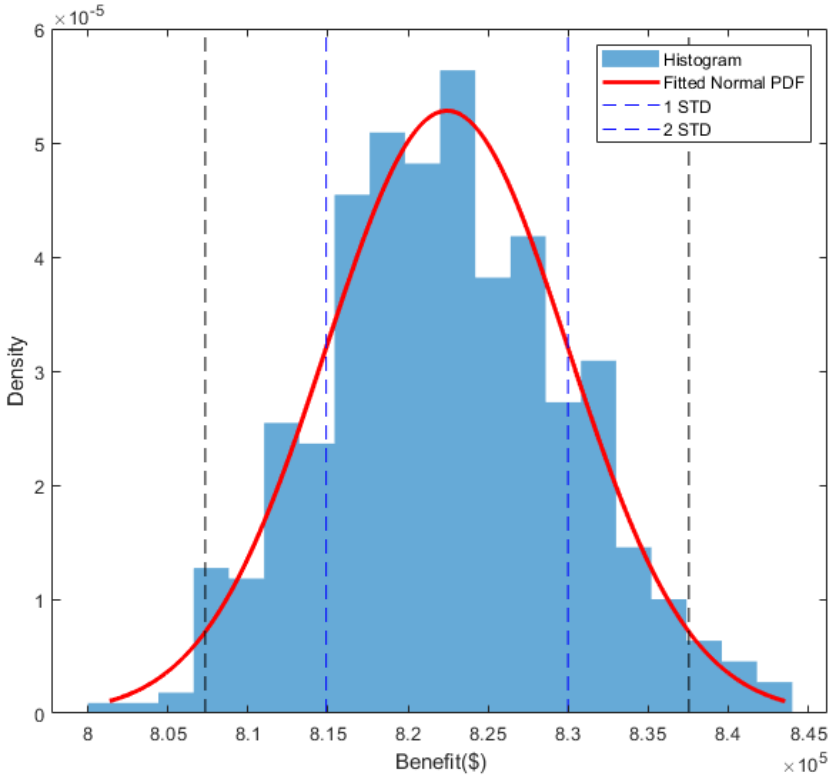


Figure 4.24: *Histogram of 500 Monte Carlo 20-year benefit samples for the selected PV–battery system (PV = 1,200 kW, battery = 3,200 kWh), with fitted Normal distribution.*

bution was 0.0235, with a p-value of 0.957, indicating that the Normal distribution provides a reasonable fit to the simulated benefit values.

The estimated 95% confidence interval for the mean 20-year benefit was \$821,773 to \$823,100. In addition, the first and second standard-deviation bands around the mean were approximately \$814,880 to \$829,990 and \$807,330 to \$837,550, respectively. Under the fitted Normal distribution, these ranges correspond to approximately 68% and 95% of individual benefit outcomes. These results indicate that the selected PV–battery configuration provides a positive and relatively stable estimated economic benefit under the simulated variations in load demand and solar irradiance. The relatively small variation in the 20-year benefit further demonstrates the robustness of the selected configuration under the considered operating and economic conditions.

4.9 Summary

This chapter developed and validated a statistically grounded framework for sizing grid-connected photovoltaic (PV) and battery energy storage systems in distribution-level applications where the utility is charged for both energy (kWh) and peak demand (kW). Our observation shows daily system peaks tended to occur either in the early morning during winter or in the early evening during summer, whereas solar irradiance peaked around noon. As a result, PV alone, even at relatively large penetrations, could not reliably reduce the billed demand. Battery storage was therefore introduced as a complementary technology capable of shifting either surplus PV or off-peak grid energy to those empirically dominant peak hours, with the goal of reducing the 95th-percentile daily peak-demand threshold. Approach 1 used a utility-specified threshold of 2000 kW, whereas Approach 2 selected battery sizes using the elbow of the peak-threshold curve.

The analysis was based on three full years (2019–2021) of hourly load demand and solar irradiance data. These data enabled two key advances. First, they allowed realistic construction of *modified* load profiles for a wide range of candidate PV nameplate sizes (200 kW to 10 MW), obtained simply by subtracting the hourly PV generation from the native load. Second, they revealed substantial day-to-day variability in the battery energy required to flatten those modified load profiles. Even for a single PV size (e.g., 2 MW), the histogram of required daily battery sizes was wide and multi-peaked; some days could be flattened with a small battery because solar output happened to align with load, while winter or cloudy days demanded several megawatt-hours of storage. This dispersion is the central practical difficulty: a battery large enough to flatten the single worst day would be uneconomic, while a battery sized for the median day would fail too often. The chapter therefore adopted a probabilistic design target: flatten 95% of days, and serve the remaining 5% from the grid.

Two complementary sizing strategies were proposed. Approach 1 treated the problem as given a peak limit (2000 kW), find all PV–battery pairs whose updated daily peaks stay below this limit with 95% probability. For each PV–battery combination on the discrete sizing grid, the updated daily peaks over the three-year dataset were collected, a scaled histogram

was formed, and three candidate probability density functions (Gamma, Lognormal, Beta) were fitted. A Kolmogorov–Smirnov goodness-of-fit test showed that, across battery sizes, the Gamma distribution consistently yielded lower KS statistics and higher p-values than the alternatives; the updated peaks are positive, right-skewed, and fairly well described by two-parameter Gamma laws. Once the best-fit Gamma PDF was identified, its 95th percentile was computed and compared with the utility target of 2000 kW. In many cases, the empirical 95th percentile was close to but not exactly 2000 kW, so a recalibration of the Gamma parameters was carried out. Plotting the resulting Gamma parameters versus battery size for a fixed PV size made it possible to read off the battery capacity that met the 2000 kW-at-95% requirement. Applying this procedure to PV sizes from 200 kW to 4 MW produced a set of technically admissible pairs. An important empirical observation was that very small PV systems (500–1000 kW) did not admit a feasible battery within the tested range (up to 10 MWh) that could push 95% of peaks down to 2000 kW; by contrast, once PV reached 1.2 MW and above, feasible battery sizes emerged and, in fact, decreased gradually with further PV growth.

Approach 2 addressed a slightly different planning question: for a given PV size, how large should the battery be before the marginal reduction in the 95% peak becomes negligible? Using the same three-year updated-peak data, 95th-percentile peak thresholds were computed for battery sizes from 1 MWh to 10 MWh. When plotted, these curves showed a characteristic structure: a steep decline in the 95% threshold for initial increases in battery size, followed by a clearly flatter region. The battery size at the point of maximum curvature (the "elbow") was declared technically optimal for that PV size. Repeating this across PV sizes yielded another set of PV–battery pairs, each one "lean" in the sense that additional storage would buy only marginal peak improvement. This second approach is attractive to utilities that prefer not to impose a hard peak limit in advance but do want to avoid oversizing the battery for small incremental gain.

Both technical approaches were then passed through a uniform **benefit analysis** over a 20-year horizon. The analysis showed several consistent patterns. First, peak-demand charges were almost twice as large, in present-worth terms, as energy charges, confirming that

demand reduction is the dominant economic lever for this utility. Second, a PV-only solution *always* reduced energy cost but reduced peak cost only modestly (5–6%), and the capital cost of PV alone was high enough to yield a negative net benefit. Third, adding storage unlocked much larger peak-demand savings (up to $\approx$ 29% relative to the baseline), and although battery operation introduced a small additional energy cost (due to round-trip losses), the demand savings more than compensated. For the PDF-based (Approach 1) solutions, the best-performing configuration was a PV of 2000 kW plus a battery of 4000 kWh, with a present-worth benefit of 812,648 USD over 20 years. For the elbow-based (Approach 2) solutions, the best configuration was PV = 1200 kW and battery = 3200 kWh, with a present-worth benefit of 820,373 USD. In both cases, the economically preferred configuration balances energy and peak-demand savings against PV and battery costs, although the two approaches use different technical sizing criteria.

To test robustness, the chapter concluded with a modified Monte Carlo validation. Because it is unrealistic to pair, for example, a winter load with a summer irradiance profile, the historical data were first clustered monthly for both load and solar (using DTW-based K-medoids and hierarchical/SOM confirmation). These clusters were then used to generate 10,000 realistic daily load–solar pairs per month, which were stitched into 10,000 synthetic annual datasets. For each Monte Carlo run, 20 distinct years were sampled (representing a 20-year project) and the full benefit model was evaluated. Repeating this 500 times produced a distribution of 20-year benefits. For the optimal system, the mean benefit was on the order of 8.1–8.2$\times 10^5$ USD, with a standard deviation of only $\sim 7.5 \times 10^3$ USD and a KS p-value near 0.95 for a Normal fit. Thus, 68% of the realizations fell in a narrow 15–16 kUSD band around the mean and 95% fell in a still moderate band, indicating that interannual variability in load and irradiance does not materially erode the economic attractiveness of the selected PV–battery system. It is important to note that the two proposed sizing approaches yielded relatively close economic benefits, although the selected PV and battery capacities were different. One reason for this similarity is that the predefined peak-demand threshold of 2000 kW in Approach 1 was close to the peak-demand threshold obtained using the elbow-based method in Approach 2. Although both approaches yielded comparable eco-

nomic benefits, the selected PV–battery configurations are not necessarily globally optimal. This is because the proposed sizing procedures are heuristic and evaluate PV and battery capacities within predefined ranges and at discrete intervals. Consequently, configurations with higher economic benefits may exist between the evaluated points. Nevertheless, the proposed approaches provide practical sizing solutions that satisfy their respective technical criteria and offer favorable economic performance.

Chapter 5

Electricity demand forecasting

Accurate electricity demand forecasting is important for distribution system state estimation (DSSE), particularly when advanced metering infrastructure (AMI) measurements are unavailable at the time of estimation. Residential demand varies with weather, occupant behavior, and daily and weekly consumption patterns. Historical averages may therefore provide inaccurate pseudo-measurements when current operating conditions differ from past observations.

This chapter develops an electricity demand forecasting framework using AMI data. The proposed Seasonal Periodic Decomposition Network (SPDNet) combines seasonal-trend decomposition and periodical decomposition to model variations in residential electricity demand. The forecasting performance is evaluated using profiles of individual customer loads and loads aggregated at the distribution transformer. Forecasting errors are also analyzed to estimate the uncertainty associated with forecast-based pseudo-measurements. The use of these pseudo-measurements in DSSE is presented in Chapter 6.

5.1 Proposed Forecasting Module

Electricity demand forecasting is essential for operation, control, and planning in distribution networks. Fig 5.1 shows a representative one-week AMI load profile for a single customer.

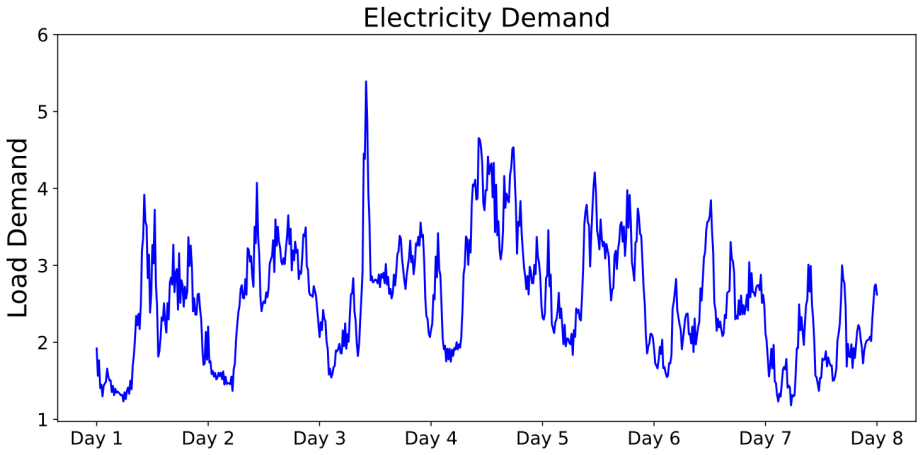


Figure 5.1: *Example of one-week residential electricity load demand.*

Four properties are immediately visible: (i) sharp, short-duration spikes caused by appliance usage; (ii) daily periodicity; (iii) weekly seasonality (weekdays versus weekend); and (iv) slow, nonstationary drift due to weather and occupant behavior. A good forecasting model for distribution-level applications should be able to represent these four characteristics simultaneously, because DSSE later relies on these forecasts as pseudo-measurements.

In recent years, different forecasting models have been developed to improve the accuracy of electricity demand prediction. These models range from traditional statistical approaches to advanced deep learning architectures, each with its own advantages and limitations. Traditional methods, such as autoregressive integrated moving average (ARIMA), seasonal ARIMA, and exponential smoothing, have been widely used to model historical trends, seasonality, and temporal correlations [64]. These methods are relatively simple, computationally efficient, and interpretable. However, their ability to capture nonlinear relationships and sudden variations in electricity demand is limited, particularly for individual residential customers.

Machine learning approaches, including multilayer perceptrons (MLPs), provide greater flexibility in modeling nonlinear relationships between historical observations and future demand. Nevertheless, conventional MLPs do not explicitly maintain temporal memory, and their ability to represent sequential dependencies depends on the input structure. Recurrent neural networks (RNNs), particularly long short-term memory (LSTM) and gated recur-

rent unit (GRU) models, were subsequently adopted to capture temporal dependencies in electricity consumption [114]. Although these models can represent nonlinear sequential patterns, their sequential computation can increase training time, and capturing long-range dependencies remains challenging.

Convolutional neural networks (CNNs) and temporal convolutional networks (TCNs) have also been investigated for electricity demand forecasting. These models use convolutional operations to extract local temporal patterns and can process input sequences in parallel [82, 85]. However, their ability to capture long-range relationships depends on the receptive field and network architecture. More recently, Transformer-based models have received considerable attention because their attention mechanisms can represent dependencies across extended historical sequences. Models such as PatchTST and iTransformer have demonstrated the potential of attention-based representations for time-series forecasting [115, 116]. Despite these advantages, Transformer architectures may require substantial computational resources, particularly for long input sequences.

Recent forecasting studies have further emphasized the importance of explicitly modeling temporal structure rather than relying only on network complexity. For example, DLinear incorporates seasonal-trend decomposition into a simple linear forecasting framework [117], while TimesNet and Times2D use two-dimensional representations to capture periodic variations in time-series data [59, 118]. These developments indicate that decomposition and periodic representations can help forecasting models identify recurring patterns while accounting for changes across different time scales.

Table 5.1 summarizes the main advantages and disadvantages of the forecasting model families discussed above.

Fig 5.2 sketches 24-hour-ahead forecasts produced by five common model families: (a) traditional time-series models (ARIMA and exponential smoothing), (b) multilayer perceptrons (MLPs), (c) recurrent models (LSTM/GRU), (d) convolutional models (1D or temporal CNN), and (e) Transformer-based models. Traditional models can represent regular trends and seasonality but may fail to capture nonlinear variations and sudden demand spikes. MLPs can learn complex mappings but depend on the historical input representa-

Table 5.1: *Advantages and disadvantages of electricity demand forecasting models.*

Model	**Advantages**	**Disadvantages**
Traditional	Simple, interpretable, and captures trends and seasonality	Limited nonlinear modeling capability
MLPs	Learns complex nonlinear relationships	No explicit sequence memory
Recurrent models	Captures sequential and temporal dependencies	Sequential computation and long-range dependency challenges
Convolutional	Captures local patterns with parallel computation	Global dependencies require larger receptive fields
Transformers	Captures long-range temporal dependencies	Potentially high computational and memory costs
Decomposition	Explicitly captures trends, seasonality, and periodic patterns	Performance depends on the decomposition strategy

tion to capture temporal relationships. Recurrent models capture sequential dependencies, although modeling long-range relationships and abrupt local changes can remain challenging. Convolutional models are effective at extracting local dependencies, but representing global patterns requires an appropriate receptive field. Transformers can capture broader temporal relationships, although this capability may come with increased computational complexity.

These limitations highlight the importance of developing forecasting models that can represent different temporal characteristics of residential electricity demand. In particular, individual AMI load profiles exhibit both regular patterns, such as daily and weekly periodicity, and irregular variations caused by changes in customer behavior and external conditions. While existing models can capture some of these characteristics, representing seasonal trends, periodic variations, and short-duration fluctuations within a unified framework remains an important consideration.

Motivated by these observations, this dissertation develops the Seasonal Periodic Decomposition Network (SPDNet), which combines seasonal-trend decomposition and periodic decomposition to model the temporal characteristics of residential electricity demand.

SPDNet consists of two main modules: the Seasonal-Trend Decomposition Module (STDM) and the Periodical Decomposition Module (PDM). The STDM decomposes the input data into trend, seasonal, and residual components to capture underlying temporal patterns effectively. The PDM transforms 1D inputs into 2D representations through frequency domain analysis. Utilizing the Fast Fourier Transform (FFT), the PDM identifies k dominant frequency components and their corresponding periods within the input data. Each pair is then used to reshape the 1D series into a corresponding 2D tensor. The two temporal dimensions of each tensor describe positions within a period and repetitions across periods. As a result, the input 1D time series is transformed into k distinct 2D tensors, each corresponding to one of the identified dominant periods. The PDM includes three submodules that are applied in parallel to each 2D tensor iteratively: (1) a short-term modeling submodule empowered by 1D convolutions to capture sharp fluctuations within short periods, (2) a Transformer-based

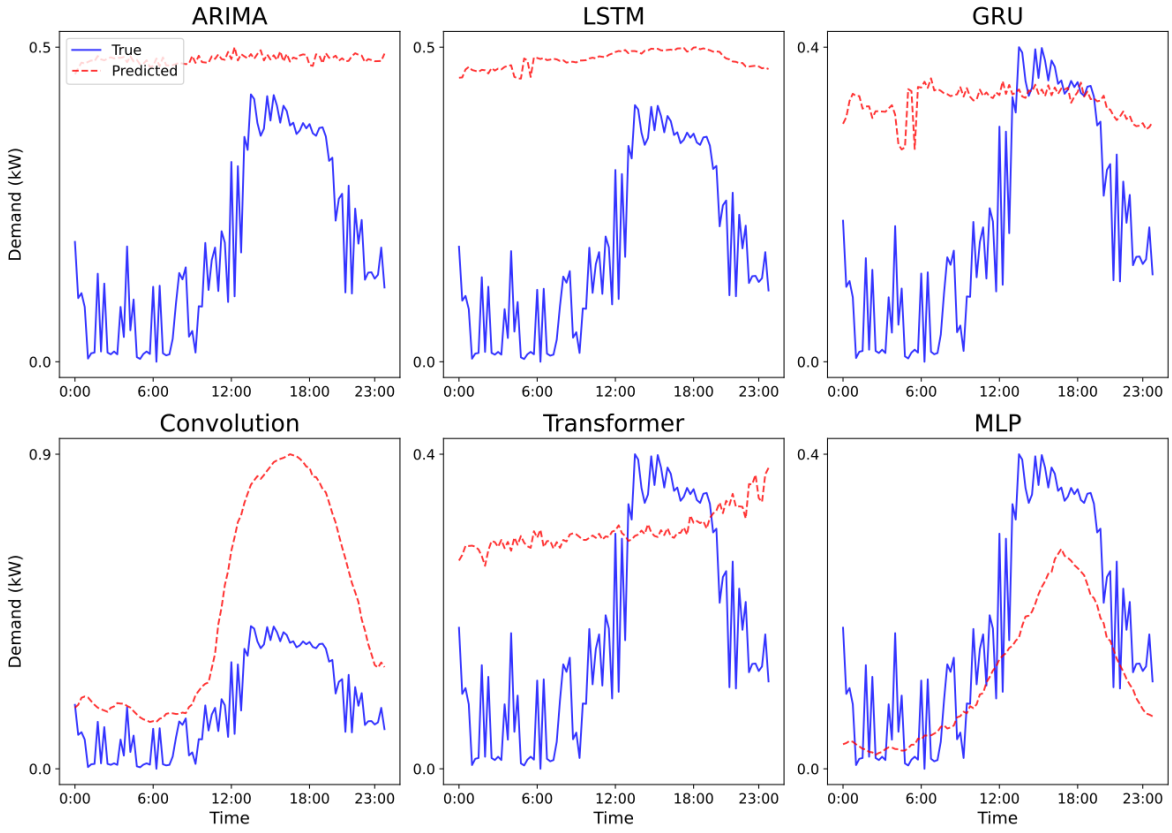


Figure 5.2: *Illustrative 24-hour-ahead forecasts from different model families.*

encoder to model global patterns within periods, and (3) a 2D convolutional submodule to discover correlations between and within periods. Fig 5.3 presents the systematic workflow of the proposed model.

5.1.1 SPDNet Methodology

The following subsections outline the steps involved in the proposed SPDNet for electricity demand forecasting.

Data Preparation and Reshaping

Electricity demand data are often used in conjunction with auxiliary variables to enhance forecast accuracy. This dataset is represented as $X = [x_1, x_2, \ldots, x_T] \in \mathbb{R}^{T \times N}$, where T denotes the length of the time series and N represents the number of variables. For effective analysis using deep learning models, the dataset must be properly formatted by defining the input and output sequences. First, a segment of the time series, referred to as the sequence length S, is selected as the input. A subsequent segment of the time series, referred to as the prediction length P, is selected as the output. By following this approach, the original data of total length T is transformed into multiple rows, where each row contains S input time points and P corresponding output. Additionally, the data is organized into batches for parallel processing during training. The samples are organized into batches of size B, resulting in input tensors with dimensions $[B, S, N]$ and output tensors with dimensions

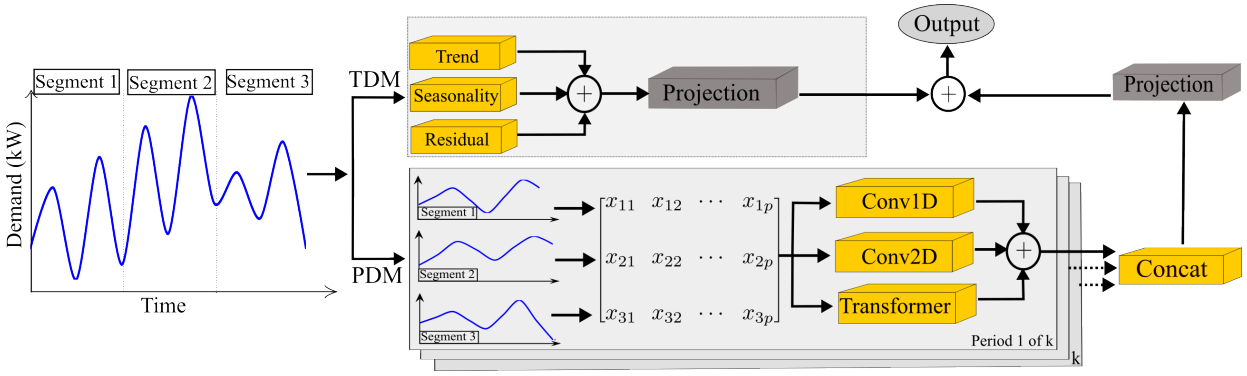


Figure 5.3: *Systematic workflow of the proposed SPDNet forecasting module.*

$[B, P, N]$.

Seasonal-Trend Decomposition Module

STDM is designed to extract the trend, seasonality, and residual components from electricity demand data. Given an input tensor $X \in \mathbb{R}^{B \times S \times N}$, the decomposition is performed as follows:

$$X_{\text{trend}}, X_{\text{seasonal}} = \text{STDM}(X) \tag{5.1}$$

The trend component X_{trend} is extracted using a one-dimensional convolution that captures the long-term trends in the time series. The trend is computed as

$$X_{\text{trend}} = \text{Conv1D}_{\text{trend}}(X) \tag{5.2}$$

Here, the $\text{Conv1D}_{\text{trend}}$ employs a large kernel size, which is suitable for capturing long-term trends. The convolution is performed along the temporal dimension S. The seasonality component captures repeating patterns that occur at specific intervals. It is extracted by applying a convolution to the detrended data, which is the original data minus the trend component, as follows:

$$X_{\text{seasonal}} = \text{Conv1D}_{\text{seasonal}}(X - X_{\text{trend}}) \tag{5.3}$$

In this case, the convolution operation $\text{Conv1D}_{\text{seasonal}}$ employs an appropriate kernel size, typically smaller than the trend kernel. The residual component X_{residual} represents the irregularities or noise that are not explained by the trend or seasonality components. It is computed by subtracting the extracted trend and seasonality components from the original input:

$$X_{\text{residual}} = X - X_{\text{trend}} - X_{\text{seasonal}} \tag{5.4}$$

The trend, seasonal, and residual components satisfy the following reconstruction identity:

$$X_{\text{STDM}} = X_{\text{seasonal}} + X_{\text{trend}} + X_{\text{residual}} \tag{5.5}$$

Since these components sum to the original input by construction, the summation alone does not create a new representation. The individual components must undergo additional processing if the decomposition is to change the learned features. To prepare the representation $X_{\text{STDM}} \in \mathbb{R}^{B\times S\times N}$ for prediction, the model employs a linear projection layer that transforms the sequence length S of the input data into the prediction length P required for forecasting as

$$\hat{X}_{\text{STDM}} = \text{LinearProjection}(X_{\text{STDM}}) \tag{5.6}$$

where $\hat{X}_{\text{STDM}} \in \mathbb{R}^{B\times P\times N}$ denotes the processed representation obtained from the STDM, which serves as the input to subsequent layers for the final forecasting.

Periodical Decomposition Module

Parallel with the STDM, the PDM is designed to extract periodic patterns from the data. The objective is to break down the 1D input tensor $X \in \mathbb{R}^{B\times S\times N}$ into multiple 2D periodic components to capture short-term variations, inter and intra-period patterns, and long-term behaviors across different periods. To achieve this, FFT is applied along the temporal dimension S of the input tensor $X \in \mathbb{R}^{B\times S\times N}$ for each batch and feature to identify the dominant periods as follows:

$$X_f = \sum_{t=0}^{S-1} X(t) e^{-2\pi i t f/S} \tag{5.7}$$

Here, X_f represents the transformed tensor in the frequency domain and f is the frequency. The magnitude A_f can be calculated as

$$A_f = |X_f| = \sqrt{\text{Re}(X_f)^2 + \text{Im}(X_f)^2} \tag{5.8}$$

For a real-valued time series, the Fourier coefficients have conjugate symmetry. Therefore, only non-negative frequencies need to be considered for identifying periodic components. The zero-frequency term represents the mean of the sequence and is excluded when selecting dominant periods.

To identify the dominant frequencies, the Fourier magnitudes are averaged over the batch and variable dimensions. The remaining frequency components are sorted in descending order of magnitude, and the first k nonzero frequency indices $f_1, f_2, \ldots, f_k$ are selected. The corresponding period lengths are obtained approximately as

$$p_i = \left\lceil \frac{S}{f_i} \right\rceil, \qquad i = 1, \ldots, k. \tag{5.9}$$

For each selected period, the sequence is padded when necessary to a length divisible by p_i. Let $n_i = \lceil S/p_i \rceil$ denote the number of period blocks after padding. The input is then arranged into a two-dimensional representation $X^i_{\text{2D}} \in \mathbb{R}^{B \times p_i \times n_i \times N}$, where p_i denotes positions within a period and n_i denotes the number of period blocks. The k resulting tensors represent the dominant periodic structures in the input sequence. Figure 5.4 illustrates this procedure.

Short-term Analysis

For short-term cyclic behavior analysis, the decomposed 2D tensors X^i_{2D} are processed to capture sharp fluctuations within the identified periods efficiently as the following step:

Reshaping: Given an input tensor $X^i_{\text{2D}} \in \mathbb{R}^{B \times p_i \times n_i \times N}$, it is reshaped by merging the B and N dimensions, resulting in $X^i_{\text{reshaped}} \in \mathbb{R}^{(B \times N) \times p_i \times n_i}$ as

$$X^i_{\text{reshaped}} = \text{Reshape}(X^i_{\text{2D}}, [B \times N, p_i, n_i]) \tag{5.10}$$

This reshaping is necessary to facilitate efficient convolutional operations over the time dimension.

1D Convolution: A one-dimensional convolution is then applied over the temporal di-

mension of the reshaped tensor X^i_{reshaped}. This operation is designed to capture short-term patterns or fluctuations within each identified period as

$$X^i_{\text{short}} = \text{Conv1D}(X^i_{\text{reshaped}}) \tag{5.11}$$

Projection: After extracting short-term patterns, the output tensor is projected to the desired prediction length P. A linear projection layer is employed to prepare the feature representations for forecasting as

$$\hat{X}^i_{\text{short}} = \text{LinearProjection}(X^i_{\text{short}}) \tag{5.12}$$

where $\hat{X}^i_{\text{short}}$ with shape $[B, P, N]$, denotes the short or local extracted features, which serve as the input to subsequent layers.

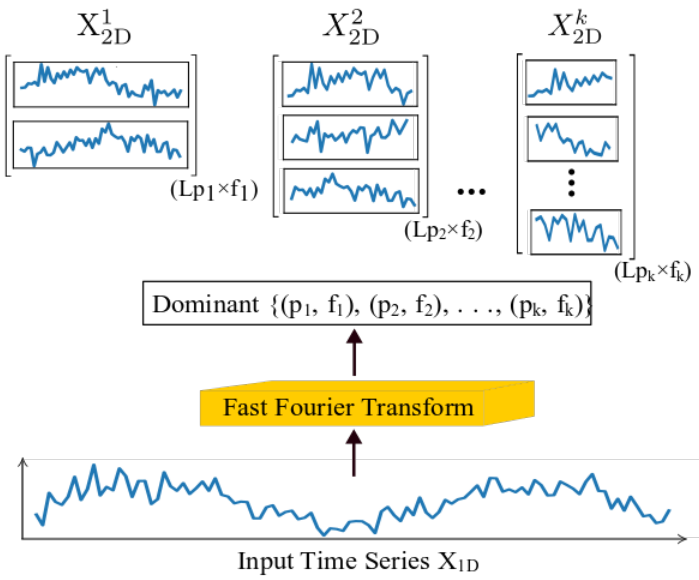


Figure 5.4: *Overview of the Periodical Decomposition Module (PDM) Framework to reshape the 1D input tensor into k distinct 2D tensors.*

Periodic Dependency Analysis

In parallel with the short-term analysis, the periodic dependency analysis module processes the 2D tensors X^i_{2D} to capture both intra-period and inter-period dependencies simultaneously. The following steps outline the procedure:

Reshaping: The periodic tensors $X^i_{\text{2D}} \in \mathbb{R}^{B \times p_i \times n_i \times N}$ undergo a two-step reshaping process. Initially, similar to the reshaping step in the short-term analysis, the B and N dimensions are merged, forming X^i_{reshaped}. Subsequently, an extra dimension is introduced to facilitate two-dimensional analysis as

$$X^i_{\text{reshaped}} = \text{unsqueeze}(X^i_{\text{reshaped}}) \tag{5.13}$$

This process produces a tensor with dimensions $[B \times N, 1, p_i, n_i]$. By adding this extra dimension, the 2D convolution is able to simultaneously capture intra-period dependencies across the p_i dimension and inter-period dependencies across the n_i dimension.

2D Convolution: A two-dimensional convolution is applied to the reshaped tensor $X^i_{\text{reshaped}} \in \mathbb{R}^{(B \times N) \times 1 \times p_i \times n_i}$ as

$$\hat{X}^i_{\text{2D}} = \text{Conv2D}(X^i_{\text{reshaped}}) \tag{5.14}$$

Projection: After 2D convolution, the tensor $\hat{X}^i_{\text{2D}}$ is processed through a projection layer to map the features to the desired prediction length P, i.e.,

$$\hat{X}^i_{\text{Periodic}} = \text{LinearProjection}(\hat{X}^i_{\text{2D}}) \tag{5.15}$$

where $\hat{X}^i_{\text{Periodic}}$ with shape $[B, P, N]$, denotes the extracted features representing periodic dependencies, which serve as the input to subsequent layers of the model.

Long-term Analysis

This module is designed to model global or long-term dependencies that influence future values. The following steps outline the procedure:

Reshaping: Similar to the reshaping step in the short-term analysis, the B and N dimensions of the input 2D tensors $X^i_{\text{2D}} \in \mathbb{R}^{B\times p_i\times n_i\times N}$ are merged, resulting in the reshaped tensor $X^i \in \mathbb{R}^{(B\times N)\times p_i\times n_i}$. This tensor is then used in subsequent processing steps.

1D Convolution: A Conv1D is applied over the temporal dimension of the reshaped tensor $X^i \in \mathbb{R}^{(B\times N)\times p_i\times n_i}$. This step is crucial for extracting and enriching temporal patterns within each period as follows:

$$X^i_{\text{conv}} = \text{Conv1D}(X^i) \tag{5.16}$$

Reshaping: After enriching input data with Conv1D, the X^i_{conv} is reshaped back to shape $[B, S, N]$ as:

$$X^i_{\text{reshaped}} = \text{Reshape}(X^i_{\text{conv}}, [B, S, N]) \tag{5.17}$$

This step realigns the tensor for embedding processes.

Reverted Embedding: This unique embedding layer processes each time series independently, operating on the rows of the reshaped tensor $X^i_{\text{reshaped}} \in \mathbb{R}^{B\times S\times N}$. Unlike traditional embedding methods that map along the feature dimension N, the reverted embedding maps along the temporal sequence S. This embedding utilizes a series of linear layers to project each time series into a higher-dimensional d_{model} as [58]

$$X^i_{\text{embedded}} = \text{RevertedEmbedding}(X^i_{\text{reshaped}}) \tag{5.18}$$

Here, X^i_{embedded} with shape $[B, d_{\text{model}}, N]$ represents the higher-dimensional feature representation of each time series. This embedding facilitates a deeper analysis by enhancing the ability of the model to discern long-term temporal dependencies, preparing the data for sophisticated processing with Transformer encoder layers.

Encoder Module: The encoder processes the embedded tensors sequentially through L stacked layers, each containing key submodules such as multi-head attention mechanisms, Feed-Forward Networks (FFN), residual connections, and normalization layers. Each component plays a crucial role in enhancing the data processing capabilities of the model. At each encoder layer ℓ, the input X^i_{embedded} is transformed through a series of operations within

the encoder. Initially, the Queries ($Q_\ell^{(h,i)}$), Keys ($K_\ell^{(h,i)}$), and Values ($V_\ell^{(h,i)}$) for each input i are computed at layer ℓ as follows:

$$Q_\ell^{(h,i)} = W_{Q_\ell}^{(h,i)} X_{\text{embedded}}^i \tag{5.19}$$

$$K_\ell^{(h,i)} = W_{K_\ell}^{(h,i)} X_{\text{embedded}}^i \tag{5.20}$$

$$V_\ell^{(h,i)} = W_{V_\ell}^{(h,i)} X_{\text{embedded}}^i \tag{5.21}$$

Here, $W_{Q_\ell}^{(h,i)}$, $W_{K_\ell}^{(h,i)}$, and $W_{V_\ell}^{(h,i)}$ are learnable weight matrices specific to each head h, each layer ℓ, and each input tensor i. The queries and keys determine the attention weights, which are applied to the values to obtain the attention output.

Next, the multi-head self-attention mechanism is employed to determine the relative importance of each time step within the sequence. For each head h, attention weights are computed by taking the dot product of the Queries $Q_\ell^{(h,i)}$ and Keys $K_\ell^{(h,i)}$, scaling the result by the inverse square root of the dimension of the keys d_k, and then applying a softmax function to normalize the scores. This operation can be mathematically expressed as [58]

$$\mathrm{A}_\ell^{(h,i)} = \text{softmax}\left(\frac{Q_\ell^{(h,i)} K_\ell^{(h,i)T}}{\sqrt{d_k}}\right) V_\ell^{(h,i)} \tag{5.22}$$

where $\mathrm{A}_\ell^{(h,i)}$ represents the attention output for head h regarding the i-th input at layer ℓ. This operation calculates how much focus each time step should place on others, allowing the model to identify relationships between different points in the sequence. Each head operates independently using a different set of learnable weights, enabling the model to capture a variety of relationships within the sequence. Then the outputs of the h attention mechanisms are concatenated over the feature dimension as

$$X_{\text{Concat},\ell}^i = \text{Concat}(A_\ell^{(1,i)}, A_\ell^{(2,i)}, \ldots, A_\ell^{(H,i)}) \tag{5.23}$$

where $X_{\text{Concat},\ell}^i$ denotes the concatenated output of the multi-head attention mechanism for the i-th input at layer ℓ. The concatenated output is then passed through a Feed-Forward

Network (FFN) to introduce non-linearity, enhancing the ability of the model to capture complex patterns, i.e.,

$$X^i_{\text{FFN},\ell} = \text{FFN}(X^i_{\text{Concat},\ell}) \tag{5.24}$$

where $X^i_{\text{FFN},\ell}$ represents the output from the FFN at layer ℓ for the i-th input. To maintain the integrity of the original input data and enhance feature integration, a residual connection is added to the FFN output. This step helps preserve information from earlier layers and facilitates deeper network training without the risk of vanishing gradients. The residual connection sums the input to the layer and the output from the FFN is

$$X^i_{\text{residual},\ell} = X^i_{\text{Concat},\ell} + X^i_{\text{FFN},\ell} \tag{5.25}$$

where $X^i_{\text{residual},\ell}$ represents the residual output at layer ℓ for the i-th input. This residual output is then normalized to stabilize the transformed output:

$$X^i_{\text{normalized},\ell} = \text{Normalization}(X^i_{\text{residual},\ell}) \tag{5.26}$$

where $X^i_{\text{normalized},\ell}$ denotes the output of the normalization process at layer ℓ for the i-th input.

The attention, feed-forward, residual, and normalization operations are repeated over L encoder layers. After all L layers have been processed, the final encoded representation X^i_{encoded} for the i-th periodic input tensor X^i is obtained. At the end, X^i_{encoded} is processed through a projection layer to map the features to the desired prediction length P:

$$\hat{X}^i_{\text{long}} = \text{LinearProjection}(X^i_{\text{encoded}}) \tag{5.27}$$

Here, $\hat{X}^i_{\text{long}}$ with shape $[B, P, N]$ denotes the extracted long-term or global dependencies, which serve as the input to subsequent layers of the model. In addition, the overall architecture of the encoder processing is depicted in Fig 5.5.

Summation

The PDM divides the 1D input tensor $X \in \mathbb{R}^{B \times S \times N}$ into k distinct 2D tensors $X^i_{2D} \in \mathbb{R}^{B \times p_i \times n_i \times N}$. Each tensor X^i_{2D} is processed through short-term analysis, periodic dependency analysis, and long-term analysis to capture the relevant features from the i-th dominant period independently.The outputs from these analyses, $\hat{X}^i_{\text{short}}$, $\hat{X}^i_{\text{Periodic}}$, and $\hat{X}^i_{\text{long}}$ should be combined to create a comprehensive representation of the different temporal scales inherent in the data. This integration is performed through a summation as

$$\hat{X}^i_{\text{PDM}} = \hat{X}^i_{\text{short}} + \hat{X}^i_{\text{periodic}} + \hat{X}^i_{\text{long}} \tag{5.28}$$

After summation, the integrated outputs $\hat{X}^i_{\text{PDM}}$ with shape $[B, P, N]$ should be fused for k different dominant periods as

$$\hat{X}_{\text{PDM_stack}} = \text{Stack}(\hat{X}^1_{\text{PDM}}, \ldots, \hat{X}^k_{\text{PDM}}) \tag{5.29}$$

where $\hat{X}_{\text{PDM_stack}}$ has shape $[B, P, N, k]$ and contains the features corresponding to the k dominant periods. Finally, to refine and consolidate these features for prediction, a linear transformation is applied to collapse the additional dimension k introduced by stacking, i.e.,

$$\hat{X}_{\text{PDM}} = \text{LinearTransform}(\hat{X}_{\text{PDM_stack}}) \tag{5.30}$$

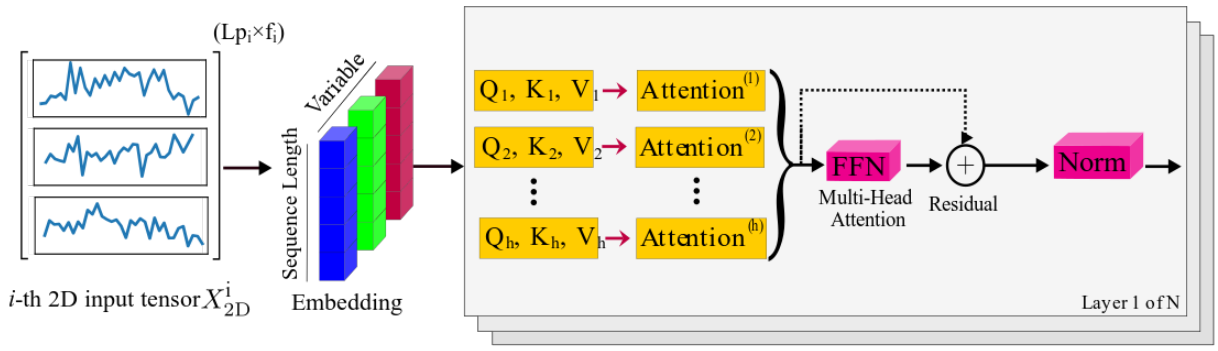


Figure 5.5: *Overall architecture of Transformer encoder processing to capture global dependencies.*

where $\hat{X}_{\text{PDM}}$ with shape $[B, P, N]$ denotes the final processed output of the PDM.

Aggregation

The final output of the PDM $\hat{X}_{\text{PDM}}$ and the output from the STDM $\hat{X}_{\text{STDM}}$ are then combined through a weighted addition to create a final feature representation for the forecasting task. This process is expressed mathematically as follows:

$$\hat{X} = \alpha_1 \hat{X}_{\text{PDM}} + \alpha_2 \hat{X}_{\text{STDM}} \tag{5.31}$$

where α_1 and α_2 are learnable weights that adjust the influence of periodical dependencies and seasonal patterns, respectively. The combined output $\hat{X}$ with shape $[B, P, N]$ encapsulates various temporal variations captured by the PDM and STDM. This representation is subsequently passed through a loss function during the training phase to optimize the model parameters for accurate forecasting.

5.1.2 Results

Data Collection and Preprocessing

The SPDNet model is evaluated using AMI load data collected from a city in Kansas, USA, from October 25, 2020, to March 10, 2024. Each dataset contains 118,356 rows with 15-minute intervals. Weather variables, including air temperature, relative humidity, wind speed, and wind direction, were also used as input features to the SPDNet model.

Because several meters reported long communication gaps or extended zero readings, a multi-stage cleaning process was applied. First, consecutive missing or zero-valued segments were detected for every load ID. Loads with gaps longer than one month were removed entirely, which eliminated 114 load IDs. Next, 31 additional load IDs with gaps longer than two weeks but shorter than one month were inspected and removed when the gap noticeably degraded sequence continuity. After this step, only meters with acceptable data completeness were retained. Remaining gaps were imputed using a two-level strategy:

- short gaps (fewer than 96 consecutive samples, i.e., less than one day at 15-minute resolution) were filled by averaging the neighboring values before and after the gap;
- long gaps (96 or more consecutive samples) were filled by averaging the corresponding periods from the previous and the next year, so that diurnal and seasonal structure was preserved.

To maintain a focus on residential-like behavior, the 9 large commercial loads were excluded. In addition, 13 loads that had no utility label but exhibited clearly residential daily and weekly patterns were retained and treated as residential. After these steps, the final dataset contained 646 loads in total: 624 residential and 22 small commercial. Finally, load IDs known to be served by the same distribution transformer were grouped together, resulting in 58 aggregated (transformer-level) load profiles that were added to the dataset for further analysis. The data were then split into 70% for training, 10% for validation, and 20% for testing. To standardize the input features, normalization was performed using the StandardScaler, which standardizes each feature by removing the mean and scaling to unit variance as

$$z = \frac{x - \mu}{\sigma} \tag{5.32}$$

where z is the normalized value, x is the original value, μ is the mean, and σ is the standard deviation of a feature. Importantly, the scaler parameters were estimated using the training set only and then applied to the validation and test sets to prevent information leakage.

Baselines

The performance of SPDNet is evaluated against 12 baseline models to ensure a comprehensive analysis across a variety of forecasting methods. This evaluation includes traditional models such as RNN, LSTM, GRU, BiLSTM, ConvLSTM, ResLSTM, and Transformer, as well as state-of-the-art techniques like Informer, Crossformer, DLinear, PatchTST, and TimesNet.

Experiment Settings

For comparison, all models were evaluated on input lengths of $S = \{96, 384, 720\}$ and prediction horizons of $P = \{1, 4, 24, 48, 96\}$, with a fixed batch size of 32. For individual residential loads, hyperparameter tuning was performed using residential load 1, and the selected parameters were subsequently applied to all other residential loads. For aggregated loads, hyperparameter tuning was performed separately using one aggregated load group, and the selected parameters were then applied to all other aggregated load groups. Additional hyperparameter tuning was not needed because the selected values provided satisfactory results for other loads. In addition, Mean Squared Error (MSE) and Mean Absolute Error (MAE) were used as performance evaluation metrics.

Forecasting Scenarios. To evaluate the forecasting module at different levels of load aggregation, two scenarios are considered. In the first scenario, the forecasting models are evaluated using individual customer load profiles to assess their ability to capture highly variable consumption patterns. In the second scenario, the demands of customers connected to the same distribution transformer are aggregated into 58 transformer-level load profiles, and SPDNet is applied to forecast the resulting demand. Because SPDNet demonstrates strong forecasting performance in the individual-load comparison, the quantitative evaluation for all 58 aggregated groups focuses on SPDNet, while a representative aggregated-load plot also compares its predictions with selected baseline models. The computational efficiency comparison is reported for the individual-load scenario.

Implementation Details

All experiments were conducted using PyTorch on a system equipped with an NVIDIA RTX A4000 GPU with 16 GB of GPU memory and an Intel Xeon E5-2680 v3 CPU to handle extensive computational demands efficiently.

Forecasting Results

Scenario 1: Individual Load Forecasting. In the first scenario, the forecasting models are evaluated on individual customer load profiles. As an example, residential load 1 is used to present a detailed comparison between SPDNet and the baseline models. Additional results for residential loads 2–6 are also presented. The computational efficiency analysis that follows also uses the individual-load scenario.

The forecasting results for residential load 1 are provided in Table 5.2. Additionally, Fig 5.6 depicts the predicted and actual electricity demand by all models for load 1 with $S = 96$ and $P = 96$. The results indicate that SPDNet achieves competitive MSE and MAE across the evaluated prediction horizons and sequence lengths, although some baseline models perform better in individual settings. At shorter prediction horizons $\{1, 4, 24\}$, SPDNet generally provides low forecasting errors. The influence of sequence length varies

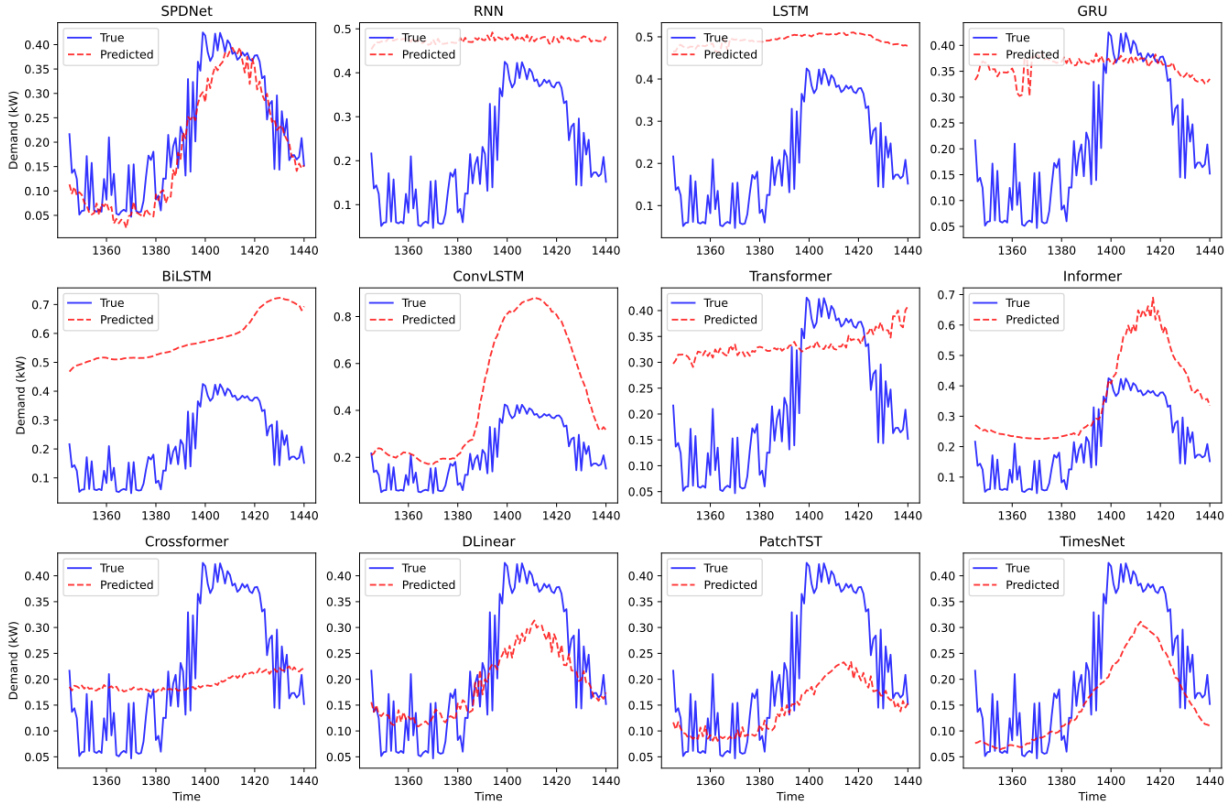


Figure 5.6: *Forecasted and actual electricity demand for load 1 at sequence length $S = 96$ and prediction horizon $P = 96$. Legends are shown only in the first subplot for clarity.*

Table 5.2: *Forecasting results for load 1 with normalized values, using sequence lengths* $S = \{96, 384, 720\}$ *and prediction horizons* $P = \{1, 4, 24, 48, 96\}$.

Model	Metric	S=96					S=384					S=720				
		Prediction Horizon					Prediction Horizon					Prediction Horizon				
		1	4	24	48	96	1	4	24	48	96	1	4	24	48	96
SPDNet(Ours)	MSE	0.010	0.024	0.052	0.056	0.064	0.010	0.023	0.050	0.055	0.062	0.010	0.024	0.048	0.054	0.057
	MAE	0.050	0.075	0.112	0.118	0.126	0.052	0.077	0.118	0.120	0.148	0.052	0.077	0.115	0.122	0.130
RNN	MSE	0.054	0.189	0.934	0.621	0.486	0.052	0.181	0.935	0.949	0.496	0.051	0.181	0.938	0.945	0.485
	MAE	0.217	0.343	0.945	0.685	0.482	0.214	0.334	0.945	0.952	0.472	0.210	0.346	0.947	0.950	0.469
LSTM	MSE	0.029	0.26	0.970	0.990	0.456	0.017	0.254	0.966	0.995	0.556	0.031	0.251	0.969	1.000	0.568
	MAE	0.142	0.35	0.962	0.972	0.511	0.100	0.353	0.960	0.974	0.525	0.148	0.355	0.962	0.977	0.534
GRU	MSE	0.017	0.161	0.843	0.398	0.364	0.015	0.177	0.844	0.381	0.477	0.016	0.160	0.856	0.320	0.448
	MAE	0.102	0.330	0.883	0.558	0.494	0.095	0.302	0.883	0.553	0.550	0.096	0.291	0.889	0.463	0.535
BiLSTM	MSE	0.020	0.087	0.830	0.803	0.475	0.020	0.084	0.821	0.870	0.472	0.020	0.084	0.817	0.731	0.423
	MAE	0.108	0.215	0.889	0.726	0.459	0.110	0.216	0.884	0.810	0.454	0.106	0.208	0.882	0.678	0.429
ConvLSTM	MSE	0.023	0.165	1.000	0.790	0.462	0.022	0.150	1.000	1.065	0.395	0.022	0.147	1.013	0.863	0.439
	MAE	0.125	0.271	0.980	0.706	0.452	0.118	0.265	0.981	1.010	0.426	0.118	0.275	0.984	0.717	0.432
ResLSTM	MSE	0.020	0.198	0.904	0.917	0.899	0.024	0.188	1.308	1.305	0.982	0.200	0.897	0.920	0.899	0.899
	MAE	0.077	0.394	0.926	0.933	0.925	0.108	0.373	1.099	1.098	0.892	0.081	0.392	0.922	0.935	0.925
Transformer	MSE	0.026	0.184	0.588	0.583	0.535	0.025	0.160	0.592	0.642	0.538	0.124	0.329	0.574	0.801	0.738
	MAE	0.108	0.255	0.737	0.734	0.700	0.100	0.342	0.733	0.764	0.702	0.266	0.488	0.545	0.766	0.697
Informer	MSE	0.036	0.062	0.563	0.762	0.763	0.027	0.074	0.440	0.730	0.639	0.111	0.194	0.524	0.335	0.819
	MAE	0.122	0.185	0.639	0.769	0.760	0.092	0.205	0.604	0.754	0.658	0.299	0.403	0.652	0.523	0.819
Crossformer	MSE	0.012	0.031	0.146	0.149	0.168	0.011	0.025	0.084	0.113	0.133	0.011	0.029	0.061	NAN	0.116
	MAE	0.067	0.114	0.348	0.346	0.375	0.055	0.087	0.234	0.284	0.328	0.060	0.109	0.175	NAN	0.239
Dlinear	MSE	0.012	0.025	0.062	0.070	0.078	0.012	0.024	0.050	0.056	0.063	0.012	0.024	0.052	0.056	0.058
	MAE	0.054	0.076	0.172	0.186	0.207	0.053	0.077	0.119	0.131	0.162	0.054	0.071	0.128	0.137	0.137
PatchTST	MSE	0.020	0.031	0.052	0.058	0.059	0.023	0.034	0.053	0.058	0.060	0.025	0.034	0.052	0.056	0.060
	MAE	0.068	0.080	0.105	0.110	0.112	0.076	0.091	0.111	0.116	0.119	0.083	0.093	0.109	0.118	0.120
TimesNet	MSE	0.012	0.027	0.051	0.058	0.060	0.014	0.030	0.052	0.056	0.058	0.016	0.030	0.054	0.064	0.058
	MAE	0.053	0.074	0.104	0.114	0.116	0.059	0.081	0.111	0.120	0.123	0.063	0.085	0.118	0.131	0.128

across prediction horizons, as shown in Table 5.2.

Among the recursive models (RNN, GRU, and LSTM), the GRU has relatively good performance across MSE and MAE, suggesting it strikes a good balance between complexity and performance. Notably, RNN outperforms LSTM in most of the experiments. BiLSTM, ConvLSTM, and ResLSTM use different architectural modifications of LSTM, and their relative accuracy depends on the prediction setting. BiLSTM enhances LSTM by processing data in both forward and backward directions, allowing it to capture dependencies from both past and future contexts. ConvLSTM introduces convolutional layers within the LSTM architecture, enabling it to capture spatial features in time series data, which is particularly beneficial for tasks where the input data includes spatial dimensions. ResLSTM integrates residual connections that help in combating the vanishing gradient problem by allowing gradients to flow through the network more effectively, leading to better performance.

PatchTST, TimesNet, and DLinear demonstrate strong performance among advanced time series forecasting models. PatchTST uses patch-based input processing with a Trans-

Table 5.3: *Performance metrics (MSE and MAE) for SPDNet and LSTM models across various prediction horizons for different loads.*

Load	Metric	SPDNet			LSTM		
		$P = 1$	$P = 4$	$P = 24$	$P = 1$	$P = 4$	$P = 24$
Load 2	MSE	0.001	0.002	0.005	0.006	0.038	0.302
	MAE	0.022	0.028	0.053	0.053	0.155	0.440
Load 3	MSE	0.002	0.004	0.011	0.008	0.011	0.024
	MAE	0.019	0.025	0.048	0.048	0.062	0.100
Load 4	MSE	0.039	0.102	0.260	0.055	0.240	0.732
	MAE	0.153	0.236	0.379	0.179	0.370	0.708
Load 5	MSE	0.010	0.026	0.072	0.019	0.084	0.219
	MAE	0.060	0.082	0.150	0.103	0.222	0.368
Load 6	MSE	0.015	0.042	0.124	0.022	0.092	0.362
	MAE	0.083	0.140	0.256	0.107	0.223	0.534

former architecture, while TimesNet processes periodic variations using two-dimensional representations. Both models obtain competitive results in several settings. Lastly, DLinear uses decomposition and linear transformations rather than an attention mechanism. Its reported performance is competitive with several more complex baseline models.

The performance of SPDNet is further evaluated across five additional residential loads. Using a sequence length of 96 and prediction horizons of {1, 4, 24}, SPDNet was trained with the identical parameters and hyperparameters utilized for load 1. The results are presented in Table 5.3. SPDNet achieves lower MSE and MAE than LSTM for the reported loads and prediction horizons.

Computational Efficiency Analysis

Fig 5.7 illustrates the average training time per epoch for all models, using a sequence length of 720 and prediction horizons {1, 4, 24, 48, 96, 192, 384, 672, 1344, 2688} on load 1. Compared to competitive models such as PatchTST, DLinear, and TimesNet, SPDNet demonstrates significantly lower average training times per epoch. Notably, as the prediction horizon increases, SPDNet maintains consistent training times. This performance underscores SPDNet's advanced computational architecture, which is pivotal for enhancing the efficiency

and scalability of forecasting models.

Scenario 2: Aggregated Load Forecasting. In the second scenario, individual load demands associated with the same distribution transformer are combined to form 58 aggregated load profiles. As SPDNet achieved favorable forecasting accuracy in the individual-load experiments, its performance is reported for all 58 aggregated groups, rather than repeating the full baseline evaluation for every group. Table 5.4 presents the MSE and MAE obtained for each group. The values are reported as provided by the aggregated-load evaluation results.

For a visual comparison with representative baseline models, Fig. 5.8 presents the actual aggregated demand and the corresponding day-ahead predictions from SPDNet, RNN, LSTM, GRU, ConvLSTM, Transformer, Informer, and PatchTST. This comparison illus-

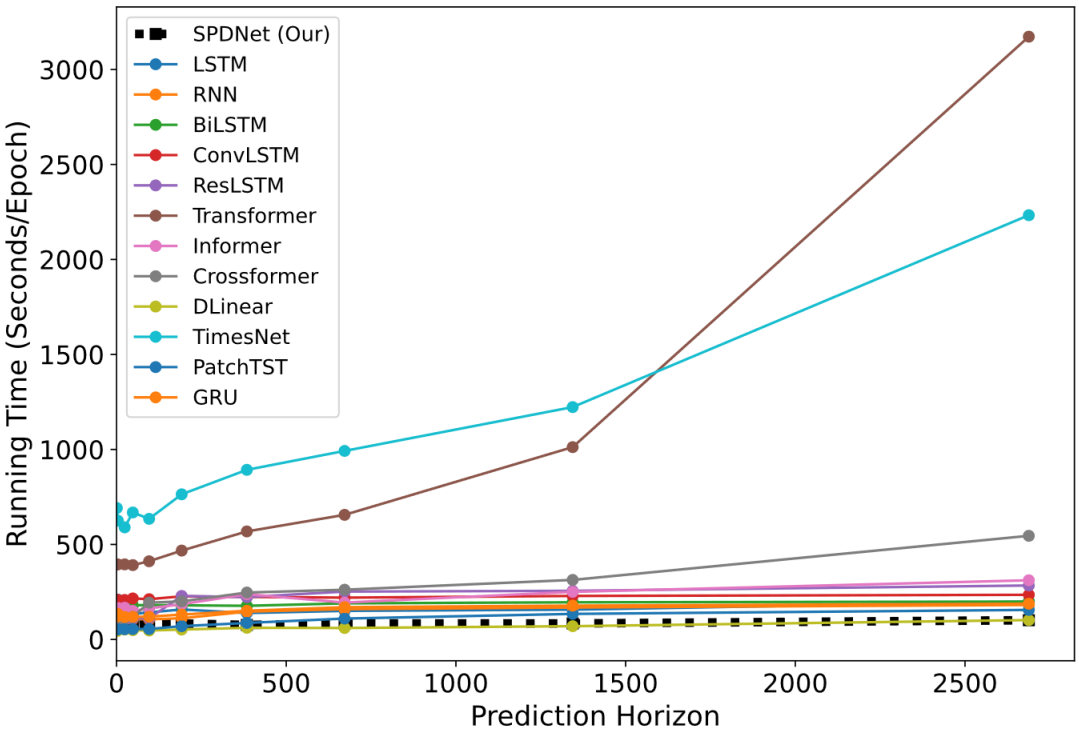


Figure 5.7: *Average training time (seconds) per epoch for different models. All experiments are trained with an input length of 720, batch size of 32, and prediction horizons of {1,4,24,48,96,192,384,672,1344,2688}.*

Table 5.4: *SPDNet forecasting accuracy for the 58 aggregated transformer-level load groups, measured using mean squared error (MSE) and mean absolute error (MAE).*

Group	MSE	MAE	Group	MSE	MAE
Group 1	0.1922	0.3197	Group 30	0.1676	0.2990
Group 2	0.1637	0.2830	Group 31	0.1588	0.3038
Group 3	0.3690	0.4527	Group 32	0.2829	0.4104
Group 4	0.2025	0.3269	Group 33	0.1703	0.2930
Group 5	0.1565	0.2829	Group 34	0.1616	0.2935
Group 6	0.1357	0.2669	Group 35	0.1623	0.2905
Group 7	0.2457	0.3635	Group 36	0.2032	0.3185
Group 8	0.1426	0.2741	Group 37	0.2428	0.3707
Group 9	0.1595	0.2789	Group 38	0.4308	0.5065
Group 10	0.1631	0.2867	Group 39	0.2157	0.3342
Group 11	0.3207	0.4290	Group 40	0.1886	0.3260
Group 12	0.2449	0.3545	Group 41	0.1750	0.3118
Group 13	0.2673	0.3602	Group 42	0.2597	0.3806
Group 14	0.1844	0.3170	Group 43	0.2192	0.3513
Group 15	0.2082	0.3401	Group 44	0.1997	0.3324
Group 16	0.1891	0.3074	Group 45	0.1954	0.3266
Group 17	0.2678	0.3901	Group 46	0.1487	0.2826
Group 18	0.2062	0.3356	Group 47	0.2147	0.3446
Group 19	0.1591	0.2927	Group 48	0.2189	0.3482
Group 20	0.2094	0.3352	Group 49	0.2228	0.3303
Group 21	0.1637	0.2889	Group 50	0.3040	0.4253
Group 22	0.1172	0.2511	Group 51	0.1525	0.2856
Group 23	0.1209	0.2540	Group 52	0.1666	0.2865
Group 24	0.1637	0.3037	Group 53	0.2578	0.3840
Group 25	0.2179	0.3498	Group 54	0.1803	0.3119
Group 26	0.1319	0.2623	Group 55	0.1689	0.2905
Group 27	0.2043	0.3321	Group 56	0.2189	0.3510
Group 28	0.1925	0.3178	Group 57	0.1881	0.3190
Group 29	0.1398	0.2672	Group 58	0.1811	0.2975

trates the forecasting trajectories of the selected methods at the aggregated-load level, while Table 5.4 provides the group-wise numerical evaluation of SPDNet.

5.1.3 Uncertainty Quantification

After generating day-ahead forecasts for all 58 aggregated load groups, the next step is to quantify the uncertainty associated with these forecasts. This step is necessary because the forecasted active power values will later be used as pseudo-measurements in the DSSE stage, and the DSSE needs not only point estimates but also a realistic measure of how much

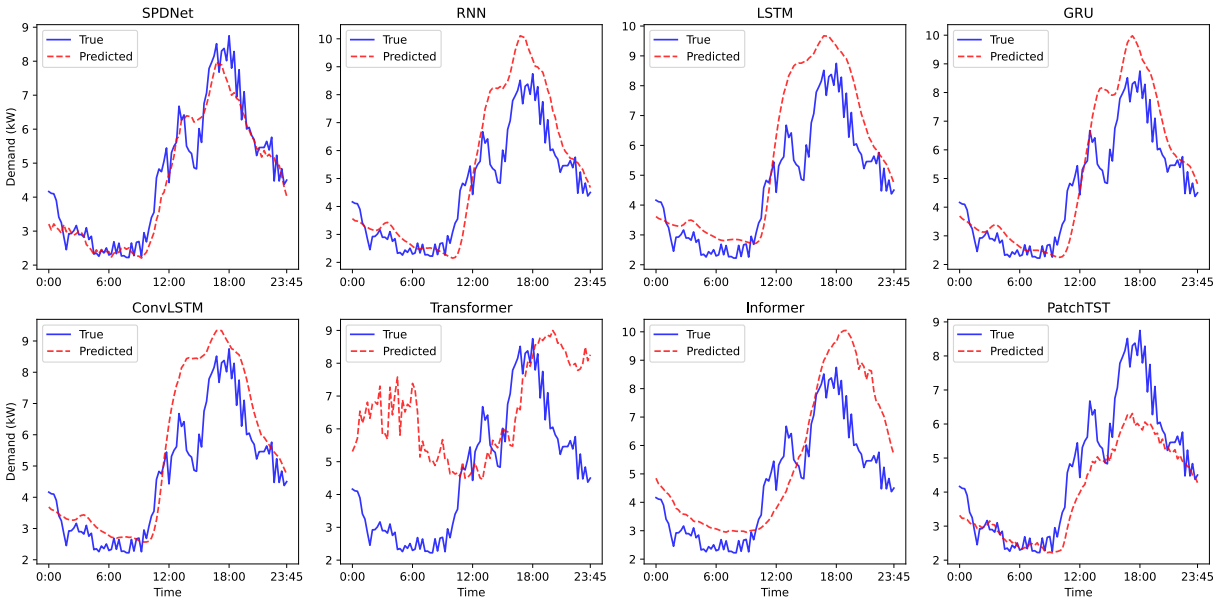


Figure 5.8: *Comparison of actual and forecasted aggregated electricity demand using SPDNet, RNN, LSTM, GRU, ConvLSTM, Transformer, Informer, and PatchTST over a 24-hour prediction horizon at 15-minute resolution.*

each value can vary. In other words, the forecasting module must provide both the mean (the forecast) and a spread (the expected variability) so that the state estimator can weight measurements properly.

For each aggregated group, the model was used to forecast the next 24 hours (96 steps at 15-minute resolution). The test set contained approximately 150 daily samples for each group. To estimate the error dispersion associated with these forecasts, the following procedure was applied:

1. For a given load group, collect all 24-hour forecasts produced on the test days. Arrange them in a matrix of size $N_{\text{days}} \times 96$, where $N_{\text{days}} \approx 150$, each row corresponds to one forecasted day, and each column corresponds to the same forecast horizon timestamp (for example, columns 1 and 2 represent the first and second 15-minute forecast steps).

2. For each forecast horizon $t = 1, \ldots, 96$, compute the sample mean and the sample standard deviation of the forecast errors across all days.

3. The pair (μ_t, σ_t) defines an empirical confidence interval for every 15-minute step.

Figure 5.9 illustrates the pointwise forecasting error, defined as

$$e_t = \hat{y}_t - y_t,$$

for two representative aggregated load groups, shown at two different time periods. The plotted errors fluctuate around zero without an obvious sustained drift. However, the absence of visible drift does not establish that the forecasts are unbiased; bias should be assessed from the empirical error means.

The error profiles provide a qualitative view of the variation across forecast horizons. A statistical test or residual-correlation analysis would be needed to establish whether the errors are independent across timestamps or across load groups.

For each forecast horizon t, the point forecast can be accompanied by the estimated error standard deviation,

$$\hat{y}_t \pm \sigma_t,$$

where σ_t is obtained by stacking all test-day forecast errors for the same timestamp and computing their sample standard deviation. This horizon-dependent treatment permits different error variances at different forecast steps. The interval shown corresponds to one standard deviation and should not be interpreted as a 95% prediction interval without additional calibration.

This representation is particularly useful for distribution system state estimation. The forecasted load $\hat{y}_t$ can be injected as a pseudo-measurement, while the corresponding variance σ_t^2 can be used as an entry on the diagonal of the pseudo-measurement covariance matrix when cross-measurement correlations are neglected.

The same uncertainty quantification procedure can also be applied to the individual-load forecasting scenario. For each individual customer, day-ahead forecast errors can be organized by prediction timestamp, and the sample standard deviation of the errors at each timestamp can be used to assign a corresponding variance to the individual-load pseudo-measurement. The uncertainty results discussed above, however, are specifically based on

the 58 aggregated load groups.

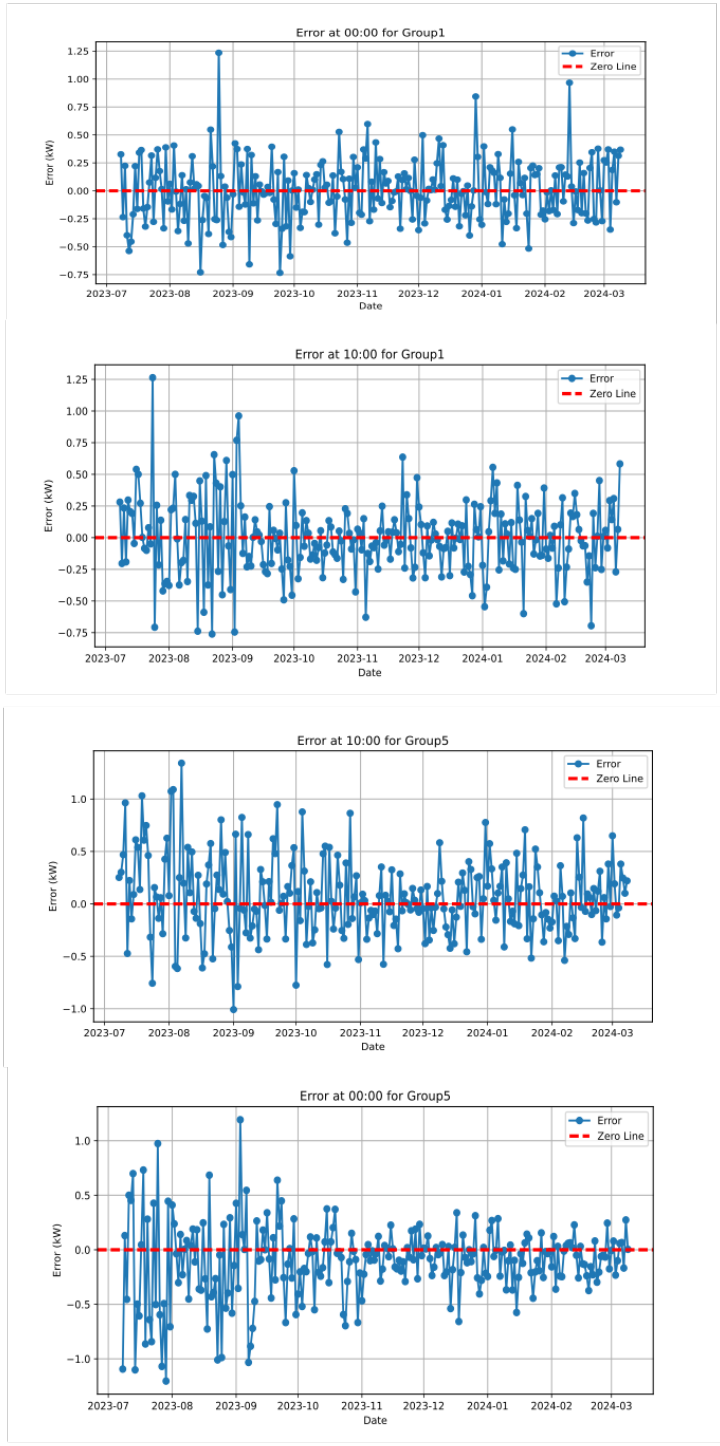


Figure 5.9: *Day-ahead forecasting errors and data-driven uncertainty bands for two representative aggregated load groups.*

5.1.4 Summary

This section developed and validated a forecasting framework that can generate high quality, uncertainty aware pseudo measurements for distribution system state estimation. Motivated by the limitations of existing load forecasting models in capturing seasonality, multi periodicity, non stationarity, and sharp short term fluctuations in residential and feeder level data, we designed a decomposition based deep architecture called Seasonal Periodic Decomposition Network (SPDNet).

SPDNet is organized in two coordinated parts. The Seasonal Trend Decomposition Module (STDM) separates the input window into trend, recurring seasonal components, and residuals by using 1D convolutions with different receptive fields. This step lets the model treat slow variations and repeated patterns differently from noisy, high frequency behavior. In parallel, the Periodical Decomposition Module (PDM) applies a Fast Fourier Transform to the input sequence, identifies the dominant periods, and reshapes the original 1D series into several 2D tensors, each aligned with one of those periods. Every 2D tensor is then processed by three parallel submodules: a short term 1D convolution to track fast changes within the period, a 2D convolution to capture intra and inter period dependencies, and a Transformer encoder to model long range temporal relations. Outputs from all dominant periods are concatenated and fused with the STDM output through learnable weights, so the final forecast uses information from all temporal scales present in the data.

The model was trained on cleaned AMI data collected from a city in Kansas, USA, comprising 118,356 time-stamped samples at 15-minute intervals. The dataset was filtered to include 646 residential loads, which were subsequently aggregated into 58 transformer-level groups. On this dataset, SPDNet was tested against a broad set of classical and state of the art baselines (RNN, LSTM, GRU, BiLSTM, ConvLSTM, ResLSTM, Transformer, Informer, Crossformer, DLinear, PatchTST, TimesNet) for several input window lengths and prediction horizons. SPDNet achieved competitive MSE and MAE across the evaluated configurations and outperformed several baseline models, although it did not obtain the lowest error in every setting.

After producing point forecasts for all 58 aggregated groups, we quantified forecast uncertainty directly from the test set. For each group we forecasted the next 24 hours over all available test days, stacked the forecast errors (forecast minus true) by horizon, and computed a sample standard deviation for every prediction timestamp. The error plots showed fluctuations around zero, and sample error variances were calculated separately for each forecast horizon. These variances describe the observed test-set errors but do not establish unbiasedness or independence. As a result, each forecast $\hat{x}(t)$ can be paired with a horizon-specific variance $\sigma^2(t)$. These forecast–variance pairs can now be used as pseudo measurements in the next layers of the data driven DSSE framework.

Chapter 6

Calibrated Fusion for Distribution System State Estimation

Advanced metering infrastructure (AMI) provides distribution utilities with high-resolution information about customer electricity consumption. Nevertheless, the availability of historical AMI records does not imply that all measurements can be collected in near real time. Meter communication restrictions, limited polling capacity, and delayed data retrieval can leave a substantial proportion of distribution-system measurements unavailable at the time of state estimation. Consequently, distribution system state estimation (DSSE) must combine the measurements that are available with estimates of the unobserved quantities. Inaccuracies in these inferred measurements can propagate through the estimation procedure and compromise the reconstructed bus-voltage magnitudes and phase angles.

Chapter 5 developed the electricity-demand forecasting methodology and evaluated its performance for individual and transformer-aggregated loads. It also established a procedure for quantifying the uncertainty of day-ahead forecasts from historical prediction errors. These results provide the inputs required by the state estimation framework developed here. Accordingly, the forecasting model, its training procedure, and its forecasting benchmarks are not repeated in this chapter. Instead, the focus is on how forecasted pseudo-measurements and their uncertainties are incorporated into DSSE and how the resulting estimates can be

improved by combining complementary estimation techniques.

The proposed framework employs two estimators. The first, Bayesian Weighted Least Squares (BWLS), combines observed and forecast-based measurements through a network-constrained nonlinear measurement model. The second, Variational Bayesian Matrix Completion (VBMC), uses correlations in incomplete AMI data to reconstruct unobserved information. Because these two procedures use different sources of information, their estimation errors may differ across the network. Maximum Ratio Combining (MRC) provides a mechanism for fusing their outputs, but its conventional variance-based formulation assumes that the component estimators are unbiased. This assumption may be violated in low-observability distribution feeders. To address this limitation, the proposed calibrated MRC framework first corrects systematic errors using observable reference data and then fuses the corrected estimates using residual-error-based weights.

Two state estimation scenarios are considered. In the first, load demands associated with customers served by a common distribution transformer are aggregated, and the resulting forecasts are used in the state estimation procedure. The associated transformer primary-side voltage is considered for evaluation at the aggregated representation. In the second, the same framework is applied using individual customer-load forecasts and their corresponding network locations. The aggregated scenario reduces the number of separate forecasting profiles required, whereas the individual-load scenario retains greater spatial resolution. This chapter develops the estimation and fusion methodology, presents the feeder-study results for the aggregated-load scenario, and compares BWLS state-estimation accuracy using aggregated and individual load forecasts.

6.1 Estimation Framework

At an estimation instant, a subset of AMI measurements is received from the distribution system, while the remaining load measurements must be reconstructed. The forecasts generated in Chapter 5 supply pseudo-measurements for the unobserved demands, together with corresponding uncertainty estimates. Two complementary state-estimation paths then pro-

cess this incomplete-information setting. BWLS uses the distribution-network measurement model and assigns weights according to measurement uncertainty. VBMC uses the structure of the partially observed AMI data to reconstruct missing information and obtain a complementary estimate. Finally, the corresponding estimates are calibrated and fused to produce the final state estimate.

The primary state variables of interest are the bus-voltage magnitudes and phase angles. For a feeder with n modeled voltage states, they are represented by

$$\mathbf{x} = \begin{bmatrix} \boldsymbol{\delta}^\top & \mathbf{v}^\top \end{bmatrix}^\top, \tag{6.1}$$

where $\boldsymbol{\delta}$ contains the bus-voltage phase angles and $\mathbf{v}$ contains the bus-voltage magnitudes. The precise number and ordering of states depend on the feeder's phase configuration and the reference-bus convention.

6.1.1 Bayesian Weighted Least Squares

BWLS estimates the system states by reconciling the available measurements with the network model. The measurement vector $\mathbf{z}$ comprises observed quantities and forecast-derived pseudo-measurements of the missing loads. It is related to the system state by

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \tag{6.2}$$

where $\mathbf{h}(\mathbf{x})$ denotes the nonlinear measurement equations based on the AC power-flow model and $\mathbf{R}$ is the measurement-error covariance matrix. Following the diagonal-error representation employed in the case study,

$$\mathbf{R} = \operatorname{diag}\big(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2\big), \tag{6.3}$$

where σ_j^2 is the variance assigned to measurement or pseudo-measurement j. For unavailable AMI loads, these variances are obtained from the forecast-error analysis described in Chap-

ter 5. In this manner, the forecasting stage determines not only the pseudo-measurement values but also their relative influence on the state estimator.

Under the assumed Gaussian measurement model and a non-informative state prior, the BWLS solution minimizes the weighted residual objective

$$\widehat{\mathbf{x}}_{\mathrm{BWLS}} = \arg\min_{\mathbf{x}} \frac{1}{2}\big[\mathbf{z} - \mathbf{h}(\mathbf{x})\big]^{\top}\mathbf{R}^{-1}\big[\mathbf{z} - \mathbf{h}(\mathbf{x})\big]. \tag{6.4}$$

The nonlinear objective is solved iteratively using the Gauss–Newton procedure. Let $\mathbf{H} = \partial\mathbf{h}/\partial\mathbf{x}$ denote the measurement Jacobian at the current state estimate. The corresponding Gauss–Newton correction can be written as

$$\Delta\mathbf{x} = \big(\mathbf{H}^{\top}\mathbf{R}^{-1}\mathbf{H}\big)^{-1}\mathbf{H}^{\top}\mathbf{R}^{-1}\big[\mathbf{z} - \mathbf{h}(\mathbf{x})\big], \tag{6.5}$$

when the inverse exists. The state estimate is updated until the selected convergence criterion is met. An approximate local state-estimation covariance is obtained from the inverse weighted information matrix,

$$\mathbf{\Sigma}_{\mathrm{BWLS}} \approx \big(\mathbf{H}^{\top}\mathbf{R}^{-1}\mathbf{H}\big)^{-1}. \tag{6.6}$$

Its diagonal entries provide approximate uncertainty measures for the estimated state components. Consequently, the BWLS path produces both a voltage-state estimate and information about its relative confidence. The approximation assumes that the assigned measurement variances adequately characterize measurement uncertainty; the diagonal measurement-error covariance matrix neglects correlations between errors.

6.1.2 Bayesian Matrix Completion

VBMC provides a complementary approach for handling incomplete AMI observations by exploiting correlations among entries of a measurement matrix. Let $\mathbf{Z}$ denote the partially observed measurement matrix and let Ω contain the indices of its available entries. The

observation model is

$$z_{ij} = x_{ij} + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2), \quad (i,j) \in \Omega, \tag{6.7}$$

where $\mathbf{X} = [x_{ij}]$ is the latent matrix to be reconstructed and $\beta_{ij} = \sigma_{ij}^{-2}$ is the observation-noise precision.

Following the hierarchical Bayesian formulation, the latent matrix is factorized as $\mathbf{X} = \mathbf{A}\mathbf{B}^\top$, where $\mathbf{A} \in \mathbb{R}^{m\times r}$ and $\mathbf{B} \in \mathbb{R}^{n\times r}$. Gaussian priors are assigned to the columns of both factors using shared precision parameters $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_r]^\top$. Each precision parameter is assigned a Gamma hyperprior. Large values of γ_k shrink the corresponding columns of both factors toward zero, thereby promoting a low-rank reconstruction. The resulting joint distribution is

$$p(\mathbf{Z}_\Omega, \mathbf{A}, \mathbf{B}, \boldsymbol{\gamma}) = p(\mathbf{Z}_\Omega \mid \mathbf{A}, \mathbf{B}) p(\mathbf{A} \mid \boldsymbol{\gamma}) p(\mathbf{B} \mid \boldsymbol{\gamma}) p(\boldsymbol{\gamma}),$$

where the observation precisions are treated as given and omitted from the conditioning notation. The posterior distribution is approximated using mean-field variational inference. Let $\mathcal{Q}$ denote the family of distributions with factorization $q(\mathbf{A}, \mathbf{B}, \boldsymbol{\gamma}) = q(\mathbf{A})q(\mathbf{B})q(\boldsymbol{\gamma})$. The variational approximation and reconstructed matrix are expressed as

$$q^\star = \arg\min_{q\in\mathcal{Q}} \mathrm{KL}(q(\mathbf{A}, \mathbf{B}, \boldsymbol{\gamma}) \,\|\, p(\mathbf{A}, \mathbf{B}, \boldsymbol{\gamma} \mid \mathbf{Z}_\Omega)), \qquad \widehat{\mathbf{X}}_{\mathrm{VBMC}} = \mathbb{E}_{q^\star}\left[\mathbf{A}\mathbf{B}^\top\right]. \tag{6.8}$$

Minimizing this divergence is equivalent to maximizing the evidence lower bound. The posterior factors are updated iteratively, yielding approximate means and covariances of the latent factors. These quantities provide both reconstructed entries and estimates of their uncertainty, as described in Chapter 3.

The reconstructed measurements are related to voltage states through the linearized power-flow relations described in Chapter 3. The uncertainty associated with the reconstructed measurements must likewise be mapped to the voltage-state quantities used for

fusion. Thus, the two estimation paths provide estimates of the same physical state,

$$\widehat{\mathbf{x}}_1 = \widehat{\mathbf{x}}_{\mathrm{BWLS}}, \qquad \widehat{\mathbf{x}}_2 = \widehat{\mathbf{x}}_{\mathrm{VBMC}}. \tag{6.9}$$

The estimation errors may differ because BWLS combines measurements and forecasted pseudo-measurements through a weighted power-flow measurement model, whereas VBMC uses a low-rank Bayesian representation to reconstruct incomplete measurements. Their different treatment of missing information motivates the fusion framework developed next.

6.2 Estimator Fusion

6.2.1 Classical MRC

Classical MRC combines two estimates according to their error variances. Consider a scalar component of the true state x_{true} and two corresponding estimates x_1 and x_2:

$$x_1 = x_{\mathrm{true}} + e_1, \qquad x_2 = x_{\mathrm{true}} + e_2. \tag{6.10}$$

The conventional derivation assumes that the errors are unbiased and mutually uncorrelated,

$$\mathbb{E}[e_1] = \mathbb{E}[e_2] = 0, \qquad \mathrm{Cov}(e_1, e_2) = 0, \tag{6.11}$$

and have variances σ_1^2 and σ_2^2. A linear fused estimate is defined by

$$x_f = w_1 x_1 + w_2 x_2, \qquad w_1 + w_2 = 1. \tag{6.12}$$

With these assumptions, its mean-squared error is

$$J(w_1, w_2) = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2. \tag{6.13}$$

The corresponding constrained minimization uses the Lagrangian

$$\mathcal{L} = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + \lambda(w_1 + w_2 - 1). \tag{6.14}$$

Differentiating with respect to the weights and enforcing the unit-sum constraint yields

$$w_1 = \frac{1/\sigma_1^2}{1/\sigma_1^2 + 1/\sigma_2^2}, \qquad w_2 = \frac{1/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}. \tag{6.15}$$

Therefore, the estimator with the smaller error variance receives the larger weight. Applying this rule componentwise to the BWLS and VBMC results produces the classical fused state estimate.

The effectiveness of classical MRC depends on the assumptions underlying the weighting formula. If either estimator has a systematic error, a low variance does not necessarily imply high accuracy. For example, an estimator may be consistently shifted away from the reference voltage while exhibiting little variation around its shifted mean. A variance-only weight can assign substantial importance to such an estimator, limiting the accuracy of the final result. This is the principal limitation addressed by the calibrated fusion method.

6.2.2 Calibrated MRC

The proposed calibrated MRC introduces a regression-based correction before the individual estimates are fused. This adjustment is motivated by the observation that BWLS and VBMC can exhibit different systematic estimation errors at different locations and operating conditions. Rather than assuming that their outputs are unbiased, the calibrated framework uses available reference data to estimate and compensate for systematic deviations.

Bias Correction

For each estimator $i \in \{1, 2\}$, an affine calibration relationship is fitted between the estimator output x_i and the corresponding reference state. The corrected output is defined as

$$\widetilde{x}_i = a_i x_i + b_i, \qquad i = 1, 2, \tag{6.16}$$

where a_i and b_i represent the multiplicative and additive calibration coefficients. For a collection of observable calibration samples, the coefficients are estimated by least squares:

$$(a_i, b_i) = \arg\min_{a,b} \sum_{t \in \mathcal{T}_{\text{cal}}} \left[x_{\text{ref}}(t) - a x_i(t) - b \right]^2, \tag{6.17}$$

where $x_{\text{ref}}(t)$ is the reference value available at calibration instant t and $\mathcal{T}_{\text{cal}}$ is the calibration-sample set. This is the correction procedure used in the proposed approach. Importantly, calibration requires reference information corresponding to the state component being corrected; the implementation must specify how such reference values are obtained from the observable subset of the network.

After calibration, the residual error is

$$r_i = \widetilde{x}_i - x_{\text{ref}}, \qquad \mu_i = \mathbb{E}[r_i], \qquad \sigma_i^2 = \text{Var}(r_i). \tag{6.18}$$

The total residual mean-squared error is then

$$q_i = \mathbb{E}[r_i^2] = \mu_i^2 + \sigma_i^2. \tag{6.19}$$

This quantity accounts for remaining bias as well as random error after calibration.

Fusion Weights

The calibrated estimates are combined as

$$x_{f,\mathrm{cal}} = w_1\widetilde{x}_1 + w_2\widetilde{x}_2, \qquad w_1 + w_2 = 1. \tag{6.20}$$

Following the weighting rule adopted in the case study, the error-variance terms in classical MRC are replaced with the residual mean-squared errors q_i:

$$w_i = \frac{1/(\mu_i^2 + \sigma_i^2)}{1/(\mu_1^2 + \sigma_1^2) + 1/(\mu_2^2 + \sigma_2^2)}, \qquad i = 1, 2. \tag{6.21}$$

The final estimate becomes

$$x_{f,\mathrm{cal}} = w_1(a_1x_1 + b_1) + w_2(a_2x_2 + b_2). \tag{6.22}$$

Compared with the conventional MRC rule, the adopted weighting scheme penalizes estimates with larger post-calibration residual biases as well as larger residual variances. The inverse-residual-MSE rule follows the proposed implementation; strict minimum-MSE optimality requires additional assumptions on the joint residual statistics because BWLS and VBMC errors can be correlated.

The procedure is applied to corresponding estimated state components so that both estimators contribute to a common voltage-state representation. The calibration parameters and fusion weights can vary with the component and with the calibration scheme selected for the implementation. To avoid information leakage, calibration parameters should be estimated from available calibration observations rather than from reference values reserved for final testing.

6.3 Case Study

6.3.1 Feeder and Measurements

The proposed framework is evaluated using a real-world distribution feeder located in Kansas, USA. The case study considers a mixed-phase 210-bus distribution network, as illustrated in Fig. 6.1. The feeder includes single-phase and three-phase sections with residential customer loads connected through distribution transformers. The electrical network is modeled in OpenDSS, where power flow analysis is performed to obtain the reference voltage magnitudes and angles at different buses and phases. These reference values are used to evaluate the accuracy of the proposed distribution system state estimation methods.

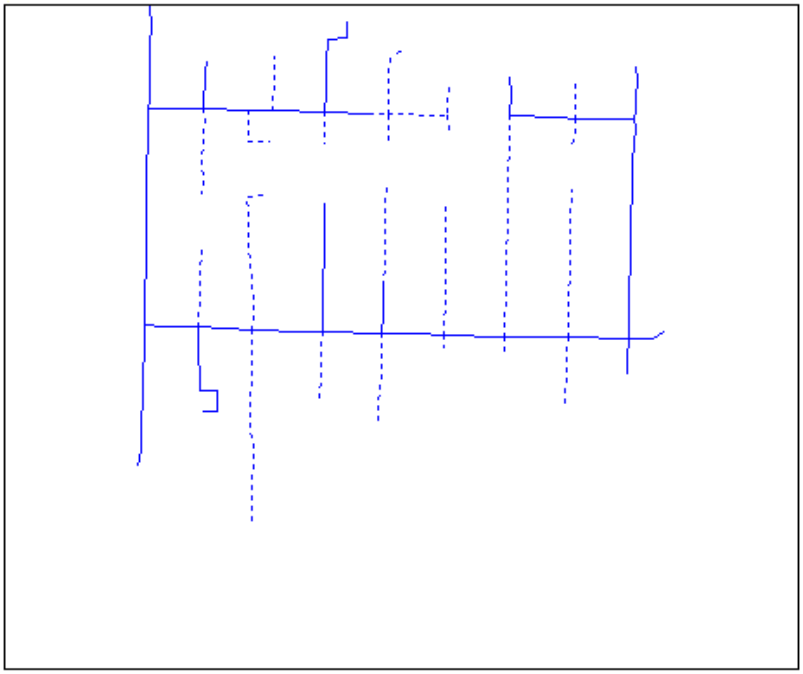


Figure 6.1: *Single-line diagram of the distribution feeder used in the case study.*

The study uses 15-minute advanced metering infrastructure (AMI) measurements col-

lected between October 25, 2020, and March 10, 2024. Following the data preprocessing procedure described in Chapter 5, 646 individual residential load profiles are retained. These loads are also aggregated according to their associated distribution transformers, resulting in 58 transformer-level load groups. Both individual and aggregated load forecasts are used to evaluate the effect of load aggregation on state estimation accuracy.

To represent a low-observability distribution network, it is assumed that only 20% of the AMI meters are available for near-real-time polling at each estimation instant, while complete AMI measurements are retrieved every 24 hours. The remaining 80% of the AMI observations are reconstructed using forecasting-based pseudo-measurements and matrix completion. The available measurements and pseudo-measurements are used in the Bayesian Weighted Least Squares (BWLS) and Variational Bayesian Matrix Completion (VBMC) estimation procedures. The resulting voltage estimates are subsequently combined using conventional and calibrated Maximum Ratio Combining (MRC) to evaluate the improvement in state estimation accuracy under limited measurement availability.

6.3.2 Scenario 1: Aggregated Loads

The first scenario represents customer demand at the distribution-transformer level. Suppose transformer g serves the customer set $\mathcal{C}_g$. The aggregated active and reactive powers at time t are

$$P_g(t) = \sum_{j \in \mathcal{C}_g} P_j(t), \qquad Q_g(t) = \sum_{j \in \mathcal{C}_g} Q_j(t), \tag{6.23}$$

with the aggregation defined at the selected electrical reference point. The forecasting analysis in Chapter 5 considers 58 aggregated load groups, and the corresponding group forecasts can be used to construct transformer-level pseudo-measurements for unavailable AMI observations.

The principal motivation for aggregation is to reduce the number of forecasting profiles. Training and executing a separate forecasting model for a large number of customer-level demand profiles can require substantial time and computing resources. Grouping loads connected to the same distribution transformer reduces the number of independent forecasting

series and the number of pseudo-measurements handled at the aggregated spatial resolution. Aggregation can also reduce the influence of highly irregular variations in individual customer demand, although any improvement in forecast accuracy or estimation accuracy must be established from the observed results rather than assumed.

For state estimation, each aggregated load is mapped to the corresponding transformer connection in the network model. The transformer primary-side voltage is used as the relevant voltage quantity when reporting transformer-level results. This distinction is necessary because the transformer secondary-side customer voltage is not generally identical to the primary voltage: differences depend on transformer characteristics and feeder operating conditions. The aggregation is therefore an intentional change in spatial modeling resolution, not merely an arithmetic reduction of the number of customer records.

The same BWLS–VBMC–calibrated-MRC pipeline is then applied to the network using the aggregated demand representation. Measurement and pseudo-measurement variances are assigned according to the data available at the aggregated level. The voltage profiles, estimator comparisons, and daily estimation errors obtained for the aggregated-load scenario are presented in Section 6.4.1.

6.3.3 Scenario 2: Individual Loads

The second scenario evaluates BWLS state estimation using individual customer-load forecasts. In this configuration, the customer-level forecasts are mapped to their corresponding locations in the distribution network, and forecast-based pseudo-measurements are incorporated with their assigned uncertainties. To represent a low-observability distribution network, 20% of the AMI meters were randomly selected at each estimation instant to provide near-real-time measurements, while complete AMI measurements were retrieved every 24 hours. The results are compared with the aggregated-load BWLS case to assess the effect of load representation on voltage-state estimation accuracy.

Individual-load forecasts retain customer-level variations that are averaged in transformer-aggregated profiles. They also require a larger number of separate forecasting series. The

numerical comparison in Section 6.4.2 reports the BWLS voltage-estimation errors for the two load representations. Computational time is not compared because corresponding timing measurements are not reported.

6.4 Results

6.4.1 Scenario 1: Aggregated Loads

The performance of the individual estimation methods and their fusion is assessed through the voltage profiles of phases A, B, and C, with the corresponding reference voltages used to evaluate the reconstructed states. The discussion first presents the results obtained using variational Bayesian matrix completion and Bayesian weighted least squares separately. The standard maximum ratio combining and calibrated MRC results are then examined to illustrate the effect of estimator fusion and bias correction.

The voltage-profile figures correspond to 13:00 on July 8, 2023, whereas the numerical snapshot comparison in corresponds to 16:00 on the same day. These operating instants are kept distinct.

The voltage-profile figures correspond to 13:00 on July 8, 2023, whereas the numerical snapshot comparison in Table 6.1 corresponds to 16:00 on the same day. These operating instants are kept distinct.

VBMC

Figures 6.2, 6.3, and 6.4 illustrate the voltage magnitudes and angles reconstructed by VBMC for phases A, B, and C, respectively. In each figure, the upper panel compares the estimated voltage magnitude with the reference profile and shows the estimated 95% uncertainty interval, while the lower panel presents the corresponding voltage angles. The observed voltage measurements are marked separately from the reference and reconstructed profiles. The VBMC magnitude estimates generally follow the shape of the reference voltages, although visible discrepancies remain at some bus locations. The width of the uncertainty band varies

along the feeder, indicating that the estimated uncertainty is not uniform across the reconstructed states. The offset between the estimated and reference magnitudes is most visible in phase A, whereas the phase-B estimate is nearly indistinguishable from the reference. In phase C, the estimate only partially captures the voltage depression over the middle bus-phase indices. The reconstructed voltage angles remain close to the reference values of approximately $0°$, $-120°$, and $+120°$ for phases A, B, and C, respectively, and their errors are not visible at the scale of the plots. These results illustrate the ability of matrix completion to recover voltage information from partial observations, while highlighting the residual errors and uncertainty of the standalone estimator.

BWLS

The BWLS results are presented in Figures 6.5, 6.6, and 6.7 for phases A, B, and C, respectively. In each figure, the upper panel shows the estimated and reference voltage magnitudes together with the observed measurements and the estimated uncertainty band, while the lower panel compares the voltage angles. BWLS incorporates the available measurements and forecast-based pseudo-measurements through the network measurement model.

In phase A, the voltage-magnitude estimate follows the general feeder trend, but departures from the reference profile are visible at several bus locations. For phase B, the estimated voltage magnitude remains close to the reference profile over most bus locations, with small deviations around the middle indices. The estimated voltage angles also follow the reference values, with limited differences along the feeder. In phase C, the estimated voltage magnitude generally follows the reference profile, although some deviations are observed around the middle bus-phase indices.

Comparison with the corresponding VBMC results in Figures 6.2, 6.3, and 6.4 shows that the two standalone estimators exhibit different spatial error patterns. These differences provide the motivation for combining their outputs rather than relying on either estimator alone.

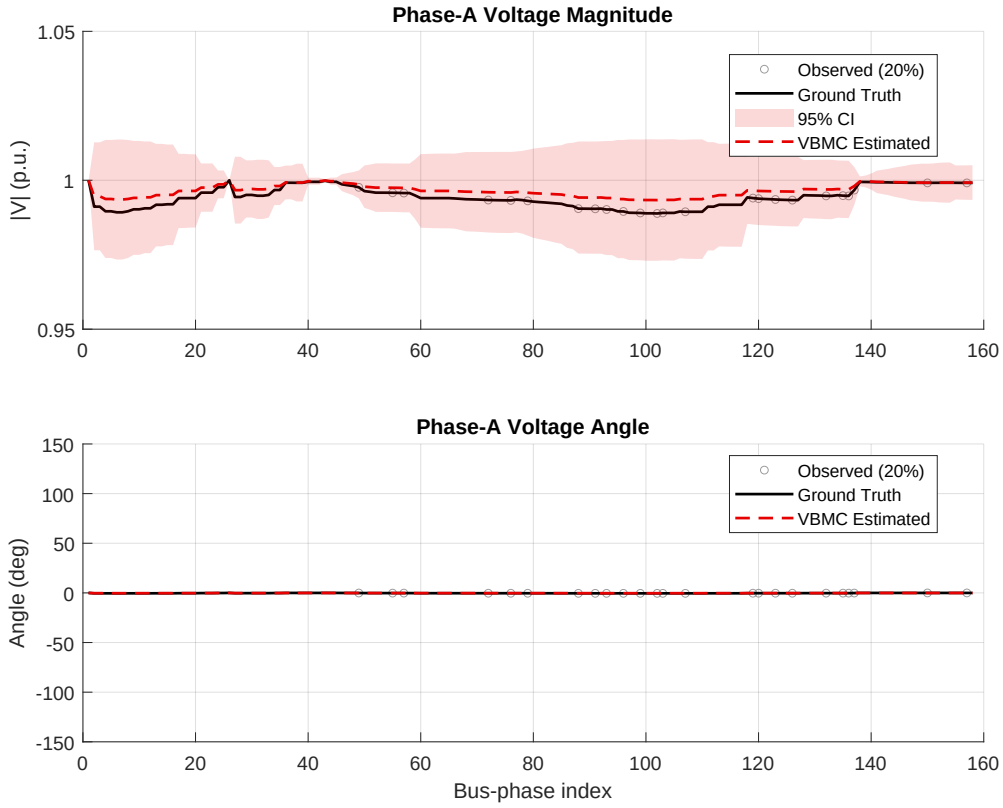


Figure 6.2: *Phase-A voltage magnitude and angle estimated by VBMC at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated 95% uncertainty interval.*

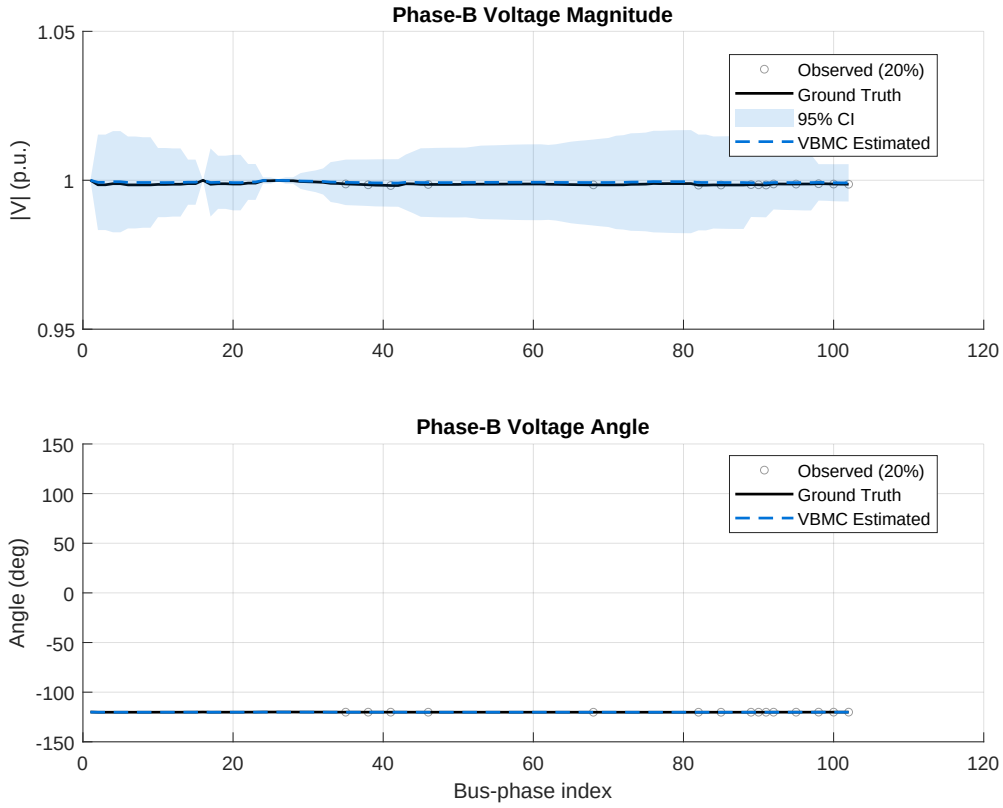


Figure 6.3: *Phase-B voltage magnitude and angle estimated by VBMC at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated 95% uncertainty interval.*

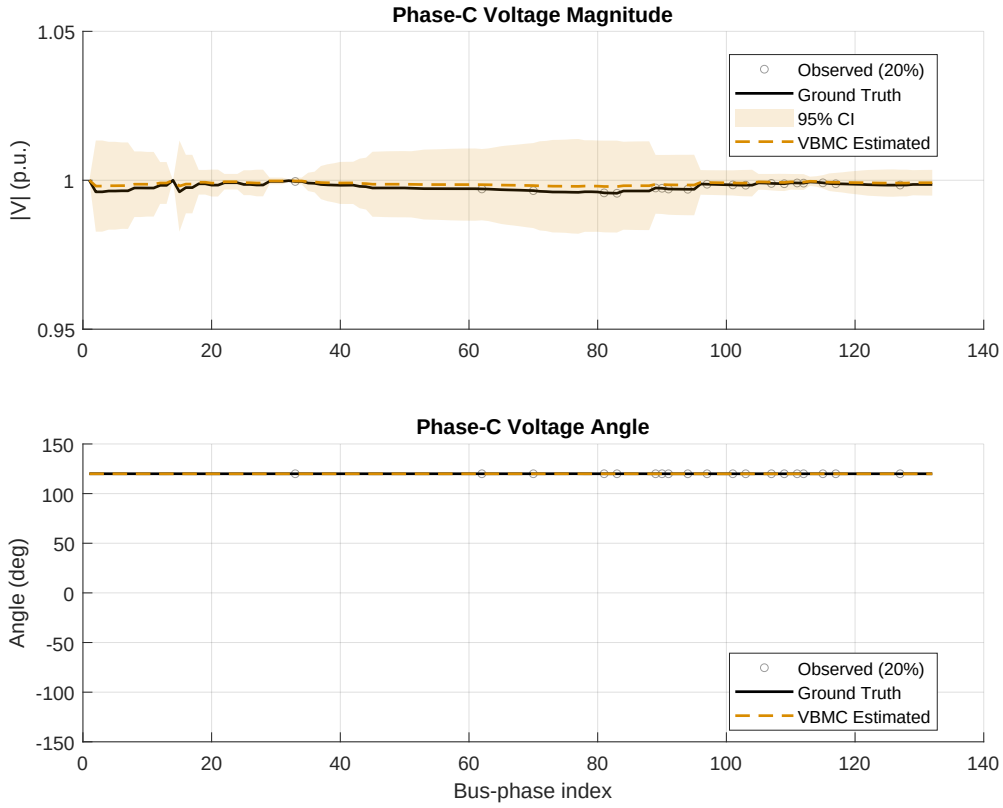


Figure 6.4: *Phase-C voltage magnitude and angle estimated by VBMC at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated 95% uncertainty interval.*

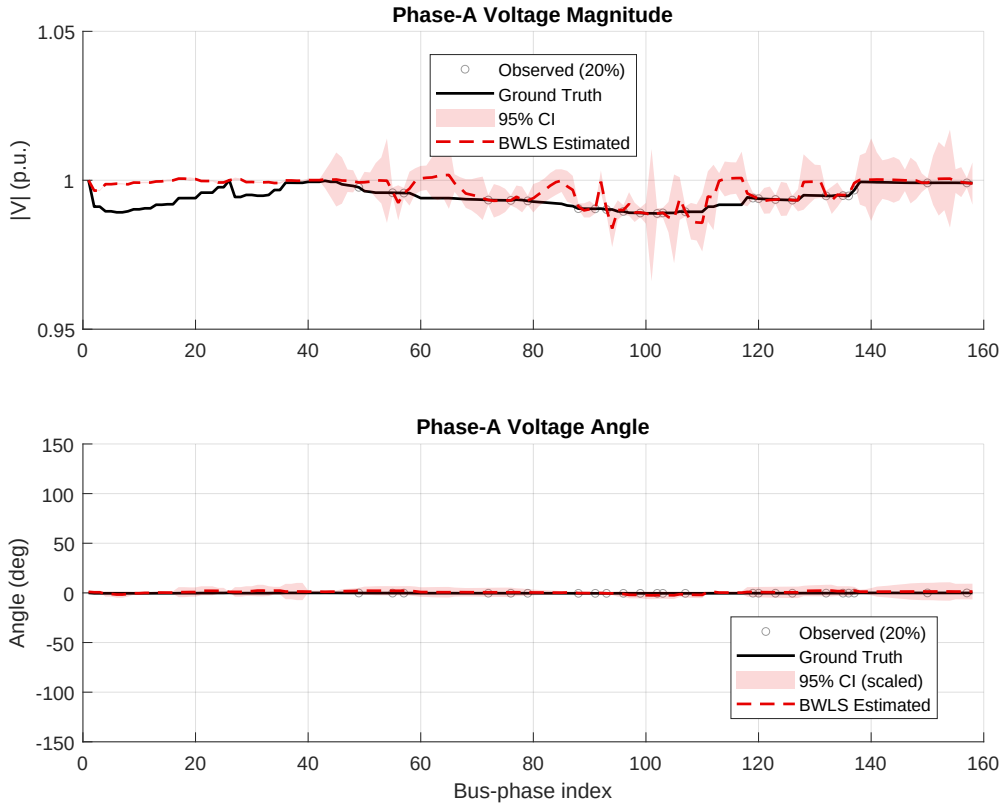


Figure 6.5: *Phase-A voltage magnitude and angle estimated by BWLS at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated uncertainty interval.*

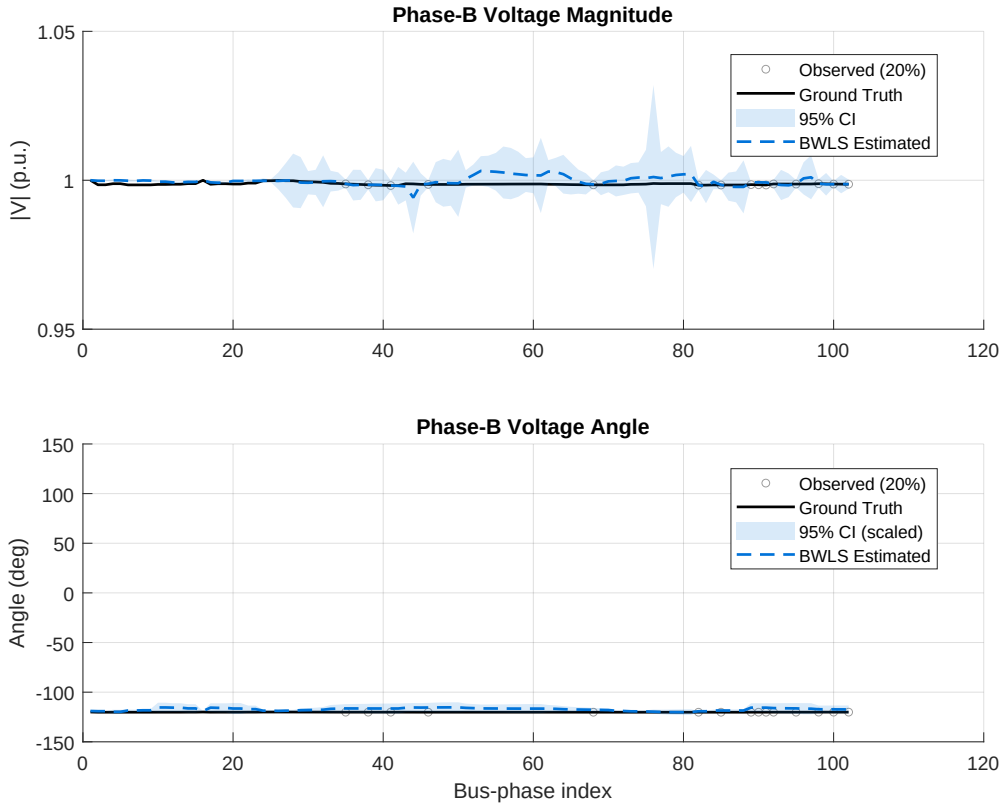


Figure 6.6: *Phase-B voltage magnitude and angle estimated by BWLS at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated uncertainty interval.*

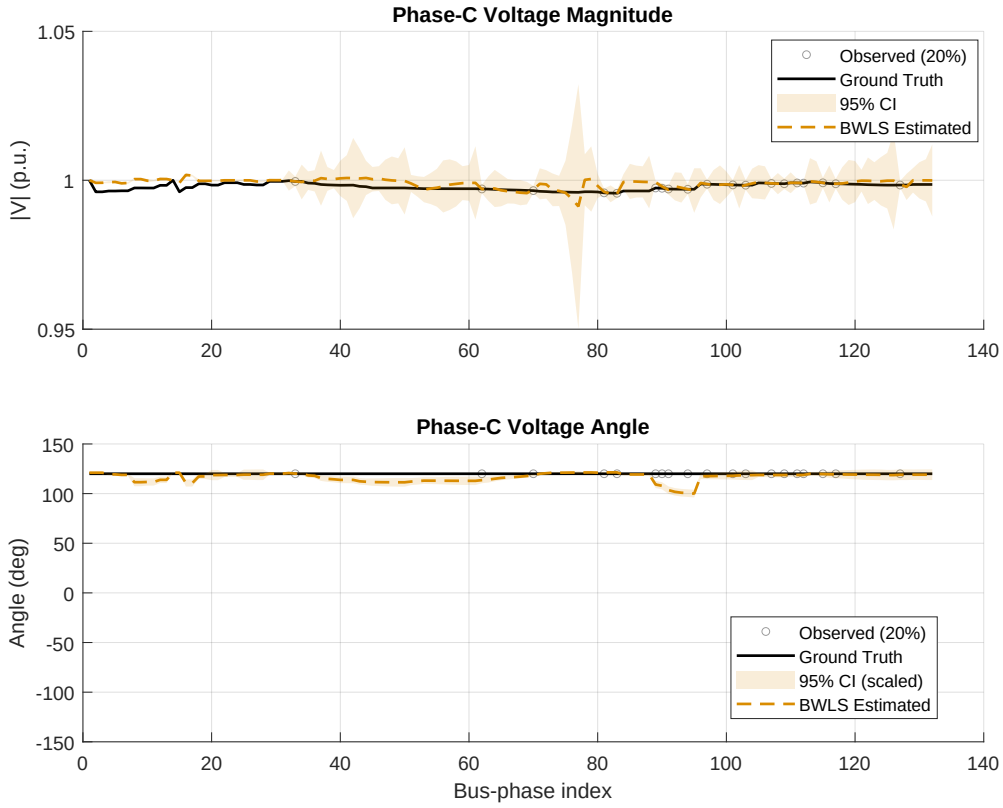


Figure 6.7: *Phase-C voltage magnitude and angle estimated by BWLS at 13:00 on July 8, 2023. The magnitude panel compares the reference, observed, and estimated voltages and displays the estimated uncertainty interval.*

Standard MRC

The standard MRC approach combines the BWLS and VBMC estimates using the variance-based weights introduced earlier. Figures 6.8, 6.9, and 6.10 compare the voltage magnitudes estimated by BWLS, VBMC, and standard MRC for phases A, B, and C, respectively, at 13:00 on July 8, 2023.

For phase A, the fused voltage profile closely follows the BWLS estimate over most bus locations, while the VBMC estimate provides a smoother voltage profile. The differences between BWLS and VBMC are particularly visible in the middle portion of the feeder, where the BWLS estimate exhibits several voltage variations that are not captured by VBMC. For phase B, the fused estimate generally follows BWLS, including the voltage depression around node index 42, although the magnitude of this deviation is reduced. Similarly, for phase C, the fused profile follows the BWLS estimate in most locations, with some differences around the voltage peaks and depressions.

Overall, the standard MRC results show that the fused estimate is influenced more strongly by BWLS at most bus locations. Although the variance-based weighting combines the information provided by both estimators, it does not explicitly account for their estimation biases. Therefore, the standard MRC approach may retain systematic errors from the individual estimators, motivating the use of calibrated MRC.

Calibrated MRC

Figures 6.11, 6.12, and 6.13 present the voltage profiles after bias-aware calibration and fusion for phases A, B, and C, respectively, at 13:00 on July 8, 2023. The figures compare the reference voltage magnitudes with the standalone BWLS and VBMC estimates and the calibrated MRC results.

For phase A, the BWLS estimate exhibits local deviations from the reference profile, particularly around the middle bus indices. The VBMC estimate provides a smoother voltage profile but does not capture all voltage variations. The calibrated MRC estimate closely follows the reference profile across most bus locations, including regions where the standalone

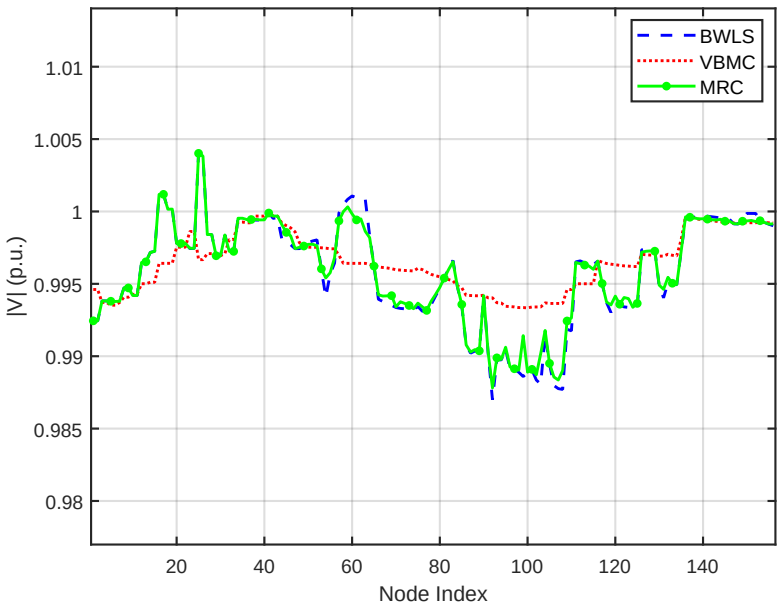


Figure 6.8: *Phase-A voltage magnitudes estimated by BWLS, VBMC, and standard MRC at 13:00 on July 8, 2023.*

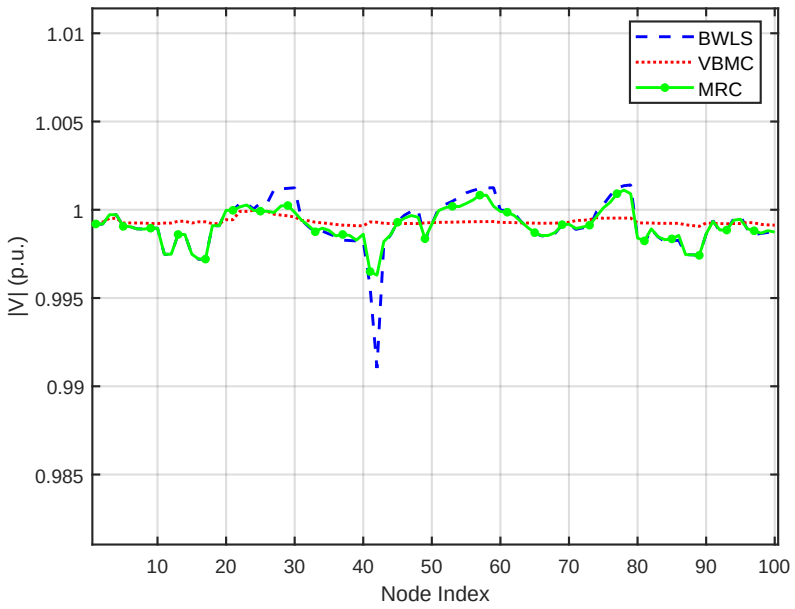


Figure 6.9: *Phase-B voltage magnitudes estimated by BWLS, VBMC, and standard MRC at 13:00 on July 8, 2023.*

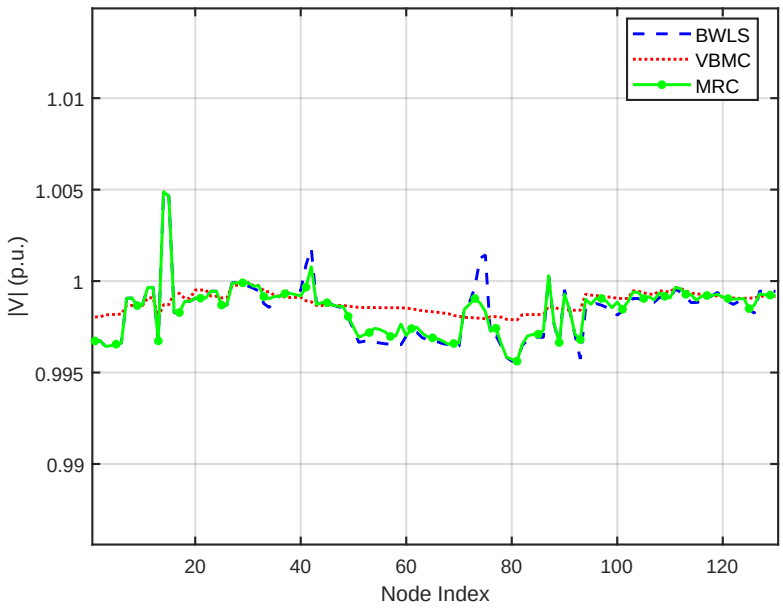


Figure 6.10: *Phase-C voltage magnitudes estimated by BWLS, VBMC, and standard MRC at 13:00 on July 8, 2023.*

estimators show noticeable differences. For phase B, the calibrated MRC estimate remains close to the reference voltage profile, while the BWLS estimate exhibits a noticeable voltage depression around node index 42. Similarly, for phase C, the calibrated MRC estimate follows the reference profile over most bus locations, reducing the influence of several local deviations observed in BWLS.

Comparison with the standard MRC results in Figures 6.8, 6.9, and 6.10 illustrates the effect of incorporating bias correction before estimator fusion. While standard MRC combines the estimates based on their reported uncertainties, calibrated MRC uses the available reference measurements to adjust the individual estimates and determine the fusion weights based on their residual errors. The resulting voltage profiles demonstrate the potential benefit of accounting for systematic estimation errors in the fusion process.

To further evaluate the performance of calibrated MRC at a different operating instant, Figures 6.14, 6.15, and 6.16 present the voltage profiles for phases A, B, and C, respectively, at 22:00 on the same day. Similar to the results at 13:00, the calibrated MRC estimates remain close to the reference voltage profiles across most bus locations. In phase A, the fused estimate follows the voltage variations along the feeder, while the standalone BWLS and VBMC estimates exhibit local deviations. For phase B, calibrated MRC closely follows the reference profile, including the region around node index 42, where BWLS shows a noticeable voltage depression. In phase C, the fused estimate also follows the reference profile and reduces the influence of local deviations observed in BWLS. These results indicate that calibrated MRC maintains good estimation accuracy at another operating instant, demonstrating its effectiveness beyond the initial snapshot.

For a separate operating snapshot at 16:00 on July 8, 2023, the numerical comparison is summarized in Table 6.1. BWLS yields a voltage-magnitude MAE of 0.00217 p.u. and RMSE of 0.00352 p.u.; the corresponding VBMC errors are 0.00356 p.u. and 0.00415 p.u. Standard MRC gives 0.00217 p.u. MAE and 0.00343 p.u. RMSE, whereas calibrated MRC achieves 0.00012 p.u. MAE and 0.00015 p.u. RMSE. On the basis of these rounded values, calibration reduces the MAE relative to standard MRC by approximately 94%. The table is included as an additional quantitative evaluation and should not be interpreted as numerical

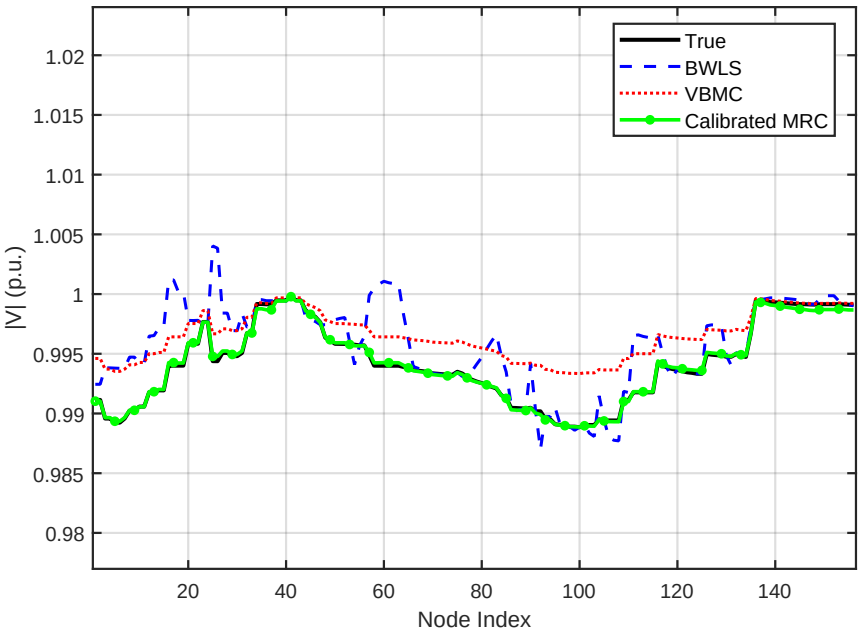


Figure 6.11: *Phase-A voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 13:00 on July 8, 2023.*

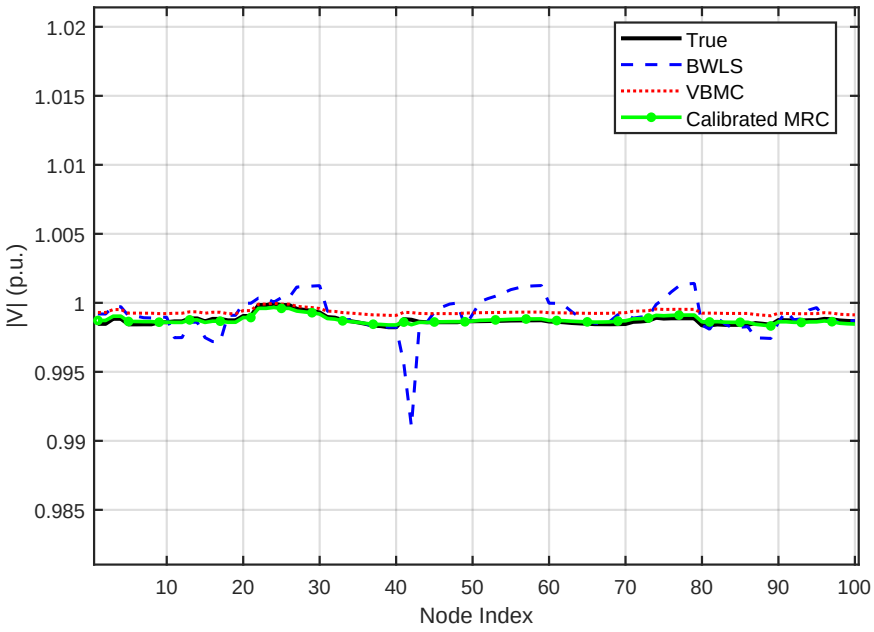


Figure 6.12: *Phase-B voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 13:00 on July 8, 2023.*

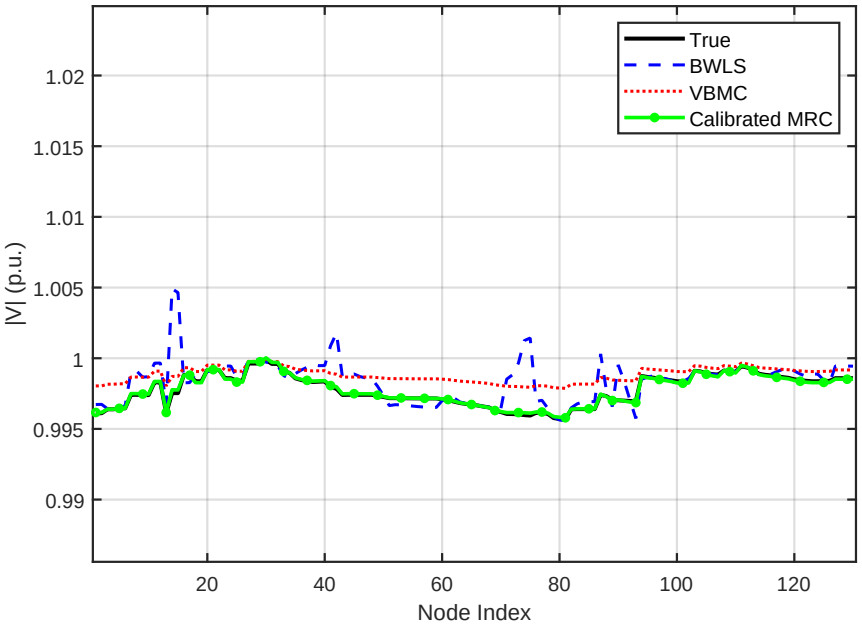


Figure 6.13: *Phase-C voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 13:00 on July 8, 2023.*

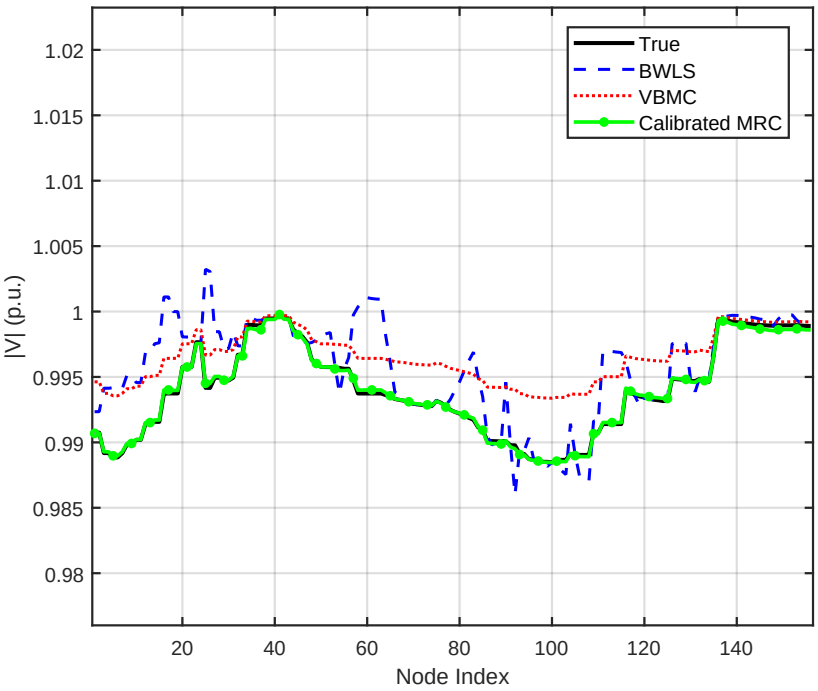


Figure 6.14: *Phase-A voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 22:00 on July 8, 2023.*

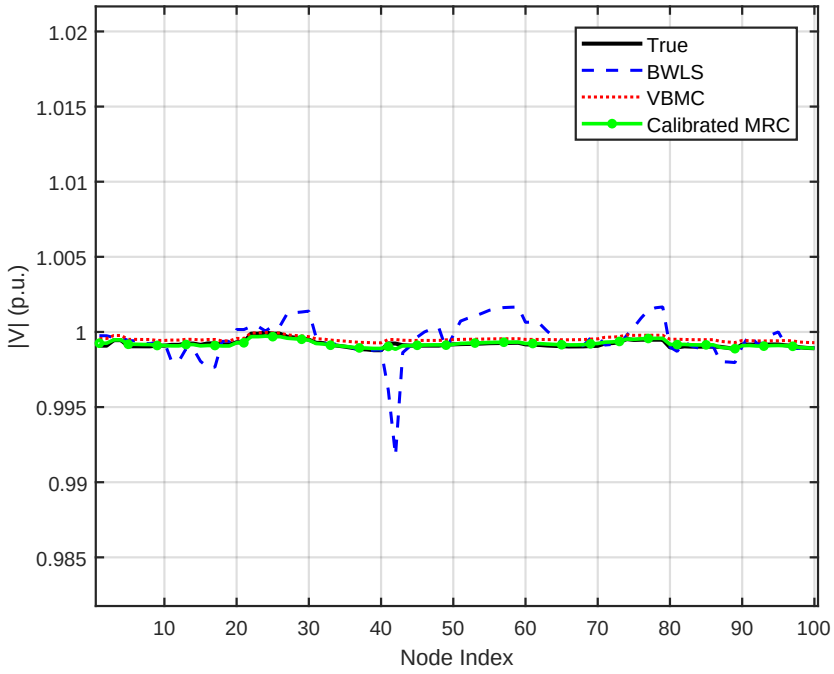


Figure 6.15: *Phase-B voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 22:00 on July 8, 2023.*

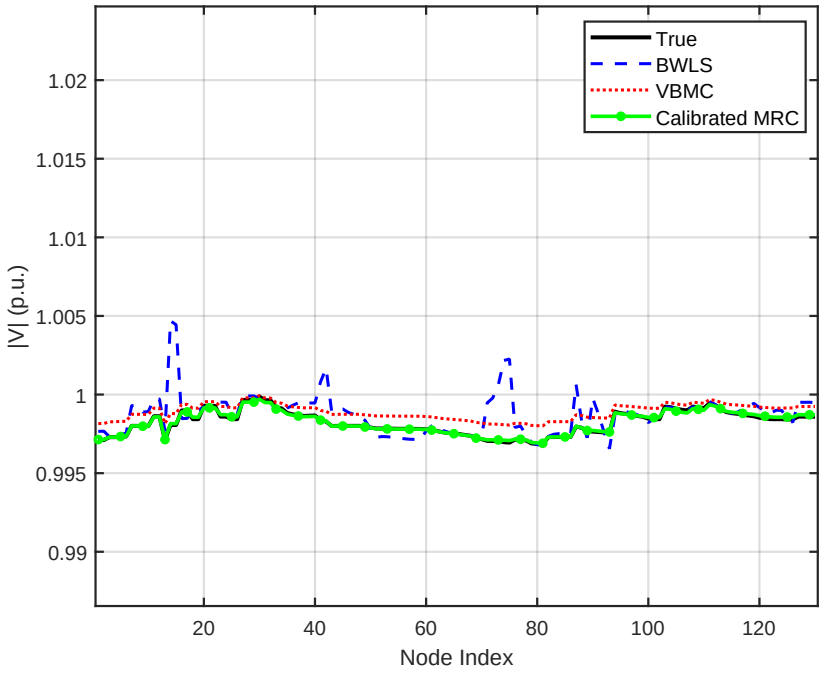


Figure 6.16: *Phase-C voltage magnitudes obtained from BWLS, VBMC, and calibrated MRC compared with the reference profile at 22:00 on July 8, 2023.*

metrics for the 13:00 and 22:00 plots.

Table 6.1: *Phase-A voltage-magnitude estimation accuracy for the feeder snapshot at 16:00 on July 8, 2023.*

Estimator	MAE (p.u.)	RMSE (p.u.)
BWLS	0.00217	0.00352
VBMC	0.00356	0.00415
Standard MRC	0.00217	0.00343
Calibrated MRC	0.00012	0.00015

Observability and Noise

The effectiveness of a state estimator in practical operation depends on both measurement availability and measurement quality. To assess these effects, the case study evaluates calibrated MRC under AMI observability levels of 5%, 10%, 20%, 30%, and 50%, with measurement-noise levels of 0%, 2%, and 5%. Table 6.2 summarizes the resulting voltage-estimation errors.

Table 6.2: *Calibrated MRC performance under different AMI observability and measurement-noise levels, in the feeder case study.*

Observability (%)	Noise (%)	MAE (p.u.)	RMSE (p.u.)
5	0	0.00007	0.00010
5	2	0.00030	0.00051
5	5	0.00447	0.00788
10	0	0.00005	0.00008
10	2	0.00039	0.00057
10	5	0.00081	0.00133
20	0	0.00005	0.00010
20	2	0.00019	0.00040
20	5	0.00055	0.00111
30	0	0.00004	0.00007
30	2	0.00015	0.00035
30	5	0.00041	0.00077
50	0	0.00003	0.00007
50	2	0.00009	0.00025
50	5	0.00025	0.00058

The results show that the proposed calibrated fusion maintains small estimation errors over several measurement conditions. As observability increases, the errors generally

decrease. Nevertheless, the combination of very low observability and relatively high measurement noise can substantially degrade performance: at 5% observability and 5% noise, the MAE rises to 0.00447 p.u. and the RMSE to 0.00788 p.u. This case illustrates that calibration does not eliminate the fundamental sensitivity of DSSE to insufficient or unreliable measurements. Thus, the results support the usefulness of calibrated fusion while also emphasizing the importance of measurement placement and data quality.

Daily Performance

The preceding voltage-profile comparisons characterize estimator behavior at selected operating instants. To investigate estimation accuracy over a full day, Fig. 6.17 presents the phase-A voltage-magnitude mean absolute error (MAE) as a function of time for BWLS, VBMC, and bias-aware MRC. The individual-estimator errors vary appreciably across the day, reflecting changes in the operating conditions and the quality of the reconstructed states. In contrast, the bias-aware fused estimate maintains a substantially smaller MAE throughout the plotted period. This temporal comparison complements the snapshot results by showing that the improvement obtained through calibration and fusion extends across the evaluated 24-hour interval.

6.4.2 Scenario 2: Individual Loads

To evaluate the effect of load aggregation on distribution system state estimation, the proposed BWLS methodology was also applied using individual customer load forecasts. Unlike the aggregated-load scenario, where customer demands were combined at the distribution-transformer level, this scenario employed separate forecasts for 646 individual loads. Since the two scenarios have different network resolutions, their voltage-state estimation accuracy was compared at the common buses on the primary side of the distribution transformers.

Table 6.3 presents the overall BWLS voltage-estimation accuracy at these common primary-side buses. The voltage-magnitude MAE increases slightly from 0.0020 p.u. for aggregated loads to 0.0022 p.u. for individual loads, while the voltage-magnitude MAPE changes from

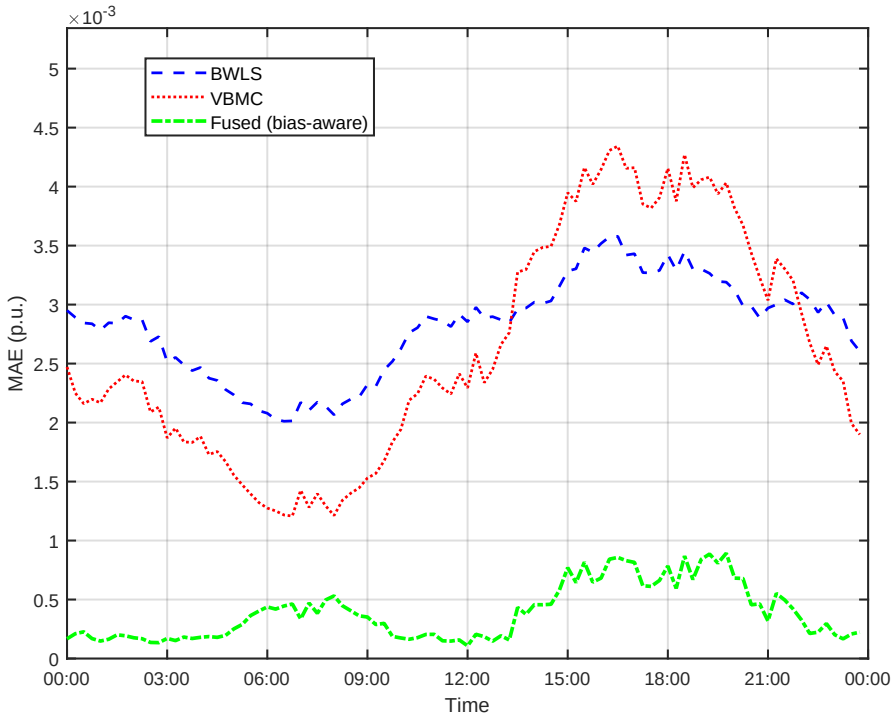


Figure 6.17: *Phase-A voltage-magnitude MAE over a 24-hour period for BWLS, VBMC, and bias-aware MRC in the aggregated-load case.*

0.20% to 0.21%. Similarly, the voltage-angle MAE remains nearly unchanged, increasing from 1.024° to 1.026°. These results indicate that both load representations provide comparable voltage-state estimation accuracy at the common primary-side buses under the evaluated operating conditions.

Table 6.3: *Overall BWLS state estimation accuracy at common primary-side buses using aggregated and individual load forecasts.*

Metric	Aggregated Loads	Individual Loads
Voltage magnitude MAE (p.u.)	0.0020	0.0022
Voltage magnitude MAPE (%)	0.20	0.21
Voltage angle MAE (°)	1.024	1.026

Table 6.4 presents the phase-wise BWLS voltage-estimation accuracy at the common primary-side buses for both scenarios. For Phase A, the voltage-magnitude MAE changes from 0.00320 p.u. to 0.00324 p.u., while the voltage-angle MAE remains unchanged at 1.170°. Phases B and C also exhibit minor differences in magnitude and angle estimation errors. These findings suggest that the additional forecasting variability associated with individual customer loads has a limited effect on BWLS estimation accuracy at the common primary-side buses.

Table 6.4: *Phase-wise BWLS state estimation accuracy at common primary-side buses using aggregated and individual load forecasts.*

Phase	Load Type	Magnitude MAE (p.u.)	Magnitude MAPE (%)	Angle MAE (°)
A	Aggregated	0.00320	0.32	1.170
A	Individual	0.00324	0.32	1.170
B	Aggregated	0.0009	0.09	0.300
B	Individual	0.0010	0.10	0.309
C	Aggregated	0.0014	0.14	1.410
C	Individual	0.0015	0.15	1.418

An important advantage of transformer-level aggregation is the reduction in the number of forecasting profiles from 646 individual loads to 58 aggregated groups, corresponding to approximately 91.0% fewer forecasting tasks. However, the aggregated-load representation does not provide voltage estimates at individual secondary-side buses. In contrast, retaining

the individual loads and their corresponding network representation enables voltage-state estimation at these locations.

To further evaluate the secondary-side estimation accuracy, calibrated MRC was applied to the individual-load scenario. Fig. 6.18 presents the voltage-magnitude MAE across 353 secondary-side bus-phase nodes, averaged over 96 timestamps on July 8, 2023. The results show that the estimation errors vary across the secondary-side nodes, with most MAE values remaining below 0.005 p.u. and a few nodes exhibiting relatively higher errors.

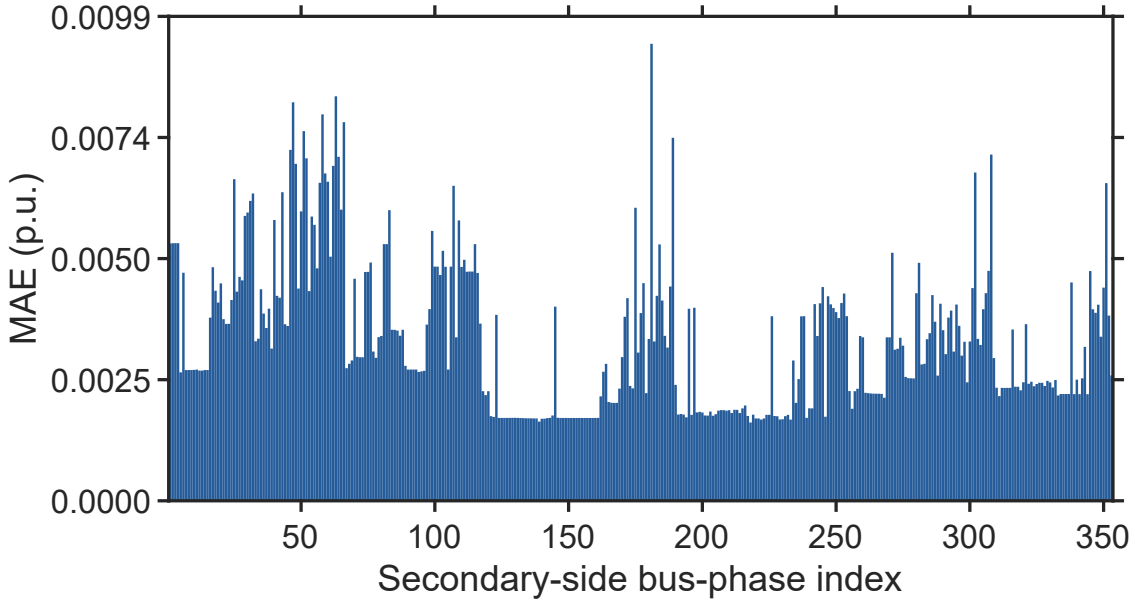


Figure 6.18: *Voltage-magnitude MAE of calibrated MRC across 353 secondary-side bus-phase nodes, averaged over 96 timestamps on July 8, 2023.*

To provide a clearer assessment of the error distribution, Fig. 6.19 presents the number of secondary-side bus-phase nodes within four MAE ranges. Among the 353 nodes, 141 have an MAE of 0.0025 p.u. or less, while 168 fall between 0.0025 and 0.005 p.u. Accordingly, 309 nodes, representing approximately 87.5% of the total, have an MAE of 0.005 p.u. or less. Of the remaining 44 nodes, 38 have an MAE between 0.005 and 0.0075 p.u., while only six exhibit errors between 0.0075 and 0.01 p.u. These results indicate that calibrated MRC provides relatively low voltage-magnitude estimation errors at most secondary-side nodes under limited real-time AMI observability.

Overall, the results demonstrate that transformer-level aggregation substantially reduces

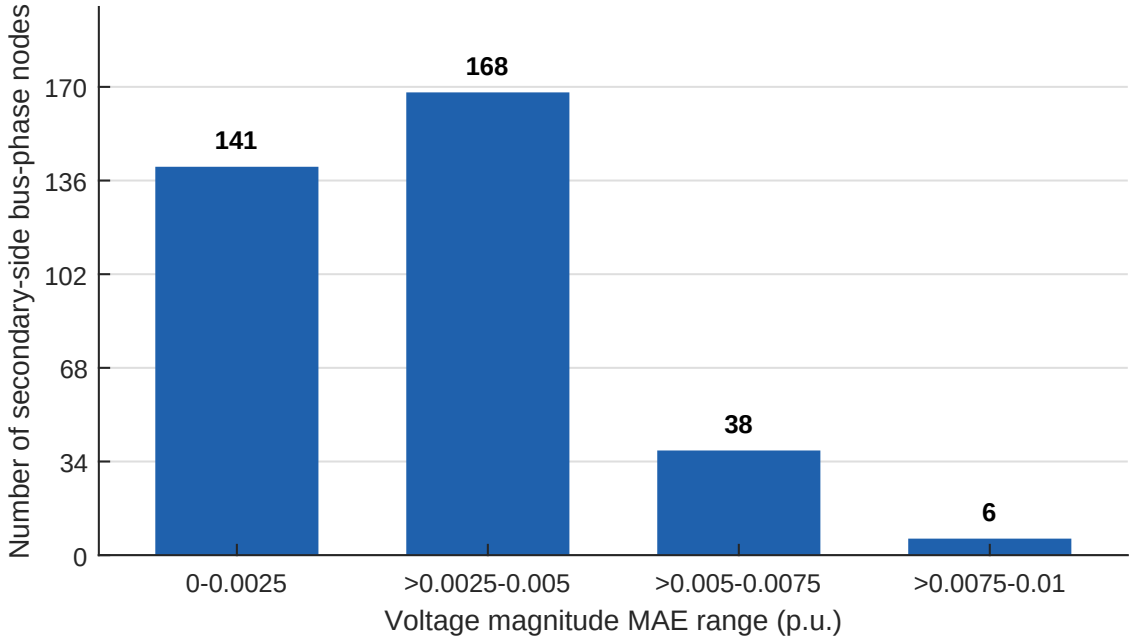


Figure 6.19: *Distribution of secondary-side bus-phase nodes across four voltage-magnitude MAE ranges using calibrated MRC, averaged over 96 timestamps on July 8, 2023.*

the number of forecasting profiles while maintaining comparable BWLS voltage-estimation accuracy at common primary-side buses. However, individual-load modeling enables secondary-side voltage estimation, providing additional visibility into the distribution network. Therefore, the selection between aggregated and individual load representations depends on the required network resolution and forecasting requirements.

6.5 Summary

This chapter developed a distribution system state estimation framework that combines two complementary estimators through calibrated maximum ratio combining. Bayesian Weighted Least Squares (BWLS) incorporates observed measurements and forecast-based pseudo-measurements into a nonlinear power-system model, with weights determined by measurement uncertainty. Variational Bayesian Matrix Completion (VBMC) exploits the relationships among partially observed AMI measurements to reconstruct missing information and provide a complementary state estimate. The resulting estimates are combined using maximum ratio combining (MRC). To address the limitations of conventional MRC

in the presence of systematic estimation errors, a calibrated fusion method was developed that corrects estimator biases and assigns fusion weights based on the residual mean-squared errors.

The proposed methodology was evaluated using a distribution feeder in Kansas under limited near-real-time AMI observability. The results demonstrated that calibrated MRC substantially improves voltage estimation accuracy compared with the individual BWLS and VBMC estimators and conventional MRC. At the evaluated operating instant, the calibrated method reduced the phase-A voltage-magnitude MAE from 0.00217 p.u. for conventional MRC to 0.00012 p.u. Furthermore, the analysis under different measurement observability and noise levels demonstrated the sensitivity of estimation performance to measurement availability and quality. The 24-hour evaluation also showed that calibrated MRC consistently achieved smaller voltage-estimation errors than the individual estimators across the evaluated operating conditions.

To investigate the effect of load representation on state estimation, two scenarios were considered using transformer-aggregated and individual customer load forecasts. The BWLS results showed comparable voltage-estimation accuracy at the common primary-side buses, with voltage-magnitude MAEs of 0.0020 p.u. and 0.0022 p.u., respectively. Transformer-level aggregation reduced the number of forecasting profiles from 646 individual loads to 58 aggregated groups, representing approximately a 91.0% reduction in forecasting tasks.

However, individual-load modeling enables voltage-state estimation at secondary-side buses, which are not represented in the aggregated-load scenario. The calibrated MRC results showed that approximately 87.5% of the 353 secondary-side bus-phase nodes had a voltage-magnitude MAE of 0.005 p.u. or less. These findings highlight the trade-off between reducing forecasting requirements through load aggregation and providing additional network visibility through individual-load modeling.

Chapter 7

Conclusion and Future Work

7.1 Conclusions

The increasing integration of renewable energy resources and advanced monitoring technologies has created new opportunities for improving the planning and operation of power distribution systems. However, uncertainty in solar generation and electricity demand, limited availability of near-real-time measurements, and the computational requirements of analyzing large volumes of metering data remain important challenges for distribution utilities. This dissertation investigated data-enabled approaches to address these challenges through two related research directions: the optimal sizing of photovoltaic (PV) and battery energy storage systems, and the enhancement of distribution-system situational awareness through electricity demand forecasting and state estimation. The proposed methodologies combine statistical modeling, optimization, deep learning, and probabilistic estimation to support both long-term planning and operational decision-making.

The first part of this dissertation, presented in Chapter 4, developed two statistical approaches for the optimal sizing of grid-connected PV–battery systems. The main objective was to identify economically viable combinations of PV generation and battery capacity that could reduce daily peak demand with a specified probability. Unlike sizing procedures that seek to flatten the largest observed peak, the proposed framework considered the distribu-

tion of updated daily peaks after integrating PV generation and battery storage. In the first approach, candidate probability distributions were fitted to the updated daily peaks, and the 95th-percentile peak demand was used to determine PV–battery combinations that satisfy a prescribed demand threshold. In the second approach, the relationship between battery capacity and the corresponding 95th-percentile peak was examined to identify the point beyond which additional storage provides only a marginal reduction in peak demand. Both approaches were coupled with an economic analysis that accounted for installation and replacement costs, operation and maintenance, energy purchases, and peak-demand charges over the project lifetime.

The proposed sizing approaches were evaluated using three years of hourly electricity demand and solar irradiance measurements collected from a small town in Kansas, USA. The results confirmed that the dominant daily peaks frequently occur outside the period of maximum solar generation, which limits the ability of PV installations alone to reduce peak-demand charges. For the prescribed 2000 kW peak threshold, the first approach identified a 2000 kW PV system with a 4000 kWh battery as the most beneficial configuration among its evaluated candidates, providing a present-worth benefit of $812,648 over 20 years. The second approach identified a 1200 kW PV system with a 3200 kWh battery, with a corresponding present-worth benefit of $820,373. To evaluate the sensitivity of these solutions to the variability of electricity demand and irradiance, a modified Monte Carlo procedure was developed. By clustering daily load and solar irradiance profiles on a monthly basis and sampling their historical relationships, the procedure generated paired scenarios that preserved their observed dependence. The resulting simulations supported the economic performance of the selected PV–battery configurations under the modeled variations in operating conditions. These findings demonstrate the importance of considering both the probabilistic characteristics of daily peaks and the utility's electricity purchasing structure when determining PV–battery sizes.

Chapter 5 addressed the challenge of estimating unavailable electricity demand measurements through data-driven forecasting. A Seasonal Periodic Decomposition Network (SPDNet) was developed to capture the temporal characteristics of electricity demand at

different time scales. The architecture combines a Seasonal–Trend Decomposition Module with a Periodical Decomposition Module. The latter identifies dominant periods through frequency analysis and processes the resulting representations using convolutional operations and a Transformer-based component. By combining information associated with seasonal variations, short-term fluctuations, periodic dependencies, and longer-term patterns, the forecasting framework generates future load estimates that can be used as pseudo-measurements in distribution system state estimation (DSSE).

The forecasting methodology was evaluated using historical advanced metering infrastructure (AMI) measurements collected from a distribution utility in a city in Kansas, USA. Following data preprocessing, 646 individual load profiles were retained and grouped into 58 transformer-level aggregated load profiles. The performance of SPDNet was examined over different input lengths and prediction horizons and compared with recurrent, convolutional, Transformer-based, and linear forecasting models. The results demonstrated favorable forecasting accuracy for the evaluated individual load profiles, while the aggregated-load analysis extended the application of SPDNet to all 58 transformer-level groups. In addition to point forecasting, the chapter investigated the uncertainty associated with day-ahead predictions. Historical forecast errors were used to estimate time-dependent dispersion and assign uncertainty information to the forecasted pseudo-measurements. This provided a connection between data-driven load forecasting and the uncertainty-weighted state estimation methods investigated in the subsequent chapter.

The state-estimation framework, presented in the chapter on calibrated fusion for DSSE, developed a method for combining two complementary estimates under limited near-real-time AMI observability. Bayesian Weighted Least Squares (BWLS) used available measurements, forecast-based pseudo-measurements, and the distribution-network model to estimate voltage magnitudes and angles. Variational Bayesian Matrix Completion (VBMC) exploited structure in the partially observed measurement data to reconstruct missing information and obtain a second estimate of the system states. Conventional Maximum Ratio Combining (MRC) was employed to combine these estimates according to their error variances. However, variance-based fusion does not fully account for systematic estimation errors. To

address this limitation, a calibrated MRC formulation was developed in which the individual estimates are corrected using reference observations and combined using weights based on post-calibration residual mean-squared errors.

The estimation framework was evaluated on a distribution feeder in Kansas under different levels of AMI observability and measurement noise. The results showed that calibrated MRC can significantly improve voltage-estimation accuracy compared with conventional fusion. For the evaluated phase-A snapshot, the voltage-magnitude mean absolute error (MAE) decreased from 0.00217 p.u. for conventional MRC to 0.00012 p.u. for calibrated MRC. The evaluation over a 24-hour period further demonstrated lower voltage-magnitude errors for the calibrated fusion method than for the individual BWLS and VBMC estimators under the examined conditions. The sensitivity analysis also showed that estimation accuracy remains dependent on the availability and quality of measurements, particularly when very low observability is combined with substantial measurement noise.

An additional investigation examined the effect of forecasting resolution on the accuracy of BWLS state estimation. Two scenarios were considered: one in which customer loads connected to the same distribution transformer were aggregated before forecasting, and another in which separate forecasts were generated for individual customer loads. Although forecasting individual customer loads is more challenging because of their irregular consumption patterns, the resulting BWLS state-estimation errors were similar in both scenarios. The overall voltage-magnitude MAE was 0.0020 p.u. using aggregated loads and 0.0022 p.u. using individual loads, while the corresponding voltage-angle MAEs were 1.024° and 1.026°, respectively. Transformer-level aggregation reduced the number of forecasting profiles from 646 to 58, a reduction of approximately 91.2% in the number of series requiring forecasts. These findings indicate that aggregation can considerably reduce the scale of the forecasting problem without materially reducing BWLS voltage-estimation accuracy at the evaluated network resolution. The reduction in the number of forecasting profiles should not, however, be interpreted as an equivalent measured reduction in computational time.

Overall, this dissertation demonstrated how data-driven and statistical approaches can support complementary planning and operational tasks in power distribution systems. The

PV–battery sizing framework provides a means of balancing peak-demand reduction and economic benefit under uncertainty. The forecasting and state-estimation frameworks make use of historical and partially available AMI measurements to improve network visibility while accounting for prediction and estimation uncertainty. Together, these contributions provide a basis for developing more practical and computationally manageable approaches to planning and monitoring modern distribution networks. The major contributions of this dissertation can be summarized as follows:

- **Statistical PV–battery sizing under uncertainty:** Developed two probabilistic sizing approaches based on the distribution of daily peak demand. The first approach determines PV–battery configurations that satisfy a predefined peak-demand threshold with 95% probability, while the second identifies battery capacities based on the diminishing returns in peak-demand reduction. Both approaches incorporate economic analysis to evaluate the financial benefits of the selected configurations.

- **Seasonal and periodic electricity demand forecasting:** Developed SPDNet, a deep learning framework that combines seasonal–trend decomposition, periodic decomposition, convolutional operations, and Transformer-based modeling to capture temporal dependencies in electricity demand. The framework was evaluated using individual and transformer-aggregated AMI load profiles and extended to quantify forecasting uncertainty for generating pseudo-measurements.

- **Calibrated fusion-based distribution system state estimation:** Developed a calibrated MRC framework that combines BWLS and VBMC estimates under limited AMI observability. By accounting for systematic estimation errors and residual uncertainties, the proposed fusion method improves voltage-state estimation accuracy compared with conventional variance-based fusion.

- **Evaluation of individual and aggregated load representations for DSSE:** Investigated the effect of load aggregation on forecasting-based BWLS state estimation. The results demonstrated that transformer-level aggregation can substantially reduce

the number of forecasting profiles while maintaining comparable voltage-estimation accuracy at the evaluated network resolution.

7.2 Future Directions

The proposed methodologies provide several opportunities for further research. Although the results demonstrate their potential for the evaluated utilities and operating conditions, additional developments are needed to improve their applicability under changing network configurations, evolving electricity demand patterns, and more complex uncertainty sources. The following directions are particularly relevant extensions of the work presented in this dissertation.

- **Extending PV–battery sizing to more detailed operational constraints.** The sizing approaches developed in Chapter 4 primarily determine battery energy capacity based on the ability to reduce updated daily peaks. Future work can incorporate explicit battery power limits, state-of-charge trajectories, charging and discharging constraints, and degradation models that depend on the number and depth of cycles. These extensions would allow the estimated economic benefit to reflect more detailed battery operating conditions. In addition, considering time-varying tariffs and alternative demand-charge structures would help establish how the preferred PV–battery combination changes across utilities with different purchasing agreements.

- **Adapting PV–battery planning to evolving load and generation conditions.** The statistical sizing framework and modified Monte Carlo procedure were developed using historical electricity demand and solar irradiance profiles. Future studies can investigate how electrification, the adoption of electric vehicles, additional distributed generation, and long-term changes in weather patterns affect the distribution of daily peaks and the required battery capacity. Incorporating these changes through multiyear scenarios would make it possible to evaluate whether a configuration selected using historical conditions continues to satisfy the prescribed peak-reduction criterion

throughout its operating life.

- **Extending forecasting across distribution-network levels.** Chapter 5 considered individual customer loads and transformer-level aggregated loads. An important extension is to investigate demand forecasting at additional spatial resolutions, including feeder, substation, and utility levels. Forecasting models can be developed to exploit relationships between these levels rather than treating every series independently. Such an approach could also enforce consistency between individual forecasts and their aggregated totals. Evaluating the trade-offs among forecasting accuracy, model complexity, and computational time would provide a more systematic basis for selecting the most appropriate forecasting resolution for different distribution-system applications.

- **Improving uncertainty characterization for forecast-based pseudo-measurements.** The uncertainty procedure developed in Chapter 5 used historical forecast errors to estimate the dispersion associated with different prediction timestamps. Future work can investigate methods that capture changes in forecasting uncertainty across seasons, load levels, and unusual operating conditions. It would also be valuable to model temporal and spatial correlations between forecast errors instead of relying only on diagonal pseudo-measurement covariance matrices. Calibration and empirical coverage studies could then be used to assess whether the assigned uncertainty intervals provide an adequate representation of actual forecasting errors before these uncertainties are incorporated into DSSE.

- **Enhancing the robustness and scalability of calibrated estimator fusion.** The state-estimation framework combined BWLS and VBMC through calibration and residual-error-based weighting. A useful extension is to investigate fusion rules that explicitly account for the covariance between estimator errors, since the two estimators may rely on overlapping measurement information. Calibration parameters could also be updated online to accommodate changes in loading, network configuration, and

sensor performance. Further experiments on larger feeders and under topology or parameter uncertainty would help establish the computational requirements and robustness of the resulting fusion methods. In particular, directly comparing aggregated-load and individual-load implementations of the complete BWLS–VBMC–MRC framework would extend the present BWLS-only assessment of aggregation effects.

- **Validation under practical measurement and communication constraints.** The evaluated DSSE framework considered limited near-real-time AMI observability and different measurement-noise levels. Future work can incorporate asynchronous meter reporting, communication delays, prolonged data loss, and corrupted measurements within a unified evaluation. Integrating additional measurements from distribution substations and other monitoring devices could improve the availability of reference information for calibration. Ultimately, testing the complete forecasting and state-estimation pipeline using continuously collected utility data would provide a clearer assessment of its computational feasibility, accuracy, and reliability in operational settings.

These directions build on the statistical planning, forecasting, and estimation methods developed in this dissertation. Their investigation would extend the proposed approaches from the evaluated case studies toward broader utility applications in which uncertainty, computational efficiency, and the availability of reliable measurements remain central considerations.

Bibliography

[1] U.S. Department of Energy. Solar futures study. Technical report, U.S. Department of Energy, Solar Energy Technologies Office, 2021. URL https://www.energy.gov/eere/solar/solar-futures-study.

[2] Manuel Betancourt Schwarz, Mathilde Veyron, and Marc Clausse. Impact of flexibility implementation on the control of a solar district heating system. In *Solar*, volume 4, pages 1–14. MDPI, 2023.

[3] Ayed Banibaqash, Ziad Hunaiti, and Maysam Abbod. Assessing the potential of qatari house roofs for solar panel installations: A feasibility survey. In *Solar*, volume 3, pages 650–662. MDPI, 2023.

[4] Shuyan Chen, Na Liu, Fan Xu, and Guodan Wei. Challenges and strategies toward future stable perovskite photovoltaics. *Solar RRL*, 7(23):2300479, 2023.

[5] Na Liu, Lina Wang, Fan Xu, Jiafeng Wu, Tinglu Song, and Qi Chen. Recent progress in developing monolithic perovskite/si tandem solar cells. *Frontiers in Chemistry*, 8: 603375, 2020.

[6] Firoz Khan, Thamraa Alshahrani, Imran Fareed, and Jae Hyun Kim. A comprehensive degradation assessment of silicon photovoltaic modules installed on a concrete base under hot and low-humidity environments: Building applications. *Sustainable Energy Technologies and Assessments*, 52:102314, 2022.

[7] Takele Ferede Agajie, Ahmed Ali, Armand Fopah-Lele, Isaac Amoussou, Baseem Khan, Carmen Lilí Rodríguez Velasco, and Emmanuel Tanyi. A comprehensive review on techno-economic analysis and optimal sizing of hybrid renewable energy sources with energy storage systems. *Energies*, 16(2):642, 2023.

[8] Adedayo Owosuhi, Yskandar Hamam, and Josiah Munda. Maximizing the integration of a battery energy storage system–photovoltaic distributed generation for power system harmonic reduction: an overview. *Energies*, 16(6):2549, 2023.

[9] Eric Tervo, Kenechi Agbim, Freddy DeAngelis, Jeffrey Hernandez, Hye Kyung Kim, and Adewale Odukomaiya. An economic analysis of residential photovoltaic systems with lithium ion battery storage in the united states. *Renewable and Sustainable Energy Reviews*, 94:1057–1066, 2018.

[10] Tobias Beck, Hendrik Kondziella, Guillaume Huard, and Thomas Bruckner. Assessing the influence of the temporal resolution of electrical load and pv generation profiles on self-consumption and sizing of pv-battery systems. *Applied energy*, 173:331–342, 2016.

[11] Ilaria Bendato, Andrea Bonfiglio, Massimo Brignone, Federico Delfino, Fabio Pampararo, Renato Procopio, and Mansueto Rossi. Design criteria for the optimal sizing of integrated photovoltaic-storage systems. *Energy*, 149:505–515, 2018.

[12] Zakaria Belboul, Belgacem Toual, Abdellah Kouzou, Lakhdar Mokrani, Abderrahman Bensalem, Ralph Kennel, and Mohamed Abdelrahem. Multiobjective optimization of a hybrid pv/wind/battery/diesel generator system integrated in microgrid: A case study in djelfa, algeria. *Energies*, 15(10):3579, 2022.

[13] Bineeta Mukhopadhyay and Debapriya Das. Multi-objective dynamic and static reconfiguration with optimized allocation of pv-dg and battery energy storage system. *Renewable and sustainable energy reviews*, 124:109777, 2020.

[14] Kumar Shivam, Jong-Chyuan Tzou, and Shang-Chen Wu. A multi-objective predictive energy management strategy for residential grid-connected pv-battery hybrid systems based on machine learning technique. *Energy Conversion and Management*, 237:114103, 2021.

[15] Yuan Zhi and Xudong Yang. Scenario-based multi-objective optimization strategy for rural pv-battery systems. *Applied Energy*, 345:121314, 2023.

[16] Ziye Song, Xin Guan, and Meng Cheng. Multi-objective optimization strategy for home energy management system including pv and battery energy storage. *Energy Reports*, 8:5396–5411, 2022.

[17] Heejung Park. A stochastic planning model for battery energy storage systems coupled with utility-scale solar photovoltaics. *Energies*, 14(5):1244, 2021.

[18] Mahshid Javidsharifi, Hamoun Pourroshanfekr Arabani, Tamas Kerekes, Dezso Sera, and Josep M Guerrero. Stochastic optimal strategy for power management in interconnected multi-microgrid systems. *Electronics*, 11(9):1424, 2022.

[19] Enyu Guo, Bo He, and Jinliang Zhang. Effects of photovoltaic panel type on optimum sizing of an electrical energy storage system using a stochastic optimization approach. *Journal of Energy Storage*, 72:108581, 2023.

[20] Nzube Ntube and Haiyu Li. Stochastic multi-objective optimal sizing of battery energy storage system for a residential home. *Journal of Energy Storage*, 59:106403, 2023.

[21] Zhuang Zheng, Xin Li, Jia Pan, and Xiaowei Luo. A multi-year two-stage stochastic programming model for optimal design and operation of residential photovoltaic-battery systems. *Energy and buildings*, 239:110835, 2021.

[22] Shuang Rong, Yanlei Zhao, Yanxin Wang, Jiajia Chen, Wanlin Guan, Jiapeng Cui, and Yanlong Liu. Information gap decision theory-based stochastic optimization for smart microgrids with multiple transformers. *Applied Sciences*, 13(16):9305, 2023.

[23] Diah Chaerani, Adibah Shuib, Tomy Perdana, and Athaya Zahrani Irmansyah. Systematic literature review on robust optimization in solving sustainable development goals (sdgs) problems during the covid-19 pandemic. *Sustainability*, 15(7):5654, 2023.

[24] M Bakhtvar and A Al-Hinai. Robust operation of hybrid solar–wind power plant with battery energy storage system. energies 2021, 14, 3781, 2021.

[25] Mehrdad Aghamohamadi, Amin Mahmoudi, and Mohammed H Haque. Two-stage robust sizing and operation co-optimization for residential pv–battery systems considering the uncertainty of pv generation and load. *IEEE Transactions on Industrial Informatics*, 17(2):1005–1017, 2020.

[26] Diederik Coppitters, Ward De Paepe, and Francesco Contino. Robust design optimization and stochastic performance analysis of a grid-connected photovoltaic system with battery storage and hydrogen storage. *Energy*, 213:118798, 2020.

[27] Wei Tian, Hsiao-Dong Chiang, and Hiroyuki Mori. Theoretical foundation of weight adjustment for weighted least squares state estimator. *IEEE Transactions on Power Systems*, 2025.

[28] Stefano Gastoni, Gianpietro Granelli, and Mario Montagna. Robust state-estimation procedure based on the maximum agreement between measurements. *IEEE Transactions on Power Systems*, 19(4):2038–2043, 2004.

[29] Efthymios Manitsas, Ravindra Singh, Bikash C Pal, and Goran Strbac. Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling. *IEEE Transactions on power systems*, 27(4):1888–1896, 2012.

[30] Y Raghuvamsi and Kiran Teeparthi. A review on distribution system state estimation uncertainty issues using deep learning approaches. *Renewable and Sustainable Energy Reviews*, 187:113752, 2023.

[31] Maria Fotopoulou, Stefanos Petridis, Ioannis Karachalios, and Dimitrios Rakopoulos. A review on distribution system state estimation algorithms. *Applied Sciences*, 12(21): 11073, 2022.

[32] Jam Angulo-Paniagua and Jairo Quirós-Tortós. Comparing chi-square-based bad data detection algorithms for distribution system state estimation. In *2020 IEEE PES Transmission & Distribution Conference and Exhibition-Latin America (T&D LA)*, pages 1–5. IEEE, 2020.

[33] Robson C Pires, A Simoes Costa, and Lamine Mili. Iteratively reweighted least-squares state estimation through givens rotations. *IEEE Transactions on Power Systems*, 14 (4):1499–1507, 1999.

[34] R Baldick, KA Clements, Z Pinjo-Dzigal, and PW Davis. Implementing nonquadratic objective functions for state estimation and bad data rejection. *IEEE Transactions on Power Systems*, 12(1):376–382, 2002.

[35] Harmohan Singh and Fernando L Alvarado. Weighted least absolute value state estimation using interior point methods. *IEEE Transactions on Power Systems*, 9(3): 1478–1484, 1994.

[36] Lamine Mili, V Phaniraj, and Peter J Rousseeuw. Least median of squares estimation in power systems. *IEEE Transactions on Power Systems*, 6(2):511–523, 2002.

[37] Yacine Chakhchoukh and Hideaki Ishii. Enhancing robustness to cyber-attacks in power systems through multiple least trimmed squares state estimations. *IEEE Transactions on Power Systems*, 31(6):4395–4405, 2016.

[38] AK Al-Othman and MR Irving. Robust state estimator based on maximum constraints satisfaction of uncertain measurements. *Measurement*, 40(3):347–359, 2007.

[39] M Majidi, M Etezadi-Amoli, and H Livani. Distribution system state estimation using compressive sensing. *International Journal of Electrical Power & Energy Systems*, 88: 175–186, 2017.

[40] SM Shafiul Alam, Balasubramaniam Natarajan, and Anil Pahwa. Distribution grid state estimation from compressed measurements. *IEEE Transactions on Smart Grid*, 5(4):1631–1642, 2014.

[41] Hazhar Sufi Karimi and Balasubramaniam Natarajan. Compressive sensing based state estimation for three phase unbalanced distribution grid. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pages 1–6. IEEE, 2017.

[42] Hazhar Sufi Karimi and Balasubramaniam Natarajan. Recursive dynamic compressive sensing in smart distribution systems. In *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5. IEEE, 2020.

[43] Shweta Dahale and Balasubramaniam Natarajan. Joint matrix completion and compressed sensing for state estimation in low-observable distribution system. In *2021 IEEE PES Innovative Smart Grid Technologies Conference-Latin America (ISGT Latin America)*, pages 1–5. IEEE, 2021.

[44] Yingchen Zhang, Andrey Bernstein, Andreas Schmitt, and Rui Yang. State estimation in low-observable distribution systems using matrix completion. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2019.

[45] Shweta Dahale, Hazhar Sufi Karimi, Kexing Lai, and Balasubramaniam Natarajan. Sparsity based approaches for distribution grid state estimation-a comparative study. *IEEE Access*, 8:198317–198327, 2020.

[46] Biswajeet Rout, Shweta Dahale, and Balasubramaniam Natarajan. Dynamic matrix completion based state estimation in distribution grids. *IEEE Transactions on Industrial Informatics*, 18(11):7504–7511, 2022.

[47] VSG Thadikemalla, Ishan Srivastava, SS Bhat, and AS Gandhi. Data loss mitigation mechanism using compressive sensing for smart grids. In *2020 IEEE international conference on power electronics, smart grid and renewable energy (PESGRE2020)*, pages 1–6. IEEE, 2020.

[48] Harold Soh and Yiannis Demiris. Spatio-temporal learning with the online finite and infinite echo-state gaussian processes. *IEEE transactions on neural networks and learning systems*, 26(3):522–536, 2014.

[49] Shweta Dahale and Balasubramaniam Natarajan. Multi time-scale imputation aided state estimation in distribution system. In *2021 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2021.

[50] Alireza Fatehi and Biao Huang. Kalman filtering approach to multi-rate information fusion in the presence of irregular sampling rate and variable measurement delay. *Journal of Process Control*, 53:15–25, 2017.

[51] Kursat Rasim Mestav, Jaime Luengo-Rozas, and Lang Tong. Bayesian state estimation for unobservable distribution systems via deep learning. *IEEE Transactions on Power Systems*, 34(6):4910–4920, 2019.

[52] Alireza Akrami and Hamed Mohsenian-Rad. Event-triggered distribution system state estimation: Sparse kalman filtering with reinforced coupling. *IEEE Transactions on Smart Grid*, 15(1):627–640, 2023.

[53] Hadis Karimipour and Venkata Dinavahi. Extended kalman filter-based parallel dynamic state estimation. *IEEE transactions on smart grid*, 6(3):1539–1549, 2015.

[54] Guanyu Tian, Yingzhong Gu, Di Shi, Jing Fu, Zhe Yu, and Qun Zhou. Neural-network-based power system state estimation with extended observability. *Journal of Modern Power Systems and Clean Energy*, 9(5):1043–1053, 2021.

[55] Narri Yadaiah and G Sowmya. Neural network based state estimation of dynamical systems. In *The 2006 IEEE international joint conference on neural network proceedings*, pages 1042–1049. IEEE, 2006.

[56] Erik Joachim Schake. *Recurrent neural network approaches to irregularly sampled data*. PhD thesis, Kauno technologijos universitetas., 2022.

[57] Chang Liu, Yuru Guo, Zhuo Wang, Yong Xu, and Chun-Yi Su. State estimation for nonuniformly sampled neural networks with hidden information. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(10):6342–6352, 2023.

[58] Yong Liu, Tengge Hu, Haoran Zhang, Haixu Wu, Shiyu Wang, Lintao Ma, and Mingsheng Long. itransformer: Inverted transformers are effective for time series forecasting. In *The Twelfth International Conference on Learning Representations*, 2024. [Online]. Available: https://openreview.net/forum?id=JePfAI8fah.

[59] Reza Nematirad, Anil Pahwa, and Balasubramaniam Natarajan. Times2d: Multi-period decomposition and derivative mapping for general time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 19651–19658, 2025.

[60] Reza Nematirad, MM Ardehali, Amir Khorsandi, and Arash Mahmoudian. Optimization of residential demand response program cost with consideration for occupants thermal comfort and privacy. *IEEE Access*, 2024.

[61] Jing Lei, Le Wang, Qingqi Pei, Wenhai Sun, Xiaodong Lin, and Xuefeng Liu. Privgrid: Privacy-preserving individual load forecasting service for smart grid. *IEEE Transactions on Information Forensics and Security*, 2024.

[62] Faezeh Bashiri, Julian L Cardenas-Barrera, and E Castillo-Guerra. A review of machine learning approaches for forecasting aggregated power demand of thermostatically controlled loads. In *2024 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 835–841. IEEE, 2024.

[63] Amir Bahador Javadi and Philip Pong. A review on symbolic regression in power systems: Methods, applications, and future directions. *Renewable and Sustainable Energy Reviews*, 224:116075, 2025. ISSN 1364-0321. doi: 10.1016/j.rser.2025.116075. URL https://www.sciencedirect.com/science/article/pii/S1364032125007488.

[64] Chafak Tarmanini, Nur Sarma, Cenk Gezegin, and Okan Ozgonenel. Short term load forecasting based on arima and ann approaches. *Energy Reports*, 9:550–557, 2023.

[65] Grzegorz Dudek, Paweł Pełka, and Slawek Smyl. A hybrid residual dilated lstm and exponential smoothing model for midterm electric load forecasting. *IEEE Transactions on Neural Networks and Learning Systems*, 33(7):2879–2891, 2021.

[66] Haixu Wu, Tengge Hu, Yong Liu, Hang Zhou, Jianmin Wang, and Mingsheng Long. Timesnet: Temporal 2d-variation modeling for general time series analysis. *arXiv preprint arXiv:2210.02186*, 2022.

[67] Reza Nematirad and Anil Pahwa. Solar radiation forecasting using artificial neural networks considering feature selection. In *2022 IEEE Kansas Power and Energy Conference (KPEC)*, pages 1–4. IEEE, 2022.

[68] Anthony Faustine, Nuno Jardim Nunes, and Lucas Pereira. Efficiency through simplicity: Mlp-based approach for net-load forecasting with uncertainty estimates in low-voltage distribution networks. *IEEE Transactions on Power Systems*, 2024.

[69] Soheil Abadifard, Sepehr Bakhshi, Sanaz Gheibuni, and Fazli Can. Dyned: Dynamic ensemble diversification in data stream classification. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 3707–3711, 2023.

[70] Reza Behnam, Amirhosein Gohari, Mohammad B Shadmand, Sertac Bayhan, and Haitham Abu-Rub. Decentralized ai-based fault detection and localization to enhance dynamic response of grid-forming inverters. In *IECON 2024-50th Annual Conference of the IEEE Industrial Electronics Society*, pages 1–6. IEEE, 2024.

[71] Lucas Barros Scianni Morais, Giancarlo Aquila, Victor Augusto Durães de Faria, Luana Medeiros Marangon Lima, José Wanderley Marangon Lima, and Anderson Rodrigo de Queiroz. Short-term load forecasting using neural networks and global climate models: An application to a large-scale electrical power system. *Applied Energy*, 348: 121439, 2023.

[72] Ahmad O Aseeri. Effective rnn-based forecasting methodology design for improving short-term power load forecasts: Application to large-scale power-grid time series. *Journal of Computational Science*, 68:101984, 2023.

[73] Salah Bouktif, Ali Fiaz, Ali Ouni, and Mohamed Adel Serhani. Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies*, 11(7):1636, 2018.

[74] Chengyu Zhang, Liangdong Ma, Zhiwen Luo, Xing Han, and Tianyi Zhao. Forecasting building plug load electricity consumption employing occupant-building interaction input features and bidirectional lstm with improved swarm intelligent algorithms. *Energy*, 288:129651, 2024.

[75] Jiang-Wen Xiao, Xue-Ying Cui, Xiao-Kang Liu, Hongliang Fang, and Peng-Cheng Li. Improved 3d lstm: A video prediction approach to long sequence load forecasting. *IEEE Transactions on Smart Grid*, pages 1–1, 2024. doi: 10.1109/TSG.2024.3458989.

[76] Nilakanta Kshetrimayum, Khumukcham Robindro Singh, and Nazrul Hoque. Pconvlstm: an effective parallel convlstm-based model for short-term electricity load forecasting. *International Journal of Data Science and Analytics*, pages 1–18, 2024.

[77] Anthony Faustine, Nuno Jardim Nunes, and Lucas Pereira. Efficiency through simplicity: Mlp-based approach for net-load forecasting with uncertainty estimates in low-voltage distribution networks. *IEEE Transactions on Power Systems*, pages 1–11, 2024. doi: 10.1109/TPWRS.2024.3400123.

[78] Christoforos Menos-Aikateriniadis, Andreas Akarepis, Isidoros Kokos, and Pavlos S. Georgilakis. Sub-hourly load forecasting for community-level flexible appliance management. In *2024 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2024. doi: 10.1109/IJCNN60899.2024.10649928.

[79] Neelesh Pratap Singh and Mahamad Nabab Alam. Short-term forecasting in smart grid environment using n-beats. Preprint, available at https://arxiv.org/abs/xxxx.xxxxx, 2024.

[80] Hengshuo Liang, Cheng Qian, Wei Yu, David Griffith, and Nada Golmie. Assessing deep learning performance in power demand forecasting for smart grid. *International Journal of Sensor Networks*, 44(1):36–48, 2024.

[81] Reza Nematirad, Mehdi Behrang, and Anil Pahwa. Acoustic-based online monitoring

of cooling fan malfunction in air-forced transformers using learning techniques. *IEEE Access*, 12:26384–26400, 2024.

[82] Sujan Ghimire, Thong Nguyen-Huy, Mohanad S AL-Musaylh, Ravinesh C Deo, David Casillas-Pérez, and Sancho Salcedo-Sanz. A novel approach based on integration of convolutional neural networks and echo state network for daily electricity demand prediction. *Energy*, 275:127430, 2023.

[83] Osaka Rubasinghe, Xinan Zhang, Tat Kei Chau, Yau Hing Chow, Tyrone Fernando, and Herbert Ho-Ching Iu. A novel sequence to sequence data modelling based cnn-lstm algorithm for three years ahead monthly peak load forecasting. *IEEE Transactions on Power Systems*, 39(1):1932–1947, 2024. doi: 10.1109/TPWRS.2023.3271325.

[84] Meng-Ping Wu and Fan Wu. Predicting residential electricity consumption using cnn-bilstm-sa neural networks. *IEEE Access*, 12:71555–71565, 2024. doi: 10.1109/ACCESS.2024.3400972.

[85] A Selim Türkoğlu, Burcu Erkmen, Yavuz Eren, Ozan Erdinç, and İbrahim Küçükdemiral. Integrated approaches in resilient hierarchical load forecasting via tcn and optimal valley filling based demand response application. *Applied Energy*, 360: 122722, 2024.

[86] Jun He, Kuidong Yuan, Zijie Zhong, and Yifan Sun. Enhancing short-term power load forecasting with a timesnet-crossformer-lstm approach. *IEEE Access*, 12:56774–56788, 2024.

[87] Jinkai Shi, Weige Zhang, Yan Bao, David Wenzhong Gao, and Zhihao Wang. Load forecasting of electric vehicle charging stations: Attention based spatiotemporal multi-graph convolutional networks. *IEEE Transactions on Smart Grid*, 2023.

[88] Jiang-Wen Xiao, Xue-Ying Cui, Xiao-Kang Liu, Hongliang Fang, and Peng-Cheng Li. Improved 3d lstm: A video prediction approach to long sequence load forecasting. *IEEE Transactions on Smart Grid*, 2024.

[89] Zhenghua Xu, Zhoutao Yu, Hexiang Zhang, Junyang Chen, Junhua Gu, Thomas Lukasiewicz, and Victor CM Leung. Phacia-tcns: Short-term load forecasting using temporal convolutional networks with parallel hybrid activated convolution and input attention. *IEEE Transactions on Network Science and Engineering*, 2023.

[90] Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 11106–11115, 2021.

[91] Jun Wei Chan and Chai Kiat Yeo. A transformer based approach to electricity load forecasting. *The Electricity Journal*, 37(2):107370, 2024.

[92] Yucong Huang, Fang Liu, Yalin Wang, Denis Sidorov, and Yong Li. Day-ahead electric load forecasting based on t2vc-informer. In *2024 43rd Chinese Control Conference (CCC)*, pages 8822–8826. IEEE, 2024.

[93] Siting Li and Huafeng Cai. Short-term power load forecasting using a vmd-crossformer model. *Energies*, 17(11):2773, 2024.

[94] Yunhao Zhang and Junchi Yan. Crossformer: Transformer utilizing cross-dimension dependency for multivariate time series forecasting. In *The eleventh international conference on learning representations*, 2023.

[95] Senfeng Cen and Chang Gyoon Lim. Multi-task learning of the patchtcn-tst model for short-term multi-load energy forecasting considering indoor environments in a smart building. *IEEE Access*, 12:19553–19568, 2024. doi: 10.1109/ACCESS.2024.3355448.

[96] Jili Fan, Wei Zhuang, Min Xia, WenXuan Fang, and Jun Liu. Optimizing attention in a transformer for multihorizon, multienergy load forecasting in integrated energy systems. *IEEE Transactions on Industrial Informatics*, 2024.

[97] Ru Li, Siyi Huang, and Hongqiang Dou. Dynamic risk assessment of landslide hazard

for large-scale photovoltaic power plants under extreme rainfall conditions. *Water*, 15 (15):2832, 2023.

[98] Felipe AP De Figueiredo, Claudio F Dias, Eduardo R De Lima, and Gustavo Fraidenraich. Capacity bounds for dense massive mimo in a line-of-sight propagation environment. *sensors*, 20(2):520, 2020.

[99] Faiza A Althubyani, Ahmed MT Abd El-Bar, Mohamad A Fawzy, and Ahmed M Gemeay. A new 3-parameter bounded beta distribution: Properties, estimation, and applications. *Axioms*, 11(10):504, 2022.

[100] Deepak Upreti, Stefano Pignatti, Simone Pascucci, Massimo Tolomio, Zhenhai Li, Wenjiang Huang, and Raffaele Casa. A comparison of moment-independent and variance-based global sensitivity analysis approaches for wheat yield estimation with the aquacrop-os model. *Agronomy*, 10(4):607, 2020.

[101] Sunghyon Kyeong, Daehee Kim, and Jinho Shin. Can system log data enhance the performance of credit scoring?—evidence from an internet bank in korea. *Sustainability*, 14(1):130, 2021.

[102] Quan Chen and Lamine Mili. Composite power system vulnerability evaluation to cascading failures using importance sampling and antithetic variates. *IEEE transactions on power systems*, 28(3):2321–2330, 2013.

[103] Eduardo Quiles-Cucarella, Adrián Marquina-Tajuelo, Carlos Roldán-Blay, and Carlos Roldán-Porta. Particle swarm optimization method for stand-alone photovoltaic system reliability and cost evaluation based on monte carlo simulation. *Applied Sciences*, 13(21):11623, 2023.

[104] Maximilian Hoffmann, Leander Kotzur, Detlef Stolten, and Martin Robinius. A review on time series aggregation methods for energy system models. *Energies*, 13(3):641, 2020.

[105] Mao Yang, Meng Zhao, Dawei Huang, and Xin Su. A composite framework for photovoltaic day-ahead power prediction based on dual clustering of dynamic time warping distance and deep autoencoder. *Renewable Energy*, 194:659–673, 2022.

[106] Fan Li, Jingxi Su, and Bo Sun. An optimal scheduling method for an integrated energy system based on an improved k-means clustering algorithm. *Energies*, 16(9): 3713, 2023.

[107] Juha Vesanto and Esa Alhoniemi. Clustering of the self-organizing map. *IEEE Transactions on neural networks*, 11(3):586–600, 2000.

[108] Leonardo Enzo Brito Da Silva and Donald C Wunsch. An information-theoretic-cluster visualization for self-organizing maps. *IEEE transactions on neural networks and learning systems*, 29(6):2595–2613, 2017.

[109] Eric U Oti, Michael O Olusola, Francis C Eze, and Samuel U Enogwe. Comprehensive review of k-means clustering algorithms. *criterion*, 12(08):22–23, 2021.

[110] Adeiza James Onumanyi, Daisy Nkele Molokomme, Sherrin John Isaac, and Adnan M Abu-Mahfouz. Autoelbow: An automatic elbow detection method for estimating the number of clusters in a dataset. *Applied Sciences*, 12(15):7515, 2022.

[111] David Criado-Ramón, Luis G Baca Ruíz, and MC Pegalajar. An improved pattern sequence-based energy load forecast algorithm based on self-organizing maps and artificial neural networks. *Big Data and Cognitive Computing*, 7(2):92, 2023.

[112] Yimin Chen, Xiaoping Liu, Xia Li, Xingjian Liu, Yao Yao, Guohua Hu, Xiaocong Xu, and Fengsong Pei. Delineating urban functional areas with building-level social media data: A dynamic time warping (dtw) distance based k-medoids method. *Landscape and Urban Planning*, 160:48–60, 2017.

[113] Pedro Pereira Rodrigues, Joao Gama, and Joao Pedroso. Hierarchical clustering of time-series data streams. *IEEE transactions on knowledge and data engineering*, 20 (5):615–627, 2008.

[114] Mobarak Abumohsen, Amani Yousef Owda, and Majdi Owda. Electrical load forecasting using lstm, gru, and rnn algorithms. *Energies*, 16(5):2283, 2023.

[115] Yuqi Nie, Nam H. Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers, 2023. URL https://arxiv.org/abs/2211.14730.

[116] Yong Liu, Tengge Hu, Haoran Zhang, Haixu Wu, Shiyu Wang, Lintao Ma, and Mingsheng Long. itransformer: Inverted transformers are effective for time series forecasting, 2024. URL https://arxiv.org/abs/2310.06625.

[117] Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. Are transformers effective for time series forecasting? In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pages 11121–11128, 2023.

[118] Haixu Wu, Tengge Hu, Yong Liu, Hang Zhou, Jianmin Wang, and Mingsheng Long. Timesnet: Temporal 2d-variation modeling for general time series analysis. arxiv 2022. *arXiv preprint arXiv:2210.02186*.