Data-Driven Simulation and Optimization of Renewable Energy Systems

by

Wenrui Ye

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Mechanical and Mechatronics Engineering

Waterloo, Ontario, Canada, 2025

© Wenrui Ye 2025

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Bala Venkatesh, Professor, Department of Electrical and Computer Engineering, Toronto Metropolitan University
Supervisor(s):	John Z. Wen Professor, Department of Mechanical Engineering, University of Waterloo
	Jatin Nathwani Professor, Department of Civil Engineering, University of Waterloo
Internal Member:	Pan Zhao Assistant Professor, Department of Mechanical Engineering, University of Waterloo

Internal Member: Xiaoyu Wu
Assistant Professor, Department of Mechanical Engineering,
University of Waterloo

Internal-External Member: Jun Liu
Professor, Department of Applied Mathematics,
University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Wenrui Ye is the sole author of Chapters 1, 2, 3 and 7 which were written under the supervision of Prof. John Z. Wen and were not published. This thesis contains parts of three manuscripts written for publication.

Research presented in Chapter 4:

Wenrui Ye, Münür Sacit Herdem, Joey Z Li, Jatin Nathwani and John Z Wen. "Formulation and data-driven optimization for maximizing the photovoltaic power with tilt angle adjustment", Energies, 15 (22) (2022), p. 8578.

Research presented in Chapter 5:

Wenrui Ye, Münür Sacit Herdem, Shucheng Huang, Wei Sun, Jun Liu, Jatin Nathwani and John Z Wen. "Component capacity optimization of a renewable energy system using data-driven two-stage algorithmic approach", Energy Conversion and Management, 312:118588, 2024.

Research presented in Chapter 6:

Wenrui Ye, Shucheng Huang, Münür Sacit Herdem, Wei Sun, Jatin Nathwani and John Z Wen. Under Review.

As the first author of these three chapters, I contributed to the conceptualization of the study design, conducted data collection, performed simulations and analyses, and prepared and submitted the manuscripts. My co-authors offered guidance throughout each stage of the research and provided feedback on the manuscript drafts.

Abstract

The transition to renewable energy systems is critical in mitigating climate change and reducing fossil fuel dependence. However, integrating these variable and intermittent sources into the existing grid raises challenges such as dynamic energy demand management and resource underutilization, leading to increased operational costs and hindering broader adoption. This thesis develops algorithms to optimize renewable energy systems, enhancing their integration and operational efficiency. The research makes a significant contribution to enhancing the utilization, reliability, and economic viability of renewable energy systems, supporting a smoother transition to sustainable energy practices.

This thesis first enhances the energy generation of photovoltaic panels by optimizing their tilt angles to maximize solar energy capture under varying environmental conditions. The machine learning models provided accurate predictions of photovoltaic output, allowing for data-driven insights into optimal system performance. A subsequent optimization process identified the best tilt angles during the operation. The results demonstrated an increase in annual energy output by up to 9.7% compared to fixed-tilt systems. This confirms that dynamic tilt adjustment is an effective strategy for maximizing photovoltaic energy generation.

The second part of the research focuses on optimizing the capacity of renewable energy system components, with a particular emphasis on energy storage systems such as batteries. This project addressed the challenge of determining the optimal capacity for each component to efficiently meet energy demands while minimizing costs. A two-stage optimization approach was applied: first, a genetic algorithm generated candidate configurations with specific capacities; second, a simulation was conducted using an energy management algorithm to evaluate the performance of these configurations. The optimized configurations led to an overall system energy independence score of 0.51 and an 18.12% higher internal rate of return, validating the effectiveness of the integrated optimization approach in capacity planning and highlighting the importance of appropriately sized energy storage in enhancing system performance.

The final part introduces an advanced energy management algorithm inspired by Model

Predictive Control, integrating both batteries and hydrogen storage to enhance renewable energy utilization. By employing time series data and transformer-based models, the system accurately predicts future energy demand. These predictions enable a rolling window optimization technique that utilizes machine learning for dynamic energy management. The inclusion of hydrogen storage allows excess renewable energy to be stored as hydrogen, providing a versatile energy carrier for applications beyond electricity and improving overall renewable energy utilization. This approach improved demand forecasting accuracy by 41.21% and increased the adjusted green hydrogen production rate from 29.54% to 54.3%, demonstrating that advanced predictive energy management strategies, combined with diverse energy storage solutions, significantly enhance system adaptability, efficiency, and renewable energy utilization.

These studies demonstrate that optimizing component configurations and energy management strategies—while integrating advanced energy storage systems like batteries and hydrogen—substantially improves the efficiency, reliability, and economic viability of renewable energy systems. The research provides valuable insights for integrating renewable energy into existing grids and supports the transition toward more sustainable and resilient energy infrastructures.

Acknowledgements

I extend my deepest gratitude to my supervisor, Professor John Z. Wen, for his unwavering guidance and invaluable insights throughout my research journey. His expertise and mentorship have been instrumental in shaping my academic development and bringing this thesis to completion. Professor Wen's dedication to excellence and patient support greatly enhanced my ability to navigate the challenges of my research. His willingness to provide assistance in various forms, whether through detailed academic discussions or life advice, has been tremendously beneficial. I am sincerely grateful for his continued support.

I am also thankful to the members of my thesis committee for their constructive feedback and essential suggestions that have significantly enriched my work. Special thanks go to Münür Herdem and Wei Sun; although not members of my committee, their expertise and guidance have been crucial in refining my research.

My gratitude extends to my colleagues, whose collaboration and insightful discussions have been invaluable. Their shared passion and intellectual curiosity have greatly enhanced the research environment.

Further, I express my deepest appreciation to my dear friends—Yaodong Cui, Haoyang Gao, Zhongkang Guo, Shucheng Huang, Yufeng Huang, Siyuan Liu, Zhenan Liu, Zilin Liu, Jianyu Pan, Xiangkang Wang, Jingyu Wen, Peizhi Wu, Jinwei Zhang, Mingchuan Zhang, Qi Zhao, Haoyang Zhou and Xueyi Zhou. Their continuous support, constant encouragement, and invaluable companionship have greatly enriched both my personal and academic life. Their presence has been fundamental to my resilience and motivation.

Finally, I want to thank every person I have met along my journey through life. Each of you, through moments of joy and trials of pain, has helped shape who I am today. You have unknowingly woven your essence into the fabric of my being, teaching me resilience in adversity and joy in simplicity. Whether you remain by my side or have drifted away, I am grateful for the impact you have had on my life. To all, for every lesson learned and every moment shared, my deepest gratitude.

Dedication

This thesis is dedicated to my parents, Mingli Gao and Nan Ye.

Their love and understanding have been the foundation of my academic journey. Through every challenge, their wisdom and support provided me with strength and resilience. They offered me the freedom to pursue my passions, while always providing a safe haven amidst the pressures of academia. Without their constant encouragement and belief in me, this work would not have been possible. I am grateful for everything they have done and for the strength they have given me throughout this journey.

Table of Contents

List of Figures xv

List of Tables xx

List of Abbreviations xxii

1 Introduction 1

1.1 Background and Motivation 1

1.2 Objective and Contributions 4

1.3 Structure of the Thesis 4

2 Literature View 6

2.1 Renewable Energy Systems 6

2.2 Renewable Energy Resources 8

2.2.1 Solar Energy and Photovoltaic Panel 9

2.2.2 Wind Energy and Wind Turbine 13

2.2.3 Impact of Renewable Energy 15

2.3 Energy Storage Systems 17

2.3.1 Battery Energy Storage Systems 20
2.3.2 Hydrogen Energy Storage Systems 23
2.4 Time Series Data Analysis 26
2.4.1 Statistical Approaches 28
2.4.2 Machine Learning Approaches 32
2.4.3 Deep Learning Approaches 33
2.5 System Optimization 45
2.5.1 Photoviltaic Panel Optimization 50
2.5.2 Capacity of Component Optimization 51
2.5.3 Energy Management 53
2.6 Summary 54

3 Extended Literature Survey **55**
3.1 System and Components Modeling 55
3.1.1 Photovoltaic Modeling 56
3.1.2 Wind Turbine Modeling 63
3.1.3 Battery Modeling 65
3.1.4 Electrolyzer Modeling 67
3.1.5 Renewable Energy System Modeling 72
3.2 Machine Learning for Data-driven Analysis 74
3.2.1 Data Preprocess 74
3.2.2 Exploratory Data Analysis 76
3.2.3 Network Design and Train 77
3.3 Optimization 79

3.3.1 Linear Programming . . . 79
3.3.2 Mixed Integer Linear Programming . . . 81
3.3.3 Genetic Algorithms . . . 82
3.4 Performance Evaluation Metrics . . . 84
3.4.1 Prediction Evaluation Metrics . . . 84
3.4.2 Economic Evaluation Metrics . . . 85
3.4.3 Sustainability Evaluation Metrics . . . 87
3.5 Summary . . . 90

4 Data-Driven Photovoltaic Tilt Angle Optimization for Power Generation 92
4.1 Contribution . . . 92
4.2 Target System and Proposed Algorithm . . . 93
4.3 Simulation Procedure . . . 96
4.3.1 Data Pre-process and Hypothesis Test . . . 97
4.3.2 Correlation Analysis . . . 100
4.3.3 Cross Validation . . . 102
4.4 Result and Discussion . . . 103
4.4.1 Energy Generation . . . 106
4.4.2 Assumptions and Limitations . . . 108
4.5 Summary . . . 109

5 Data-Driven Two-stage Component Capacity Optimization 111
5.1 Contribution . . . 111
5.2 Target System and Proposed Algorithm . . . 112

5.2.1 Target System . . . 113
5.2.2 Conceptual Framework . . . 114
5.2.3 Performance Metrics . . . 116
5.3 Simulation Procedure . . . 118
5.3.1 First Stage Optimization . . . 118
5.3.2 Second Stage Optimization . . . 121
5.3.3 Interrelation of Decision Variables and Objectives . . . 125
5.4 Result and Discussion . . . 126
5.4.1 Capacity of Major Components . . . 129
5.4.2 Decision Framework . . . 131
5.4.3 Energy Management . . . 135
5.4.4 Comparative Analysis . . . 138
5.4.5 Assumptions and Limitations . . . 141
5.5 Summary . . . 143

6 Prediction-Based Energy Management 145

6.1 Contribution . . . 145
6.2 Target System and Proposed Algorithm . . . 146
6.2.1 Target System . . . 146
6.2.2 The Predictive-based Energy Management Algorithm . . . 149
6.2.3 Bus Hydrogen Demand Modelling . . . 151
6.3 Energy Management Optimization . . . 152
6.4 Result and Discussion . . . 157
6.4.1 Machine Learning for Energy Demand Prediction . . . 158

6.4.2 Energy Management 160
6.4.3 Comparative Analysis 163
6.4.4 Assumptions and Limitations 170
6.5 Summary 172

7 Conclusion and Future Work **174**
7.1 Conclusion 174
7.2 Future Work 175

Letter of Copyright Permission **178**

References **179**

Appendices **197**

A Optimization Result Validation **198**

List of Figures

2.1 **Classification of Energy Storage System**, plot based on [1] 18

2.2 **Long Short-Term Memory (LSTM)s Structure** 37

2.3 **Visualization of the Self-attention Computation Process** 42

2.4 **The Transformer Architecture** [2] 44

2.5 **Optimization for Renewable Energy Systems** 46

3.1 **One-diode Model** 60

3.2 **Procedure for Machine Learning Modeling** 75

4.1 **(a) Schematic of Energy Flow (b) Functional Relationship of Model Parameters** 94

4.2 **Flow Chart of the Proposed Methodology** 95

4.3 **Pairwise Plot of NIST Data Set (Raw)** 97

4.4 **Two Dimensions Raw Data of NIST Photovoltaic Data Set - Irradiance vs Power Output** 98

4.5 **Correlation Matrix of Pre-processed Data** 101

4.6 **Irradiance vs Ambient Temperature vs Module Temperature by Five Parameters Model** 103

4.7 **Two Dimensions Processed Data of NIST Photovoltaic Data Set - Irradiance vs Module Temperature** 103

4.8 **Performance of the Neural Network on Training Set** 105

4.9 **Performance of the Neural Network on Testing Set** 105

4.10 **Optimization Results and Optimal Tilt Angles for Different Time Intervals** 108

5.1 **Schematic of Target System with Key Variables** 113

5.2 **Framework of the Proposed Two-stage Optimization Approach** Note: The center image was created with the assistance of DALL·E 2 . . . 115

5.3 **Algorithm** (a) Algorithm Flowchart of the Proposed Two-stage Optimization Approach (b) Concept of Pareto Frontier in Multi-objective Optimization 119

5.4 **Relation of Variables and Optimization Objectives** 125

5.5 **Energy Demand Profiles Based on the Original Dataset** (a) Monthly Energy Demand Profile: Detailed Consumption Analysis by Equipment for the Building (b) Annual Energy Demand Profile: Proportional Analysis by Equipment (c) Daily Energy Demand Profile: Zoomed-in View (d) Hourly Energy Demand Profile: Zoomed-in View (e) Annual Hourly Energy Demand Profile Note: The simulation assumed the existence of five identical buildings and added white noise. As a result, the energy demand profile used in the simulation is approximately five times the energy represented in this figure. 127

5.6 **Meteorological Resource Overview from the Original Dataset** (a) Monthly Solar Energy Profile: Breakdown of GHI, DNI, and DIF (daytime data only) (b)–(g) Monthly Distributions Derived from Hourly Measurements of: (b) Global Horizontal Irradiation (c) Direct Normal Irradiation (d) Diffuse Horizontal Irradiation (e) Wind Speed (f) Temperature (g) Atmospheric Pressure 128

5.7 **Evolution of Genetic Algorithm Search for IRR Optimization** (a) Genetic Algorithm Search Epoch 1: Optimal IRR Achievement (b)–(e) Genetic Algorithm Search Epoch 2-5 130

5.8 **Evolution of Genetic Algorithm Search for EIS Optimization** (a) Genetic Algorithm Search Epoch 1: Optimal EIS Achievement (b)–(e) Genetic Algorithm Search Epoch 2–5 131

5.9 **Genetic Algorithm Convergence: Fitness Across Generations and Epochs** (a) IRR (b) EIS 132

5.10 **Results of Optimization: Pareto Front and Additional Indicators** (a) Epsilon Pareto Front in Multi-objective Optimization (b) Achievable EIS by Given Initial Investment (c)–(k) Comparative Evaluation of Normalized Performance Metrics for Configurations Selected from Pareto Front, Indexed 1–9. 133

5.11 **Simulation Results for Configuration 4: Load Demand Energy Supply Breakdown** (a)Hourly Energy Supply Dynamics: PV, BESS, Grid (Zoomed View) (b)Yearly Energy Supply Percentages (c)Weekly Energy Supply Percentages (d)Monthly Energy Supply Percentages 136

5.12 **Energy Management Simulation Results for Configurations 4: Renewable Energy Utilization** (a)Hourly Renewable Energy Utilization Dynamics: PV, BESS, Grid (Zoomed View) (b)Weekly Utilization Percentages (c)Monthly Utilization Percentages. 137

5.13 **Comparison of Technical Performance of Pareto-optimal Configurations** (a) Load Demand Energy Supply (b) RE Utilization 138

5.14 **Energy Management Simulation Results for Pareto-optimal Configurations: BESS Utilization** (a) Monthly BESS Utilization Rate for Optimized System Configuration (b) Monthly BESS SOH for Optimized System Configuration 139

5.15 **Technical Performance Analysis of System Configurations Determined by Primary Optimization/Genetic Algorithm** (a) GEIF: Monthly Statistical Distribution (b) GEIF: Monthly Trend Analysis (c) UF: Monthly Statistical Distribution (d) UF: Monthly Trend Analysis (e) SLF: Monthly Statistical Distribution (f) SLF: Monthly Trend Analysis 141

6.1 **Components Overview of Target System** 147

6.2 **Schematic of Target System with Key Variables** 148

6.3 **Conceptual Framework of the Proposed Energy Management Algorithm**(a) Optimization Algorithm Structure (b) Data Stream Integration 150

6.4 **Hydrogen Demand Profile** (a) Workday (b) Weekend and Holiday . . . 152

6.5 **Building Energy Demand Prediction** (a) Predicted Energy Demand at Time T (b) Predicted Energy Demand at Time T+3 Hours (c) Predicted Energy Demand at Time T+6 Hours (d) Predicted Energy Demand at Time T+9 Hours (e) Monthly Comparison of Network Prediction Performance (f) Overall Network Prediction Performance Comparison 160

6.6 **Energy Management Simulation Results for the Hydrogen Energy Storage System** (a) Hydrogen Demand and Supply Dynamics from Hour Index 1000 to 1072 (b) Detailed Breakdown of Energy Supply by Source in MWh (c) Proportional Distribution of Energy Supply by Source (d) Detailed Breakdown of Hydrogen Supply in MWh (e) Proportional Distribution of Hydrogen Supply 162

6.7 **Energy Management Simulation Results for the Building Energy Supply** (a)-(f) Optimized Energy Supply at Time T based on Predicted Energy Demand with 5-Hour Rolling Horizon Shifting (g) Weekly Energy Supply Breakdown (h) Monthly Energy Supply Breakdown 164

6.8 **Processed Indicator Comparison** 166

6.9 **Comparison of Energy Management Simulation Results for the Hydrogen Energy Storage System with Different Configurations** (a) 72-Hour Hydrogen Dynamics (Config. 2) (b) Energy Supply Distribution (Config. 2) (c) Hydrogen Supply Distribution (Config. 2) (d) 72-Hour Hydrogen Dynamics (Config. 3) (e) Energy Supply Distribution (Config. 3) (f) Hydrogen Supply Distribution (Config. 3) 167

6.10 **Energy Management Simulation Results for the Battery Energy Storage System** (a) Energy Flow (b) Usage Frequency (c) State of Health 168

6.11 **Energy Flow within the System**(a) Configuration 1 (b) Configuration 4 (c) Configuration 5 171

List of Tables

2.1 Efficiencies, Advantages, and Disadvantages of Photovoltaic Technologies . . . 11

2.2 Comparison of Solar Irradiance Models . . . 14

2.3 Comparison of Battery Technologies . . . 21

2.4 Comparison of Electrolysis Technologies . . . 25

2.5 Comparison of RNN, LSTM, and Transformer Models . . . 34

4.1 Hyperparameters of the Neural Network . . . 103

4.2 Simulation Results . . . 104

4.3 Optimal Tilt Angle of Power-based and Irradiance-based Monthly Optimization . . . 109

5.1 Parameter Value . . . 120

5.2 Correlation Matrix of Indicators . . . 130

5.3 Pareto-optimal Configurations and Indicators . . . 134

6.1 Configuration Comparison . . . 158

6.2 Training Hyper-parameters . . . 159

6.3 Comparison of Prediction Performance Across Different Models . . . 161

6.4 **Summary of Indicator Values** 165

List of Abbreviations

AC Alternating Current 20, 113, 148, 149, 162

AE Alkaline Electrolysis 24–26

AHU Air Handling Unit 126

ANN Artificial Neural Network 51, 93, 109

AR Autoregressive 28–31, 33, 77

ARIMA Autoregressive Integrated Moving Average 28, 30, 31, 33, 38, 53

BESS Battery Energy Storage System 19–21, 23, 46, 48, 52, 65, 67, 113, 114, 116–118, 122–124, 135, 137–142, 147–149, 153, 155, 161–163, 168

DC Direct Current 12, 20, 113, 148, 149, 162

DCF Discounted Cash Flow 86

DIF Diffuse Horizontal Irradiation 126

DNI Direct Normal Irradiation 126

DOD Depth of Discharge 66, 124

DPD Dynamic Population Decay 112, 120, 121

EDA Energy Data Analysis 74, 77

EIS Energy Independent Score 87, 116, 118, 119, 121, 125, 129–133, 135, 165, 198

EoT Equation of Time 57

ESS Energy Storage System 3, 6, 7, 17–21, 48, 51, 52, 54, 73, 147

GA Genetic Algorithm 51, 52, 55, 79, 82, 91

GDP Gross Domestic Product 1

GEIF Grid Energy Interaction Factor 87, 116, 130, 140, 165, 167, 170

GHI Global Horizontal Irradiation 126

GHR Green Hydrogen Ratio 88, 89, 162, 165–167

HAWT Horizontal-Axis Wind Turbine 15

HESS Hydrogen Energy Storage System 19, 20, 24, 26, 48, 147–149, 151, 158, 160, 163–165, 170

HSSS Hydrogen Supply Sufficiency Score 89, 90, 165, 168–170

IRR Internal Rate of Return 86, 87, 116–119, 121, 125, 129, 130, 133, 135, 140, 198

k-NN k-Nearest Neighbors 32, 33

LCOE Levelized Cost of Energy 3, 9, 50

LFP Lithium Iron Phosphate 22, 23

LH2 Liquid Hydrogen 24

LLM arge Language Model 176

LP Linear Programming 55, 79–81

LSTM Long Short-Term Memory xv, 33, 34, 37–40, 77, 158, 159, 161, 169, 171

MA Moving Average 28–31, 33, 77

MAE Mean Absolute Error 84, 85

MILP Mixed Integer Linear Programming 55, 79, 81, 82, 91, 112, 114, 115, 124, 149

ML Machine Learning 3–5, 48, 146, 157, 158, 163, 169, 170

MLP Multilayer Perceptron 33–35, 77, 103

MPC Model Predictive Control 53

MPPT Maximum Power Point Tracking 12

MSE Mean Squared Error 78, 159, 161

NES Normalized Energy Shortage 89, 90, 165, 170

NIST National Institute of Standards and Technology 95, 97–100, 102

NOCT Nominal Operating Cell Temperature 61

NPV Net Present Value 85, 86, 116, 117, 135

NSF Normalized Shortage Frequency 89, 90, 165, 169, 170

NSGA-II Non-Dominated Sorting Genetic Algorithm II 51–53, 112, 114, 119, 121, 129

PEM Proton Exchange Membrane 24–26, 67, 68

PSO Particle Swarm Optimization 51, 52

PV Photovoltaic 6, 9, 10, 12, 13, 16, 22, 24, 46, 47, 49–53, 55, 56, 72, 90, 91, 93–96, 105–107, 113, 114, 116–118, 122, 126, 135–140, 146, 147, 151, 154, 155, 174, 198

RE Renewable Energy 2–9, 13, 15–17, 19, 20, 22–26, 46–49, 51, 53, 54, 73, 84, 88, 114, 135–140, 146–148, 151, 161

RES Renewable Energy System 1, 3–8, 10, 16, 17, 19, 20, 22–26, 45, 50–55, 67, 72, 79, 84, 87, 90, 91, 111–115, 121, 124, 130, 132–135, 138, 139, 141, 142, 145, 146, 151, 153, 160, 163, 174, 176

RF Random Forests 32, 33

RMSE Root Mean Squared Error 78, 85, 105

RNN Recurrent Neural Network 33–37, 40, 42, 45

SLF Satisfied Load Fraction 88, 116, 130, 140

SOC State of Charge 67, 73, 123, 149, 155

SOE Solid Oxide Electrolysis 24–26

SOH State of Health 67, 124, 138, 169

STC Standard Test Conditions 61

SVM Support Vector Machines 32, 33

TRNSYS TRNSYS Simulation Software 95, 96, 104, 105

UF Utilization Factor 88, 116, 140, 165, 167, 170

UTC Coordinated Universal Time 57

VAWT Vertical-Axis Wind Turbine 15

WT Wind Turbine 6, 55, 72, 90, 91, 146, 151, 154, 155

Chapter 1

Introduction

This chapter provides a comprehensive foundation for the thesis. Section 1.1 discusses the background and motivation, focusing on the economic, technological and environmental drivers that emphasize the importance of Renewable Energy System (RES) optimization. Section 1.2 outlines the primary objective of the research, emphasizing the key contributions related to system modelling, system development, and improvements in energy optimization strategies. Finally, Section 1.3 provides an overview of the thesis structure, describing the logical flow of the subsequent chapters and highlighting the interconnections between the main topics.

1.1 Background and Motivation

Energy independence is crucial for a nation's modernization, economic security, and resilience. It promotes technological advancement as demand for energy continues to grow, despite improvements in efficiency [3]. Energy independence mitigates risks from volatile energy prices and economic fluctuations, particularly for nations reliant on imported energy. Historically, energy consumption is closely linked to global real Gross Domestic Product (GDP), with fluctuations impacting inflation due to dependencies on affordable energy

for competitive production and efficient logistics. Furthermore, energy independence bolsters national resilience by ensuring a stable, secure energy supply, reducing reliance on external sources, and minimizing geopolitical tensions. This stability is vital for maintaining economic activities and supporting critical infrastructure in challenging times, thereby fostering a sustainable and prosperous future.

However, the pursuit of energy independence is increasingly challenged by the surging global demand for energy and the continued reliance on fossil fuels. As the global population and economic activities expand, there has been a corresponding surge in the demand for energy within human societies over recent decades. Projections indicate that by 2050, both the global population and energy demands are expected to increase by approximately one-quarter and three-quarters compared to 2023, respectively [4]. Despite the critical role of energy as a primary driver of economic prosperity, the predominant reliance on fossil fuels, which constituted 81% of the total energy supply as of 2019 [5], presents significant challenges. The combustion of these fossil fuels emits a variety of byproducts, including COx, SOx, and NO_x, which are major contributors to climate change, environmental degradation and significant health concerns [6]. These environmental and health impacts present substantial barriers to the sustainable development of societies and the sustained growth of global economies.

The reliance on conventional fossil-based energy sources thus necessitates an urgent transition towards more sustainable energy solutions, such as Renewable Energy (RE) sources—including solar, wind, and hydrogen energy. RE sources like solar and wind power emerge as essential solutions to this global challenge, offering the potential to meet surging energy demands while mitigating environmental impacts. Hydrogen energy, particularly when produced using renewable resources (green hydrogen), acts as a promising energy carrier. It can provide heat through direct consumption or generate electricity via fuel cells, with water as the only byproduct.

To effectively harness the potential of sustainable energy and achieve energy independence, addressing the intermittency of RE sources such as solar and wind is crucial. This variability in energy production can lead to grid instability and inefficient utilization. The

RES offers a comprehensive solution by integrating seamlessly with existing grid infrastructures. These systems not only harness energy from renewable sources but also incorporate Energy Storage System (ESS) to store excess energy for later use [7]. Additionally, RES can connect to the local grid [8] with minimal hardware upgrades, allowing the grid to serve as a backup during shortages and enabling energy trade to enhance RE utilization. This integration effectively mitigates the challenges of RE intermittency, promoting a more stable and efficient energy landscape.

Additionally, the implementation of an effective energy management algorithm is crucial for RES to enhance the utilization rates of RE and achieve cost-effectiveness, particularly as the Levelized Cost of Energy (LCOE) of RE remains higher than that of energy derived from fossil fuels. The energy management algorithm is designed to optimize the distribution and consumption of energy, ensuring alignment between energy availability and user demand patterns. By managing how energy is consumed, the algorithm not only guarantees efficient use of RE but also maintains grid stability amid the variability of RE production. Furthermore, it optimizes user energy costs, striving to maximize the utilization of RE while minimizing expenditure, thereby addressing economic and efficiency challenges in the energy sector.

Given the challenges discussed above, the primary motivation of this research is to enhance the utilization of green energy resources through the development of a RES tailored to the energy demands of a mid-scale community. This study will first focus on accurately modelling the dynamics and interactions of various RE sources to reflect their real-world performance and integration capabilities. It then seeks to optimize the capacity of these sources, ensuring they can meet energy demands and operate efficiently under various scenarios. Lastly, the research will incorporate an advanced energy management strategy, utilizing Machine Learning (ML) algorithms for precise demand prediction and optimization algorithms to effectively balance supply and demand. This strategy is designed to achieve economic efficiency, ensure grid stability, and foster sustainable energy development, positioning RE as a dependable solution for energy needs.

1.2 Objective and Contributions

The main objectives of the research presented in this thesis are centred around developing, optimizing, and managing a grid-connected RES to support the energy demands of a mid-scale community. The detailed objectives of this research are outlined as follows:

- Modeling the RES Components with a ML Algorithm: This objective focuses on using ML techniques to accurately simulate the behaviour and interactions of RES components such as solar panels and hydrogen storage systems under diverse environmental and operational conditions. The goal is to create predictive models that enhance the understanding and operational efficiency of these components.
- Optimizing the Size of Components in the RES: This objective involves determining the optimal size and configuration of the RES components by calculating the most efficient and cost-effective scale of RE resources and energy storage units. The optimization will consider technical and financial indicators to ensure that the RES can efficiently and economically meet energy demands.
- Developing an Advanced Energy Management Strategy: The final objective is to develop a sophisticated energy management strategy that effectively integrates the hydrogen-based RES with the existing power grid. This strategy will utilize ML algorithms for precise demand forecasting and optimization algorithms to balance energy supply with demand. The aim is to enhance the utilization rate of RE sources and ensure stable, efficient, and sustainable energy delivery.

1.3 Structure of the Thesis

The rest of the thesis is organized as follows: Chapter 2 provides background information on RESs, offering an in-depth literature review of RE and energy storage systems. This chapter also introduces the ML algorithms and optimization techniques used in the study.

Chapter 3 presents an extended literature review of the general methodology employed in this thesis, including the mathematical modelling of the main components in the RES. It details the ML techniques used, including model implementation and architecture, alongside the optimization techniques.

Chapters 4, 5, and 6 present the three main projects of this research. Chapter 4 explores and validates the use of ML algorithms in modelling the system, specifically focusing on optimizing the tilt angle of solar panels. The ML based modelling technique will be employed in subsequent projects. Chapter 5 discusses the optimization of the main components for a typical RES, which is crucial in the early stages before the system is built. This optimization is based on stakeholder requirements to determine the actual size of the system. Chapter 6 develops a robust energy management algorithm to enhance the RE utilization rate using advanced ML and optimization techniques.

Chapter 7 summarizes the work, offering conclusions and outlining potential research avenues in the field.

Chapter 2

Literature View

This chapter builds upon the discussions in the previous chapter by offering a detailed review of the current literature related to the RES and reviewing the relevant methodologies. Section 2.1 begins with an introduction to the concept of RES, addressing their existing challenges and potential solutions. Section 2.2 explores the current state of RE technologies and relevant modeling techniques, focusing on the Photovoltaic (PV) panel for solar energy and the Wind Turbine (WT) for wind energy. Section 2.3 discusses ESSs, emphasizing their crucial role in enhancing system efficiency and stability. The chapter then transitions to time series data analysis in Section 2.4, where the application of predictive models in energy systems is presented. Section **??** examines deep learning approaches for data prediction, providing insights into advanced forecasting methods. Finally, Section 2.5 reviews various system optimization techniques, highlighting their relevance within the research framework and emphasizing the importance of optimizing RES for improved performance.

2.1 Renewable Energy Systems

RES consists of four main components: generation, storage, distribution, and consumption. Generation relies on RE resources such as solar, wind, and biomass, selected based

on power stability, responsiveness to demand fluctuations, and costs, including both initial investments and operational expenses. ESS are essential for balancing supply and demand by storing excess energy during periods of high generation and releasing it when production is low, addressing the intermittent nature of resources like solar and wind. On the distribution side, integrating RE into existing networks is complex due to the variable and decentralized nature of these resources. This integration requires advanced monitoring, control, and communication technologies to efficiently manage energy flow and maintain grid stability. Finally, in the consumption phase, RE supports a range of applications across residential, commercial, and industrial sectors. By lowering energy costs, enhancing sustainability, and aiding in regulatory compliance, RES contribute substantial benefits across these sectors.

These systems are typically classified into two main types: stand-alone systems and grid-connected systems, each serving distinct purposes. Stand-alone RESs are often deployed in remote or isolated areas where access to centralized grids is limited or non-existent. By leveraging local RE resources and utilizing storage solutions, these systems provide reliable power independent of a larger grid. Stand-alone systems address challenges such as energy access and autonomy in off-grid locations, contributing to sustainable development in rural and remote communities [9]. In contrast, grid-connected RESs integrate directly with existing power grids, enhancing the grid's capacity to incorporate RE [10]. These systems support various strategies such as demand response, which adjusts energy consumption based on real-time supply and demand; peak shaving, which reduces peak load demand on the grid; and energy trading, where storage systems buy electricity during periods of low demand and sell it during peak demand. By taking these strategies, grid-connected systems help maintain grid stability and improve operational efficiency, ensuring a more resilient and responsive energy infrastructure.

There are several significant challenges while RESs play a crucial role in promoting RE utilization. Those challenges including energy storage and management, grid integration, and economic and policy barriers. First, the intermittent nature of RE sources poses substantial difficulties in maintaining a balance between supply and demand. Energy storage solutions, such as batteries, are crucial for storing and releasing energy as needed.

However, issues like limited lifespan and degradation highlight the need for storage systems that are not only cost-effective but also scalable and durable. Second, integrating RES into existing grid infrastructures requires adaptation to the bidirectional flow of energy, unlike the traditional one-way flow. The grid must be flexible enough to accommodate fluctuating energy inputs, necessitating advanced technologies and infrastructure upgrades to manage this variability efficiently. Finally, economic and policy-related challenges present barriers to widespread RES adoption. Financial incentives, regulatory frameworks, and market structures significantly influence RES deployment. Overcoming regulatory limitations on grid access and decentralizing energy supply will require a stable policy environment that supports long-term sustainability goals.

Current literature on RES emphasizes advancements in energy storage, artificial intelligence, and optimization techniques. These technologies are essential for improving the efficiency, reliability, and scalability of RES, addressing critical challenges such as storage limitations, grid integration, and variable energy supply. Additionally, evaluation tools such as RETScreen play a significant role in initially assessing the technical and financial feasibility of renewable energy projects. RETScreen is a software tool developed by the Canadian government; and it is widely used for benchmarking, conducting feasibility studies and evaluating project performances. It employs simulation models to assess key factors such as energy production, costs, emissions, financial viability, and risks. Additionally, RETScreen offers a comprehensive database that supports decision-makers in evaluating the feasibility and performance of renewable energy systems. Details on these technologies will be discussed in the following sections.

2.2 Renewable Energy Resources

RE refers to energy sources that are perpetually replenished by natural processes. The principal RE resources include solar, wind, hydro, biomass, and geothermal energy. These renewable, carbon-free energy sources provide several advantages over traditional, carbon-based energy sources. Firstly, unlike fossil fuels, which are concentrated in specific regions,

RE sources are globally accessible. Secondly, RE does not produce greenhouse gas emissions during the production and consumption stages, making it essential for transitioning to a low-carbon economy. Thirdly, the long-term costs of RE are lower than those of conventional carbon-based energy due to decreasing technology costs and economies of scale. Among these RE sources, solar and wind energy are particularly promising due to their rapid development and increasing cost-effectiveness.

Solar and wind energy are increasingly recognized as essential for achieving the ambitious net-zero emissions goal by 2050. These RE sources contribute significantly to the decarbonization of the energy sector. Over the past decade, the deployment of these technologies has experienced exponential growth; global solar energy capacity has increased by approximately 22% annually [11], while wind energy has seen an annual growth rate of approximately 11.4% [12] [13]. By the mid-21st century, it is anticipated that wind power will contribute over 30% to global electricity production, supporting numerous countries' objectives towards achieving carbon neutrality. Furthermore, wind energy could significantly mitigate CO2 emissions, potentially reducing them by up to 14,871 Mt by 2050 and 32,864 Mt by 2100 [14]. Looking ahead, solar energy alone is expected to form the cornerstone of the global electricity grid by 2050, potentially supplying nearly 40% of global electricity demand [15]. Solar and wind energy are projected to account for approximately 70% of the global electricity supply.

2.2.1 Solar Energy and Photovoltaic Panel

Solar energy, recognized for its sustainability and potency, has witnessed a significant increase in adoption, driven by technological advancements in PV systems and a notable reduction in production costs. The LCOE for solar PV technology, which was approximately $0.05 per kilowatt-hour in 2022, is projected to decrease further to $0.014 per kilowatt-hour by 2050, underscoring the economic viability of this RE source [11]. This form of energy is primarily captured through PV cells that convert sunlight directly into electricity via the PV effect, a phenomenon first identified in 1839. In these cells, sunlight excites electrons within semiconductor materials, generating an electric current that

is then harnessed by the cell's electrodes for practical use. This process of direct energy conversion, which eliminates the need for mechanical parts or chemical reactions, offers a clean and efficient energy transition, aligning with global sustainability goals.

PV systems exhibit remarkable versatility across various applications, enhancing their utility and adoption in both residential and commercial settings. These systems are effectively integrated into building designs through Building Integrated Photovoltaics [16], contributing both aesthetically and functionally to the structures. Large-scale applications, such as utility grids, benefit from the stability that PV systems provide, supplemented by advanced inverters and RES technologies. In remote or rural areas, stand-alone systems [17] paired with batteries deliver essential energy solutions. Moreover, PV technology supports sustainable practices in transportation and agriculture, such as powering electric vehicle charging stations [18] and facilitating agrivoltaic systems [19] that combine crop production with electricity generation.

Since the development of the first practical PV cell at Bell Labs in 1954, continuous improvements have been made, particularly in the areas of material efficiency, and integration with other technologies. These developments are further influenced by governmental policies and market dynamics.

PV cell materials are classified based on their properties and applications. Silicon-based cells, predominantly monocrystalline and polycrystalline silicon, are the market leaders because they balance efficiency and cost. Monocrystalline cells provide high efficiency due to their pure, uniform crystalline structure, though they are more expensive [20]. Conversely, polycrystalline cells are less expensive and slightly less efficient. Additionally, thin-film solar cells such as amorphous silicon, cadmium telluride, and copper indium gallium selenide offer flexibility and lower costs but with reduced efficiencies [21]. Emerging materials like perovskite and tandem cells, still mainly experimental due to stability issues, along with biohybrid and organic photovoltaic, which struggle with efficiency and durability, represent the frontier of research in this field [22]. The summary and efficiency of these materials are comprehensively detailed in Table 2.1.

Table 2.1: **Efficiencies, Advantages, and Disadvantages of Photovoltaic Technologies**

Technology	η (%)	Advantages	Disadvantages
Monocrystalline Silicon [23]	15-26	High efficiency; stable; long lifespan	Higher cost; energy-intensive production
Polycrystalline Silicon [23]	13-18	Lower cost; simpler production	Lower efficiency; somewhat shorter lifespan
Amorphous Silicon [24]	6-10	Very flexible; lower production costs	Lower efficiency; less durable
Cadmium Telluride [25]	10-18	Good efficiency; faster production	Toxic materials used; disposal concerns
Copper Indium Gallium Selenide [26]	12-20	High efficiency; flexible applications	Complex manufacturing; cost issues
Perovskite Solar Cells [27]	Up to 25	High efficiencies; potential for low-cost production	Stability concerns; environmental impact
Tandem Solar Cells [28]	Above 30	Very high efficiency; broad spectrum absorption	High cost; complex manufacturing
Organic Photovoltaic Cells [29]	3-17	Low-cost production; flexible uses	Lower efficiency; longevity issues

Photovoltaic Modeling

PV technology captures and converts solar energy into electricity, utilizing PV cells that transform solar irradiance into Direct Current (DC) electricity. Each PV array comprises multiple panels, with each panel consisting of several PV cells. These arrays are adaptable to various settings, including household rooftops and commercial open fields. Adjustments such as tilting the panels and employing Maximum Power Point Tracking (MPPT) algorithms significantly enhance the energy conversion efficiency of these systems, as validated by numerous studies [30] [31] [32].

The effectiveness of PV systems hinges on precise modeling of solar irradiance and the subsequent energy output calculations. The modeling process, essential for maximizing PV system performance, typically involves estimating the solar irradiance that each panel receives and calculating the expected energy output from this data. Such modeling is crucial for tailoring PV systems to their specific environmental contexts and is commonly achieved through mathematical approaches that allow for nuanced predictions and operational planning.

Various modeling approaches are employed in literature to optimize and predict the performance of PV panels under different environmental conditions:

- **Empirical models** are based on experimental data and observations. These models are straightforward but may not capture complex scenarios with high accuracy.
- **Analytical models** derive from fundamental principles that describe the behaviors of system components. These models offer enhanced accuracy over empirical models by integrating theoretical insights into system operations.
- **Numerical models** tackle complex mathematical equations that characterize the PV system's behavior using numerical techniques. While these models are the most flexible and accurate, they require significant computational resources to execute.

Further detailing the modeling of solar irradiance, it's crucial to understand its three main components: direct, diffuse, and reflected irradiance. Each component affects the to-

tal energy received by a PV panel and, therefore, its performance. Direct irradiance is the unfiltered solar radiation that reaches the earth, diffuse is scattered by atmospheric particles, and reflected irradiance varies with ground characteristics. These can be measured directly or estimated through sophisticated mathematical models, setting the groundwork for accurate system modeling.

To refine the approach to modeling PV systems, various specific irradiance models have been developed to cater to different environmental and operational needs. Models such as Hay and Davies, Perez, and others each have their tailored applications and associated strengths and challenges. For instance, the Hay and Davies model is optimal for clear skies, whereas the Perez model excels in varying weather conditions. Such diversity in models allows for targeted strategies in system design, optimizing the interaction between irradiance and panel technology. Detailed comparisons of those models are listed in Table 2.2.

Finally, after establishing the incident irradiance using an anisotropic model, the 5-parameter model comes into play to estimate the power output of the PV panel. This model delineates the panel's current-voltage characteristics through key parameters like photocurrent, saturation current of the diode, and resistances, which are crucial for assessing performance under varied conditions. The integration of the irradiance model with the 5-parameter output model facilitates a comprehensive simulation of the PV system's behavior from sunlight exposure to electrical output. This integrated approach is essential for designing and optimizing PV systems, ensuring they perform optimally in their respective environments and yield maximum energy efficiency.

The detailed discussion of the mathematical expression, including the Liu and Jordan model and the five-parameter model, will be presented in the next chapter.

2.2.2 Wind Energy and Wind Turbine

Wind energy represents another significant RE resource with the potential to substitute conventional fossil fuels and expedite the transition to net zero emissions. In 2019, wind

Table 2.2: **Comparison of Solar Irradiance Models**

Model	Advantages	Disadvantages
Hay and Davies [33]	Suitable for clear skies, uses simple adjustments based on sun's position	Less accurate in cloudy or varied weather conditions
Perez [34]	Handles varying weather conditions, detailed analysis of diffuse and direct components	Requires extensive atmospheric data, high computational cost
Klucher [35]	Includes effects of terrain reflectivity, good for areas with significant ground reflection	Less effective in non-reflective terrain or complex urban areas
Reindl [36]	Ideal for urban environments, integrates multiple radiation components	High complexity and computational demand
Liu and Jordan [37]	Simple, low computational demand, good for preliminary studies	Lacks accuracy in directional irradiance, not suitable for detailed analysis

power generated 1427 TWh, significantly influencing the global energy landscape [5]. This output has consistently risen from 2005 to 2019, driven by technological improvements and resulting cost decreases. Notably, countries within the Organization for Economic Co-Operation and Development are the main producers, followed by China [38]. However, compared to the total energy consumption of 606 EJ in 2019, the contribution from wind energy, which equates to 5.14 EJ, remains relatively minor. Despite this, the potential of wind energy technology remains significant, presenting a feasible route toward a more sustainable future.

Wind turbines, essential for harvesting wind energy, are classified based on three key aspects: orientation, capacity, and application. Starting with orientation, wind turbines

are divided into Horizontal-Axis Wind Turbine (HAWT) and Vertical-Axis Wind Turbine (VAWT). HAWTs, due to their established technology and cost-effectiveness, are widely used and more efficient, particularly suitable for non-urban areas. In contrast, VAWTs, which are typically smaller, find preference in urban settings due to their adaptability to limited spaces [39].

Moving to capacity, wind turbines are segmented into small, medium, and large based on their power output [40]. Small turbines, generating under 100 kW, are ideal for residential or small agricultural settings. Medium turbines, with outputs between 100 and 1000 kW, are utilized predominantly in commercial applications. Large turbines, capable of producing more than 1000 kW, are designated for utility-scale electricity generation.

Lastly, regarding application, wind turbines are categorized as either offshore or onshore. Offshore wind turbines, often employing HAWT designs, are strategically installed in oceanic settings to harness the stronger and steadier winds available offshore [41]. Such placement not only enhances energy yield but also significantly mitigates visual and auditory disturbances, in addition to minimizing land use impact. Despite these advantages, offshore turbines require substantially higher initial investments and incur greater ongoing maintenance costs. These increased expenditures are mainly due to their remote, maritime locations, which are exposed to harsh environmental conditions, and the larger dimensions of offshore turbines compared to their onshore counterparts, which have higher construction costs. Conversely, onshore wind turbines are installed on land and are typically less costly in terms of both installation and maintenance. Although smaller and subjected to less consistent wind speeds than offshore turbines, onshore installations often face significant challenges related to visual and noise pollution.

2.2.3 Impact of Renewable Energy

The increasing integration of RE sources into the global energy landscape presents unique challenges to traditional electrical grid systems. These challenges stem primarily from the inherent intermittency and variability of RE sources. In contrast to conventional energy

sources such as fossil fuels that deliver a steady and predictable power output, production from RE sources fluctuate significantly, depending heavily on meteorological conditions.

This variability introduces a critical challenge: the intermittency of power supply [42]. For example, solar and wind energies do not generate power consistently throughout the day or under varying weather conditions. This leads to periods of significant overproduction or underproduction in relation to the demand. During times of strong wind or intense sunlight, the excess electricity can overwhelm the grid, potentially leading to overloads. Moreover, if this surplus energy cannot be efficiently exported or utilized immediately, it may result in curtailment and the wastage of valuable energy. Conversely, during periods of low wind or cloud cover, the power generated may not suffice to meet the demand, further complicating efforts to balance the load on the grid. This variability necessitates robust management strategies to maintain grid stability and efficiency.

Furthermore, the integration of these RE sources impacts essential grid operations such as frequency and voltage control [43]. Conventional power grids are largely dependent on synchronous generators that contribute to inertial response, crucial for maintaining frequency and voltage within safe operational limits. However, RE technologies like PV systems and wind turbines lack inherent inertial responses, which can lead to potential instabilities in grid frequency and voltage, particularly during abrupt generation changes.

To address these challenges, it is essential to develop a comprehensive understanding of how RE geneators interact within broader RESs, not just within electrical grids. The advancement of new technologies and management strategies is crucial. The deployment of energy storage solutions, such as batteries and hydrogen storage [44], becomes not only beneficial but necessary for constructing resilient and efficient systems capable of accommodating the dynamic nature of RE. These systems must be designed to manage a diverse mix of energy types [45], ensuring stability and efficiency across varying operational conditions.

2.3 Energy Storage Systems

ESSs are fundamental to modern RESs, playing a pivotal role in the integration and optimization of energy resources. Serving as crucial bridges between energy generation and consumption, ESSs ensure that energy supply reliably meets demand by temporarily storing energy in specific forms and releasing it as needed. This capability enhances effective energy management and increases the efficiency of energy utilization. Additionally, by storing surplus RE, ESSs reduce waste and contribute to overall system sustainability. In the context of RES, ESSs have diverse applications across multiple scales, significantly advancing energy technologies.

On a macro-scale, ESSs function as buffers to absorb unstable generated energy from RE sources like solar and wind, which are inherently intermittent due to their dependence on unstable weather conditions. By storing excess energy during periods of high generation and releasing it during periods of low generation or high demand, ESSs prevent curtailment and enhance the efficiency of variable energy generation. This approach reduces overall energy costs by replacing expensive and environmentally harmful load-following plants and peak units in conventional energy systems. Additionally, the distributed nature of ESSs within RES diminishes the need for new transmission infrastructure.

On a micro-scale, ESSs play a role in energy management by enabling strategies such as load-shifting, peak shaving, energy arbitrage, and facilitating demand response. Load-shifting involves transferring energy consumption from peak demand periods to off-peak times by storing energy when demand or prices are low and utilizing it during high-demand periods. This strategy not only smooths out individual energy demands but also optimizes the use of available energy resources. Peak shaving refers to reducing the amount of energy drawn from the grid during peak usage times, thereby lowering peak demand charges and alleviating stress on the grid infrastructure. Through energy arbitrage, consumers can capitalize on fluctuating energy prices by purchasing and storing electricity when prices are low and using or selling it back to the grid when prices are high. Additionally, ESSs facilitate demand response by enabling consumers to adjust their energy consumption patterns in response to real-time signals or incentives from utility providers. By integrating

these approaches, ESSs not only reduce reliance on the local grid during peak times but also contribute to cost savings for consumers and enhance overall grid stability and efficiency [46].

With the advancement and maturation of energy storage technologies, ESSs have been applied in various fields. ESSs can be classified into five categories based on the technology or energy conversion process [47]: chemical, electrochemical, mechanical, electromagnetic, and thermal.

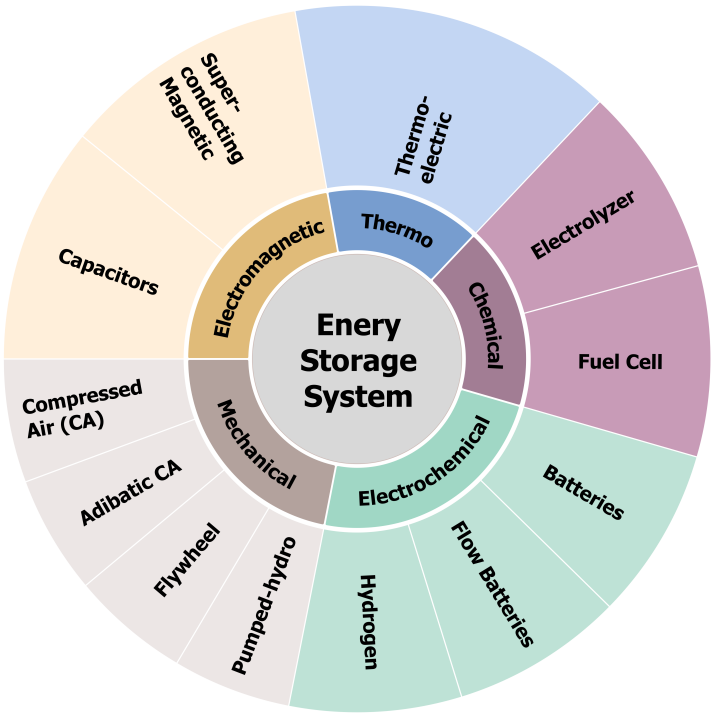


Figure 2.1: **Classification of Energy Storage System**, plot based on [1]

Given the diverse range of energy storage technologies, determining the most appropriate technology for a specific application requires a rigorous assessment of several critical factors. These factors include energy capacity, cycle life, continuous charge and discharge capability, response time, conversion efficiency, and long-term cost. The selection process for an ESS must take these considerations into account, as they are fundamental to aligning

the technology with the operational demands of the application. Accordingly, ESS applications are categorized into three principal types: energy management, bridging power, and power quality, each defined by distinct functional requirements.

- **Energy Management**: ESSs used for energy management need to continuously discharge for several hours, with an energy capacity ranging from 100kW to 1GW, depending on the system scale. Technologies such as electrolyzers, fuel cells, high-energy batteries, pumped hydro storage, compressed air energy storage, and thermal energy storage are suitable for this application. These systems enable shifting energy usage over longer periods, aiding in balancing supply and demand on a larger scale.
- **Bridging Power**: For bridging power applications, ESSs require faster response times than those used for energy management, with response times ranging from seconds to a few minutes and capable of several thousand life cycles. Continuous discharge of up to one hour is required. Conventional batteries are considered a good choice for this type of application, providing immediate backup power during short-term outages or transitions between power sources.
- **Power Quality**: ESSs for power quality must have rapid response capabilities, generally less than one second, with discharge durations up to 10 seconds and energy capacities ranging from 1kW to 100kW. Energy storage technologies involved in this application include flywheels, capacitors, and superconducting magnetic energy systems. These systems help maintain voltage stability, reduce flicker, and provide reactive power support to enhance the overall quality of power supplied.

For small to medium-scale RES, ESSs are primarily used for energy management purposes. This focus arises from the need to balance RE generation and consumption within localized systems, ensuring reliability and efficiency. Therefore, this study concentrates on ESS applications for energy management, with a particular emphasis on two of the most promising technologies: the Battery Energy Storage System (BESS) and the Hydrogen Energy Storage System (HESS). The subsequent sections will dive deeper into the

specifics of BESS and HESS, exploring their potential applications and benefits in energy management.

2.3.1 Battery Energy Storage Systems

BESS is a crucial subsystem in the RES, enhancing flexibility and reliability in integrating RE. By using batteries to store electrical energy, it supports the daily operation of the RES, ensuring a consistent and efficient power supply. Its high energy density, minimal energy efficiency loss, rapid response time, scalability, and deployment flexibility make the battery one of the most widely used ESS for the RES. These characteristics enable BESS to effectively manage fluctuations in energy generation and demand. Therefore, BESS is an ideal candidate for enhancing the reliability of RE supply, supporting energy management, and guaranteeing power quality.

A BESS typically comprises three main components: batteries, inverters, and control systems [48]. Batteries are the core energy storage units, operating through electrochemical reactions during charging and discharging processes. Battery technology significantly impacts the performance of BESS, influencing efficiency, cost-effectiveness, and suitability for various applications [49]. Inverters are devices that convert the DC electricity stored in batteries into Alternating Current (AC) electricity at specific voltages (typically 110V to 240V) for use upon discharge. Additionally, inverters facilitate the reverse process of converting AC to DC during charging. Finally, the control system manages and regulates the operation of the BESS, executing and supervising the charging and discharging processes based on system constraints and energy management commands.

BESS can be classified based on their functionality into two categories: Power-oriented Systems and Energy-oriented Systems. Power-oriented systems are mainly used for frequency regulation, providing backup power, and voltage support [50] [51]. These systems are required to deliver high power output over a short time frame, so the batteries need a high charge/discharge rate relative to their energy capacity and must respond rapidly. In contrast, Energy-oriented Systems primarily support energy management strategies and improve the overall efficiency of the energy system by enhancing the utilization of RE

[52][53]. Energy-oriented BESS are required to perform continuous charge and discharge over extended periods, which necessitates batteries with high energy capacity.

The battery technology can significantly impact the performance of the BESS. The mainstream technology can be divided into four categories: Lead-Acid Batteries, Nickel-Based Batteries, Lithium-Ion Batteries, and Flow Batteries [54]. In addition, the performance of those technologies can be measured from multiple dimensions: energy density, charge/discharge speed, cost and lifespan. Different technology has different advantages and limitations that influence its application in ESSs. Table 2.3 provides a comparison of these technologies across multiple performance dimensions, including energy density, charge/discharge speed, cost, and lifespan. Each technology has distinct advantages and limitations that affect its application in ESSs.

Table 2.3: **Comparison of Battery Technologies**

Criteria	**Lead-Acid**	**Nickel-Based**	**Lithium-Ion**	**Flow**
Energy Density (Wh/kg)	Low (30–50)	Moderate (40–80)	High (100–265)	Low (10–30)
Charge/Discharge Speed	Moderate	Moderate	Fast	Slow
Cost ($/kWh)	Low ($100–$200)	Moderate ($200–$500)	Higher ($300–$700)	High ($500–$800)
Lifespan (k Cycles)	Low (0.5–1)	Moderate (1–2)	High (2–10)	Very High (10–20)

Lead-acid batteries, particularly conventional valve-regulated lead-acid batteries, are among the oldest and most mature technologies in energy storage. Their most notable advantage is their low initial cost, making them suitable for applications such as backup power and uninterruptible power supplies. However, they have relatively low volumetric

and gravimetric energy densities compared to other battery types, meaning they store less energy per unit volume or weight. An advanced variant, the advanced lead-acid battery, incorporates ultra-capacitors to enhance performance. While advanced lead-acid batteries offer improved capabilities over standard lead-acid batteries, they are still in the demonstration phase and are significantly more expensive than typical valve-regulated lead-acid batteries.

Nickel-based batteries, including nickel-metal hydride (NiMH) and nickel-cadmium (NiCd) batteries, are known for their good performance at low temperatures and high tolerance to overcharging. NiMH batteries possess a high energy density; however, their high self-discharge rate limits their widespread use in PV systems and other RE applications. NiCd batteries, on the other hand, have been largely phased out due to the toxicity of cadmium, which has led to bans in many countries. Consequently, NiCd batteries are not considered viable candidates for use in RESs.

Lithium-ion batteries have gained prominence due to their high energy density, efficiency, and long cycle life, making them ideal for a wide range of applications from consumer electronics to electric vehicles and RE integration. The lithium battery family includes various chemistries such as lithium nickel manganese cobalt oxide, Lithium Iron Phosphate (LFP), lithium nickel cobalt aluminum oxide, and lithium titanate oxide. These batteries offer superior performance characteristics, including low self-discharge and minimal maintenance requirements. Despite their higher initial cost and the need for thermal management to mitigate risks like thermal runaway, lithium-ion batteries are widely adopted due to their efficiency and scalability.

Lastly, flow batteries, such as vanadium redox flow and zinc-bromine flow batteries, are recognized for their scalability and long operational life. They have the unique ability to decouple power and energy capacity, allowing for flexible system design where energy storage capacity can be increased independently of power output by enlarging the electrolyte tanks. This makes them suitable for large-scale energy storage applications requiring long-duration discharge. However, flow batteries typically have lower energy densities and higher system complexities, leading to higher capital costs and making them less practical

for small-scale or space-constrained applications.

This research uses LFP batteries because they avoid the significant drawbacks of other technologies when implementing energy-oriented BESS in RES. Lead-acid batteries have low energy density and a short cycle life, leading to frequent replacements and higher long-term costs, which is impractical for applications requiring a reliable, sustained energy supply. Nickel-based batteries, such as NiCd and NiMH, present environmental hazards due to toxic materials like cadmium and suffer from high self-discharge rates and memory effects, making them less efficient for RE storage. Flow batteries, while offering long lifespans and scalability, are still under development, have high initial costs, and possess low energy densities that require large physical spaces. Additionally, their exceptionally long lifespan exceeds the typical service life of RE installations (18 to 25 years), and with rapid technological advancements, more cost-effective solutions may emerge before the end of their operational life. Therefore, LFP batteries are chosen as they mitigate these issues, making them the most suitable option for developing efficient, safe, and sustainable energy-oriented BESS within RES.

2.3.2 Hydrogen Energy Storage Systems

Hydrogen is increasingly recognized as a next-generation energy source as its significance stems from several key advantages: high energy density, minimal emissions, and versatility across various applications [55]. These attributes make hydrogen an integral part of our society's shift towards sustainability, particularly in the domain of low-carbon transportation. Hydrogen's role extends to fuel cells [56], which convert it into electricity for powering electric vehicle motors, and direct combustion for powering aircraft [57], with water vapour being the only emission. Lubeck et al. assessed the life cycle of hydrogen-powered public buses, underscoring the significantly lower environmental and global warming impacts of hydrogen and electric buses compared to diesel ones [58]. Liu et al. have explored the potential of liquid hydrogen-fueled turbojet engines, establishing hydrogen as a clean combustion fuel and a valuable source of cold exergy for power generation [59]. Meanwhile, Hoelzen et al. have assessed the economic implications of hydrogen infrastructure in avi-

ation. Their findings suggest that the viability of hydrogen-powered aircraft is contingent on the cost of green Liquid Hydrogen (LH2) infrastructure, potentially leading to operating cost increases of 10-70% for short-range and 15-102% for medium-range aircraft due to LH2 costs [60]. It is crucial to recognize that the environmental impact of hydrogen production varies significantly, depending on its source: green hydrogen is produced using RE, blue hydrogen via natural gas, and grey hydrogen from coal [61]. Consequently, integrating HESS within RESs is crucial for fostering the production of green hydrogen, aligning with global sustainability goals.

RES with HESS has been widely explored in existing literature and proven it can surpass the system only with electricity. Dong et al. developed a stochastic optimization algorithm to optimize a RES with HESS and sell hydrogen for profit, achieving 80% energy cost saving and 100% RE utilization efficiency [62]. Wang et al.developed an optimization model for RES with PV and HESS on the city level, showing that electrolytic hydrogen can effectively store excess PV energy and reduce CO2 emissions, with hydrogen potential ranging from 0.236 to 9.795 megatons in ten Chinese cities, surpassing electrochemical storage in efficiency and cost [63]. Bellocchi et al. modelled a RES with electricity, gas and hydrogen, then performed techno-economic optimization to assess green-hydrogen pathways in Italy, finding hydrogen most effective in fuel cell electric vehicles and reducing CO2 emissions by up to 49% at an 8% cost increase [64]. Khalili et al. applied robust multi-objective optimization in Iran's electricity market, reducing retailer costs by 5.32% and electrolyzer power consumption by 14.19%, effective peak-shaving and load-shifting capabilities of hydrogen storage[65]. Li et al. developed a hierarchical scheduling method for RES with shared electricity-hydrogen storage using a Stackelberg game approach. This method, augmented by time-of-use hydrogen pricing, cut operating costs by 9.96% and boosted energy utilization by 2.97% compared to electric-only storage systems. [66].

In RES, HESS primarily produces hydrogen through the electrolysis of water, separating it into hydrogen and oxygen. The resulting gases are either compressed into high-pressure gas or liquefied into LH2. Three mainstream technologies dominate water electrolysis: Alkaline Electrolysis (AE), Proton Exchange Membrane (PEM) electrolysis, and Solid Oxide Electrolysis (SOE) [67][68][69].

AE, the most mature technology, has been widely used in industry for decades. This method employs solutions such as sodium hydroxide ($NaOH$) and potassium hydroxide (KOH) to conduct hydroxide ions (OH^-) from the cathode to the anode. Alkaline Electrolyzers, made from cost-effective materials like stainless steel and nickel, are the most economical option. However, they exhibit lower efficiency, typically between 60% and 70%, and their performance can be significantly affected by the solution concentration, temperature, and current density. Additionally, the long response time of AEs renders them unsuitable for use as energy storage components in RES to handle intermittent RE supply.

Table 2.4: **Comparison of Electrolysis Technologies**

Parameter	**AE**	**PEM**	**SOE**
Principle	Solution	Membrane	Solid ceramic
Maturity	Mature	More recent	Emerging
Efficiency	60-70%	Up to 80%	Over 80%
Current Density	Low	High	Medium
Cost	Low	High	Medium
Operational Temperature	40-80°C	20-90°C	700-800°C
Compactness	Moderate	High	Low
Response Time	Slow	Fast	Very Slow
Stability	High	Moderate	Low
Maintenance Cost	Medium	High	High
Machine Cost	Low	High	Medium

PEM Electrolysis, a more advanced technology, utilizes a solid polymer electrolyte membrane to transfer protons from the anode to the cathode, effectively separating hydrogen and oxygen. PEM operates at higher current densities, allowing for a more compact design than AE for equivalent hydrogen output levels. Moreover, its rapid startup and

response time make PEM an ideal component for managing intermittent RE in RES. The efficiency of PEM ranges from 70% to 80%, slightly higher than AE. However, PEM's higher costs, due to the use of precious metals such as platinum and iridium as catalysts, and the challenges associated with membrane durability, are notable drawbacks.

SOE operates at temperatures exceeding 700 degrees Celsius and employs a ceramic electrolyte instead of a membrane to conduct ions and separate hydrogen and oxygen. The high operational temperature of SOE offers significant advantages, including enhanced efficiency (exceeding 80%) and reduced catalyst costs, as it can utilize waste heat from industrial processes and geothermal energy. Nevertheless, the high operational temperatures present challenges related to material stability, thermal management, slow startup, and high maintenance requirements, making SOE more suitable for RES that utilize geothermal energy and possess waste heat recovery systems.

These three technologies exemplify the progression in hydrogen production methods, each with its inherent advantages and limitations and compared in Table 2.4. As technologies and materials advance, efficiencies are expected to improve, and costs are anticipated to decrease. This research employs PEM as the electrolyzer in its HESS because PEM's rapid response time effectively manages intermittent RE generation, and its compact design is advantageous for community use.

2.4 Time Series Data Analysis

Time series data is a prevalent concept in various fields, including finance, economics, meteorology, and the energy sector, and consists of data points collected at consistent time intervals. This data type is distinct from other types due to three unique characteristics: trend, seasonality, and cyclicity. The trend reflects the long-term progression in the data, indicating whether values increase or decrease over time, and showing the underlying direction of movement. Seasonality refers to recurring patterns that occur at regular time intervals, such as daily, monthly, or yearly. Unlike seasonality, which has fixed and predictable cycles, cyclicity involves patterns that occur at irregular intervals and are

influenced by broader, unpredictable factors.

In energy consumption data, the trend is shaped by factors such as technological advancements and consumer behaviour. For instance, the adoption of high-efficiency technologies may reduce overall energy consumption if consumption patterns remain stable. Conversely, increased consumption due to the convenience offered by these technologies could lead to a rising trend. Seasonality in energy consumption data is often driven by environmental changes; for example, energy use increases in winter for lighting as daylight hours are shorter than in summer. Cyclicity in the energy sector is influenced by factors such as shifts in production technology, regulatory changes, and innovations in energy storage and distribution, all of which occur unpredictably.

In addition to the three fundamental characteristics of time series data—trend, seasonality, and cyclicity—conventional statistical models often require the data to meet certain assumptions and properties. These assumptions include homoscedasticity, stationarity, and independence of errors.

Homoscedasticity refers to the requirement that the variance of the time series remains constant over time. In practice, energy consumption data rarely satisfy this assumption due to shifts in demand patterns and external factors, leading to heteroscedasticity. When heteroscedasticity is present, data transformations may be needed to achieve homoscedasticity.

Stationarity implies that the statistical properties of the data, such as mean, variance, and autocorrelation, remain constant over time. Non-stationary data, which may exhibit trends or seasonality, need to be addressed through techniques like differencing or detrending before being fed into the model.

Another important assumption is the independence of errors, which refers to the behaviour of residuals—the differences between observed and predicted values. The errors should resemble white noise, meaning they are uncorrelated, normally distributed with a mean of zero, and have constant variance. If this assumption is violated, it indicates that the model has not fully captured the underlying patterns in the time series, potentially reducing the model's accuracy.

2.4.1 Statistical Approaches

Statistical models, including Autoregressive (AR), Moving Average (MA), and Autoregressive Integrated Moving Average (ARIMA), play a critical role in time series analysis by enabling the examination of linear relationships in data. This section explores how these models, which are predicated on assumptions of data stationarity and linearity, offer structured methodologies for forecasting. It highlights the limitations of these models in addressing the complexities of real-world data. This section will cover each model's core features: AR's reliance on previous values, MA's utilization of past error terms, and ARIMA's capability to incorporate non-stationary data, demonstrating their advantages and optimal application contexts.

Autoregressive Model

The AR model is a fundamental statistical method used to forecast time series data by leveraging the dependencies between current and past observations. The core idea of the AR model is that the future value of a time series can be predicted as a linear function of its previous values, reflecting the temporal correlation inherent in the data sequence.

The mathematical expression of an AR model of order p, denoted as AR(p), is given in Equation 2.1.

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t \tag{2.1}$$

Here, Y_t represents the current value of the time series at time t , each ϕ (ranging from ϕ_1 to ϕ_p) is a coefficient that quantifies the influence of past values on the current value, and ϵ_t is the residual term, accounting for the unpredictability not captured by the past observations. This model structure allows the AR model to capture and utilize the underlying patterns within time series data effectively.

While AR models have been extensively studied and widely adapted due to their simplicity and effectiveness in capturing short-term dependencies in time series data, they still

face several challenges. First, AR models struggle with data that exhibit trends, seasonality, or non-stationary behaviors, as these characteristics can lead to biased or inefficient results. However, those behaviors are normal in meteorology data and energy consumption data.

Second, the sensitivity to the order of the model poses another significant challenge. The choice of the order of the AR model is critical; an inappropriate selection can lead to underfitting, where the model does not capture all the dynamics of the data, or overfitting, where the model is too closely fitted to the specific data points of the training set, reducing its generalizability and performance on new data.

Finally, AR models cannot capture complex relationships, especially in contexts like energy consumption data where multiple interacting factors influence the outcomes. This limitation makes AR models less effective in environments where the interactions between variables are significant and where variables influence each other in non-linear or intricate ways.

Moving Average Model

MA models are another fundamental class of statistical methods used for time series prediction. These models assume that the current value of a series is a linear combination of past white noise error terms. A Moving Average model of order q, denoted as MA(q), is demonstrated in Equation 2.2.

$$Y_t = \mu + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t \tag{2.2}$$

In Equation 2.2, Y_t represents the current value of the time series at time t. The mean of the series, denoted as μ, is often assumed to be zero in a zero-mean MA model. The coefficients of the MA model, $(\theta_1, \theta_2, \ldots, \theta_q)$, quantify the influence of lagged error terms on the current value. These coefficients reflect how past white noise error terms, $(\epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_{t-q})$, which represent random shocks from previous time points, impact the current series value. Additionally, ϵ_t is the current error term, also modelled as a white

noise process, introducing new, unpredicted information that affects the series at time t. This setup allows the MA model to effectively account for and smooth out random fluctuations in the time series data.

MA models are extensively utilized due to their proficiency in handling random variations and smoothing noisy data. This characteristic is particularly beneficial in financial markets for smoothing stock prices or in economic data analysis where unexpected events need to be accounted for.

Despite their utility in smoothing noisy data and handling random variations, MA models face several limitations. They inherently possess a 'finite memory,' limited only to a few periods defined by q, which might not adequately capture long-term dependencies within the time series. Like AR models, MA models require the series to be stationary; non-stationary data, such as those displaying trends or seasonal patterns, need to be differenced or transformed before modeling. Furthermore, with an increasing order q, MA models grow in complexity, complicating their interpretation and fitting, particularly as identifying the appropriate model order becomes more challenging. This aspect can limit their effectiveness in scenarios where the historical values themselves contain crucial information for forecasting, restricting their applicability in complex environments where understanding long-term historical dependencies is vital.

Autoregressive Integrated Moving Average Model

The ARIMA model combines the concepts of both the AR and MA models. It is widely applied in time series forecasting, particularly when data exhibit trend or non-stationary behavior, which AR and MA models alone cannot effectively address. ARIMA consists of three main components: AR, differencing to remove non-stationarity, and MA. These elements work together to model time series data by accounting for both the temporal relationships between past observations and the adjustments needed to handle non-stationarity, making ARIMA a more flexible approach for complex time series.

The general expression of an ARIMA model is denoted as ARIMA(p, d, q), where p is the number of autoregressive terms, d is the degree of differencing to make the series

stationary and q is the number of moving average terms. The mathematical expression is given in Equation 2.3.

$$\Delta^d Y_t = \phi_1 \Delta^d Y_{t-1} + \cdots + \phi_p \Delta^d Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} \tag{2.3}$$

Where:

- Y_t is the original time series at time t
- $\Delta^d Y_t = \Delta^{d-1} Y_t - \Delta^{d-1} Y_{t-1}$, $\Delta^1 Y_t = Y_t - Y_{t-1}$, are the differenced time series
- $(\phi_1, \ldots, \phi_p)$ are AR coefficients and $(\theta_1, \ldots, \theta_q)$ are MA coefficients
- ϵ_t represents the error terms.

The ARIMA model is specifically designed to address the challenges of non-stationarity, dependencies, and randomness in time series data, which frequently occur in real-world datasets, especially in the energy sector. The model first employs differencing, a technique that subtracts the previous observation from the current one, effectively stabilizing the mean by eliminating broader-scale changes. This transformation is critical for removing trends and making the data stationary, a necessary condition for effective time series analysis.

However, differencing alone does not capture the predictive power inherent in the past values of the time series. To address this, the ARIMA model incorporates AR components, which account for the relationships between observations and their lagged values, thus utilizing past data to improve future forecasts.

Lastly, time series data often experience random shocks—unpredictable events that cause deviations from expected patterns. These residual effects are not fully addressed by differencing and autoregression. The MA component of ARIMA handles these random shocks by modeling the noise or error terms from previous predictions. By averaging these residuals, the MA component smooths their impact, refining the model's accuracy and enhancing its ability to adapt to unexpected changes in the data. This comprehensive approach improves the model's predictive performance in complex real-world scenarios.

2.4.2 Machine Learning Approaches

Beyond traditional statistical methods, machine learning has emerged as a promising alternative for time series prediction, demonstrating significant potential to capture complex patterns in time series data. The literature identifies several key machine learning methods for this purpose, including Support Vector Machines (SVM), Random Forests (RF), and k-Nearest Neighbors (k-NN). Each of these approaches has unique advantages and limitations, making them suitable for specific types of time series data.

SVMs are particularly effective for modeling complex nonlinear patterns. They use kernel functions to map data into higher-dimensional spaces, where nonlinear patterns become more separable and thus easier to predict. This flexibility allows SVMs to handle various forms of nonlinearity, making them well-suited for small datasets with intricate relationships. However, applying SVMs to large datasets is computationally demanding due to the resource requirements of kernel transformations. Additionally, SVMs require careful tuning of kernel parameters, as these parameters directly affect prediction accuracy.

RFs are ensemble models that combine multiple decision trees to generate predictions. Each tree contributes independently to the final decision, and using multiple trees reduces the risk of overfitting, enhancing the model's robustness compared to individual decision trees. In time series prediction, RFs can capture nonlinear dependencies within the data and handle multivariate time series where multiple variables interact. However, RFs lack temporal awareness, meaning they cannot capture the sequential or ordered nature inherent in time series data. This limitation reduces their effectiveness in modeling complex temporal patterns. Furthermore, as the number of trees in the ensemble grows, the computational cost increases significantly, which can be a constraint in practical applications.

The k-NN algorithm, though simpler than SVM and RF, is occasionally used in time series prediction by leveraging the similarity between current observations and historical data. It assumes that patterns in the current data resemble those in past data, allowing it to identify the closest historical observations for making predictions. However, k-NN does not account for temporal order, making it less suitable for high-dimensional data where the sequence of observations is crucial. Additionally, its reliance on finding similar past

patterns limits its usefulness when the time series data are highly dynamic or irregular, as these conditions reduce the relevance of historical analogs.

In conclusion, while SVM, RF, and k-NN each offer valuable capabilities for time series prediction, they also have specific limitations. SVMs are adaptable to nonlinearity and useful for small, complex datasets but require significant computational resources and careful parameter tuning. RFs provide robustness and can capture nonlinear dependencies but lack temporal awareness and can be computationally intensive with large ensembles. k-NN offers simplicity and interpretability through similarity-based prediction but is less effective for high-dimensional, sequential data. Together, these methods expand the range of machine learning tools available for time series analysis beyond traditional statistical approaches.

2.4.3 Deep Learning Approaches

Deep learning models, as subsets of machine learning approaches, include architectures such as Recurrent Neural Network (RNN), LSTM, and Transformers are transforming time series prediction, offering a marked improvement over traditional statistical methods. This section examines how these advanced models excel in capturing complex, nonlinear patterns and dependencies that evade simpler statistical models like AR, MA, and ARIMA, constrained by their assumptions regarding data stationarity and linearity. By exploring the capabilities of each model—from RNNs' ability to process sequences through memory to LSTMs' enhancement for long-term dependencies, and Transformers' innovative use of self-attention mechanisms—the review will introduce their respective advantages and practical applications in the field. Additionally, a comparative analysis of these deep learning models is provided in Table 2.5.

Recurrent Neural Network

Feedforward neural networks, including Multilayer Perceptron (MLP), encounter substantial challenges when handling sequential data due to their inherent structural limitations.

Table 2.5: **Comparison of RNN, LSTM, and Transformer Models**

Metric/Model	cRNN	LSTM	Transformer
Computational Efficiency	Sequential processing; limited parallelism	Complex operations; slower than RNN	Highly parallelizable; efficient with GPUs
Handling Long Sequences	Struggles due to vanishing gradients	Good at capturing long-term dependencies	Excellent with long-range dependencies
Training Difficulty	Difficult due to vanishing gradients	Moderate; complex structure	Requires extensive parameter tuning
Memory Requirement	Moderate	High; more parameters	Very high; self-attention over sequences
Advantages	Simplicity; easy to implement	Handles long-term dependencies; mitigates vanishing gradients	Scalability; handles long-range dependencies well
Disadvantages	Poor at long-term dependencies; vanishing gradients	Computationally intensive; harder to parallelize	High resource requirements; complex architecture
Parallelizability	Low	Low	High
Model Interpretability	Low	Low to moderate	Moderate; attention weights provide insights

These networks are not equipped to manage sequences of variable lengths or to capture temporal dependencies that are critical for analyzing time series data. The root of this limitation lies in the static architecture of MLPs, which process each input independently, without the capability to store or reference earlier states. As a result, MLPs cannot account for the impact of prior inputs on subsequent outputs.

The RNN was developed specifically to address the challenges and limitations inherent in processing sequential data, notably those identified in models like the MLP. This innovation in neural network design marks a pivotal advancement in sequence processing technologies. RNNs feature a unique architecture that includes loops, allowing them to retain information from previous inputs and use it in processing subsequent inputs. This capability allows RNNs to preserve a memory of past inputs and to learn temporal relationships across various time steps, thus enabling a more nuanced understanding of sequential dynamics. RNNs have demonstrated superior performance in applications such as language modeling and time series analysis, where leveraging historical data is crucial for achieving high accuracy.

In RNNs, the nodes in a hidden layer are interconnected, and the input to the hidden layer includes the output of the input layer and the output of the hidden layer from the previous time stamp. In theory, RNNs can process arbitrary lengths of time series data. However, to reduce the complexity of RNNs, researchers often assume that current states are only related to a few previous states (i.e., short-term states).

At each time step t, the RNN takes an input vector x_t and computes the hidden state h_t based on the current input, the previous hidden state h_{t-1} , and internal parameters. The hidden state serves as the network's memory, capturing information from all previous inputs processed up to time t. The update equation for the hidden state is given by Equation 2.4:

$$h_t = \phi(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \tag{2.4}$$

In Equation 2.4, W_{xh} and W_{hh} are weight matrices for the input-to-hidden and hidden-to-hidden connections, respectively; b_h is a bias vector; and ϕ is a nonlinear activation function, such as sigmoid or tanh, which introduces nonlinearity into the network processing.

The output y_t at each time step t is computed from the current hidden state h_t as follows shown in Equation 2.5:

$$y_t = W_{hy} h_t + b_y \tag{2.5}$$

Here, W_{hy} is the weight matrix that connects the hidden state to the output, and b_y is the corresponding bias vector. This output configuration can be tailored for various tasks, from classification to complex sequence modeling, depending on the network's intended application.

Unfolding an RNN across the time steps of an input sequence provides a clearer understanding of its operations, as each temporal instance is shown as a distinct but connected unit within a larger computational graph. This expanded view helps illustrate the parameter sharing across different time steps, a fundamental feature that allows the RNN to generalize across sequences of variable lengths and structures.

Mathematically, unfolding an RNN can be represented as a series of function compositions, with each function corresponding to the state update and output computation at each time step:

$$h_t = \phi(W_{xh} x_t + W_{hh} \phi(W_{xh} x_{t-1} + \ldots)) \tag{2.6}$$

Conventional RNNs encounter several critical limitations that impair their ability to model long-term dependencies effectively. These include the vanishing gradient problem, the exploding gradient problem, and inherent memory retention limitations. Each of these challenges contributes to the difficulty RNNs face in maintaining and processing information over long sequence durations.

- **Vanishing Gradient Problem:** Gradients during backpropagation diminish exponentially as the sequence length increases, leading to minimal weight updates for early inputs. This causes the network to lose information about the initial part of the sequence, hindering its ability to connect distant events.
- **Exploding Gradient Problem:** Conversely, gradients can increase exponentially, resulting in very large updates to the weights. This instability often leads to a divergent training process where the model fails to converge, making it difficult to train effectively.

- **Memory Retention Limitations:** The architecture of conventional RNNs relies on a single hidden state to store all past information, which becomes overwhelmed with longer input sequences. As new inputs arrive, older information is overwritten, which limits the network's ability to remember long-term dependencies.

To address these limitations, LSTM networks introduce as an advanced RNN and significantly enhances the capability for times series prediction.

Long Short Term Memory

LSTM networks enhance conventional RNNs by introducing specific gating mechanisms that effectively manage the flow of information. These mechanisms address the issues of vanishing and exploding gradients, thereby preserving long-term dependencies in sequential data.

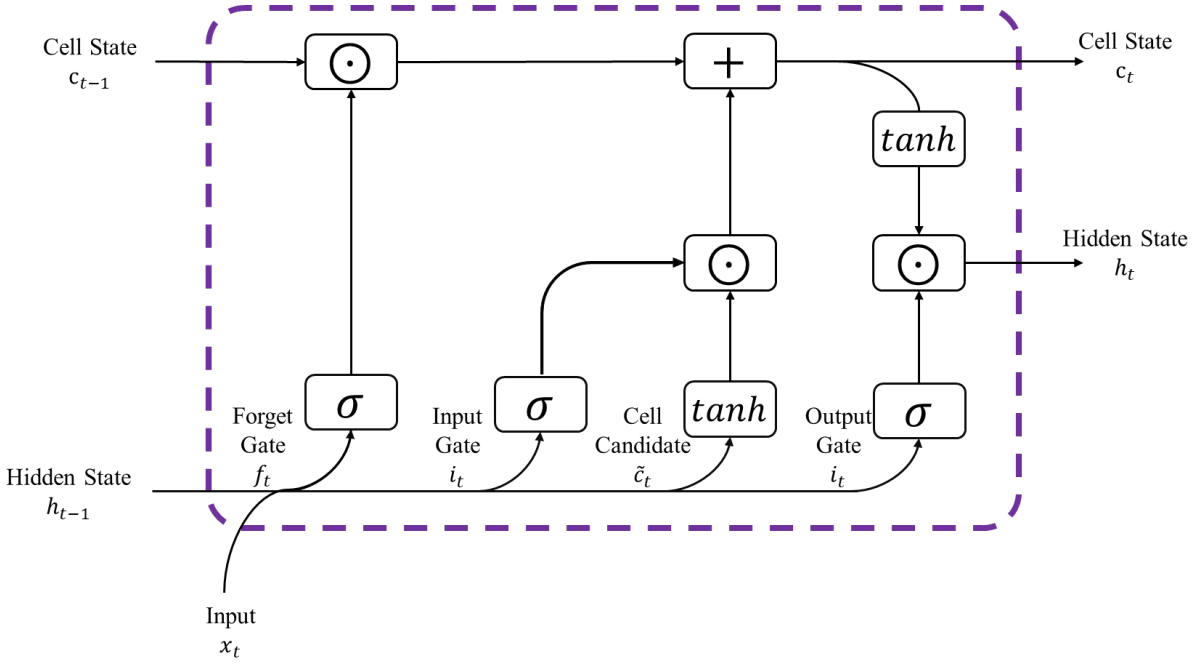


Figure 2.2: **LSTMs Structure**

With many studies, LSTMs have been proven to be able to handle long-term time series problems and it has much better performance than other ML structures. In the literature, several studies have been conducted on the application of LSTM models to improve the accuracy of load forecasting in the power industry. Xu et al. [70] proposed a modified two-layer LSTM neural network for joint forecasting of active and reactive loads in building energy systems, which outperformed a single LSTM-based approach. Mughees et al. [71] developed a deep bidirectional LSTM-based sequence-to-sequence (S2S) regression approach for day-ahead peak demand forecasting, achieving better performance than several other benchmark models. Zang et al. [72] proposed a novel day-ahead residential load forecasting method that utilizes feature engineering, pooling, and a hybrid deep learning model based on LSTM and self-attention mechanism, which was shown to be effective and superior to other benchmark methods through case studies. Gul et al. [73] focused on mid-term electricity consumption prediction for a metropolitan city using a time-series dataset and compared the performance of ARIMA and CNN-Bi-LSTM models, showing promising accuracy after tuning hyperparameters. These studies demonstrate the potential of deep learning models for improving load forecasting in the power industry.

An LSTM cell comprises several components, each governed by precise mathematical equations that control information flow. The key elements include the forget gate, the input gate, the cell candidate, the cell state update, the output gate, and the hidden state. Each component plays a crucial role in the cell's functionality. The architecture of the LSTM cell is demonstrated in Figure 2.2.

The forget gate, denoted as f_t, determines the extent to which information from the previous cell state c_{t-1} is retained or discarded. It is defined by Equation 2.7:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{2.7}$$

In Equation 2.7, σ represents the sigmoid function, ensuring that the output of the forget gate ranges between 0 (information is completely forgotten) and 1 (information is fully retained). The parameters W_f, U_f, and b_f are the weights and bias specific to the forget gate, influencing how it processes the current input x_t and the previous hidden state

h_{t-1}.

The input gate, represented by i_t, controls the incorporation of new information into the cell state. Its operation is described by Equation 2.8:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{2.8}$$

The cell candidate, denoted as $\tilde{c}_t$, provides potential new values for updating the cell state. It is calculated using Equation 2.9:

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{2.9}$$

In this equation, tanh refers to the hyperbolic tangent function, producing outputs in the range of -1 to 1. The parameters W_c, U_c, and b_c are specific to the cell candidate, interacting with x_t and h_{t-1}.

The cell state update, denoted as c_t, integrates the previous cell state and the new candidate values to refresh the cell state. This process is defined by Equation 2.10:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{2.10}$$

In Equation 2.10, $\odot$ signifies element-wise multiplication. This operation is pivotal for selectively updating the cell state based on the outputs from the forget and input gates.

The output gate, denoted as o_t, determines which portions of the cell state are transmitted as output. Its functionality is captured by Equation 2.11:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{2.11}$$

Finally, the hidden state h_t, which serves as the output of the LSTM cell for the current timestep, is computed using Equation 2.12:

$$h_t = o_t \odot \tanh(c_t) \tag{2.12}$$

This equation indicates that the hidden state is the element-wise product of the output gate's activation o_t and the hyperbolic tangent of the updated cell state $\tanh(c_t)$. This operation effectively filters and transmits the necessary information for future predictions or sequence processing.

Each component within the LSTM cell is meticulously designed to selectively update, retain, or forget data. The activation functions σ and tanh normalize the gate outputs and candidate values, which helps stabilize the updates applied to the cell state. Consequently, LSTMs are particularly effective for tasks that require managing long-term dependencies in sequential data.

Moreover, the gating mechanisms inherent in LSTMs play a critical role in mitigating the vanishing and exploding gradient problems common in traditional RNNs. By regulating the flow of information through the cell state, these gates maintain more consistent gradients during backpropagation through time. The forget gate prevents the accumulation of irrelevant information, which could lead to exploding gradients, by selectively discarding unnecessary data. Simultaneously, the input gate and cell candidate ensure that only pertinent new information is incorporated into the cell state, normalized by the tanh activation function to keep values within a bounded range. The output gate further refines the information passed to subsequent layers, preventing the propagation of vanishing or exploding gradients. This careful modulation of information flow allows LSTMs to preserve essential information over long sequences, effectively addressing gradient-related challenges and enhancing their capability to learn long-term dependencies.

Self-attention Mechanism and Transformer

Self-attention mechanisms have fundamentally transformed the processing of sequential data by enabling neural networks to focus selectively on relevant features without relying on sequential memory of inherent characteristics. Initially proposed by researchers at Google Brain [2], this architecture allows each element within an input sequence to efficiently attend to all other elements. By addressing limitations inherent in RNNs and LSTMs—such as the necessity for sequential processing, memory constraints on long se-

quences, and slow training processes—self-attention mechanisms have been widely adopted in fields like natural language processing and image recognition. In these domains, data can be tokenized or segmented into smaller units, such as words in language or pixels in images, facilitating more effective analysis and prediction. This segmentation transforms language and images into sequential data where the position of each data point is crucial, leading to varying outcomes based on different input sequences [74].

At the core of the attention mechanism are the Query (Q), Key (K), and Value (V) vectors, each assigned to inputs within the sequence. The Query vector Q_i for input X_i, defined in Equation 2.13, acts as a probe representing the current input for which information is sought. It queries the sequence to extract relevant information from other data points. The Key vector K_i, shown in Equation 2.14, assesses how each piece of data relates to the Query, while the Value vector V_i, detailed in Equation 2.15, contains the actual data that the Query aims to access. Figure 2.3 visualizes this computation process.

$$Q = X \cdot W_Q, \quad X = [X_1, X_2, X_3, \ldots, X_i] \tag{2.13}$$

$$K = X \cdot W_K \tag{2.14}$$

$$V = X \cdot W_V \tag{2.15}$$

By computing the dot product between the Query and Key vectors, the attention score A is obtained, as shown in Equation 2.16. This score determines the degree of relevance or 'attention' a particular data point has concerning the current input. The relevance established by the attention score dictates how the data in the Value vector influences the output, with more relevant data exerting a greater impact. The attention scores are normalized using the softmax function, as per Equation 2.17, which stabilizes the learning process by ensuring the scores sum to one, forming a probability distribution.

$$A = Q \cdot K^T \tag{2.16}$$

$$A' = \text{softmax}\left(\frac{A}{\sqrt{D_{out}}}\right) \tag{2.17}$$

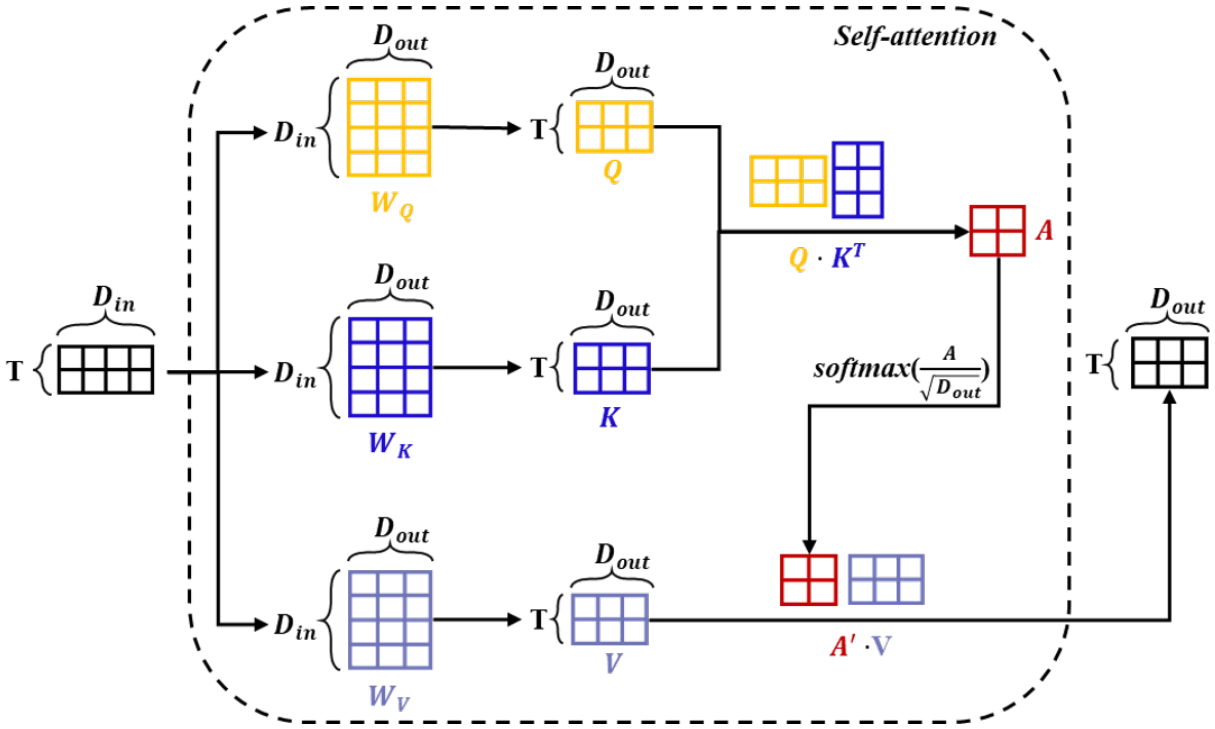


Figure 2.3: **Visualization of the Self-attention Computation Process**

This mechanism enables the model to dynamically weigh the importance of different elements when generating representations, capturing dependencies regardless of their positional distance within the sequence. By overcoming the memory constraints typical of RNNs, self-attention facilitates immediate and lossless retrieval of information and effectively handles long-range dependencies. The inherent parallel computation not only enhances computational efficiency but also significantly augments the model's capacity to focus on pertinent information across extensive input sequences, culminating in the output expressed by Equation 2.18.

$$Y = A' \cdot V \tag{2.18}$$

While the self-attention mechanism effectively improves the robustness and reliability of sequential data prediction, a single-head self-attention network is limited in capturing the variety of dependencies within the input data. To address this limitation, the multi-head self-attention mechanism has been proposed. This approach allows the network to capture various dependencies by using multiple sets of attention scores in parallel, enabling a more comprehensive understanding of sequential data, which is crucial for data with long-range dependencies.

Building upon the self-attention mechanism, the Transformer architecture, illustrated in Figure 2.4, has significantly advanced sequential data prediction, particularly in fields such as natural language processing, voice recognition, and time series prediction. The Transformer is primarily composed of two fundamental components: the encoder and the decoder. The encoder processes the input sequential data, capturing both the characteristics of individual data points and the internal relationships among them. Subsequently, the decoder takes the encoder's output—which encapsulates the important information—and produces precise and contextually appropriate results.

The encoder blocks, depicted on the left side of Figure 2.4, each contain two sub-layers. The first sub-layer is a multi-head attention mechanism that enables the network to focus on different parts of the input sequence, capturing the relevance of each data point with others in the sequence. This is critical for understanding the internal dependencies of time-sensitive patterns in sequential data. The second sub-layer is a fully connected feed-forward network that processes the output from the multi-head attention block. Each sub-layer is followed by a normalization layer, and residual connections are employed around each sub-layer. This design helps the encoder avoid the gradient vanishing problem as the model complexity increases, improving efficiency in training large Transformer models.

In contrast, the decoder blocks, shown on the right side of Figure 2.4, replicate the encoder structure with an additional sub-layer: the masked multi-head attention mechanism. This mechanism modifies the attention matrix of the standard attention layer by setting the attention scores for future tokens to a large negative value, effectively masking them. Consequently, after applying the softmax function in Equation 2.17, the normalized

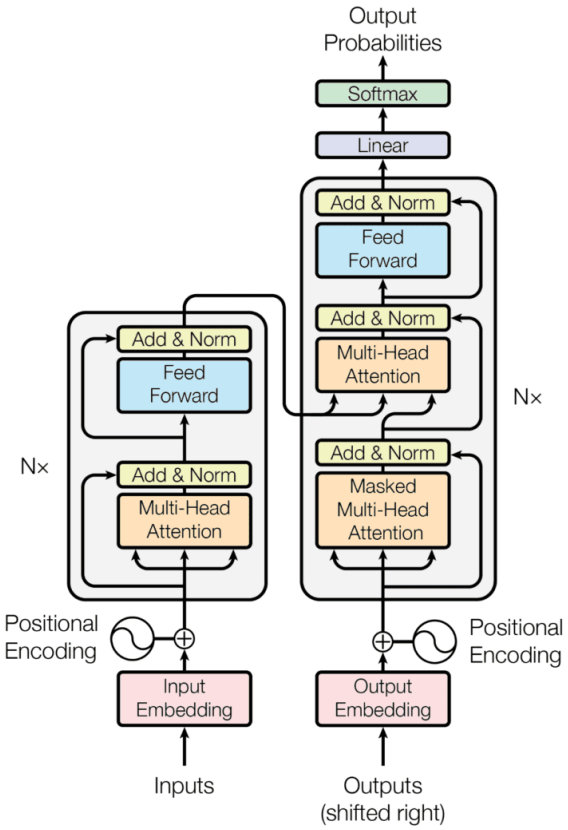


Figure 2.4: **The Transformer Architecture** [2]

attention scores for future tokens become zero. This ensures that the attention mechanism does not consider information from future tokens when generating the current output, preserving the causality essential for prediction tasks.

Following the masked multi-head attention sub-layer, the decoder integrates a second multi-head attention layer that attends to the encoder's output. This setup enables the model to align and adapt the encoder representations with the target sequence, enhancing the contextual relevance of predictions. Similar to the encoder, the decoder includes a feed-forward network, with each sub-layer followed by normalization and residual connections. This consistent architecture between the encoder and decoder helps maintain a balance in

learning and exchanging contextual information, which is crucial for accurate sequential data prediction.

Beyond the encoder and decoder components, the Transformer incorporates a positional encoding mechanism essential for capturing the sequence order of the input data. Unlike RNNs that process input data sequentially, the parallel processing nature of self-attention layers lacks inherent ability to account for the order of the input sequence. The positional encoding mechanism addresses this by adding unique positional information to each data point within the input sequence, using functions such as sine and cosine or learned embeddings.

The architectural integration of self-attention mechanisms, multi-head attention, and positional encoding within the Transformer offers a robust framework for sequential data prediction in a parallel and efficient manner. This architecture not only enhances the scalability and adaptability of neural networks but also improves model accuracy and training efficiency across various tasks, particularly those involving complex and long-range dependencies in sequential data.

2.5 System Optimization

The optimization of RESs is a critical process that involves across two main stages: the system design stage and the operational stage. Each stage is essential for maximizing efficiency, reliability, and performance in meeting energy demands sustainably and cost-effectively. Figure 2.5 categorizes RES optimization by stage and focus. The X-axis represents stages, with the system design stage on the left and the operational stage on the right. The Y-axis shows optimization focus, with efficiency at the top and cost at the bottom. This layout clarifies how optimization strategies are applied across stages, balancing efficiency and cost.

In the system design stage, the primary goal is to establish a robust structural foundation that ensures operational efficiency and economic viability over the system's lifetime. This involves two main considerations: configuration optimization and cost-efficiency plan-

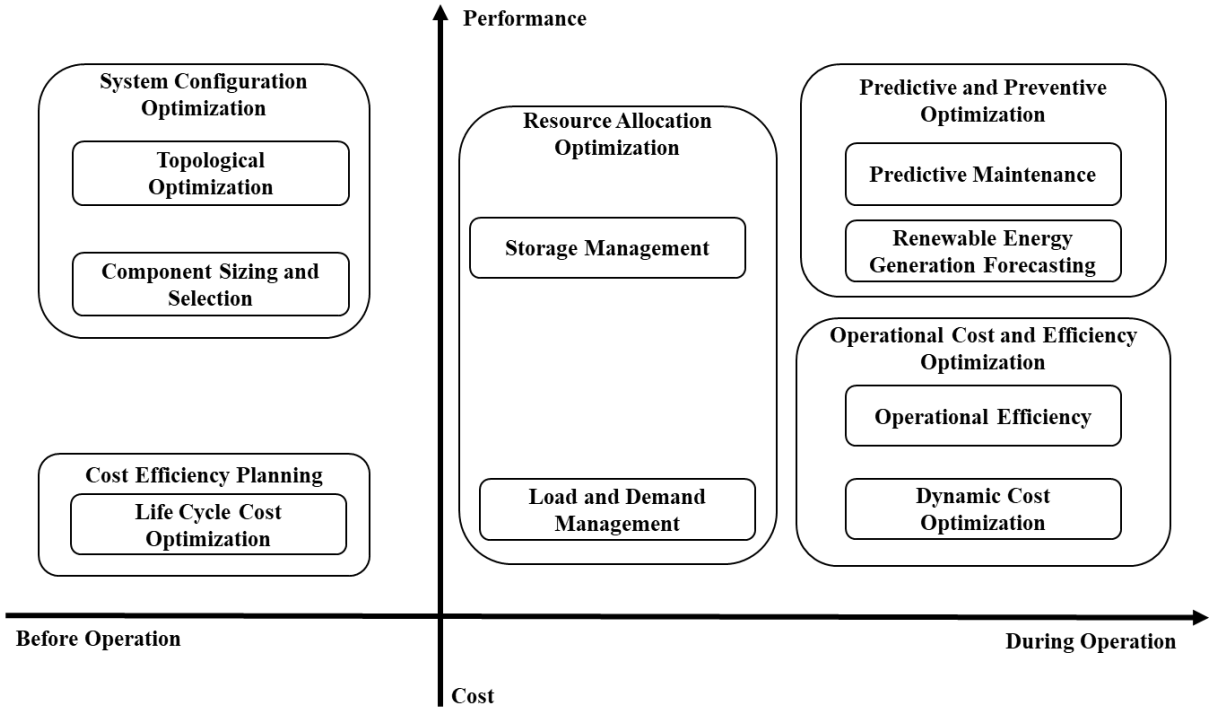


Figure 2.5: **Optimization for Renewable Energy Systems**

ning. These aspects form the basis of a well-integrated system that balances energy performance, initial investment, and long-term sustainability.

System configuration optimization includes selecting appropriate components and arranging them in a layout that maximizes RE capture and enhances system resilience. The process of component sizing and selection determines the optimal sizes and types of system elements [45], such as PV panels and BESS, to meet anticipated energy demands. While larger components offer greater flexibility in managing energy resources, they require a higher initial capital investment [7]. The choice of technology also impacts efficiency; for example, the conversion efficiency of PV panels or the storage capacity of batteries directly affects both energy output and cost. Therefore, sizing and selecting components involves carefully balancing flexibility with capital requirements, often necessitating trade-offs between cost and efficiency. This optimization aligns with user-defined priorities, ensuring

that the system meets specific energy performance goals while meeting financial constraints.

Topological optimization within system configuration focuses on spatial arrangements that optimize energy capture and improve integration with other components [75]. For PV systems, which commonly employ fixed tilt angles, this involves determining the orientation [76], tilt angle[77], and spatial layout of the panels to maximize solar irradiance capture. Each of these factors can significantly affect energy output, especially when the system is constrained to a fixed structure. By adjusting these parameters, topological optimization enhances energy output from the renewable source, contributing to overall system efficiency without incurring additional component costs. This approach also aligns spatial design with the RE resource potential, ultimately supporting sustainable and effective energy generation throughout the system's operational life.

Cost-efficiency planning is also integral to the design stage, primarily implemented through lifecycle cost analysis [78]. This analysis seeks to minimize the overall cost of energy production over the system's lifespan by carefully selecting configurations and components that balance initial investment, operational expenses, maintenance costs, and anticipated energy yield. By addressing these financial factors from the outset, lifecycle cost analysis enables informed decisions on component selection and system layout that optimize economic performance over time. This approach ensures that the system not only meets energy demands but does so in a financially sustainable manner, balancing long-term cost-effectiveness with reliability and performance.

By integrating system configuration with cost-efficiency planning, the design stage develops an optimized structure that minimizes initial costs while supporting effective and reliable operation. This strategic foundation ensures that the system meets both immediate and long-term performance expectations, achieving an efficient balance between investment and operational value.

In the operational stage, optimization mainly focuses on three primary categories: resource allocation optimization, predictive and preventive optimization, and operational cost and efficiency optimization. These approaches manage energy resources effectively, ensuring supply-demand balance, minimizing operational costs, and extending component

lifespan [79].

Resource allocation optimization primarily addresses load and demand management. This includes techniques such as real-time demand response [80], load forecasting [81] [82], and strategic scheduling to efficiently align energy supply with demand fluctuations and maintain system stability. The main objective is to balance supply and demand, reducing costs by avoiding energy use during peak times. By shifting consumption away from high-demand periods, the system can limit dependency on external grid imports at peak prices, thus reducing overall operational costs. Additionally, resource allocation optimization tailors energy management based on user usage patterns, strategically distributing energy to minimize unnecessary expenditures.

Another critical aspect of resource allocation is storage management, which focuses on optimizing the operation of energy storage systems such as BESS and HESS. Effective storage management involves carefully scheduling charge and discharge cycles, as well as electrolyzer usage, to ensure resource availability while reducing operational costs and preventing rapid degradation of storage components. This optimization extends the lifespan of storage systems by limiting high-stress usage patterns, enhancing energy availability in storage, and minimizing replacement costs over the system's lifespan. Consequently, resource allocation optimization in storage not only improves energy availability but also contributes to cost savings through reduced maintenance and replacement needs.

The second category, predictive and preventive optimization, focuses on predicting system demands and adjusting operations to prevent disruptions. Predictive optimization is primarily used in maintenance, employing ML models and historical data to forecast potential failures in critical components such as RE generators and ESSs. By identifying potential failure in advance, this approach significantly reduces the risk of unplanned downtime and mitigates the costly consequences associated with power outages or energy shortages, particularly in isolated or off-grid energy systems. Predictive maintenance provides users with sufficient time to prepare for necessary repairs or component replacements, thereby enhancing system reliability and minimizing unexpected operational interruptions.

Another aspect of predictive optimization is RE generation forecasting. This involves

predicting the inherent variability of renewable resources like solar and wind power and dynamically adjusting system operations based on forecasted outputs. By proactively managing these fluctuations, renewable generation forecasting enhances the stability and reliability of energy output. This results in a more consistent RE supply that demands less dependency on energy storage systems and reduces reliance on local grid support. Such optimization enables a smoother RE generation process, minimizing operational disruptions and improving the overall efficiency of the energy system.

The final category, operational cost and efficiency optimization aims to maximize system efficiency while minimizing expenses associated with energy conversion and consumption. Operational efficiency optimization focuses on reducing energy losses that occur during conversion and distribution by using advanced energy management algorithms. By aligning energy generation more closely with demand patterns, these algorithms can minimize conversion losses and transmission inefficiencies, ensuring that generated energy is effectively utilized and waste is reduced. In addition, it also includes RE generation optimization. For instance, in PV systems, adjusting the tilt angle, orientation, or direction of PV panels can maximize solar energy capture under varying conditions [83]. Similarly, in wind energy systems, optimizing blade angles enhances energy generation by aligning with prevailing wind conditions. These dynamic adjustments allow for continuous alignment with resource availability, thereby maximizing energy output and improving overall system performance.

Dynamic cost optimization further enhances operational efficiency by leveraging strategies such as time-of-use pricing and demand-side management. This involves adjusting energy consumption to off-peak times when energy prices are lower, directly reducing operational costs. Dynamic cost optimization is highly dependent on the structure of local electricity pricing and often requires supportive policies to be most effective. The performance of this strategy is influenced by local regulations, societal norms, and grid management policies, all of which determine the flexibility available to adjust energy use patterns. Thus, while operational efficiency optimization directly enhances system performance, dynamic cost optimization tailors consumption to economic incentives, leveraging external conditions to achieve lower operating costs within the existing energy pricing framework.

In summary, the optimization of RESs is a multifaceted process that spans both the design and operational stages. By systematically addressing configuration, cost-efficiency, resource allocation, predictive maintenance, and operational strategies, it is possible to develop and maintain RESs that are efficient, reliable, and economically sustainable.

The following three subsections provide an in-depth review of optimizations relevant to this research. These include tilt-angle optimization, component size optimization, and energy management for load and demand management. Each subsection examines the methodologies and applications that contribute to enhancing performance, efficiency, and reliability within RESs.

2.5.1 Photoviltaic Panel Optimization

Optimizing the tilt angle of PV panels is essential for maximizing solar energy capture and enhancing solar-to-electricity conversion efficiency. Studies have shown that proper tilt angle adjustments can significantly increase the incident solar irradiance on PV panels [84][85][86].

Raina et al. [84] reported that bifacial PV modules installed at optimal orientations could collect up to 84.76% more irradiance. Ashutosh et al. [87] achieved an 18.03% increase in energy harvest by optimizing the tilt angle monthly for a PV plant in India. Alghamdi [88] improved the energy output of roof-mounted PV systems by 15.9% using an unsupervised artificial neural network for Net LCOE optimization. Mansour et al. [89] found that monthly tilt angle optimization could yield 4.2% more power than yearly adjustment. Hassan et al. [90] determined monthly optimal tilt angles for maximum solar irradiance across 20 locations in Iraq.

Discrepancies in optimization outcomes arise due to factors such as using irradiance instead of power generation as the optimization objective, not explicitly evaluating the correlation between tilt angle and energy conversion efficiency, and differences in climate zones studied.

Most existing studies rely on mathematical models, which have limitations including

strict assumptions and high computational costs. Artificial Neural Network (ANN) offer improved estimation accuracy and can overcome these limitations [91, 92, 93]. Shboul [94] used ANN to predict hourly solar radiation with higher accuracy and efficiency. Agbulut et al. [95] demonstrated precise forecasting of daily global solar radiation using machine learning algorithms. Villegas-Mier et al. [96] achieved faster convergence and more stable performance in PV power system optimization using ANN.

Challenges in deploying ANN include the black-box nature of the models and manual design of architecture and hyperparameters. Optimizing these aspects can enhance performance over unoptimized models.

This study proposes an optimization framework in Chapter 4 applicable to PV panels in any region within the climate zone. The focus is on how the trade-off between incident solar irradiance and conversion efficiency affects power production. A novel data-driven approach examines the relationship between tilt angle and conversion efficiency, cross-validated with a reliable mathematical model to mitigate the black-box risk of the neural network.

2.5.2 Capacity of Component Optimization

The operations of RE generators and ESSs are closely integrated. ESSs store excess energy from RE generators, essential for continuous energy supply and economic benefits. In grid-connected RES, ESSs act as buffers to smooth out production fluctuations and serve as additional energy sources when needed [97]. Determining optimal capacities involves simulating system operations and implementing energy management strategies using updated data.

Capacity optimization in RES is challenging due to the non-convex nature of the problem. Traditional gradient-based methods are often ineffective [98], necessitating sophisticated optimization techniques. Recent research employs heuristic algorithms, utilizing financial or technical indicators as objective functions. Genetic Algorithm (GA) [99], Non-Dominated Sorting Genetic Algorithm II (NSGA-II) [100] and Particle Swarm Optimiza-

tion (PSO) [101] are among the methods used.

These algorithms have been applied to diverse scenarios. Mohamed et al. optimized a grid-connected PV-wind system using PSO [102]. Ali et al. used PSO for load segmentation, noting that prioritizing higher loads increases energy cost [103]. The HOMER software has been used for economic and performance simulations in RES with PV and battery systems. Campana et al. assessed the cost-effectiveness of Li-ion batteries in commercial buildings with PV energy [104]. Ali et al. developed a customized program for faster RES optimization without sacrificing accuracy compared to traditional simulation software like HOMER [105]. Hassan et al. combined hardware models with NSGA-II to identify cost-effective configurations, reducing energy costs by 30% with battery storage [106]. Wu et al. minimized life-cycle cost in a PV-battery system using mixed-integer nonlinear programming, finding time-of-use tariffs more beneficial than flat-rate tariffs [107].

To enhance optimization and operational efficiency, a two-stage methodology is adopted. The first stage generates system design parameters and component capacities. The second stage enhances system performance based on established capacities. This approach links capacity setting with performance optimization and was initially applied to power systems for reliability. Liu et al. optimized power system capacity using GA for cost minimization and reliability maximization, supported by a rule-based operational strategy [108]. Their findings indicate that incorporating multiple ESSs significantly enhances system reliability. He et al. used evolutionary algorithms with a simplified rule-based scheduling strategy to optimize the levelized cost of energy for systems equipped with different ESSs [109]. Yang et al. minimized loss of power supply probability and levelized cost of energy using NSGA-II and an off-design model, recommending the integration of concentrated solar power across most regions of China [110]. Zahra et al. minimized energy costs in a RES with BESS and fuel cells using a hybrid PSO-GA approach and a rule-based energy management strategy [7]. Al-Quraan et al. optimized stand-alone RES capacity using a two-layer model combining mixed-integer linear programming and GA, effectively minimizing total cost and energy loss [111].

Despite extensive research, gaps remain in addressing crucial aspects of RES operation

and integration. The lack of a comprehensive decision-making framework that includes external constraints and stakeholder preferences indicates a need for advanced research approaches. Current literature relies heavily on traditional simulation and rule-based strategies, lacking advanced optimization techniques that could enhance performance by better addressing regional variability and grid complexities. This study proposes a novel two-stage capacity optimization strategy, detailed in Chapter 5.

2.5.3 Energy Management

Energy management strategies have been extensively explored and have proven to be highly effective in enhancing the overall system performance in daily operations. Existing research predominantly focuses on optimizing techno-economic indicators, including energy costs, system costs, carbon emissions, grid dependency, and addressing uncertainty in forecasting during the energy management process. The primary approaches employed are deterministic, heuristic, and machine learning. The deterministic approach includes techniques such as linear programming [112], mixed-integer linear programming [113], and non-linear programming. Evolutionary approaches encompass methods like NSGA-II [114], Golden Jackal Optimization [115], and Prairie Dog Optimization [116]. Machine learning approaches are represented by Deep Reinforcement Learning [117].

To manage the uncertainty of energy supply and demand, current literature integrates stochastic processes [117] and Markov decision processes [118] into their optimization algorithms. Specifically, Vergine et al. developed a Stochastic Unit Commitment for a hybrid RES, employing Markov processes to model uncertainties in wind and PV power outputs and an ARIMA model to predict demand, thereby optimizing operational costs and enhancing the management of RE integration [119]. Alvarez et al. utilized a two-stage stochastic programming model to manage uncertainties in wind and PV production within RES, focusing on optimizing energy storage as the sole flexible resource to minimize operational costs and maintain energy balance [120]. Yao et al. presented a multi-level optimization model for Integrated Energy Systems that employs Conditional Value at Risk and Model Predictive Control (MPC)to manage uncertainties and optimize economic, en-

vironmental, and energy consumption objectives, demonstrating enhanced performance through feedback correction compared to traditional optimization approaches [121].

While existing literature has extensively documented various models focusing on energy management system optimization and uncertainty management, a significant gap remains in effectively integrating advanced machine learning approaches with optimization algorithms to diminish uncertainty. This deficiency is crucial as it impacts the ability to capture the dynamic interactions between system components and accommodate real-world variabilities, which are essential for advancing the deployment of smart RESs. This study proposes a novel energy management algorithm that introduces a new RES modeling strategy combined with an advanced predictive-based energy management algorithm, as presented in Chapter 6.

2.6 Summary

This chapter provides an in-depth review of recent advancements in RESs, with a focus on essential components like RE generators and ESSs. It examines the application of time series data and deep learning, as well as system optimization, which is vital for developing efficient smart RESs. Additionally, it includes a comparative analysis of various algorithms and system configurations, emphasizing their impact on enhancing system design and functionality. Finally, this chapter concludes by discussing ongoing challenges in system optimization.

Chapter 3

Extended Literature Survey

This chapter introduces the methodology employed in this research, offering a detailed description of the system modeling, machine learning techniques, optimization methods, and performance evaluation metrics. Section 3.1 presents the system modeling process, covering mathematical models for PV, WT, battery, and electrolyzer systems, which provide a solid technical foundation for the overall system modeling. Section 3.2 discusses machine learning techniques applied for data-driven analysis, including data preprocessing, exploratory data analysis, and the design and training of neural networks for system state estimation and energy prediction. The chapter proceeds to optimization methods in Section 3.3, examining different approaches such as Linear Programming (LP), Mixed Integer Linear Programming (MILP), and GA, which are applied to various optimization tasks to enhance system performance. Lastly, Section 3.4 outlines the performance evaluation metrics, specifying the criteria used to evaluate prediction accuracy, economic performance, and the sustainability of the system.

3.1 System and Components Modeling

Mathematical models are essential in the RESs modeling by providing a theoretical and physical foundation for understanding the behavior of different system components. These

models not only help represent the operational characteristics of each system but also offer essential support for developing data-driven approaches. Through precise modeling, these models enable the integration of machine learning techniques, which can be used to perform regression and enhance predictive capabilities. Eventually, the combination of mathematical modeling and machine learning allows for more precise optimization of system performance, leveraging both theoretical understanding and data-driven insights to maximize efficiency and improve decision-making.

3.1.1 Photovoltaic Modeling

The PV system modeling integrates three key models to estimate solar energy output. The solar position model first determines the sun's position in the sky at a given time. Based on this position and the ground horizontal solar irradiance, the Liu-Jordan model calculates the total irradiance incident on the solar panels. Finally, the five-parameter model uses this total irradiance, along with additional meteorological data such as temperature, to calculate the overall energy output of the solar panels. These models work together to provide a comprehensive estimation of the system's performance.

Solar Position Model

The solar position model is used to locate the Sun's position relative to the Earth's specific region. Before doing so, it is crucial to correct the standard time. The difference between the solar time and standard time is caused by the angle between the ecliptic plane and the equator, as well as the elliptical shape of the Earth's orbit. Inputs of this model are the day number within a year n, longitude L, standard time t, and the output of this model is hour angle ω.

First of all, Equation 3.1 calculates the fractional year γ in radians with the given standard time and the day number. During a leap year, 366 will replace 365 in the first term.

$$\gamma = \frac{2\pi}{365} \times (n - 1 + \frac{\mathrm{t} - 12}{24}) \tag{3.1}$$

In order to correct the standard time, this study needs assistance from two terms: the difference between standard time and mean solar time; and the difference between the longitude of the simulation position and the meridian. The difference between standard time and mean solar time, also known as Equation of Time (EoT), is caused by the Earth's rotation. A positive value means that the standard time is faster than the true solar time, and a negative value means that the standard time is slower than the true solar time. Equation 3.2 is an empirical equation used to estimate EoT. The absolute value of EoT could be more than 15 minutes on certain days. Therefore, the effect of EoT cannot be ignored during the true solar time calculation. Meanwhile, the difference between the longitude of the simulation position and the meridian comes from the Coordinated Universal Time (UTC). Based on UTC, every 15 degrees of longitude L is a new time zone TZ. However, UTC only gives the standard time exactly on that meridian. Hence, this model must calculate the standard time for the region near the meridian. With the help of Equation 3.3, this study was able to calculate the solar time L_{st} of any location on the Earth.

$$\begin{aligned} EoT = 229.18 \times (&0.000075 + 0.001868\cos(\gamma) \\ &- 0.032077\sin(\gamma) - 0.014615\cos(2\gamma) \\ &- 0.040849\sin(2\gamma)) \end{aligned} \tag{3.2}$$

$$L_{st} = t + EoT + 4[L - 15\Delta\mathrm{TZ}] \tag{3.3}$$

Finally, this study calculated the hour angle ω with the true solar time. In Equation 3.4, ω defines 0 degrees at 12 PM L_{st}, negative before 12 PM and positive after 12 PM. In the next model to be discussed, the hour angle will be used to calculate the extraterrestrial radiation on the horizontal surface.

$$\omega = 15^{\circ} \times (\mathrm{L_{st}} - 12) \tag{3.4}$$

Liu-Jordan Model

The Liu and Jordan Model [122] will use the output from the solar position model to decompose the solar radiation and calculate the total solar radiation incident on a tilted surface. The Liu and Jordan Model is a type of isotropic model that calculates the hourly total solar radiation incident on a tilt surface I_t through three components: hourly direct beam radiation on a tilted surface $I_{b,t}$, hourly diffuse radiation on a tilted surface $I_{d,t}$ and hourly reflective radiation on a tilted surface $I_{r,t}$. Equation 3.5 and 3.6 show the basic concept of the model. Equation 3.5 decomposes the global radiation on the horizontal plane I_h into two components: hourly direct beam radiation on the horizontal plane I_b, and hourly diffuse radiation on the horizontal plane I_d. Then, Equation 3.6 uses I_b, I_d and I_h to multiply the relative inclination factor to calculate the total radiation incident on a tilted surface. Inclination factors R_b, R_d and R_r are related to the tilt angle of the photovoltaic panel. Thus, the optimization variable, the tilt angle of the photovoltaic panel will determine the value of inclination factors and affect the solar radiation incident on a tilted surface at this step. A high solar radiation incident on a tilted surface will lead to a high cell temperature and decrease the efficiency of the photovoltaic panel. This model uses n, ω, β, ϕ and I_h as inputs to calculate I_t.

$$I_h = I_b + I_d \tag{3.5}$$

$$I_t = I_{b,t} + I_{d,t} + I_{r,t} = I_b \times R_b + I_d \times R_d + I_h \times R_r \tag{3.6}$$

I_d highly correlates with the atmosphere's clearness. The clearer the atmosphere, the less solar radiation will be diffused. Hence, one can use the clearness index to measure the clearness of the atmosphere. This study uses k_T to represent the hourly clearness index and calculate clearness of the atmosphere by comparing the solar radiation at the bottom

of the atmosphere (radiation received on the horizontal plane) to the theoretical maximum solar radiation that can be received on a horizontal plane (extraterrestrial radiation on the horizontal surface). Equation 3.7 shows how the clearness index is calculated: the more radiation received on the ground, the higher the clearness index.

$$k_T = \frac{I_h}{I_0} \tag{3.7}$$

In Equation 3.7, I_h will be measured, or provided by the sensor located on the ground. This study will use this value as an input for the model during the simulation. Meanwhile, hourly extraterrestrial radiation on the horizontal surface I_0 can be calculated through Equation 3.8. In this Equation, G_{sc} is the solar constant, which is $1367W/m^2$, with up to 1% error. ϕ is the latitude and is related to the simulation location, and ω is the hour angle, the output of the solar position model. δ is the declination, which can be calculated by Equation 3.9.

$$\begin{aligned} I_o = \frac{12 \times 3600}{\pi} G_{sc} \left(1 + 0.033 \cos \frac{360n}{365}\right) \\ \times \left[\cos\phi \cos\delta \left(\sin\omega_2 - \sin\omega_1\right)\right. \\ \left. + \frac{\pi\left(\omega_2 - \omega_1\right)}{180} \sin\phi \sin\delta\right] \end{aligned} \tag{3.8}$$

$$\delta = 23.45 \cdot \sin\left[360 \cdot \left(\frac{284 + n}{365}\right)\right] \tag{3.9}$$

After calculating the clearness index, this study uses Equation 3.10 to estimate I_d. With I_d, this study can decompose global radiation on the horizontal plane and calculate total radiation with three inclination factors.

$$\frac{I_d}{I_h} = \begin{cases} 1.0 - 0.09k_T & \mathrm{k}_T \in [0, 0.22] \\ 0.95 - 0.16k_T + 4.39k_T^2 & \mathrm{k}_T \in (0.22, 0.8] \\ -16.64k_T^3 + 12.34k_T^4 & \\ 0.165 & \mathrm{k}_T \in (0.8, 1] \end{cases} \tag{3.10}$$

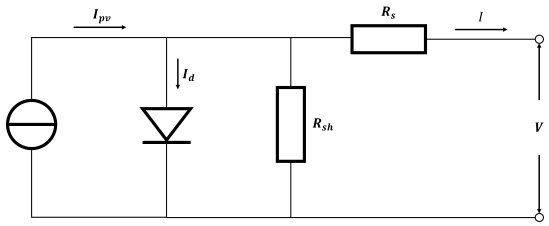


Figure 3.1: **One-diode Model**

The inclination factors, R_d, R_d and R_r, are given by Equations 3.11, 3.12 and 3.13, respectively. Inclination factors are related to the tilt angle, latitude, hour angle and declination. Reflective coefficient, ρ, in Equation 3.13 depends on the surface of the ground. Generally, for dry concrete surfaces, the value of the reflective coefficient is approximately equal to 0.2.

$$R_b = \frac{\cos(\phi - \beta)\cos\delta\cos\omega + \sin(\phi - \beta)\sin\delta}{\cos\phi\cos\delta\cos\omega + \sin\phi\sin\delta} \tag{3.11}$$

$$R_d = \frac{1 + \cos\beta}{2} \tag{3.12}$$

$$R_r = \frac{1 - \cos\beta}{2} \times \rho \tag{3.13}$$

Five-parameter Model

The five-parameter model is a mathematical model used to estimate the photovoltaic cell's I-V output in the ideal experimental condition. Figure 3.1 represents an equivalent circuit of a photovoltaic cell. Two components located on the left side represent a p-n junction. Once solar energy reaches the photovoltaic cell, the built-in electric field produces a photovoltaic current I_{pv}. Then, electrons are forced to flow in opposite directions and produce current i_d after the new electric field is created. The resistance R_{sh} represents the leakage of the P-N junction, and R_s represents the power loss of the cell.

$$I = I_{pv} - I_s \left[\exp\left(\frac{V + R_s I}{f N_s V_t} \right) - 1 \right] - \frac{V + R_s I}{R_{sh}} \tag{3.14}$$

Equation 3.14 is commonly used in the literature to describe the one-diode model which is shown in Figure 3.1. I and V are the output current and voltage of the photovoltaic cell, respectively. N_s is the number of cells in a panel, and V_t (thermal voltage) is a constant number. Parameters f (modified ideality factor), I_s (diode's saturation current), I_{pv} , R_s and R_{sh} are known as the five parameters. These five parameters are crucial in calculating the power output of the photovoltaic cell, and they can be estimated through available technical data from manufacturers.

First of all, Equation 3.15 describes how to calculate thermal voltage through Boltzmann's constant k, the absolute temperature T of the p-n junction and the charge of the electron q.

$$V_t = \frac{kT}{q} \tag{3.15}$$

Before calculating the five parameters, it is crucial to calculate open circuit voltage V_{oc} and short circuit current I_{sc}. These two variables depend on solar irradiance and cell temperature. HOMER[123], a photovoltaic analysis platform, implemented Equation 3.16 to estimate the cell temperature. Inputs of Equation 3.16 include ambient temperature T_a, ambient temperature at Nominal Operating Cell Temperature (NOCT) $T_{a,noct}$, cell temperature at NOCT $T_{c,noct}$, solar radiation G, solar radiation at NOCT G_{noct}, the efficiency of the photovoltaic array at its maximum power point under Standard Test Conditions (STC) $\eta_{mp,STC}$, the temperature coefficient of power α_P and the product of the radiation transmission coefficient through the covering system and the cell absorption coefficient $\tau\alpha$. As per the suggestion made by Duffie and Beckman[124], $\tau\alpha$ is equal to 0.9. With the assumption that the coefficient of heat transfer h to the surroundings is constant, Equation 3.16 can significantly reduce its complexity. The heat transfer coefficient to the surroundings will depend on meteorological conditions such as wind speed V_w and relative humidity RH.

$$T_c = \frac{T_a + (T_{c,noct} - T_{a,noct}) \left(\frac{G}{G_{noct}}\right) \left[1 - \frac{\eta_{mp,STC}\left(1-\alpha_P T_{c,STC}\right)}{\tau\alpha}\right]}{1 + (T_{c,noct} - T_{a,noct}) \left(\frac{G}{G_{noct}}\right) \left(\frac{\alpha_P \eta_{mp,STC}}{\tau\alpha}\right)} \tag{3.16}$$

Once the cell temperature is known, V_{oc} and I_{sc} can be calculated with varying solar radiation and ambient temperatures through Equation 3.17 and Equation 3.18. In these two equations, $V_{oc,ref}$ is the open-circuit voltage at reference condition, $I_{sc,ref}$ is the short circuit current at reference condition, $\mu_{V,oc}$ is the temperature coefficient of open-circuit voltage and $\mu_{I,sc}$ is the temperature coefficient of the short circuit current.

$$V_{oc}(G,T) = V_{oc,ref} + fN_S V_t \ln\left(\frac{G}{G_{ref}}\right) + \mu_{V,oc}\left(T_c - T_{c,STC}\right) \tag{3.17}$$

$$I_{sc}(G,T) = I_{sc,ref}\frac{G}{G_{ref}} + \mu_{I,sc}\left(T_c - T_{c,STC}\right) \tag{3.18}$$

R_{s0} and R_{sh0} can be estimated through the slope at open-circuit voltage V_{oc} and slope at short circuit current I_{sc}, respectively. R_{s0} and R_{sh0} also represent the initial state of R_s and R_{sh}, respectively. Equations 3.19, 3.20 and 3.21 describe this process. These two parameters can be estimated through the slope of the current-voltage curve. R_{sh} is assumed to be a constant number because it represents the leakage and is related to the P-N junction itself. Equation shows how to calculate R_s. Because losses in electrodes and semiconductors are related to the the operating condition, the value of R_s depends on other parameters.

$$R_{s0} = -\left(\frac{dV}{dI}\right)\Big|_{V=V_{oc}} \tag{3.19}$$

$$R_{sh} = R_{sh0} = -\left(\frac{dV}{dI}\right)\Big|_{I=I_{sc}} \tag{3.20}$$

$$R_s = R_{s0} - \left[\frac{fNV_t}{I_s}\exp\left(-\frac{V_{oc}}{fN_sV_t}\right)\right] \tag{3.21}$$

Equations 3.22 and 3.23 show how to calculate parameters I_s and I_{pv}. Parameter f is assumed to be a constant number. Thus, the radiation, ambient temperature and heat transfer coefficient determine V_{oc} and I_{sc}. In addition, these two parameters will change the values of the five parameters. The photovoltaic panel will operate at the maximum power point, and the five parameters change the maximum power point through Equation 3.14.

$$I_s = \left(I_{sc} - \frac{V_{oc}}{R_{sh}}\right)\exp\left(-\frac{V_{oc}}{fN_sV_t}\right) \tag{3.22}$$

$$I_{pv} = I_{sc}\left(1 + \frac{R_s}{R_{sh}}\right) + I_s\left[\exp\left(\frac{I_{sc}R_s}{fN_sV_t}\right) - 1\right] \tag{3.23}$$

3.1.2 Wind Turbine Modeling

Wind turbines convert the kinetic energy of wind into mechanical power, which can then be transformed into electricity. Understanding the fundamental principles governing wind turbine operation is essential for optimizing their efficiency and performance. This involves analyzing the relationships between wind speed, turbine geometry, and the physical laws that dictate energy extraction from the wind.

One of the key equations in this analysis is the wind power equation, which describes the total power available in the wind passing through the area swept by the turbine blades:

$$P_w = \frac{1}{2}\rho_{air}A_{tb}v_w^3 \tag{3.24}$$

In this equation, P_w represents the power available in the wind, ρ_{air} is the air density, A_{tb} is the swept area of the turbine blades, and v_w is the wind speed. The swept area A is calculated using the formula $A_{tb} = \pi r_{tb}^2$, where r_{tb} is the radius of the rotor. The term

$\frac{1}{2}\rho_{air}v_w^3$ quantifies the kinetic energy per unit volume of the air flowing through the rotor's swept area.

Building upon the wind power equation, Betz's Law establishes a theoretical maximum for the proportion of wind power that can be extracted by a turbine, limiting it to approximately 59.3% of the total available power. This limitation is expressed through the power coefficient C_p, which measures the efficiency of the turbine in converting wind power into mechanical power. The actual power output P_{out} of the turbine is given by Equation 3.25:

$$P_{\text{out}} = C_p \frac{1}{2}\rho_{air} A_{tb} v_w^3 \tag{3.25}$$

Here, C_p is constrained by Betz's limit, with a theoretical maximum value of about 0.593. This coefficient accounts for the inefficiencies inherent in the energy conversion process and varies depending on the turbine design and operating conditions.

Wind turbines operate efficiently within specific wind speed ranges, characterized by three critical thresholds: the cut-in speed, the rated speed, and the cut-out speed. The cut-in speed is the minimum wind speed at which the turbine begins to generate power; below this speed, the turbine remains idle, and no electricity is produced. The rated speed is the wind speed at which the turbine reaches its nominal power output, operating at peak efficiency and within design limits. The cut-out speed is the maximum wind speed at which the turbine can safely operate; beyond this speed, the turbine shuts down to prevent mechanical damage due to excessive wind loads. Shutdown procedures may involve feathering the blades or yawing the rotor to minimize wind exposure.

Understanding these operational thresholds is essential for modeling how the power output of a wind turbine varies with wind speed. Below the cut-in speed, when the wind speed is less than $v_{\text{cut-in}}$, the turbine does not generate power ($P(v) = 0$). Between the cut-in and rated speeds, the power output increases cubically with wind speed, as indicated by $P(v) \propto v^3$. This relationship arises from the wind power equation and reflects the increasing kinetic energy available in the wind. At the rated speed, the turbine produces its nominal power output (P_{nom}), operating at optimal efficiency and within design limits.

Above the rated speed and below the cut-out speed, the turbine's power output is regulated and maintained at P_{nom} despite higher wind speeds. This regulation prevents mechanical stress and ensures safety by capping the power output through control mechanisms such as adjusting blade pitch or rotor speed. When wind speeds exceed $v_{\text{cut-out}}$, the turbine ceases operation ($P(v) = 0$) to avoid damage.

Combining these concepts, the relationship between wind speed and turbine power output is mathematically represented by Equation 3.26:

$$P(v) = \begin{cases} 0 & \text{if } v < v_{\text{cut-in}} \text{ or } v > v_{\text{cut-out}} \\ C_p \frac{1}{2} \rho_{air} A_{tb} v_w^3 & \text{if } v_{\text{cut-in}} \leq v \leq v_{\text{rated}} \\ P_{\text{nom}} & \text{if } v_{\text{rated}} < v \leq v_{\text{cut-out}} \end{cases} \tag{3.26}$$

In this function, $P(v)$ denotes the power output at a given wind speed v, C_p is the power coefficient reflecting turbine efficiency, ρ is the air density, and A is the swept area of the blades. The critical wind speeds $v_{\text{cut-in}}$, v_{rated}, and $v_{\text{cut-out}}$ define the operational boundaries, while P_{nom} represents the turbine's nominal power output.

This model illustrates how wind turbines adapt to varying wind conditions through power control mechanisms. By regulating the power output based on wind speed thresholds, turbines maximize energy capture within safe operating limits. This approach maintains structural integrity and extends the operational lifespan of the turbine. Incorporating these factors into modeling is crucial for designing efficient and reliable wind energy systems that can effectively harness the inherently variable wind environment.

3.1.3 Battery Modeling

Batteries are crucial components in BESS, where precise modelling enhances simulation efficiency and optimization accuracy. According to Fotouhi et al. [125], battery modelling approaches can be broadly categorized into three types: mathematical models, electrochemical models, and electrical equivalent circuit network models. This research focuses

on the mathematical modelling approach, particularly the model based on the System Advisor Model [126], a software tool that simulates both lead-acid and lithium-ion batteries using detailed voltage, thermal, and lifetime sub-models [127].

Lifetime model Battery aging is caused by two primary factors: time (calendar aging) and usage (cycle aging). Batteries have both a calendar life, typically five to ten years, and a limited number of charge-discharge cycles before their performance deteriorates significantly. Therefore, accurately modelling battery degradation is essential for predicting battery lifespan and scheduling replacements.

In this model, the Depth of Discharge (DOD) is a critical parameter, calculated by integrating the energy charged into the battery over its capacity:

$$DOD = \frac{\int_{t=0}^{t=T} E_{\text{char}}\, dt}{\text{Battery capacity}} \tag{3.27}$$

The auxiliary DOD, DOD_{Aux}, is determined using the SOC range:

$$DOD_{\text{Aux}} = \frac{SOC_{\max} - SOC_{\min}}{\text{Battery capacity}} \tag{3.28}$$

The calendar aging factor and cycle aging factor quantify the degradation due to time and usage, respectively, and are calculated using:

$$\text{Calendar aging factor} = \frac{1}{L_{\text{BESS,cal}} \times 24 \times 365} \tag{3.29}$$

$$\text{Cycle aging factor} = \frac{1}{e^{\left(\frac{\log DOD - \log 771.51}{-0.604}\right)} - 45300} \tag{3.30}$$

The overall aging factor, $F_{\text{BESS\_age}}$, is determined by taking the maximum of the calendar aging factor and the cycle aging factor:

$$F_{\text{BESS\_age}} = \max\left(\text{Cycle aging factor}, \text{Calendar aging factor} \times 24\right) \tag{3.31}$$

The State of Health (SOH) is an indicator of the battery's remaining capacity and is calculated using:

$$SOH = \frac{1}{0.3} \times \frac{\text{Capacity}_{\text{Aux}}}{\text{Battery capacity}} - \frac{7}{3} \tag{3.32}$$

The battery capacity is updated to account for degradation:

$$\text{Capacity}_{\text{Aux}} = \text{Battery capacity} \times (1 - 0.3 \times (1 - SOH + F_{\text{BESS,age}})) \tag{3.33}$$

The State of Charge (SOC) is then updated:

$$SOC_{\max} = SOC_{\min} + DOD_{\text{Aux}} \times \text{Capacity}_{\text{Aux}} \tag{3.34}$$

This battery modelling approach is adapted from methodologies established by Angenendt et al. [128] and Holger et al. [129]. The model dynamically updates the BESS 's state at each time step based on optimization outcomes. It incorporates a detailed assessment of degradation by comparing both calendar and cycle life to calculate the SOH. The model stipulates that BESS replacement within the RES is necessary once the SOH falls below 70%, considering the costs associated with replacement and its implications on the BESS 's capacity and performance during charging and discharging processes.

3.1.4 Electrolyzer Modeling

Electrolyzers are devices that convert electrical energy into chemical energy by driving the electrolysis of water to produce hydrogen and oxygen gases. Modeling the electrolyzer involves understanding the electrochemical reactions, thermodynamic principles, and kinetic parameters that govern its operation. The modelling approach is from bentvelsen's work [130].

In the PEM electrolyzer, the anode and cathode are where the reaction happens—this is where water gets split into hydrogen and oxygen. The anode takes on the task of oxidizing

water, which means it pulls the water apart into oxygen, protons, and electrons. These electrons then make a journey through an external circuit all the way to the cathode. Over at the cathode, these electrons have a crucial role: they reduce the protons, which have moved across the PEM, turning them into hydrogen gas. The PEM itself is quite the standout, selectively allowing only protons to pass through while keeping the gases separate, which makes the whole setup more efficient.

At the anode, water is oxidized to produce protons, oxygen gas, and electrons:

$$\text{Anode Reaction:} \quad \mathrm{H_2O}(l) \rightarrow 2\mathrm{H^+}(g) + \frac{1}{2}\mathrm{O_2}(g) + 2e^- \quad E^\circ = -1.23\,\mathrm{V} \tag{3.35}$$

At the cathode, protons are reduced to form hydrogen gas:

$$\text{Cathode Reaction:} \quad 2\mathrm{H^+}(g) + 2e^- \rightarrow \mathrm{H_2}(g) \quad E^\circ = 0\,\mathrm{V} \tag{3.36}$$

The overall reaction combines the anode and cathode processes:

$$\text{Total Reaction:} \quad \mathrm{H_2O}(l) \rightarrow \mathrm{H_2}(g) + \frac{1}{2}\mathrm{O_2}(g) \tag{3.37}$$

The theoretical cell voltage E° assumes no heat loss within the electrolyzer cell and is calculated using the standard electrode potentials:

$$E^\circ = E^\circ_{\text{cathode}} - E^\circ_{\text{anode}} = 0\,\mathrm{V} - (-1.23\,\mathrm{V}) = 1.23\,\mathrm{V} \tag{3.38}$$

The reversible cell voltage E_{Rev} can be derived from the Gibbs free energy change ΔG°:

$$E_{\text{Rev}} = \frac{-\Delta G^\circ}{2F} \tag{3.39}$$

where F is the Faraday constant, calculated as:

$$F = N_A e = (6.022 \times 10^{23}\,\mathrm{mol}^{-1})(1.602 \times 10^{-19}\,\mathrm{C}) = 96485\,\mathrm{C/mol} \tag{3.40}$$

Here, N_A is Avogadro's number, and e is the elementary charge. The relationship between Gibbs free energy ΔG°, enthalpy ΔH°, and entropy ΔS° is:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \tag{3.41}$$

The enthalpy change ΔH° is then:

$$\Delta H^\circ = \Delta G^\circ + T\Delta S^\circ \tag{3.42}$$

The standard entropy change ΔS° for the reaction is calculated from the standard molar entropies of the reactants and products:

$$\Delta S^\circ = S^\circ_{\mathrm{H_2}} + \frac{1}{2}S^\circ_{\mathrm{O_2}} - S^\circ_{\mathrm{H_2O}} = 130.68 + 0.5 \times 205.15 - 69.95 = 163.3\,\mathrm{J/(mol\,K)} \tag{3.43}$$

Using ΔH°, the thermoneutral voltage E is calculated:

$$E = \frac{\Delta H^\circ}{2F} = 1.48\,\mathrm{V} \tag{3.44}$$

The reversible cell voltage at non-standard conditions is given by the Nernst equation:

$$E_{\mathrm{Rev}} = E^\circ + \frac{RT}{2F}\ln\left(\frac{[p_{\mathrm{H_2}}][p_{\mathrm{O_2}}]^{1/2}}{[p_{\mathrm{H_2O}}]}\right) \tag{3.45}$$

where R is the universal gas constant, T is the absolute temperature, and p_i represents the partial pressures of the species.

The cell efficiency is defined as:

$$\text{Cell Efficiency} = \frac{E_{\text{Cell}}}{E_{\text{Total}}} \times 100\% \tag{3.46}$$

The total cell voltage E_{Total} includes overpotentials due to activation, ohmic, and concentration losses:

$$E_{\text{Total}} = E_{\text{Cell}} + E_{\text{Act}} + E_{\text{Ohm}} + E_{\text{Conc}} \tag{3.47}$$

The activation overpotential E_{Act} accounts for the energy barriers at the electrodes:

$$E_{\text{Act}} = \frac{RT_A}{2\alpha_A F} \sinh^{-1}\left(\frac{i}{2i_{0,A}}\right) + \frac{RT_C}{2\alpha_C F} \sinh^{-1}\left(\frac{i}{2i_{0,C}}\right) \tag{3.48}$$

In this equation, T_A and T_C are the temperatures at the anode and cathode, α_A and α_C are the anodic and cathodic charge transfer coefficients, i is the current density, and $i_{0,A}$ and $i_{0,C}$ are the exchange current densities.

The concentration overpotential E_{Conc} arises due to concentration gradients:

$$E_{\text{Conc}} = \sum \frac{RT}{zF} \ln\left(\frac{C_1}{C_0}\right) \tag{3.49}$$

Here, C_1 and C_0 are the concentrations at the electrode surface and in the bulk, and z is the number of electrons transferred.

The mass flow rates of hydrogen and oxygen produced are calculated using Faraday's laws:

$$\dot{m}_{\text{H}_2} = \frac{M_{\text{H}_2} I_{\text{cell}}}{zF} \tag{3.50}$$

$$\dot{m}_{\text{O}_2} = \frac{M_{\text{O}_2} I_{\text{cell}}}{zF} \tag{3.51}$$

where M_{H_2} and M_{O_2} are the molar masses, and I_{cell} is the cell current.

The cell current is determined from the input power and total cell voltage:

$$I_{\text{cell}} = \frac{P_{\text{in}}}{E_{\text{Total}}} \tag{3.52}$$

The reversible cell voltage as a function of cell temperature T_{cell} is approximated by:

$$E_{\text{Rev}}(T_{\text{cell}}) = 1.5148 - 1.5421 \times 10^{-3} T_{\text{cell}} + 9.523 \times 10^{-5} T_{\text{cell}} \ln T_{\text{cell}} + 9.84 \times 10^{-8} T_{\text{cell}}^2 \tag{3.53}$$

The activation overpotential as a function of cell current is modeled as:

$$E_{\text{Act}}(I_{\text{cell}}) = 0.0514 I_{\text{cell}} + 0.2798 \tag{3.54}$$

The ohmic overpotential is proportional to the cell current:

$$E_{\text{Ohm}}(I_{\text{cell}}) = 0.09 I_{\text{cell}} \tag{3.55}$$

The cell efficiency η_{cell} is:

$$\eta_{\text{cell}} = \frac{E_{\text{Rev}}}{E_{\text{Rev}} + E_{\text{Act}} + E_{\text{Ohm}}} \tag{3.56}$$

The power per cell is calculated from the total power divided by the number of cells:

$$P_{\text{cell}} = \frac{P_{\text{elec}}}{N_{\text{cells}}} \tag{3.57}$$

The cell current is then:

$$I_{\text{cell}} = \frac{P_{\text{cell}}}{(E_{\text{Rev}} + E_{\text{Act}} + E_{\text{Ohm}})\eta_{\text{cell}}} \tag{3.58}$$

The volumes of hydrogen and oxygen produced are:

$$V_{\mathrm{H_2}} = \frac{M_{\mathrm{H_2}}}{\rho_{\mathrm{H_2}}} \frac{I_{\mathrm{cell}}}{zF} \tag{3.59}$$

$$V_{\mathrm{O_2}} = \frac{M_{\mathrm{O_2}}}{\rho_{\mathrm{O_2}}} \frac{I_{\mathrm{cell}}}{zF} \tag{3.60}$$

By integrating these equations, the electrolyzer model calculates the cell voltage, efficiency, current, and production rates of hydrogen and oxygen based on the input power and operating conditions. This comprehensive approach allows for the optimization of electrolyzer performance, essential for efficient energy conversion in renewable energy systems.

3.1.5 Renewable Energy System Modeling

The mathematical models developed for each component—PV panels, WTs, batteries, and electrolyzers—are designed to represent the fundamental physical and operational characteristics of each part within a RES. To simulate the system as a whole, these models are interconnected, enabling a dynamic and comprehensive understanding of energy flows and interactions across all components. By integrating these models, the system-level simulation captures the impact of each component's behavior on overall performance, creating a cohesive framework that reflects real-world operating conditions.

While mathematical models provide a strong theoretical foundation, the approach within this research is adaptable. Some components are represented by machine learning models or a hybrid of both methods to optimize computational efficiency and responsiveness to complex patterns. Mathematical models deliver a precise, physics-based framework, essential for understanding the basic operations of each component. However, machine learning models can replace mathematical models when the focus shifts toward computational efficiency, enabling quicker predictions without sacrificing accuracy. This hybrid approach allows the system to balance theoretical accuracy with practical demands, depending on each component's specific characteristics and requirements.

In this integrated framework, each component model plays a specific role based on its unique characteristics. The PV and WT models calculate energy generation by using

meteorological inputs, such as solar irradiance and wind speed, to predict power output at each time step. These RE generation models produce a dynamic energy supply that varies with environmental conditions. This calculated energy supply then feeds into the broader system model, providing the basis for determining how to meet energy demand.

The energy demand and supply balance are managed through dynamic interactions between the RE geneators, ESS and load, requiring continuous updates to the state of each component. For the battery model, this involves tracking parameters such as SOC, current capacity, charging efficiency, and factors related to aging. These variables represent the battery's real-time capability to store and supply energy and provide insights into its long-term performance within the system. Similarly, the electrolyzer model continuously updates its state based on available surplus energy, adjusting parameters that impact hydrogen production and storage readiness. Each component's state reflects a combination of physical and operational factors, dynamically adapting to the changing conditions of renewable generation and load demand.

The optimization algorithm plays a central role in coordinating these interactions by dynamically calculating the optimal energy allocation across the load, storage, and potential grid export. Using inputs from the component models and additional factors such as grid energy prices and operational constraints, the algorithm determines whether energy should be directed to the load, stored, or exported. This approach allows the integrated system model to account for each component's state and performance in real time, adjusting energy flow accordingly to maintain efficiency and stability across the entire system.

Through this integration, the mathematical models for each component work together to provide an accurate and flexible system-level simulation. This interconnected model enables a continuous, time-stepped simulation that updates the system's state based on real-world factors, ensuring a comprehensive representation of the renewable energy system's behavior. This cohesive modeling framework is essential for understanding the interactions between components, optimizing energy usage, and evaluating the system's overall performance under various operating scenarios.

3.2 Machine Learning for Data-driven Analysis

The data-driven approach underpinning this research commenced with the essential stage of data collection, which serves as the foundation for all subsequent analytical efforts. The specific objectives of the study dictated the type of datasets gathered, for instance, photovoltaic power generation data from solar panels and wind speed measurements from meteorological stations. By strategically aligning the data collection process with the defined goals of the research, we ensured the acquisition of pertinent data essential for conducting substantive analyses.

Following the successful collection of data, our methodology progressed through three critical stages: data cleaning, Energy Data Analysis (EDA), and the design and training of a neural network. These stages are sequentially illustrated in Figure 3.2, which details the comprehensive process employed from initial data handling to the application of advanced machine learning techniques.

3.2.1 Data Preprocess

The initial focus was on data cleaning to prepare the raw data for accurate analysis. This process started with addressing structural errors within the dataset. Time-step inconsistencies were standardized to achieve uniformity across recordings, whether hourly or daily. Corrections were made to data type misassignments; variables incorrectly stored as integers but representing categorical data were converted to appropriate categorical types, and all date and time fields were standardized into datetime formats. These adjustments were essential for maintaining the dataset's logical integrity and facilitating precise temporal analyses.

Once structural corrections were in place, the removal of outliers was undertaken to enhance data quality. Statistical methods, particularly the calculation of Z-scores, were employed to identify and eliminate data points that significantly deviated from the mean. This step ensured that anomalies did not skew the analysis or modeling results.

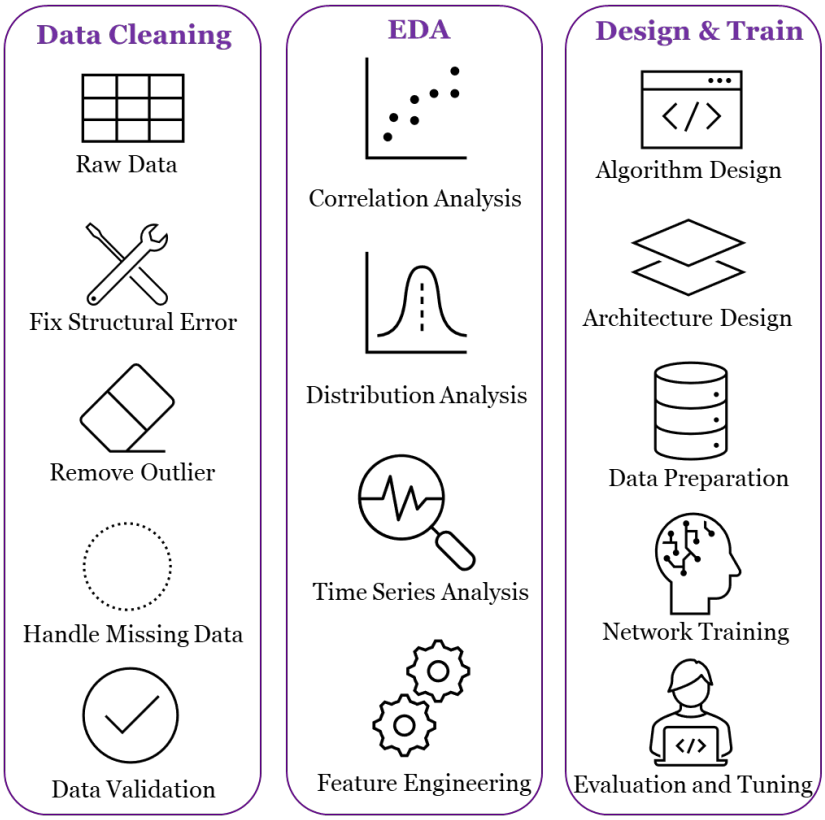


Figure 3.2: **Procedure for Machine Learning Modeling**

Handling missing data is a critical step in data cleaning, employing various imputation techniques to address gaps in numerical datasets. Common methods include replacing missing values with the mean or median, depending on the variable's distribution, to minimize bias. Additionally, interpolation or using values from temporally adjacent data points, such as the previous day at the same time, can be effective, especially in time-series data where maintaining temporal integrity is crucial. These approaches ensure the dataset's completeness and reliability, supporting robust analysis and modeling.

The final phase of data cleaning involved a rigorous data validation process to ensure the dataset's consistency and accuracy. This was achieved by identifying and removing duplicate records using key identifiers and conducting comprehensive data integrity checks.

Furthermore, the validation of data could be substantiated by cross-referencing with existing literature and by evaluating the dataset collection process itself. This evaluation included verifying the adherence to established real-world physical principles and using cross-validation techniques among related variables. Such methods helped to detect and resolve any conflicting information, significantly enhancing the reliability and robustness of the dataset. This thorough validation approach ensures that the data is not only clean but also credible and well-suited for high-quality analysis and modeling based on a deep understanding of the underlying phenomena.

This systematic and meticulous approach to data cleaning established a robust foundation for the subsequent stages of exploratory data analysis and machine learning.

3.2.2 Exploratory Data Analysis

With a clean and reliable dataset, the next stage involved exploratory data analysis to uncover underlying patterns and relationships among different variables. Correlation analysis was performed to identify the strength and direction of relationships between variables. By calculating correlation coefficients, significant associations were detected, informing the selection of features for the neural network model. Understanding these correlations helped in determining which variables had the most influence on the target outcomes, thereby optimizing the model's predictive capabilities.

Distribution analysis provided insights into the statistical properties of the data, such as mean, median, variance, and skewness. Examining the distribution of each variable allowed for the detection of any anomalies or biases that could affect the analysis. Histograms and probability density functions were utilized to visualize the distributions, facilitating a deeper understanding of the data's behavior and guiding any necessary transformations to achieve normality or homoscedasticity.

For time series data, additional analyses were conducted to capture temporal dynamics. Time series analysis involved examining trends, seasonal patterns, and cyclic behaviors inherent in the data. Stationarity tests, such as the Augmented Dickey-Fuller test, were

applied to determine whether the time series had constant statistical properties over time. Ensuring stationarity was crucial for accurate modeling and forecasting. High-accuracy statistical methods, including AR and MA, were used to identify dependencies and inform the development of appropriate models.

Feature engineering played an integral role in enhancing the model's performance. By leveraging mathematical equations and real engineering understanding, new variables were created to capture complex relationships within the data. For instance, in photovoltaic power generation analysis, features such as the angle of solar incidence, temperature coefficients, and efficiency ratios were derived. These engineered features encapsulated domain-specific knowledge, allowing the neural network to learn from more informative inputs and improve prediction accuracy.

3.2.3 Network Design and Train

In the progression from EDA to model development, defining the research objectives is crucial, as it determines whether the focus is on regression, classification, or another predictive modelling task. This decision influences the selection of the appropriate machine learning framework and algorithm, which depends heavily on whether the estimation is predictive or explanatory and on the data type—be it standard or time-series.

For time-series data that exhibits significant temporal relationships, algorithms such as LSTMs or Transformers are ideal due to their proficiency in handling sequential data and capturing temporal dependencies. Alternatively, for tasks involving straightforward regression or classification, a MLP is suitable. MLPs can model complex relationships through multiple layers and non-linear activation functions, making them effective for diverse problems, including those with non-linear and intricate feature interactions. This selection process ensures the model not only fits the data characteristics but also aligns with the research goals, thereby optimizing both performance and interpretability.

Once the algorithm was selected, designing the correct architecture involved determining the number of layers, the number of neurons in each layer, and other structural

components of the network. Initial hyperparameters were set to establish a baseline model. Hyperparameters are configuration settings that govern the learning process and the structure of the neural network but are not learned from the data. They include parameters like learning rate, batch size, number of epochs, activation functions, and optimization algorithms. Setting these hyperparameters correctly is crucial, as they significantly impact the model's ability to learn from the data and generalize to unseen instances.

Preparing the data to suit the chosen algorithm and architecture was a vital step. This involved normalizing or standardizing the data to ensure that all features contributed equally to the learning process and to accelerate convergence during training. For time-series data, time encoding techniques were applied to capture temporal information effectively. This could include converting timestamps into cyclical features to reflect periodic patterns or creating lagged features to provide the model with historical context. Ensuring that the data was appropriately formatted and scaled was essential for the neural network to learn meaningful patterns and relationships.

With the data prepared, the network training commenced by feeding the input data into the model and allowing it to adjust its weights and biases to minimize a defined loss function. Evaluation metrics such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) were used to assess the model's performance. These metrics quantified the average squared difference between the predicted and actual values, providing a measure of the model's prediction accuracy. Lower values of MSE and RMSE indicated better performance.

Based on the evaluation results, hyperparameters were fine-tuned to enhance the model's performance. This iterative process involved adjusting hyperparameters like the learning rate, modifying the network architecture by adding or removing layers or neurons, and experimenting with different activation functions or optimization algorithms. Fine-tuning is essential to optimize the model's learning process, prevent overfitting or underfitting, and improve generalization to new data.

Through this meticulous approach of selecting an appropriate algorithm, designing a suitable network architecture, preparing the data effectively, and rigorously training and

evaluating the model, a robust neural network was developed. This comprehensive process ensured that the final model was well-suited to the data, accurately captured the underlying patterns, and aligned with the research objectives. The result was a reliable and effective tool capable of making meaningful predictions and contributing valuable insights to the study.

3.3 Optimization

Different optimization algorithms are utilized for various tasks in RES optimization, each with its own advantages and limitations. LP is ideal for optimization tasks with linear relationships between variables, enabling efficient solutions when constraints are clearly defined. MILP adds the capability to handle both continuous and discrete variables, offering greater flexibility at the cost of increased computational complexity. GA, inspired by evolutionary processes, are particularly effective for addressing complex, non-linear problems and exploring large solution spaces. These algorithms are selected based on the specific optimization requirements, ensuring an optimal balance between system performance and computational efficiency.

3.3.1 Linear Programming

LP is a fundamental technique in mathematical optimization, dealing with optimizing a linear objective function subject to linear equality and inequality constraints. Since its development by George B. Dantzig in 1947, LP has become a mature and widely applied optimization method. The strength of LP lies in its ability to model and solve complex problems involving numerous variables and constraints efficiently, especially with the aid of powerful modern computers.

A standard LP problem can be mathematically formulated as Equation 3.61:

$$\begin{aligned} &\min_x obj = f^T x \\ &\text{s.t.} \quad \begin{cases} Ax \leq b \\ A_{eq}x = b_{eq} \\ l_b \leq x \leq u_b \end{cases} \end{aligned} \tag{3.61}$$

In this formulation, x represents the vector of decision variables to be determined. The objective function $f^T x$ is a linear function of these variables, where f is a known coefficient vector reflecting the cost, profit, or other measures associated with each variable. The term f^T denotes the transpose of f, ensuring proper matrix multiplication with the vector x.

The constraints include linear inequality constraints $Ax \leq b$, where A is a known matrix of coefficients, and b is a known vector, representing the limitations or capacities within which the system must operate, such as resource constraints or maximum allowable levels. Linear equality constraints $A_{\text{eq}}x = b_{\text{eq}}$, with A_{eq} being a known matrix and b_{eq} a known vector, enforce exact conditions that the solution must satisfy, such as balance equations or conservation laws. Variable bounds $l_b \leq x \leq u_b$, where l_b and u_b are the lower and upper bounds on the decision variables, respectively, define the range for each variable, often based on physical, practical, or policy considerations.

The goal is to find the vector x that minimizes the objective function while satisfying all the constraints. This involves navigating a feasible region defined by the intersection of all constraints to locate the optimal solution.

The advantages of LP include its efficiency in solving large problems, where LP algorithms, such as the simplex method and interior-point methods, are highly efficient and can handle large-scale problems with thousands of variables and constraints. LP is well-understood mathematically, providing guarantees about the existence and optimality of solutions under certain conditions. The linear nature of the objective function and constraints allows for straightforward interpretation of results and sensitivity analysis, aiding in decision-making processes.

However, LP also has limitations. It requires both the objective function and all constraints to be linear, limiting its applicability to problems where relationships are inherently

nonlinear. LP does not naturally accommodate problems requiring integer solutions, such as the number of units produced, which cannot be fractional, necessitating the use of integer programming or mixed-integer programming techniques. LP solutions can be highly sensitive to the coefficients in the objective function and constraints, where inaccurate data can lead to misleading results. While LP is efficient for large sparse problems, it can become computationally intensive for large dense problems where the coefficient matrices lack sparsity.

LP is a powerful optimization tool with a solid theoretical basis and practical efficiency for a wide range of linear problems. Its effectiveness is contingent upon the accurate linear modeling of the problem, and while it offers many advantages, practitioners must be mindful of its limitations and ensure that LP is an appropriate method for the problem at hand.

3.3.2 Mixed Integer Linear Programming

MILP extends the capabilities of LP by incorporating integer constraints into the variable definitions, enhancing the model's applicability to scenarios where variables represent inherently discrete quantities, such as numbers of trucks in a fleet or machines in a factory. This inclusion of integer constraints allows MILP to address a wider array of practical optimization problems than LP, which assumes that decision variables can take continuous values.

The fundamental difference between LP and MILP lies in the nature of the decision variables. In LP, all decision variables are continuous, facilitating simpler computational procedures and allowing problems to be solved efficiently using polynomial-time algorithms such as the simplex method or interior-point methods. In contrast, MILP complicates the computational landscape by requiring that some or all of the decision variables take integer values only.

Solving MILP often involves using combinations of LP solutions with advanced strategies like branch-and-bound, branch-and-cut, or heuristic methods. These techniques are

designed to efficiently navigate the complex solution spaces typical of MILP problems, which combine linear constraints with the non-linear complexities introduced by integer variables.

$$\begin{aligned} \min_x \quad & c^T x \\ \text{subject to} \quad & Ax \leq b, \\ & Ex = d, \\ & l \leq x \leq u, \\ & x_j \in \mathbb{Z}, \quad \forall j \in J \end{aligned} \tag{3.62}$$

The mathematical formulation of a MILP problem can be stated as follows: minimize or maximize a linear objective function subject to linear equality and inequality constraints, with the additional stipulation that certain decision variables must take integer values. Specifically, the problem is defined by an objective function $c^T x$, subject to constraints $Ax \leq b$ and $Ex = d$, where x denotes the vector of decision variables, c, A, and E are matrices detailing the coefficients of the objective and constraint functions, and b and d are vectors representing the limits of these constraints. Furthermore, the variables are bounded by $l \leq x \leq u$, and a subset of these variables indexed by J is restricted to integer values.

MILP significantly broadens the scope of linear optimization techniques to include a variety of real-world applications that involve discrete decision-making processes. While the inclusion of integer variables enhances the practical relevance of MILP, it also introduces considerable computational challenges. As such, the application of MILP, particularly in complex industrial, commercial, and public sector scenarios, often relies on the availability of advanced computational tools and algorithms to find optimal or near-optimal solutions within reasonable time frames.

3.3.3 Genetic Algorithms

GA constitute a class of search heuristics within the evolutionary algorithm family, devised to identify optimized solutions by emulating the processes of natural selection. Originat-

ing from the seminal work of John Holland in the 1970s, these algorithms harness mechanisms akin to biological evolution—selection, crossover, mutation, and inheritance—to solve problems.

The operation of a genetic algorithm commences with an initial population of potential solutions, which are often encoded as strings of bits, although variations exist depending on the specific requirements of the problem. These potential solutions, referred to as individuals or chromosomes, are evaluated using a fitness function tailored to the problem's objectives. This evaluation mimics natural selection by measuring how effectively each solution achieves the desired outcomes.

As the algorithm progresses through successive generations, it employs a fitness-based selection process to determine which individuals will reproduce, thereby passing advantageous traits to the next generation. The integration of genetic operators such as crossover and mutation facilitates the generation of new individuals, thereby enriching the genetic diversity of the population. Crossover merges attributes from two parent solutions to produce offspring, while mutation introduces random genetic variations, enhancing the algorithm's ability to explore a broader solution space.

Despite their stochastic nature, genetic algorithms are designed to incrementally improve the quality of solutions. They typically conclude when a set number of generations have been produced, a satisfactory level of fitness has been achieved, or other specific termination conditions have been met.

The principal strength of genetic algorithms lies in their versatility and robustness, making them particularly suitable for tackling complex optimization challenges that are intractable for conventional methods. These challenges often include problems with nonlinear, multimodal, or ambiguous properties. The probabilistic elements of genetic algorithms reduce the likelihood of becoming ensnared in local optima, a common issue with gradient-based optimization approaches.

Nevertheless, genetic algorithms are not without their drawbacks. They can be computationally demanding, especially for large-scale or complex problems. The rate at which they converge to a solution can be slow, and the calibration of parameters—like popula-

tion size, mutation rate, and crossover rate—is critical but can be intricate and generally requires thorough empirical experimentation to refine.

In conclusion, genetic algorithms provide a robust methodology for solving diverse optimization problems by replicating evolutionary principles. Their effectiveness, however, is heavily dependent on the careful configuration of algorithm parameters and the balance they maintain between exploring new solutions and exploiting known ones.

3.4 Performance Evaluation Metrics

This section introduces the performance evaluation metrics used to assess the effectiveness of the RES optimization. It is divided into three parts: prediction evaluation metrics, economic performance metrics, and sustainability evaluation metrics. Prediction evaluation metrics measure the accuracy of the system's forecasting capabilities, ensuring reliable energy demand and supply predictions. Economic performance metrics evaluate the cost-efficiency of the system, focusing on minimizing operational expenses while maximizing returns. Sustainability evaluation metrics assess the environmental impact of the system, ensuring that the RE solutions align with long-term sustainability goals. These metrics provide a comprehensive assessment of the system's performance, helping to validate the optimization strategies and their real-world applicability.

3.4.1 Prediction Evaluation Metrics

Mean Absolute Error

The Mean Absolute Error (MAE) is a widely used statistical measure to evaluate the accuracy of prediction models. It calculates the average magnitude of errors between paired observations expressing the actual values and predictions. The MAE is particularly useful because it quantifies the error in the same units as the data, making the interpretation straightforward and intuitive. The formula for MAE is expressed as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3.63}$$

Here, n represents the number of observations, y_i denotes the actual values, and $\hat{y}_i$ are the predicted values. This metric is especially beneficial in scenarios where it is important to avoid large errors; MAE treats all errors linearly, meaning that all individual differences are weighted equally in the average.

Root Mean Square Error

The RMSE is another robust metric used to measure the accuracy of predictions made by a model. RMSE is the square root of the average of squared differences between the actual and predicted values. This metric gives a relatively high weight to large errors, which often makes it more appropriate for many predictive tasks where larger errors are particularly undesirable.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3.64}$$

In this formula, y_i represents the actual values, $\hat{y}_i$ are the predicted values, and n is the number of observations. Due to its formulation, RMSE is sensitive to outliers and can be particularly useful when large errors are significantly more critical to diagnose and correct. Both MAE and RMSE are crucial in the field of data science for validating models and ensuring the reliability of predictive insights.

3.4.2 Economic Evaluation Metrics

Net Present Value

The Net Present Value (NPV) is a fundamental financial metric used to assess the viability and profitability of investment projects by quantifying the difference between the present

value of cash inflows and outflows over the project's duration. In the context of renewable energy systems, the NPV calculation takes into account both the initial investment and the expected future cash flows, which include revenues from energy sales and cost savings, as well as ongoing operational and maintenance costs.

The NPV is calculated using the Discounted Cash Flow (DCF) method, where future cash flows are discounted to their present value using a specific discount rate, typically reflecting the investor's required rate of return or the cost of capital. The formula for NPV is as follows:

$$NPV = \sum_{y=1}^{y=Y} \frac{\left(R_{ESav}^{(y)} + R_{ESol}^{(y)} - C_{O\&M}^{(y)} - C_{Rep}^{(y)}\right)}{(1+r)^{y-1}} - Inv \tag{3.65}$$

In this Equation 3.65, $R_{ESav}^{(y)}$ and $R_{ESol}^{(y)}$ represent the annual savings from reduced energy consumption and revenues from energy sold back to the grid, respectively. $C_{O\&M}^{(y)}$ and $C_{Rep}^{(y)}$ are the annual costs for operation, maintenance, and replacement of system components. The initial investment (Inv) covers the installation costs of all relevant system components. Y is the lifespan of the system, generally between 15 to 25 years.

Internal Rate of Return

The Internal Rate of Return (IRR) is another critical financial metric, used in conjunction with NPV, to evaluate the efficiency of an investment. The IRR is the discount rate at which the NPV of all cash inflows and outflows equals zero, effectively making it a break-even point from a financial perspective. It provides a measure of an investment's yield as a percentage, making it easy to compare with other potential investments or the cost of capital.

$$r = IRR \text{ if } NPV = 0 \tag{3.66}$$

The IRR is particularly useful as it allows investors to understand the rate of return at which the total present value of costs equals the total present value of benefits. For

renewable energy projects, which often require significant upfront investments but yield returns over long periods, the IRR is instrumental in decision-making processes, offering a clear indicator of the project's potential profitability over its operational lifetime.

3.4.3 Sustainability Evaluation Metrics

Energy Independence Score

The Energy Independent Score (EIS) is a technical indicator used to measure and quantify the degree of energy independence or self-sufficiency of a RES. The EIS is calculated through Equation 3.67 based on the proportion of the total energy demand of the system (E_{Load}) that is supplied by the system itself, as opposed to being imported from the grid (E_{Im}). The total energy supplied by the system is difference between E_{Load} and E_{Im}. The system will reach its upper boundary of EIS when all generated renewable energy is utilized. Therefore, the EIS can also indirectly reflect the carbon emission rate and efficiency of the RES.

$$EIS = 1 - \frac{\sum_{t=1}^{T} E_{Im}^{(t)}}{\sum_{t=1}^{T} E_{Load}^{(t)}} \tag{3.67}$$

Grid Energy Interaction Factor

The Grid Energy Interaction Factor (GEIF), as defined in Equation 3.68, is a crucial metric that quantifies the system's reliance on the electrical grid. It captures the balance between energy surplus and deficit by considering all energy flows between the grid, ESS, and the load. This factor provides insights into how well the system manages its energy in relation to grid dependency, highlighting the efficiency of energy surplus or deficit management.

$$GEIF = \frac{E_{RE\_Grid} + E_{ESS\_Grid} + E_{Grid\_ESS} + E_{Grid\_Load}}{E_{Load}} \tag{3.68}$$

Utilization Factor

The Utilization Factor (UF), presented in Equation 3.69, measures the effectiveness of renewable energy utilization in satisfying the load demand. It calculates the ratio of energy directly used by the load from RE sources to the total energy produced by these sources, including energy sent to the grid, stored, or curtailed. This indicator assesses the efficiency of renewable energy usage, emphasizing how effectively the generated energy is harnessed to meet the demand.

$$UF = \frac{E_{RE\_Load}}{E_{RE\_ESS} + E_{RE\_Grid} + E_{RE\_Load} + E_{RE\_Curt}} \tag{3.69}$$

Satisfied Load Fraction

Satisfied Load Fraction (SLF), detailed in Equation 3.70, evaluates the proportion of the load demand met directly by renewable energy sources. This metric reflects the sustainability and clean energy reliance of the system by illustrating the percentage of the total load covered by green energy.

$$SLF = \frac{E_{RE\_Load}}{E_{Load}} \tag{3.70}$$

Green Hydrogen Ratio

The Green Hydrogen Ratio (GHR) is designed to measure the sustainability of hydrogen production through the use of renewable energy. This ratio calculates the percentage of the total energy used in the hydrogen production process that comes from renewable sources, providing a direct measure of environmental sustainability. The GHR is defined by the following equation:

$$GHR = \frac{E_{RE\_HESS}}{E_{HESS}} \times 100 \tag{3.71}$$

In Equation 3.71, E_{RE_HESS} is the energy input from renewable sources that are directly utilized in the hydrogen production process. $E_{HESS}$is the total energy consumed by the hydrogen production system, regardless of source.

A higher GHR value indicates a greater reliance on renewable energy sources in the production of hydrogen, aligning with environmental sustainability goals and reducing the carbon footprint of hydrogen production. Conversely, a lower GHR points to a greater use of non-renewable energy sources and highlights areas for potential improvement in the integration of renewable technologies.

Hydrogen Supply Sufficiency Score

The Hydrogen Supply Sufficiency Score (HSSS) has been developed to assess the efficiency of hydrogen supply systems, incorporating two standardized metrics that measure the frequency and magnitude of hydrogen supply shortfalls. This indicator provides a quantitative analysis to guide strategic decisions aimed at optimizing system performance and reliability.

The first component of the HSSS is the Normalized Shortage Frequency (NSF), which reflects the frequency of supply shortfalls relative to the total operational period. It is calculated as follows:

$$NSF = \frac{F_S}{T} \tag{3.72}$$

In this equation, F_S represents the Shortage Frequency, or the number of time intervals where the hydrogen supply is insufficient to meet demand, and T denotes the total number of observed intervals.

The second component, Normalized Energy Shortage (NES), measures the total volume of unmet hydrogen demand compared to the total demand throughout the assessment period, providing a measure of the magnitude of shortages:

$$NES = \frac{S_{Short}}{\sum_{t=1}^{T} D_t} \tag{3.73}$$

Here, S_{Short} stands for the cumulative shortfall in hydrogen supply, highlighting areas where demand consistently outstrips supply. D_t represents the hydrogen demand at each time interval t over the period being assessed.

These metrics are combined to form the HSSS, calculated by weighting each component according to its impact on overall system performance:

$$HSSS = \frac{1}{\alpha \times NSF + \beta \times NES + 0.01} \tag{3.74}$$

In Equation 3.74, α and β are weights assigned to NSF and NES, respectively, ensuring that their sum equals one to balance the contribution of both metrics. The addition of 0.01 in the denominator prevents division by zero and ensures that the HSSS remains well-defined, even if both NSF and NES are zero. This adjustment sets the HSSS to scale between 1 and 100, where a score of 100 represents an ideal state with no supply shortfalls, and a lower score indicates more significant operational deficiencies.

3.5 Summary

This chapter presents a comprehensive methodology for modeling and optimizing RESs, focusing on key components such as PV panels, WTs, batteries, and electrolyzers. It presents the application of machine learning techniques, particularly deep learning, for data-driven analysis and prediction, enhancing the efficiency and reliability of RESs. The chapter also explores various optimization methods, including linear programming, mixed-integer linear programming, and genetic algorithms, highlighting their roles in improving system design and functionality. Performance evaluation metrics are discussed to assess prediction accuracy, economic viability, and sustainability, ensuring the development of efficient and smart RESs.

Chapter 4 focuses on component modeling, utilizing the PV model from Section 3.1 to provide theoretical support. It also employs the approach from Section 3.2 to design a machine learning model that estimates PV energy output.

Chapter 5 implements a two-stage optimization approach for the RES. It uses the PV model and battery model from Section 3.1 to model the RES. In the first stage, it applies GA from Section 3.3 to determine the component sizes for the RES, and in the second stage, it uses MILP from the same section to manage energy allocation. Indicators from Section 3.4 are used to evaluate system performance based on specific system configurations.

Chapter 6 combines the PV model, WT model, battery model, and electrolyzer model from Section 3.1 to simulate the RES. It also uses the approach from Section 3.2 to design the machine learning model. For energy management, it employs MILP from Section 3.3, and indicators from Section 3.4 are used to assess the performance of different optimization approaches.

Details of how these methods interact within each project are provided in the respective chapters for each project.

Chapter 4

Data-Driven Photovoltaic Tilt Angle Optimization for Power Generation

This chapter focuses on the optimization of the photovoltaic panel tilt angle for maximizing energy generation. Section 4.1 summarizes main contributions and novelties of this project. Section 4.2 introduces the target system and the proposed algorithm for optimization. Section 4.3 outlines the simulation procedure, covering essential steps such as data preprocessing, hypothesis testing, correlation analysis, and cross-validation to ensure the robustness of the modeling and optimization process. The results and their implications are discussed in Section 4.4, evaluating the effectiveness of the proposed algorithm and the insights gained from the optimization process. Finally, Section 4.5 summarizes the key findings.

4.1 Contribution

The main contributions of this project are as follows:

- A novel data-driven methodology is introduced to explain the intricate relationship between tilt angle and conversion efficiency in photovoltaic systems. The conclu-

sions are rigorously cross-validated against a reliable mathematical model, enhancing the robustness and credibility of the findings. Unlike conventional neural networks that require expert-driven feature extraction from ultra-high-dimensional datasets to avoid irrelevant inputs, this approach mitigates the 'black box' risk inherent in neural networks through comprehensive cross-validation, thereby improving transparency and reliability.

- A robust ANN model is developed to accurately estimate the power output of photovoltaic panels, utilizing power output as the optimization objective. By applying ANN architecture optimization techniques, the model's accuracy is significantly enhanced. The well-trained ANN demands relatively low computational resources during operation, relying on simple matrix manipulations rather than complex integral computations. Furthermore, the trained network exhibits broad applicability across similar climate zones, and its rapid response capabilities facilitate real-time control of the tilt angle, thereby improving system adaptability and efficiency.

- This study investigates how hourly optimization and varying optimization horizons impact overall optimization performance, particularly when accounting for energy conversion efficiency. Valuable insights are provided into the temporal dynamics of optimization in photovoltaic systems, addressing an area that remains inadequately explored.

4.2 Target System and Proposed Algorithm

Figure 4.1 (a) presents the schematic of the energy flow of this study, which focuses on optimizing the tilt angle of PV panels to maximize power output. Solar irradiance reaches the tilted surface of a PV panel through three primary pathways: directly from the sun, after diffusion by clouds, or following reflection from the ground. The PV panel converts this solar radiation into electrical energy, and adjusting the panel's tilt angle allows control over the amount of incident radiation, directly influencing the power output.

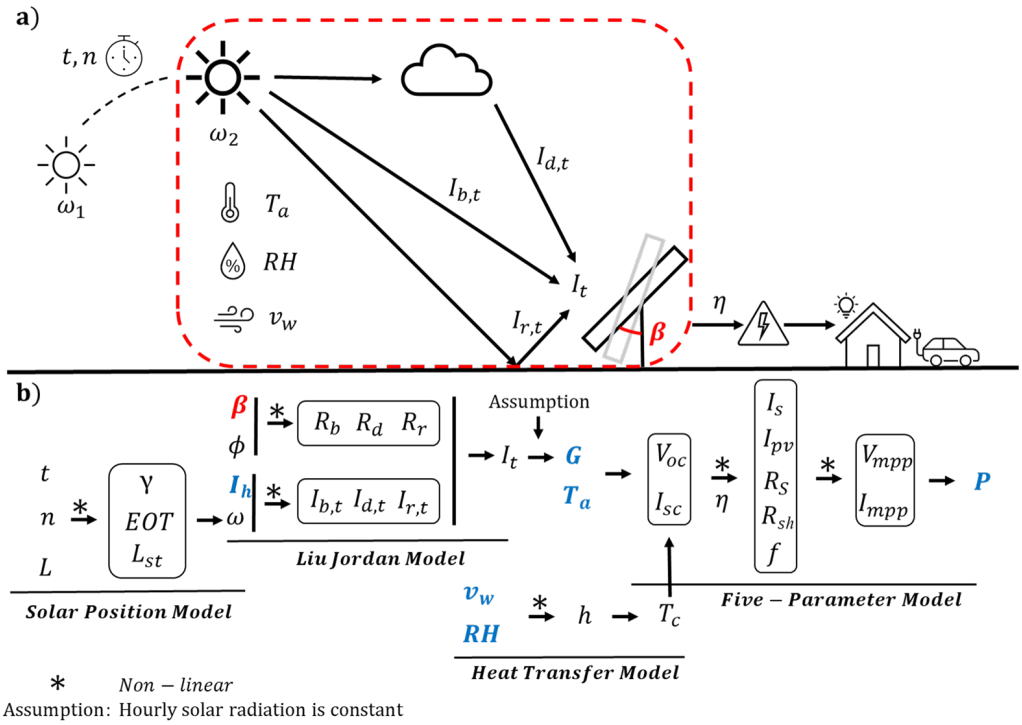


Figure 4.1: **(a) Schematic of Energy Flow (b) Functional Relationship of Model Parameters**

Optimizing the tilt angle involves calculating the solar position, the incident radiation, and the power output of the PV panel for different tilt angles. Based on these calculations, the optimal tilt angle is determined to maximize power output. The required models for this study include a solar position model to locate the sun based on time, the Liu-Jordan model to calculate solar irradiance incident on the tilted surface, and a model to estimate the power output of the PV panel. Due to limitations in the traditional five-parameter power output model, a neural network is introduced to improve estimation accuracy.

Figure 4.1 (b) illustrates the functional relationships among the model parameters. The solar position model uses time t and longitude L to calculate the sun's hour angle ω. The Liu-Jordan model then uses the tilt angle β, latitude ϕ, and global radiation on the horizontal plane I_h to calculate the radiation incident on the tilted surface I_t.

The optimization variable, tilt angle β, affects I_t at this stage. Finally, inputs such as radiation G, ambient temperature T_a, wind speed v_w, and relative humidity RH are used by the neural network to estimate the power output P of the PV panel. Sections 3.1.1 provide detailed explanations of the solar position model, the Liu-Jordan model and the five-parameter model.

Variables marked in blue were extracted from external data sources: the National Institute of Standards and Technology (NIST) dataset and the TRNSYS Simulation Software (TRNSYS) simulation platform. One set of variables (G, T_a, v_w, RH, P) was extracted from the NIST dataset for analysis and to train the neural network. Another set of variables (I_h, T_a, v_w, RH) was obtained from TRNSYS for simulation purposes.

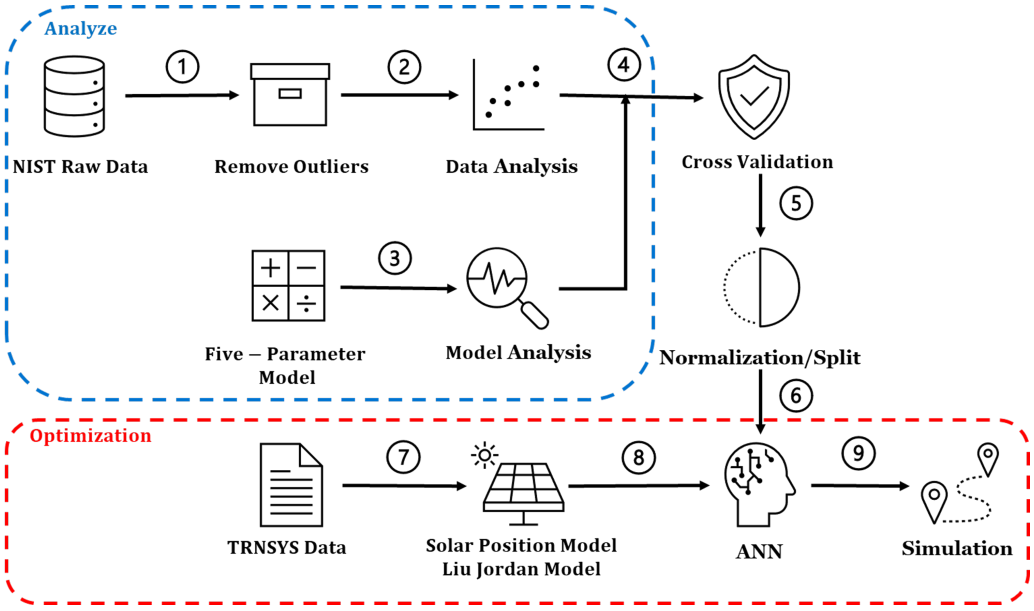


Figure 4.2: **Flow Chart of the Proposed Methodology**

Figure 4.2 illustrates the proposed methodology. The study begins with preprocessing the raw data from the NIST photovoltaic dataset, removing outliers to ensure data quality. A statistical analysis is performed on the NIST data to identify crucial input parameters for the neural network. The traditional five-parameter model is analyzed to determine

essential parameters from a physical standpoint, and conclusions from the statistical and mathematical analyses are cross-validated to ensure consistency.

The data is then normalized and shuffled to prevent the neural network from memorizing trend patterns, and the processed data is split into training and testing sets. The neural network is trained with the training set, and its performance is evaluated using the test set. Data is extracted from TRNSYS and prepared for simulation.

Using the data from TRNSYS, the solar position relative to a specific location on Earth is calculated, expressed through the hour angle ω. The Liu-Jordan model is then used to calculate the total solar radiation incident on the tilted surface for different tilt angles. The trained neural network is employed to estimate the power output, and a brute-force algorithm is used to determine the optimal tilt angle that maximizes the power output.

The optimization process is performed for different time intervals, including one hour, one month, and one year, to observe how the optimal tilt angle changes over time. The results are analyzed, and uncertainties are discussed to provide insights into the performance of the methodology and potential areas for future research.

By following this comprehensive methodology, the study aims to develop a robust approach for optimizing PV panel tilt angles, and enhancing energy production under varying environmental conditions.

4.3 Simulation Procedure

This study used a neural network to estimate the power output of a photovoltaic panel. Neural networks are data-driven models that can overcome the limitations of conventional mathematical models. Mathematical models often face a trade-off between accuracy and computational cost. They tend to make assumptions and simplify problems to ensure relatively high accuracy with low computational cost requirements during real-world applications. However, this means that mathematical models could fail when calculating scenarios outside the scope of their assumptions. The five-parameter model, for instance,

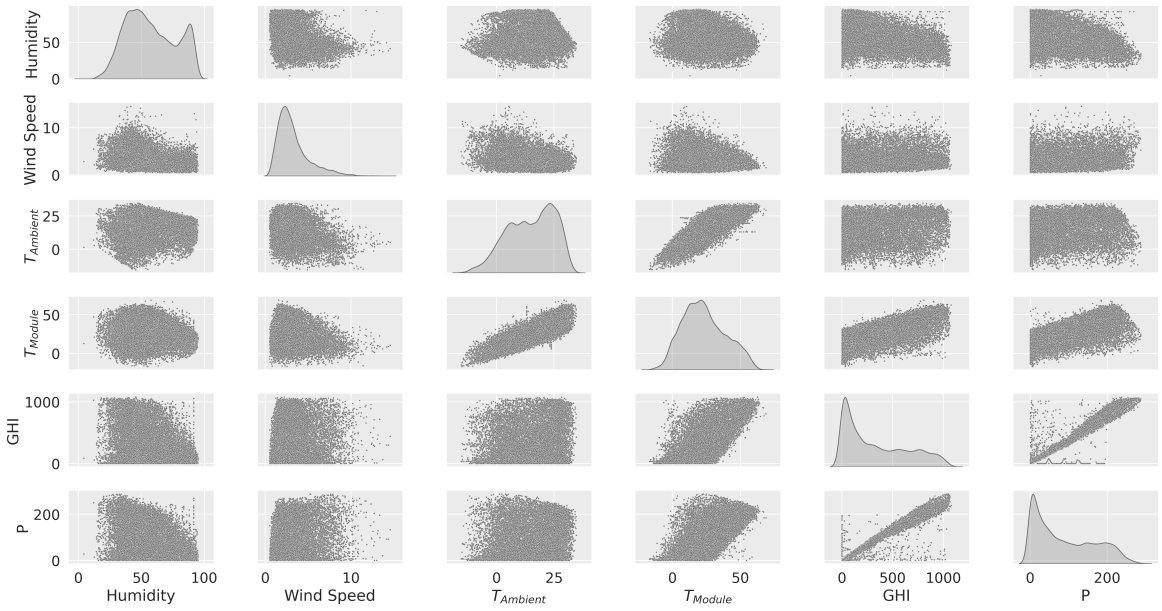


Figure 4.3: **Pairwise Plot of NIST Data Set (Raw)**

does not consider hidden patterns such as the effect of anti-reflection coatings. Although modeling such hidden patterns with mathematical models might not significantly improve accuracy, it would dramatically increase computational requirements and decrease real-time computational efficiency. However, neural networks can perform sophisticated regression and overcome the imperfections of mathematical models because the number of neurons in a specific neural network is fixed. Therefore, neural networks can model complex systems without excluding hidden patterns. This study used statistical analysis to determine inputs for the neural network and cross-validated them with the mathematical model to guarantee that the inputs have physical meaning.

4.3.1 Data Pre-process and Hypothesis Test

The NIST[131] photovoltaic dataset includes data collected from photovoltaic arrays and weather stations located at the NIST Campus in Gaithersburg, Maryland, USA. The dataset includes thousands of data points, each with multiple dimensions, sufficient for

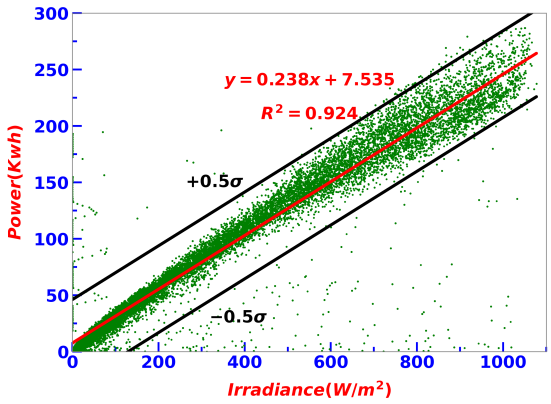


Figure 4.4: **Two Dimensions Raw Data of NIST Photovoltaic Data Set - Irradiance vs Power Output**

training a neural network to estimate a photovoltaic array's power output. This study used six of the most important features and visualized the data. Figure 4.3 shows the pairwise plot of the six-dimensional data: relative humidity (%), wind speed (m/s), ambient temperature (T_{Ambient} in °C), module temperature (T_{Module} in °C), global horizontal irradiance (GHI in W/m^2), and power output (P in kW) extracted from the NIST dataset. The plots on the diagonal show distributions of the single variables, and scatter plots visualize the relationships between two variables. From the distribution plots, we can conclude that any location in the same climate zone and following a similar meteorological pattern could use this dataset for neural network training purposes because this dataset provides sufficient data. Besides, from the scatter plots, we can easily find some variables that demonstrate a strong linear relationship with other variables; for example, GHI and P. However, we cannot conclude that variables that do not show a linear relationship are useless at this step. The pairwise plot provides a first impression of the NIST dataset; this study still needs to perform follow-up data analysis and cross-validate the conclusions with the mathematical model to guarantee the reliability of extracted features.

Nevertheless, some problems exist with using the data-driven approach to predict photovoltaic power output. Firstly, the dataset used for developing the model usually includes outliers that deviate from regular data points, caused by various reasons such as record-

ing and measuring errors. Removing outliers from the dataset is crucial to ensure that these data points will not confuse the neural network. In order to remove outliers from multi-dimensional data, this study chose two of the most critical dimensions: irradiance and power output. The GHI is equal to irradiance on the surface since photovoltaic panels in this dataset are horizontal.

Figure 4.4 visualizes the data points in the NIST data set. All data points with negative power output values beyond 0.5 standard deviations were removed. The data points with a negative power output value are not realistic and are believed to be caused by measuring errors, thus they must be removed. The performance of the photovoltaic panel at the deficient irradiance level has already been investigated by Sebastian et.al.[132]. The black bold line in Figure 4.4 was calculated through the ordinary least squares regression approach. The value of R square (R^2) is 0.924, which means 92.4 % of the variation of the dependent variable (power output) can be explained through the independent variable (GHI). At the same time, the correlation between the two variables is more than 0.96 before data pre-processing, and it shows a highly linear relationship between the two variables. The research of Alzahrani et.al.[133] can support this result since they found that irradiance is a determining factor for photovoltaic power output. Moreover, the data points are more concentrated around the fitted line in the lower GHI range. This phenomenon is caused by other factors, such as module temperature since these factors will affect the efficiency of photovoltaic panels.

The hypothesis testing examined if independent variables can explain the dependent variable in the multiple linear regression. The dependent variable is the power output, and the independent variables are meteorological variables. Equation 4.1 shows the null hypothesis and alternative hypothesis for hypothesis testing. The null hypothesis H_0 states that slope coefficients $b_1, b_2, ..., b_n$ (n is the index of independent variables) of multiple linear regression equals zero, and the dependent variable is not significantly affected by independent variables. If the null hypothesis is rejected, the study can conclude that slope coefficients are significantly different than zero. After performing the hypothesis test for slope coefficients, all p-values were lower than 0.05. Thus, this study has at least 95% confidence that slope coefficients are significantly different than zero.

$$H_0 : b_1, b_2, ..., b_n = 0$$
$$H_a : b_1, b_2, ..., b_n \neq 0 \tag{4.1}$$

4.3.2 Correlation Analysis

Correlation analysis is one of the fundamental approaches in data analysis. Correlation represents how strong the linear relationship is between two variables. If the absolute value of the correlation coefficient is closer to 1, this means the relationship between two variables is more linear and thus more correlated. However, correlation analysis is only effective when analyzing linearity in relationships. It cannot be used to identify other relationships, such as quadratic relations. Thus, correlation analysis is not a perfect data analysis approach. However, the result of correlation analysis is still valuable to help researchers to develop their models. Figure 4.5 shows correlation analysis results for the main variables of the NIST data set.

The correlation matrix provides evidence that GHI strongly correlates with the power output, and this relationship has already been investigated in the previous section. Thus, solar irradiance is the most critical parameter to calculate energy generated.

The correlation between energy output and module temperature is more than 0.67, demonstrating a strong correlation. Additionally, the correlation between GHI and module temperature is 0.71. The relationship between input variables should be independent if they are used as neural network inputs. The complexity of the regression analysis increases if input variables are highly dependent on each other. The correlation between the ambient temperature and the module temperature is 0.8, and the correlation between GHI and module temperature is 0.7. This study chose the ambient temperature as the second input. Furthermore, relative humidity has both positive and negative effects on the power output of the photovoltaic panel. Water vapor in the air affects the power output of the photovoltaic panel in two ways: the moisture on the panel and the air temperature. After the water vapor accumulates on the photovoltaic panel, it can reflect a portion of

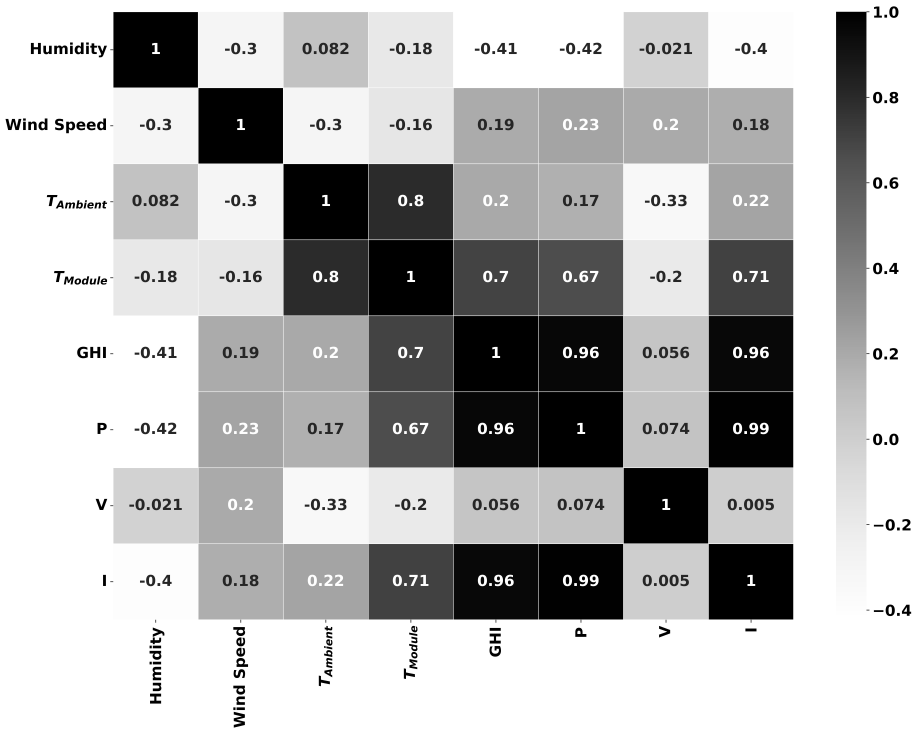


Figure 4.5: **Correlation Matrix of Pre-processed Data**

the received solar irradiance. However, the water vapor can increase the heat transfer coefficient and help the photovoltaic panel exchange heat. Thus, a high relative humidity level can improve the power output of the photovoltaic panel by helping it to decrease the module temperature. The correlation analysis found that the negative effect dominates the positive effect, which means the correlation between humidity and power output was -0.42. This negative correlation tells us that high relative humidity will decrease power output. Finally, the wind speed causes a minor increase in the power output since the two variables have a correlation of 0.18. The high wind speed will reinforce the convection effect and help cool down the photovoltaic panel by changing the heat transfer coefficient.

Based on the results of the correlation analysis, this study used four variables as inputs: relative humidity, wind speed, ambient temperature and irradiance. Even though these four input variables are essential in the correlation analysis, they still require cross-validation

with a mathematical model to ensure they are physically meaningful.

4.3.3 Cross Validation

The mathematical model can cross-validate the input variables of the neural network with the validated input variables determined by the correlation analysis. Figure 4.6 shows the relationship between ambient temperature, GHI and module temperature as calculated by the mathematical model. In contract, Figure 4.7 shows the same relationship that was extracted from the NIST data set. However, in Figure 4.7 the ambient temperature is fixed. In Figure 4.7, the black dash line is a linear regression line that can fit all data, and the black line is a line calculated by the mathematical model when ambient temperature equals the average ambient temperature of the data set. From these two figures, it can be surmised that the ambient temperature determines the intercept since the average ambient temperature in the data set is close to $11°C$ and is equal to the ambient temperature used to calculate the blue line. However, the blue line and red line slopes are 0.022 and 0.036, respectively. The red line and the blue line share a similar intercept, indicating that the module temperature mainly depends on the ambient temperature. Therefore, the conclusion from the correlation analysis is verified.

On the other hand, the slopes of the two lines are different; this phenomenon is caused by the mathematical model, which only considers the effects of irradiance and ambient temperature. The five-parameter model assumes that the heat transfer coefficient is constant. In other words, the model does not consider the effect of wind and relative humidity. The heat transfer coefficient cannot significantly impact module temperature during low irradiance. However, the module temperature increases rapidly with increasing irradiance, and the heat transfer coefficient becomes a crucial variable. The reason described above could also explain the phenomenon that the difference between the mathematical model and regression model increase with the irradiance. Thus, humidity, wind speed, ambient temperature and GHI are four input variables used to estimate power output. Based on the previous analysis, they have been proven to have a high correlation with the power output and low correlation between each other.

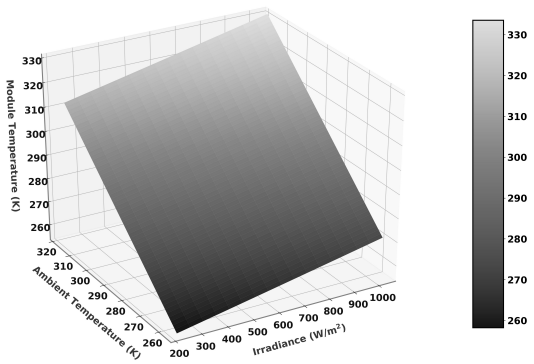


Figure 4.6: **Irradiance vs Ambient Temperature vs Module Temperature by Five Parameters Model**

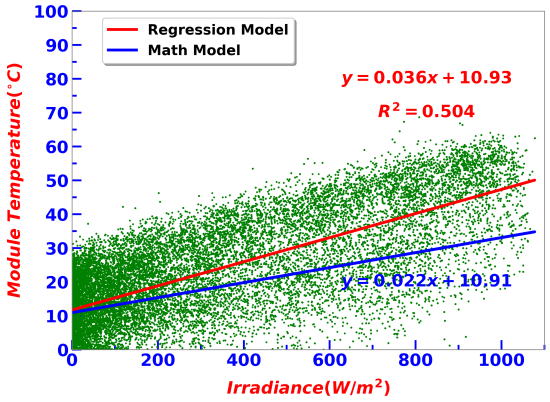


Figure 4.7: **Two Dimensions Processed Data of NIST Photovoltaic Data Set - Irradiance vs Module Temperature**

Table 4.1: **Hyperparameters of the Neural Network**

Hyperparameter	Value
Layers	5
Neurons	[4, 32, 64, 16, 1]
Batch Size	64
Epochs	30
Dropout	0.2
Decay	0.8

4.4 Result and Discussion

After determining the four inputs for the neural network, the next step was to design its architecture. A five-layer MLP was established, comprising one input layer, three hidden layers, and one output layer. The dataset used for training was extracted from

Table 4.2: **Simulation Results**

Month	Hourly (MWh)	Monthly (MWh)	Annually (MWh)	Base (MWh)
Jan	2.84	2.78	2.69	2.19
Feb	3.98	3.93	3.93	2.92
Mar	6.33	6.27	6.27	5.17
Apr	9.09	8.75	8.75	8.10
May	12.87	10.69	10.46	10.48
Jun	14.02	11.17	10.73	11.08
Jul	13.02	10.42	10.09	10.28
Aug	11.07	10.09	10.05	9.52
Sep	9.04	8.98	8.86	7.50
Oct	7.97	7.88	7.44	5.61
Nov	4.64	4.58	4.31	3.29
Dec	2.48	2.43	2.34	1.88

TRNSYS, a transient systems simulation software validated in the literature. The data were divided into a training set (80%) and a testing set (20%), with the order shuffled to prevent the neural network from memorizing sequential patterns inherent in the data trend—specifically, the daily power output rising from zero at night to a peak around noon and then decreasing. Random search was employed to optimize hyperparameters, enhancing the neural network's accuracy. Although the specifics of the optimization process are beyond the scope of this study, techniques such as learning rate decay were utilized to maintain a high learning speed initially while ensuring convergence in later stages. To mitigate overfitting—a scenario where the neural network performs well on training data but poorly on new data—the dropout technique was applied, randomly deactivating a proportion of neurons during each training iteration. The optimized hyperparameters are

presented in Table 4.1.

Figure 4.8 and Figure 4.9 illustrate the neural network's performance on the training and testing sets, respectively. Due to the large volume of test data, a representative sample of data points was randomly selected for visualization. The x-axis represents values estimated by the neural network, while the y-axis indicates the actual values from the dataset. These figures demonstrate that the neural network predicts power output with relatively low error, as most data points lie near the diagonal. For a quantitative evaluation of the prediction accuracy, the RMSE was used instead of visualization. The neural network achieved an RMSE of 0.04 on the training set and 0.06 on the testing set, which aligns with RMSE values reported in the literature [134].

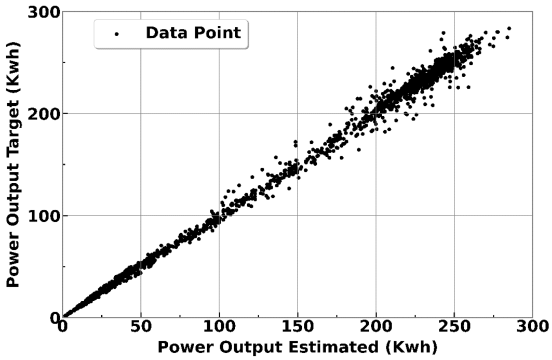


Figure 4.8: **Performance of the Neural Network on Training Set**

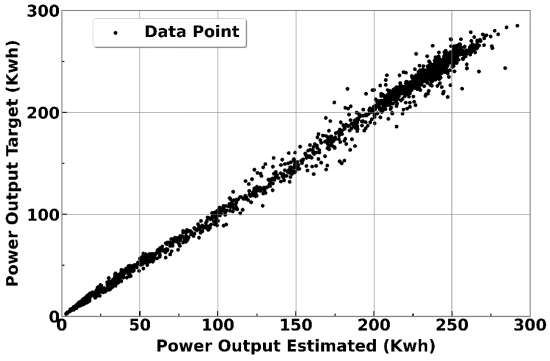


Figure 4.9: **Performance of the Neural Network on Testing Set**

After validating all mathematical models and training the neural network, simulations were performed to optimize the tilt angle of the PV panel based on power output. Ottawa, Canada $(45.38°N, 75.72°W)$ was selected as the location for the simulation. The following assumptions were made during the simulation: the hourly meteorological conditions were considered constant; sufficient space was assumed in the photovoltaic farm to neglect shading effects; the PV panels were assumed to use monocrystalline silicon solar cells; and the neural network was presumed to capture all internal patterns.

The meteorological data were extracted from TRNSYS, ensuring reliability without the

need for additional processing to detect outliers or handle missing data. The simulation process proceeded as follows: first, the meteorological data were used as inputs to calculate the radiation on the tilted surface. Subsequently, the trained neural network was employed to calculate the power output of the PV panel for each tilt angle. The optimal tilt angle was then determined over a range from 0 to 90 degrees, in 1-degree increments, using a brute-force search algorithm. An hourly model was used for all optimization horizons, and hourly outcomes were integrated to obtain monthly and annual optimization results. Finally, the tilt angle that yielded the maximum power output was selected as the optimal angle, and the harvested energy was calculated accordingly.

4.4.1 Energy Generation

The simulation results show that, compared to horizontally placed panels, optimizing tilt angles annually, monthly, and hourly can achieve 9.7%, 12.74%, and 24.78% more energy, respectively. Table 4.2 and Figure 4.10 present the optimization results. Figure 4.10 visualizes the benefits of the power-based tilt angle optimization. In this figure, the base case is the flat-placed photovoltaic panels. The blue, orange, and green bars represent the excess energy harvested compared to the base case for the hourly, monthly, and annual tilt angle optimizations, respectively. For example, hourly optimization can help the photovoltaic panels collect almost 25% more energy than the base scenario. The red line indicates the total amount of energy harvested during monthly optimization, and the energy harvested under other optimization horizons can be derived accordingly. Table 4.2 supplements Figure 4.10, presenting the energy harvested during different scenarios on a monthly basis.

An important observation from the simulation results is that the optimization provides more substantial percentage increases in energy harvested during the winter months, exceeding 40% compared to horizontally placed panels. This significant improvement is attributed to the lower baseline solar irradiance in winter, where optimal tilt adjustments can more effectively capture oblique sunlight. In contrast, during the summer months, the optimization achieves at least a 10% increase in energy harvested. Although the percentage

improvement is smaller, the absolute amount of additional energy collected is substantial due to the higher solar irradiance during summer. This highlights that the optimization is beneficial throughout the year, with both relative and absolute gains depending on seasonal conditions.

Moreover, the benefit of the optimization remains significant during the summer since it yields a substantial absolute increase in energy harvested, despite the percentage improvement being around 10%. The proposed approach can provide 16.23% more energy than the conventional rule-based optimization method, which sets the tilt angle equal to the latitude. Table 4.3 provides the optimal tilt angles from the monthly optimization and compares them with irradiance-based optimization. Notably, compared to the irradiance-based optimization, the monthly optimal tilt angles are smaller in power-based optimization. This suggests that factors beyond simply maximizing irradiance, such as the efficiency of the PV panels at different angles and temperatures, play a crucial role in maximizing power output. The neural network's ability to capture these complex interactions between meteorological variables and PV performance enables a more precise optimization than traditional methods.

These results provide compelling evidence of the value of power-based tilt angle optimization. The effective application of the neural network in this study underscores the potential of machine learning techniques in renewable energy optimization. By accurately modeling the nonlinear relationship between environmental conditions and power output, the neural network facilitates the determination of optimal tilt angles that enhance energy harvesting. This data-driven approach can adapt to varying conditions and system characteristics, offering a versatile tool for optimizing photovoltaic installations. The significant improvements in energy output achieved through relatively small adjustments in tilt angle emphasize the importance of fine-tuning system parameters to maximize efficiency and sustainability.

Furthermore, the study demonstrates that even standard practices in PV system installation may not fully exploit the potential energy yield. The conventional method of setting the tilt angle equal to the latitude does not account for the dynamic interplay be-

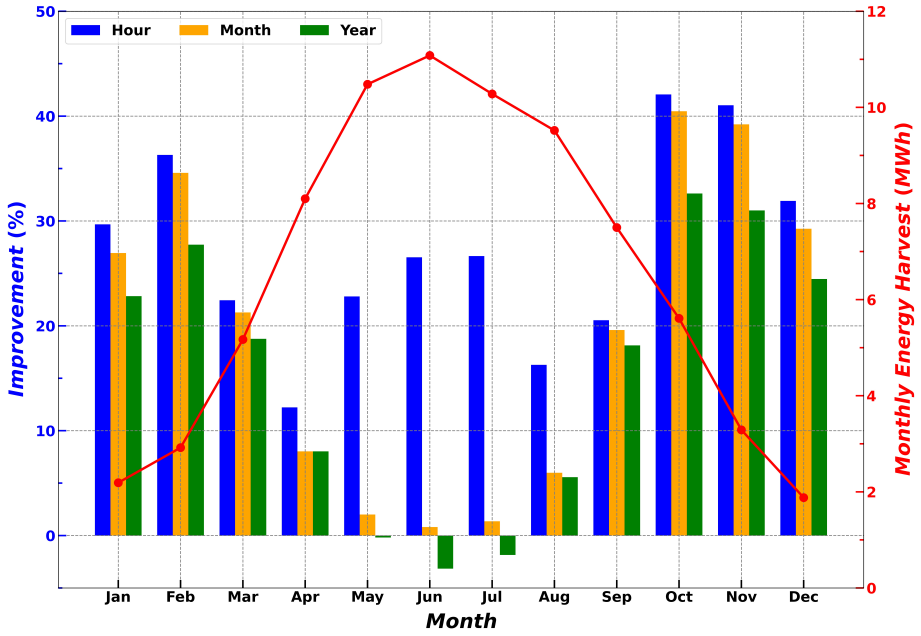


Figure 4.10: **Optimization Results and Optimal Tilt Angles for Different Time Intervals**

tween various factors affecting power output. By employing a power-based optimization approach, it is possible to achieve considerable energy gains over conventional methods. This highlights the importance of incorporating advanced optimization techniques in the design and operation of photovoltaic systems to fully leverage their capabilities.

4.4.2 Assumptions and Limitations

This study has two limitations. The most important limitation is that we assumed that the hourly irradiance was constant. Such an assumption must be made in order to perform the dynamic simulation. The optimization result might have an error if the irradiance changes significantly within an hour; however, our study provided an optimization approach. This limitation could be solved by measuring irradiance with sensors during real-time optimization. Another limitation is that it is essential to obtain a relative data set before using

Table 4.3: **Optimal Tilt Angle of Power-based and Irradiance-based Monthly Optimization**

Month	Power-based	Irradiance-based
Jan	51°	56°
Feb	55°	56°
Mar	46°	47°
Apr	29°	31°
May	15°	17°
Jun	9°	12°
Jul	12°	14°
Aug	24°	26°
Sep	42°	42°
Oct	56°	56°
Nov	58°	61°
Dec	52°	56°

the proposed optimization approach. When the photovoltaic technology or climate zone changes, the neural network must be retrained; otherwise, the ANN cannot accurately calculate power output. However, the training methodology is already provided in this study. It is easy to collect data and retrain ANN with the proposed approach.

4.5 Summary

This study developed a new data-driven power-based optimization approach to help solar farms harvest more energy during operation. This study proved that power-based optimiza-

tion will provide a different tilt angle than the conventional irradiance-based optimization because of the trade-off between irradiance and efficiency. This study also investigated the return on different optimization horizons and found that hourly optimization can significantly decrease the green energy premium. Despite higher initial investment on adjustable panels, the optimization could significantly improve the capital return in the long term in two important ways. First, the solar farm using this optimization approach could generate 22.43% more power with the same number of photovoltaic panels, resulting in greater cash flow during the operation. Second, fewer photovoltaic panels mean less land usage and lower maintenance costs, which means a lower initial investment. Panels with more advanced module materials could improve the benefits of the tilt angle optimization further.

Chapter 5

Data-Driven Two-stage Component Capacity Optimization

This chapter presents the project for optimizing the capacity of components in a RES. Section 5.1 summarizes main contributions and novelties of this project. Section 5.2 introduces the target system, outlines the conceptual framework, and defines the performance metrics used for evaluation, which are adapted from Chapter 3 to suit this specific optimization problem. Section 5.3 describes the simulation procedure, detailing the two-stage optimization process and the interrelation between decision variables and objectives, which are crucial for achieving optimal system performance. The results are analyzed in Section 5.4, with a focus on the capacity of major components, the decision-making framework, system performance, and a comparative analysis of different optimal system configurations. This section also discusses the assumptions and limitations that influence the optimization outcome. Finally, Section 5.5 summarizes the key findings from the component capacity optimization.

5.1 Contribution

The main contributions of this project are as follows:

- A novel modeling approach that addresses scalability and adaptability issues in energy system management by capturing the economic interactions between RES and the grid with enhanced realism. It effectively overcomes the limitations of models in literature by incorporating local regulations and wholesale electricity prices, thereby refining energy management practices.
- The integration of a modified NSGA-II with a MILP strategy, introducing a Dynamic Population Decay (DPD) mechanism to accelerate the optimization process.
- Introducing a novel approach that extends proven retail-level methodologies to the wholesale domain, enabling effective interactions with large-scale grid operations. This adaptation significantly enhances the management of power and information flows, improving efficiency across different scales. The proposed model bridges the gap between retail and wholesale systems, showcasing a significant advancement in grid interaction.
- Introducing of a decision-making framework utilizing multi-objective optimization via the Pareto-adaptive Epsilon dominance algorithm. This innovation allows for the customization of energy solutions to specific user requirements, with human experts making the final decision based on their preference among various objectives.

5.2 Target System and Proposed Algorithm

This section primarily presents the methodology behind the development of a novel two-stage optimization framework for RES. It begins with Section 5.2.1, which provides an overview of the target system. Following this, Section 5.2.2 details the proposed two-stage RES capacity optimization framework. Finally, Section 5.2.3 outlines the performance metrics employed in this study for optimization and decision-making purposes.

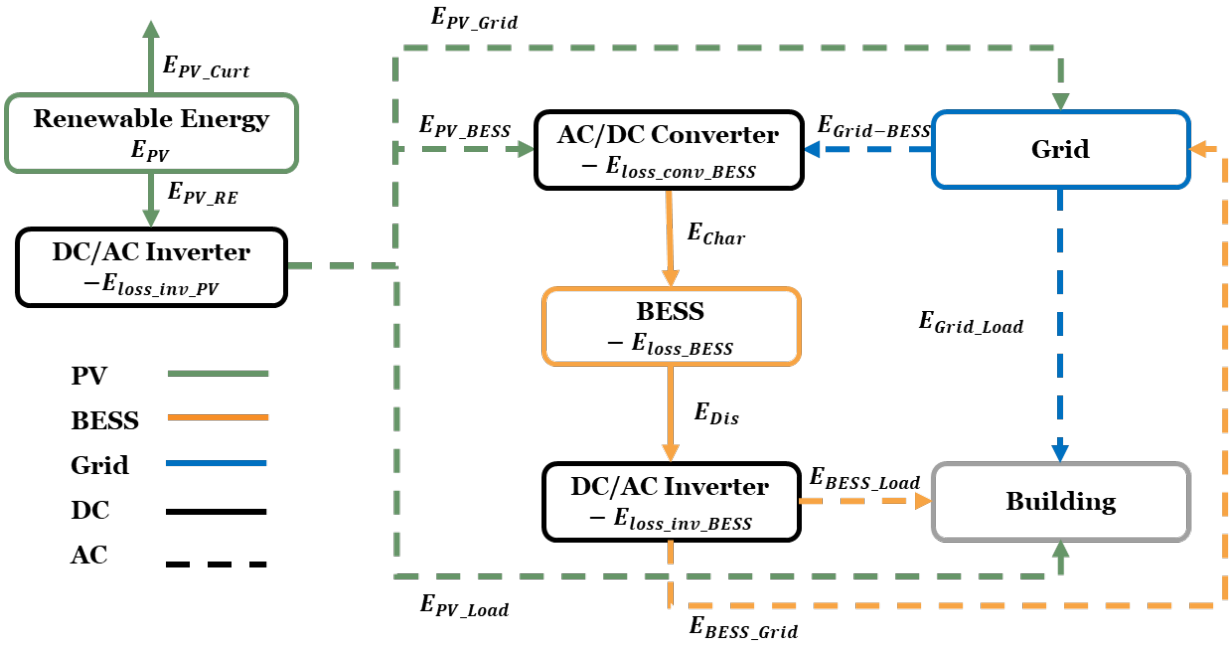


Figure 5.1: **Schematic of Target System with Key Variables**

5.2.1 Target System

The target system used in this study is demonstrated in Figure 5.1, which presents a detailed schematic illustrating the interaction among its three main components: the RES with PV arrays, the BESS, and the grid. Energy produced by the PV array (E_{PV}) is outputted to the RES (E_{PV_RE}), with surplus energy not consumed by the RES being curtailed (E_{PV_Curt}). This energy, once transformed in a DC/AC inverter from DC to AC) at a uniform voltage, leads to an energy loss ($E_{Loss\_Inverter\_BESS}$). After conversion, the energy may directly meet the energy demands of the load (E_{PV_Load}), be reserved in the BESS for later usage (E_{PV_BESS}), or be exported to the grid (E_{PV_Grid}) for credit. The BESS requires low-voltage DC for charging. As a result, any AC energy intended for storage in the BESS undergoes an AC/DC conversion, leading to another loss ($E_{Loss\_Converter\_BESS}$). This converted energy (E_{Char}) is then stored in the BESS. When discharged, the battery energy (E_{Dis}) experiences inherent losses (E_{loss_BESS}) and another DC/AC inversion. This

energy can then either supply the building's energy demand (E_{BESS_Load}) or be sold to the grid (E_{BESS_Grid}). In scenarios where the RE or BESS fails to meet the energy demands, the grid supplies energy directly to the building (E_{Grid_Load}), or energy can be imported into the BESS (E_{Grid_BESS}) during times of lower energy price, optimizing for later consumption. These electricity prices are influenced by the fluctuations in the wholesale electricity market, which in turn reflect the dynamic relation between demand and supply of electricity. Typically, these lower prices occur during late night to early morning hours, a period characterized by a substantial decrease in energy demand relative to supply. This strategic timing allows for the import of electricity at reduced costs, optimizing the system for later consumption while ensuring economic efficiency.

5.2.2 Conceptual Framework

The concept of the two-stage RES capacity optimization framework deployed in this study is demonstrated in Figure 5.2. The proposed framework comprises a system capacity optimization stage on the left side, serving as the primary optimization, and a year-long energy management simulation on the right side, which is used as the secondary optimization. The system capacity optimization is based on the NSGA-II and tackles the computational challenges encountered by the extensive searching space required for component capacity optimization. In phase one, the primary optimization randomly initializes a group of candidate solutions, each containing specific sizes for PV panels and a BESS. These sizes serve as system parameters, directly influencing the renewable energy generation profiles and energy storage ability used in the energy management simulation. The second phase involves preparing data, including an enhanced load profile derived from aggregating and introducing volatility to individual building profiles, making the simulation more close to a real-world scenario. Details on this enhancement are clarified further in the later section. The third phase formulates MILP problems based on these inputs. During the fourth stage, the MILP solver optimizes the system using predefined decision variables, objective functions, and constraints. Meanwhile, a BESS aging model will simultaneously update the BESS available capacity after aging and the state of BESS. The fifth phase evaluates

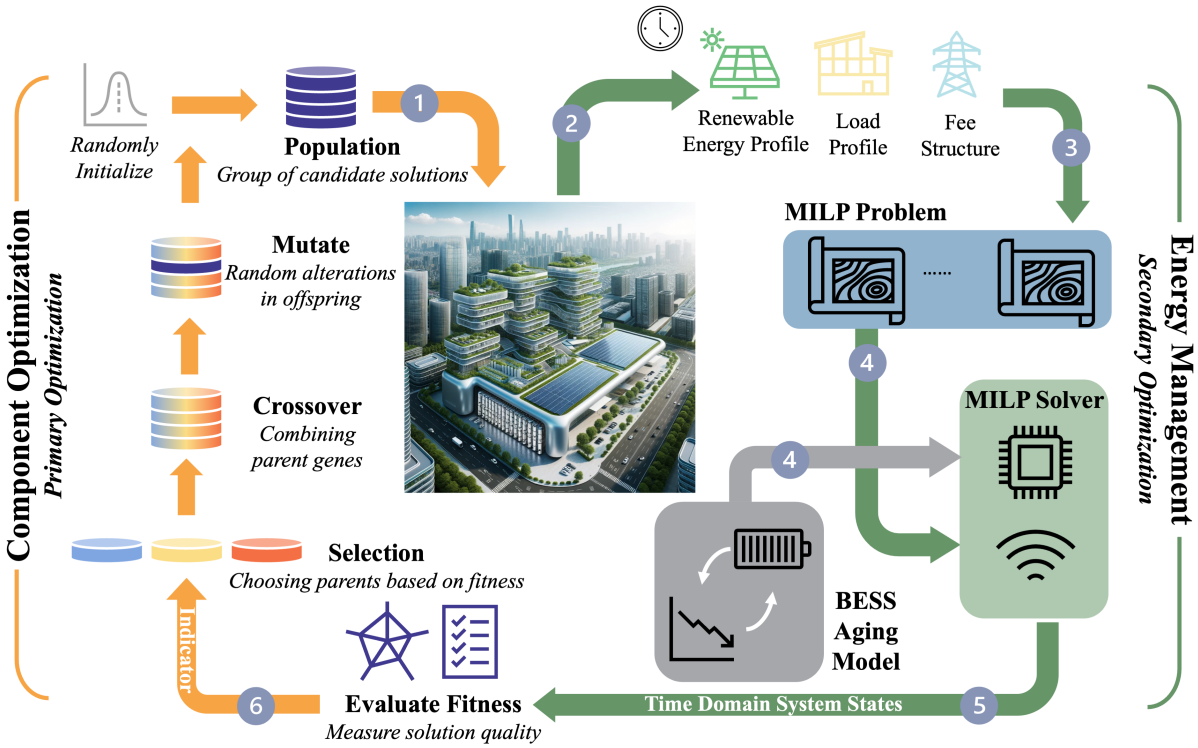


Figure 5.2: **Framework of the Proposed Two-stage Optimization Approach** Note: The center image was created with the assistance of DALL·E 2

the quality of each candidate solution through calculated indicators and is used as fitness based on simulation outcomes. In the last phase, the top-performing candidates are then selected as parents to form the basis for the next generation of candidate solutions through crossover and mutation processes. Through this iterative loop, the candidate solutions progressively evolve, converging towards the optimal solution and achieving the study's aim of efficient system capacity optimization.

Building upon the two-stage RES capacity optimization framework outlined above, the integration of sophisticated scheduling optimization within our MILP-based energy management framework plays a crucial role. By utilizing 24-hour forecasts of energy prices and demand, our system conducts load-shifting and energy arbitrage. This ensures efficient

energy use by charging the BESS during low-price periods and discharging during high-price periods or peak demand times. This strategy not only optimizes energy consumption for cost savings but also supports grid stability, exemplifying our commitment to innovative, cost-effective, and sustainable energy management solutions.

5.2.3 Performance Metrics

This study utilizes a set of financial and technical indicators to assess system performance. The financial indicators are IRR, NPV, and initial investment (Inv). The technical indicators include EIS, GEIF, UF, and SLF. IRR and EIS are chosen as the primary optimization objectives for their system-level relevance, with IRR focusing on profitability and EIS on energy independence. The other indicators, including NPV, initial investment, GEIF, UF, and SLF, are used for evaluating the performance of the optimized results, providing insights into financial viability and component-level efficiency.

5.2.3.1 Financial Indicators

The financial indicators, IRR and NPV, are discussed in Section 3.4.2. The following section provides a more detailed mathematical formulation tailored to this project.

Cash inflows include energy bill savings (R_{ESav}) and revenue from energy sold back to the grid (R_{ESol}). R_{ESav} is calculated hourly by subtracting energy imported from the grid (E_{Im}) from energy demand (E_{Load}), then multiplying by the import energy price (π_{Im}). It is summed up yearly. R_{ESol} is the yearly sum of hourly revenues, calculated by multiplying exported energy to the grid (E_{Ex}) by the export energy price (π_{Ex}). Both E_{Ex} and E_{Im} are determined hourly through the secondary optimization model. The 18-year optimization horizon is chosen in this study based on the expected payback period [135] for PV panels and their degradation rate, ensuring a realistic assessment of financial viability and operational efficiency.

$$NPV = \sum_{y=1}^{y=18} \frac{\left(R_{ESav}^{(y)} + R_{ESol}^{(y)} - C_{O\&M}^{(y)} - C_{Rep}^{(y)}\right)}{(1+r)^{y-1}} - Inv, r = IRR \text{ if } NPV = 0 \quad (5.1)$$

$$R_{ESav}^{(y)} = \sum_{t=1}^{t=8760} (E_{Load}^{(t)} - E_{Im}^{(t)}) \cdot \pi_{Im}^{(t)} \quad (5.2)$$

$$R_{ESol}^{(y)} = \sum_{t=1}^{t=8760} E_{Ex}^{(t)} \cdot \pi_{Ex}^{(t)} \quad (5.3)$$

The outflow cash consists of the operation and maintenance costs ($C_{O\&M}$) and replacement costs (C_{Rep}). The $C_{O\&M}$ shown in Equation 5.4 includes three parts: the operation and maintenance costs of PV, BESS, and other components of the system such as inverters and converters. Therefore, Equation 5.4 to Equation 5.6 shows the details of how to calculate those costs. The calculation approaches used in Equation 5.5 to Equation 5.7 are identical for those three equations: the $C_{O\&M}$ is a proportion ($\omega_{O\&M}$) of the initial capital cost of the relative component (capacity of component Cap multiplied by unit capacity cost π_{Cap}). While the replacement costs (C_{Rep}) in Equation 5.8, which account for the cost of replacing the battery, are adjusted by the battery price and discounted at the battery discount rate (r_B) based on historical data.

$$C_{O\&M}^{(y)} = C_{O\&M,PV}^{(y)} + C_{O\&M,BESS}^{(y)} + C_{O\&M,Others}^{(y)} \quad (5.4)$$

$$C_{O\&M,PV}^{(y)} = Cap_{PV} \cdot \pi_{Cap,PV} \cdot \omega_{O\&M,PV} \quad (5.5)$$

$$C_{O\&M,BESS}^{(y)} = Cap_{BESS} \cdot \pi_{Cap,BESS} \cdot \omega_{O\&M,BESS} \quad (5.6)$$

$$C_{O\&M,Others}^{(y)} = Cap_{Others} \cdot \pi_{Cap,Others} \cdot \omega_{O\&M,Others} \quad (5.7)$$

$$C_{Rep}^{(y)} = Cap_{BESS} \times \frac{\pi_{Cap,BESS}}{(1 + r_{BESS})^{y-1}} \tag{5.8}$$

The initial investment of the system includes the investment in PV, BESS, and others relatively, calculated in Equation 5.9 by multiplying the capacity of each component with the relative unit price.

$$\begin{aligned} Inv = Cap_{PV} \times \pi_{Cap,PV} + Cap_{BESS} \times \pi_{Cap,BESS} \\ + Cap_{Others} \cdot \pi_{Cap,Others} \end{aligned} \tag{5.9}$$

5.3 Simulation Procedure

This section gives the detail of capacity optimization and the scheduling energy management, offering a comprehensive insight into each stage. Additionally, it explores the interrelationship between the two stages of optimization, elucidating how variables interplay and influence each other throughout the optimization process.

5.3.1 First Stage Optimization

Primary optimization is formulated as a multi-objective optimization problem, integrated with an energy management algorithm, to ensure efficient utilization of the system's component capacities. Defined by Equation 5.10, the primary optimization is formulated to determine the optimal sizes for PV arrays and BESS. The objective function aims to maximize two metrics: IRR and EIS. These objective functions are subject to various equality and inequality constraints.

The first two constraints in Equation 5.10 indicate that both IRR and EIS are functions of decision variables: capacities of PV (Cap_{PV}) and BESS (Cap_{BESS}). The third equation

sets user-specific budget constraints for initial investments, while the fourth equation defines thresholds on the Cap_{PV} user energy consumption profile. These constraint values, as shown in Table 5.1 , are estimated based on the past user's load profile. Finally, the last equation indicates Cap_{BESS} exceeding Cap_{PV} is not cost-effective due to diminishing returns, as supported by [136].

$$
\begin{aligned}
\max \quad & IRR, EIS \\
\text{s.t.} \quad & IRR = f_1(Cap_{PV}, Cap_{BESS}) \\
& EIS = f_2(Cap_{PV}, Cap_{BESS}) \\
& Inv \leq Inv_{Req} \\
& Cap_{PV} \leq Cap_{PV\_max} \\
& Cap_{BESS} \leq Cap_{PV}
\end{aligned}
\tag{5.10}
$$

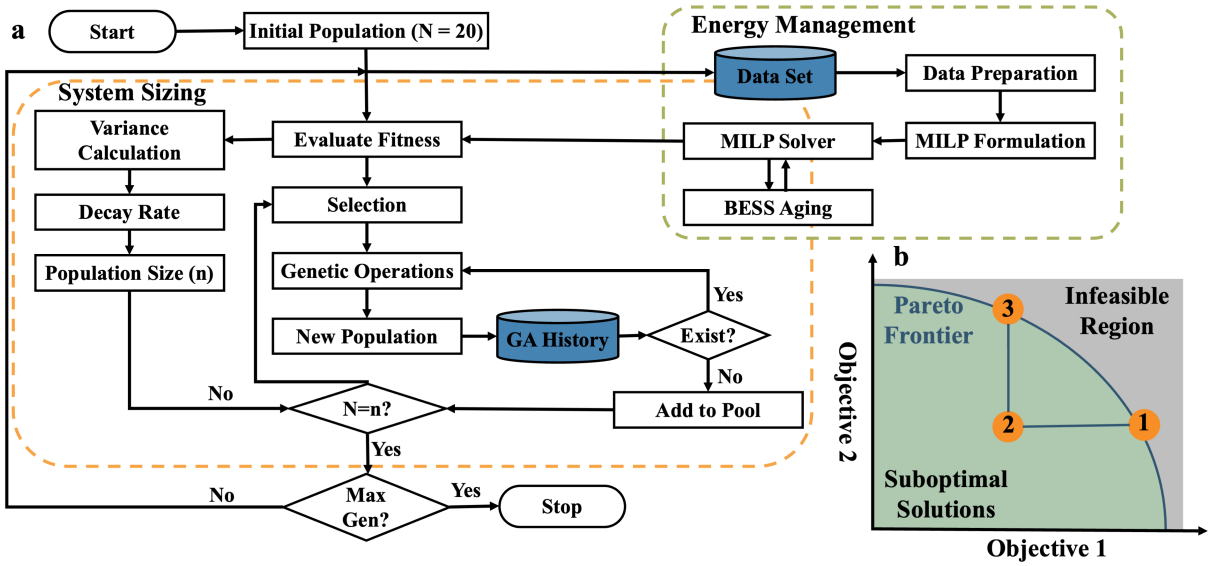


Figure 5.3: **Algorithm** (a) Algorithm Flowchart of the Proposed Two-stage Optimization Approach (b) Concept of Pareto Frontier in Multi-objective Optimization

In the two-stage optimization framework detailed earlier, this study employs a NSGA-II

for the primary optimization task demonstrated in Figure 5.3 (a). The process begins with the initialization of the population, a critical step where individuals—each representing a potential solution—are generated randomly following a normal distribution. This method of initialization ensures a diverse and representative starting point for the energy management system and facilitates the comprehensive evaluation of fitness for each candidate solution.

Table 5.1: **Parameter Value**

Variable Name	**Value**	**Unit**	**Source**
$\pi_{Cap,PV}$	1300	kW/USD	From literature [137]
$\pi_{Cap,BESS}$	150	kWh/USD	Bloomberg [138]
$\omega_{O\&M,BESS}$	0.5	%	From literature [136]
$\omega_{O\&M,Others}$	1	%	From literature [136]
$\omega_{O\&M,PV}$	1	%	From literature [136]
r_{BESS}	4	%	Bloomberg [138]
r	2	%	Federal Reserve [139]
Cap_{PV_max}	1000	kW	-
M	100K	-	-
USD/EUR	1.07	-	Exchange Rate
Crossover rate	0.7	-	Experiment
Population	20	-	Experiment
Mutation rate	0.1	-	Experiment
Generation	12	-	Experiment

During the optimization process, populations are selected for multi-objective evolution through fitness calculations based on Pareto efficiency and diversity. This selection process is driven by a crossover operation that averages the capacities of parent solutions, promoting the generation of balanced offspring and facilitating the effective evolution of the population towards optimal solutions.

Concurrently, the DPD mechanism operates alongside genetic operations. The DPD algorithm assesses the variance in fitness scores across the population to dynamically adjust

the population size. By evaluating the distribution of fitness scores, the DPD mechanism determines the need for population size adjustments—reducing the population size when a lower variance is observed. This indicates that the population is converging towards similar fitness values, prompting the DPD to decrease the population size to focus the search on the most promising solutions. This strategy of dynamically adjusting the population size, based on the fitness variance, facilitates an optimal balance between the exploration of new solutions and the exploitation of existing ones, continuously refining the solutions until the maximum generation is reached. The 'NSGA-II History' is instrumental throughout this iterative process, providing ongoing insights that guide the exploration towards achieving the study's objective of efficient system capacity optimization and management. This comprehensive approach, integrating multi-objective selection with the DPD mechanism and a thoughtfully randomized initial population, ensures a thorough exploration of potential solutions in alignment with the study's goal.

The final stage of the multi-objective optimization is identifying optimal solutions with the Pareto frontier, demonstrated in Figure 5.3 (b) and detailed in [140]. This frontier includes optimal solutions where each point is a Pareto optimal. The diagram categorizes solutions into "Sub-optimal Solutions" (green region), the "Pareto Frontier", and the "Infeasible Region". Points 1 and 3, highlighted in orange, illustrate the concept of Pareto efficiency. Even though fitness of point 2 equals points 1 and 3 in one objective, it is excluded from the frontier because it is inferior on another objective.

5.3.2 Second Stage Optimization

The secondary optimization aims to calculates optimal time domain variables, supports the calculation of IRR and EIS and provides an effective energy management strategy for the RES simulation. As the secondary optimization in this study, its primary objective is to minimize the energy cost of the system. To achieve this, the secondary optimization algorithm for energy management has been formulated as a single-objective optimization problem which based on Pena-Bello et al.'s work [141]. The secondary optimization is defined by a series of mathematical equations, ranging from 5.11 to 5.14. The objective

function is subject to constraints in four categories: basic (Equation 5.12), energy balance (Equation 5.13), loss balance (Equation 5.14), and power balance (Equation 5.15), ensuring optimal time-domain variable calculation.

Equation 5.11 defines the objective function, which aims to minimize the electricity bill. This is achieved through a balancing act between importing and exporting energy. Specifically, when energy is imported ($E_{Im}^{(t)}$) from the grid at import cost ($\pi_{Im}^{(t)}$), it incurs a cost. Conversely, exporting surplus energy ($E_{Ex}^{(t)}$) to the grid at the export price ($\pi_{Ex}^{(t)}$) earns credit, reducing the electricity cost. Therefore, the final due of the electricity bill is computed by the difference between cost and credit. The $\pi_{Ex}^{(t)}$ used in this study is based on the wholesale price from the futures market [142]. And the $\pi_{Im}^{(t)}$ calculated as a premium added on the $\pi_{Ex}^{(t)}$ as the approach used in literature [136].

$$\min \sum_{t=0}^{t=8760} E_{Im}^{(t)} \cdot \pi_{Im}^{(t)} - E_{Ex}^{(t)} \cdot \pi_{Ex}^{(t)} \tag{5.11}$$

Equation 5.12 outlines the basic constraints, ensuring the system will not concurrently charge ($Char$) and discharge (Dis) the BESS, as well as simultaneously import and export energy. In these equations, the terms E_{BESS_max} and E_{Grid_max} represent the rated power of the BESS and the grid, respectively, while B denotes a binary integer that enforces these constraints.

$$\begin{aligned}
&-E_{BESS_max} * B_{\text{Dis}} \leq E_{\text{Dis}} \leq E_{BESS_max} * (1 - B_{\text{Char}}) \\
&-E_{BESS_max} * B_{\text{Char}} \leq E_{\text{Char}} \leq E_{BESS_max} * (1 - B_{\text{Dis}}) \\
&B_{\text{Char}} + B_{\text{Dis}} = 1 \\
&-E_{Grid_max} * B_{\text{Ex}} \leq E_{\text{Ex}} \leq E_{Grid_max} * (1 - B_{\text{Im}}) \\
&-E_{Grid_max} * B_{\text{Im}} \leq E_{\text{Im}} \leq E_{Grid_max} * (1 - B_{\text{Ex}}) \\
&B_{\text{Im}} + B_{\text{Ex}} = 1
\end{aligned} \tag{5.12}$$

Equation 5.13 represents energy balance constraints. These equations ensure the energy balance across system elements such as PV, BESS, load, and grid while considering the

energy loss. Additionally, they restrict the battery's SOC from exceeding its aged capacity (Cap_{BESS_aged}).

$$
\begin{aligned}
&E_{PV} = E_{PV\_Load} + E_{PV\_BESS} + E_{PV\_Grid} + E_{PV\_Curt} + E_{loss\_Inverter\_PV} \\
&E_{Load} = E_{PV\_Load} + E_{BESS\_Load} \cdot \eta_{Inverter\_BESS} + E_{Grid\_Load} \\
&E_{Im} = E_{Grid\_BESS} + E_{Grid\_Load} \\
&E_{Ex} = E_{PV\_Grid} + E_{BESS\_Grid} \cdot \eta_{Inverter\_BESS} \\
&E_{Char} = E_{PV\_BESS} + E_{Grid\_BESS} - E_{Loss\_Converter\_BESS} \\
&E_{Char} \leq SOC_{max} \cdot Cap_{BESS} - SOC_{t-1} \cdot Cap_{BESS} \\
&E_{Dis} = E_{BESS\_Load} + E_{BESS\_Grid} \\
&E_{Dis} \leq SOC_{t-1} \cdot Cap_{BESS} - SOC_{min} \cdot Cap_{BESS} \\
&SOC_t \cdot Cap_{BESS\_aged} = SOC_{t-1} \cdot Cap_{BESS\_aged} + E_{Char} - E_{Dis} - E_{Loss\_BESS}
\end{aligned}
\tag{5.13}
$$

Equation 5.14 introduces the loss balance constraints. These equivalent constraints are employed to compute energy losses specifically in the inverter, converter, and BESS.

$$
\begin{aligned}
&E_{Loss\_Converter\_BESS} = (E_{PV\_BESS} + E_{Grid\_BESS}) \cdot (1 - \eta_{Converter\_BESS}) \\
&E_{loss\_Inverter\_BESS} = (E_{BESS\_Load} + E_{BESS\_Grid}) \cdot (1 - \eta_{Inverter\_BESS}) \\
&E_{loss\_Inverter\_PV} = (E_{PV\_Load} + E_{PV\_BESS} + E_{PV\_Grid}) \cdot \frac{1 - \eta_{Inverter\_PV}}{\eta_{Inverter\_PV}} \\
&E_{Loss\_BESS} = (E_{Grid\_BESS} + E_{PV\_BESS}) \cdot \eta_{Converter\_BESS} \cdot (1 - \eta_{BESS})
\end{aligned}
\tag{5.14}
$$

Equation 5.15 demonstrates the power balance constraints. These constraints guarantee that the power remains within the physical limits of the system components.

$$
\begin{aligned}
&P_{\text{Converter\_BESS}} \cdot dt \geq E_{PV\_BESS} + E_{Grid\_BESS} \\
&P_{\text{Inverter\_BESS}} \cdot dt \geq E_{BESS\_Load} + E_{BESS\_Grid} \\
&P_{\text{Inverter\_PV}} \cdot dt \geq \frac{E_{PV\_Load} + E_{PV\_BESS} + E_{PV\_Grid}}{\eta_{Inverter\_PV}}
\end{aligned}
\tag{5.15}
$$

In addition to the MILP, the energy management algorithm integrates a BESS aging model following methodologies from Holger et al. [129] and Angenendt et al. [128]. This model updates BESS's state based on each time step's optimization outcomes, assessing degradation by comparing calendar and cycle life and calculating SOH. BESS replacement is scheduled in the RES when SOH falls below 70%, accounting for replacement costs and affecting BESS's capacity and performance in charging and discharging.

The Aging Factor $F_{BESS,age}$ of the BESS is calculated as described by Equation 5.16, by taking the maximum between the cycle aging factor and the calendar aging factor. The cycle aging factor is represented as the inverse of the exponential function of the logarithm of the DOD divided by -2.168. This expression accounts for cycle aging based on the DOD. The calendar aging factor is the inverse of the product of the battery's calendar life $L_{BESS,cal}$, 24 hours, and 365 days, which computes the calendar aging considering the total operational hours in a year. The use of the maximum function ensures that the most significant aging effect, whether cycle or calendar, is considered in the model.

$$F_{BESS,age} = max(\frac{1}{\frac{e^{log(DOD)}}{-2.168}}, \frac{1}{L_{BESS,cal} \cdot 24 \cdot 365}) \tag{5.16}$$

The aged capacity of the battery is computed by reducing the original capacity Cap_{BESS} based on the $F_{BESS,age}$ and the SOH, as shown in Equation 5.17.

$$Cap_{\text{BESS-aged}} = Cap_{BESS} \cdot (1 - 0.3 \cdot (1 - SOH + F_{BESS,age})) \tag{5.17}$$

Finally, the SOH is updated by Equation 5.18 based on $F_{BESS,age}$ and $Cap_{\text{BESS,aged}}$, ensuring that the model accurately reflects the current health and operational efficiency of the BESS.

$$SOH = \frac{1}{0.3} \cdot \frac{Cap_{\text{BESS-aged}}}{Cap_{BESS}} - \frac{7}{3} \tag{5.18}$$

5.3.3 Interrelation of Decision Variables and Objectives

The IRR and EIS serve as financial and technical indicators, respectively, as mentioned in the primary optimization section. The decision variables of optimization are closely related to these indicators, with their complex relationships demonstrated in Figure 5.4. The Cap_{PV} determines the volume of renewable energy E_{PV} available, whereas the Cap_{BESS} sets the upper boundary for the system's charging and discharging capabilities, influencing both E_{Char} and E_{Dis}. These capacities, in turn, have chain effects on subsequent parameters within the energy management phase, ultimately impacting the aggregate energy imported to and exported from the grid. Subsequently, these parameters have a direct influence on the technical indicator, EIS, and an indirect effect on the financial indicator, IRR, through the cost structures and revenue streams generated by the corresponding system.

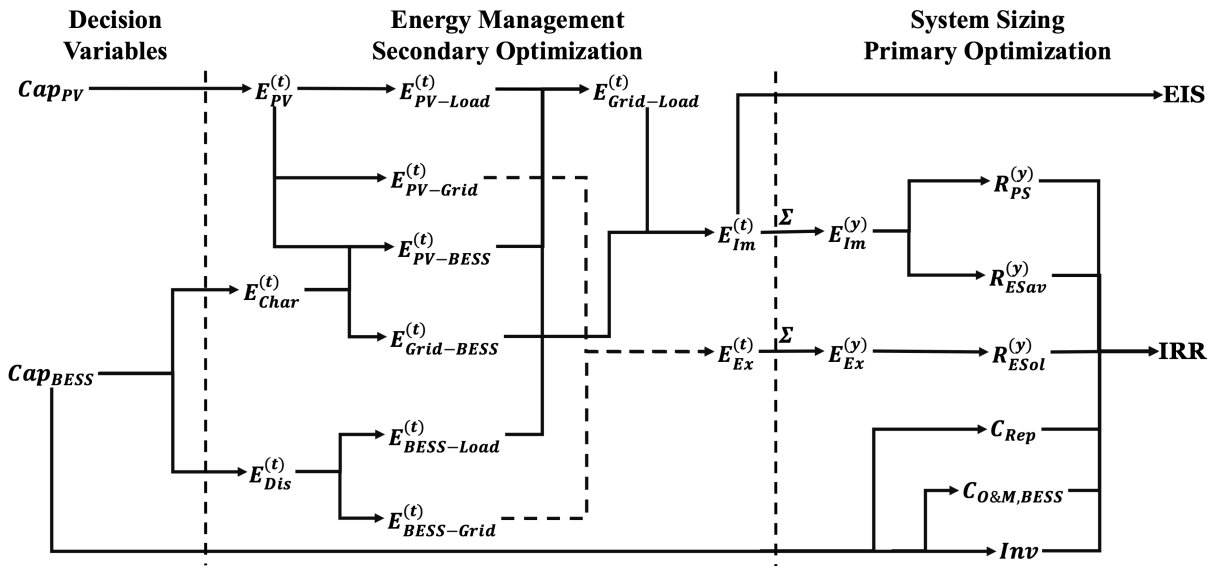


Figure 5.4: **Relation of Variables and Optimization Objectives**

5.4 Result and Discussion

To validate the proposed system and methodology, a case study has been conducted on a commercial building in the Netherlands, utilizing real-world data rather than simulated inputs. The data used in this study has two sets: the energy consumption profile of a commercial building located in the Netherlands and the corresponding meteorological profile. The energy consumption profile is crucial and will be directly used for optimization as the user's energy consumption after enhancement. At the same time, meteorological data is employed to estimate renewable energy outputs. The existing literature has collected and validated the data [143].

Figure 5.5 illustrates the energy demand profile of a commercial building in the Netherlands, essential for system capacity optimization and energy management simulations in this project. Figure 5.5 (a) shows monthly energy consumption by equipment, with four main systems: Air Handling Unit (AHU), chiller, humidifier, and lights and others. AHU and lighting and others demands are consistent year-round, while chillers and humidifiers show seasonal patterns—higher chiller use in summer and increased humidifier use in winter due to temperature-related humidity changes. Figure 5.5 (b) displays annual energy demand proportions, with lighting and others being the largest at 47.9%, followed by AHU at 38.1%. Chillers and humidifiers contribute less, at 8.1% and 5.8% respectively. Figures 5.5 (c) and (d) indicate higher usage on workdays, peaking during daytime hours and decreasing at night. Lastly, Figure 5.5 (e) provides a broader view of energy trends.

Figure 5.6 offers a comprehensive insight into the meteorological conditions crucial for renewable energy sources in the nearby weather station of the building. Figure 5.6 (a) shows the monthly solar energy profile with Global Horizontal Irradiation (GHI), Direct Normal Irradiation (DNI), and Diffuse Horizontal Irradiation (DIF), highlighting increased solar energy from April to August. The data suggests clearer days from May to July. Box plots in Figures 5.6 (b)–(g) display monthly variations in GHI, DNI, DIF, wind velocity, temperature, and atmospheric pressure. Wind velocity is consistent year-round, but solar energy is more dominant near the building. Temperatures are within a proper range for PV panel efficiency [144], and atmospheric pressure remains steady throughout the year.

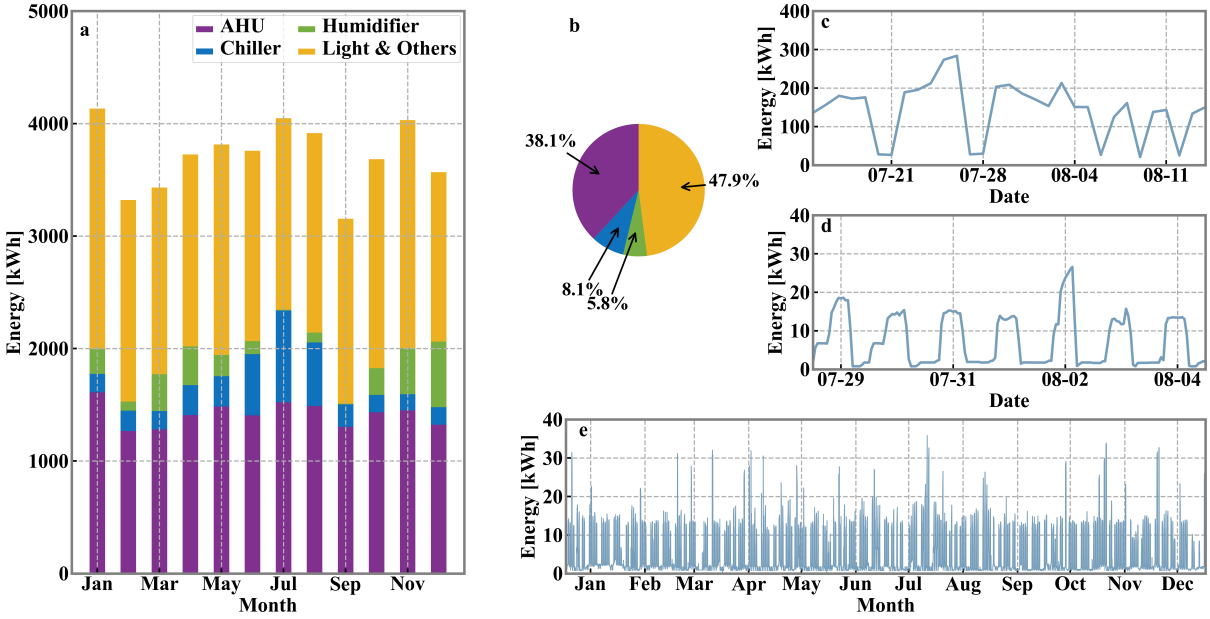


Figure 5.5: **Energy Demand Profiles Based on the Original Dataset** (a) Monthly Energy Demand Profile: Detailed Consumption Analysis by Equipment for the Building (b) Annual Energy Demand Profile: Proportional Analysis by Equipment (c) Daily Energy Demand Profile: Zoomed-in View (d) Hourly Energy Demand Profile: Zoomed-in View (e) Annual Hourly Energy Demand Profile Note: The simulation assumed the existence of five identical buildings and added white noise. As a result, the energy demand profile used in the simulation is approximately five times the energy represented in this figure.

In the data preprocessing stage, the primary goal was to improve the accuracy and credibility of the dataset. To achieve this, all missing data were filled using an interpolation approach. Additional refinement procedures included identifying and removing outliers, defined as values beyond the 99.999th percentile, to maintain data integrity without compromising its continuity. Finally, the original data, logged on a minute scale, was resampled to an hourly scale. This adjustment was crucial to aligning the data's resolution with the optimization time horizon.

As mentioned in the framework section, the target system was expanded and scaled

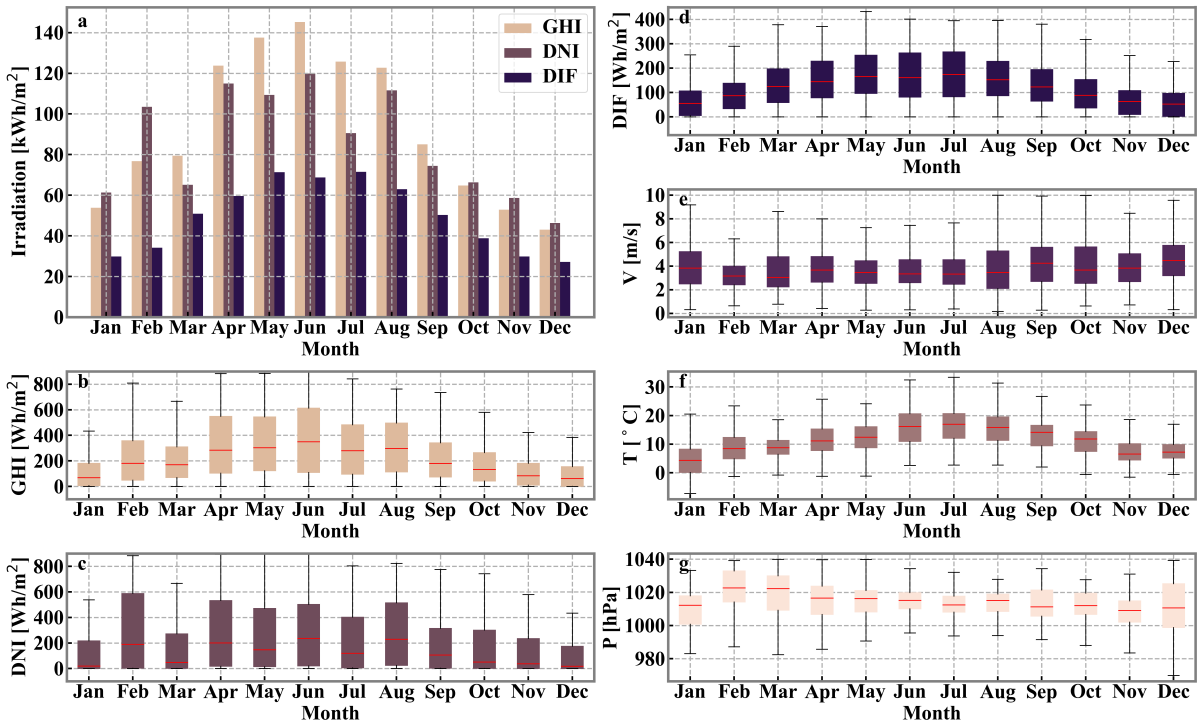


Figure 5.6: **Meteorological Resource Overview from the Original Dataset** (a) Monthly Solar Energy Profile: Breakdown of GHI, DNI, and DIF (daytime data only) (b)–(g) Monthly Distributions Derived from Hourly Measurements of: (b) Global Horizontal Irradiation (c) Direct Normal Irradiation (d) Diffuse Horizontal Irradiation (e) Wind Speed (f) Temperature (g) Atmospheric Pressure

up to include five similar buildings, operating under the assumption of consistent energy consumption patterns across these buildings. To simulate the variable nature of real-world energy consumption, independent and identically distributed white noise was incorporated into the load data for each building. Specifically, for each day, a random variable was generated following a standard normal distribution (N(0,1)) and added to the load values of that day. This study adopted white noise due to its uncorrelated nature and constant variance. The amplitude of the white noise was calibrated to 10% of the dataset's standard deviation to balance between introducing realistic variability and preserving the integrity of

the data patterns. By incorporating this variability, the optimization results are expected to be more robust and adaptive to changing and unpredictable real-world scenarios.

5.4.1 Capacity of Major Components

Figure 5.7 (a)–(e) offers a detailed visualization of NSGA-II's optimization of IRR over five epochs and twelve generations each. The shrinking distribution of points indicates the NSGA-II's refinement and convergence to optimal solutions. Figure 5.7 (a) highlights the epoch with the best fitness. The figure also includes blue ellipses representing the algorithm's exploration space, derived from the covariance matrix's eigenvalues and eigenvectors, illustrating search variance and direction. Figure 5.8 similarly explores the NSGA-II's optimization for EIS; the only difference lies in their respective convergence areas. These results validate the experimental design choice of five epochs and twelve generations, as significant convergence was observed after the eighth generation.

A detailed view of fitness across generations for each epoch is presented in Figure 5.9. Figure 5.9 (a) shows the objective IRR, and Figure 5.9 (b) focuses on the objective EIS. Orange lines demonstrate the fitness trends of individual epochs, with a dashed black line marking the average fitness and a solid black line highlighting the epoch with the best fitness. An upward fitness trend stabilizes after 6–8 generations, indicating the NSGA-II's improvement and convergence. Minor fluctuations in the NSGA-II's performance are caused by its mutation process, providing insights into the algorithm's exploratory behavior within the search process.

After the optimization process, the Pareto-adaptive Epsilon-dominance algorithm was used to select points from the optimization results. Given the real-world nuances where the objective values of two points might not be exactly the same, a small value, epsilon, was introduced as a tolerance. Figure 5.10 (a) visualizes this process and the Pareto frontier. Points from all past generations have been plotted, not just those from the last generation, ensuring that potential Pareto optimal solutions from earlier generations aren't omitted. Blue points show NSGA-II optimizations for IRR and orange points for EIS. Black diamonds, numbered 1–9, mark nine configurations lying on the Pareto frontier.

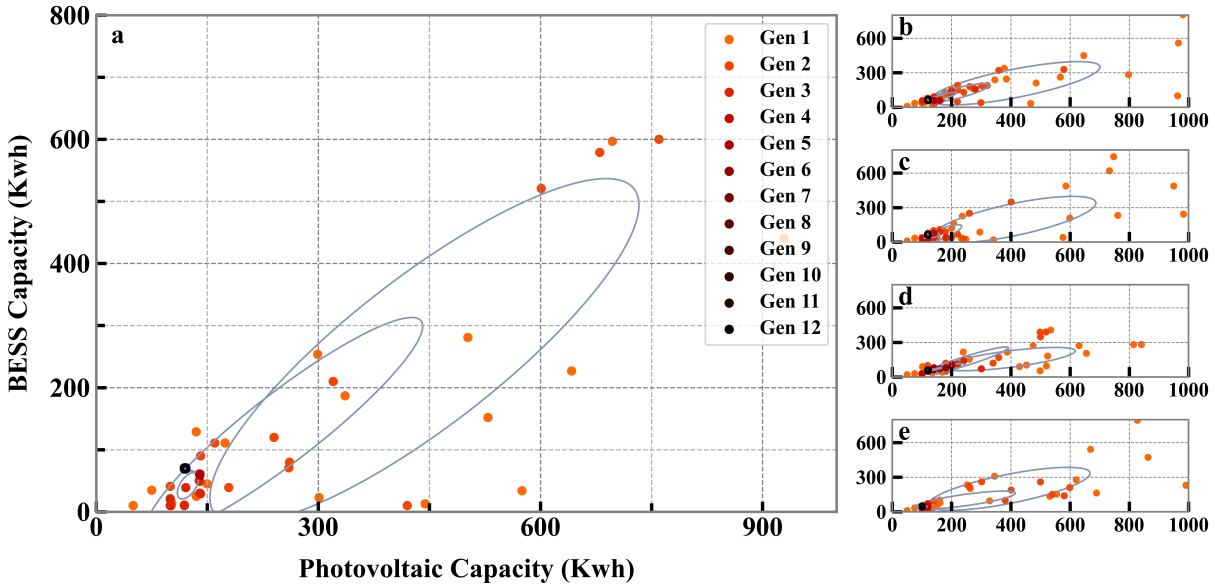


Figure 5.7: **Evolution of Genetic Algorithm Search for IRR Optimization** (a) Genetic Algorithm Search Epoch 1: Optimal IRR Achievement (b)–(e) Genetic Algorithm Search Epoch 2-5

Table 5.2: **Correlation Matrix of Indicators**

	IRR	EIS	GEIF	UF	SLF
IRR	1.00	-0.96	-0.98	1.00	-0.80
EIS	-0.96	1.00	0.91	-0.97	0.94
GEIF	-0.98	0.91	1.00	-0.98	0.72
UF	1.00	-0.97	-0.98	1.00	-0.83
SLF	-0.80	0.94	0.72	-0.83	1.00

The Pareto Frontier in RES optimization highlights a crucial trade-off between energy independence and economic returns. A strong negative correlation between IRR and technical indicators like EIS, GEIF, and SLF, as shown in Table 5.2, confirms this trade-off. The capacity and performance of those nine configurations are detailed in Table 5.3. In addition, the correlation between the financial indicator IRR and other technical indicators

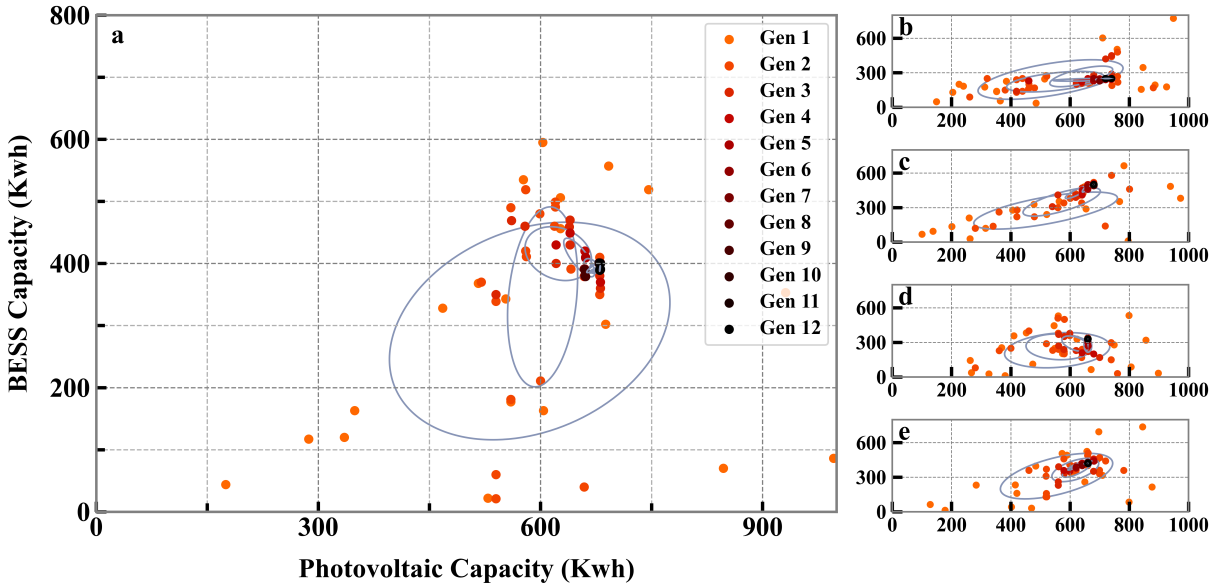


Figure 5.8: **Evolution of Genetic Algorithm Search for EIS Optimization** (a) Genetic Algorithm Search Epoch 1: Optimal EIS Achievement (b)–(e) Genetic Algorithm Search Epoch 2–5

is shown in Table 5.2.

Moreover, Figure 5.10 (b) reveals how initial investment and EIS are related. Radar plots (c)–(k) illustrate normalized performance metrics for these configurations, providing a complete view of each configuration.

5.4.2 Decision Framework

The decision framework, as demonstrated by the analysis in Figure 5.10, guides decision maker in making decisions with a structured approach for making informed choices. This approach integrates a mathematical-based decision process with the accommodation of unique preferences, laying the groundwork for a comprehensive decision-making strategy.

At the central of this framework is the mathematical-based decision process, which

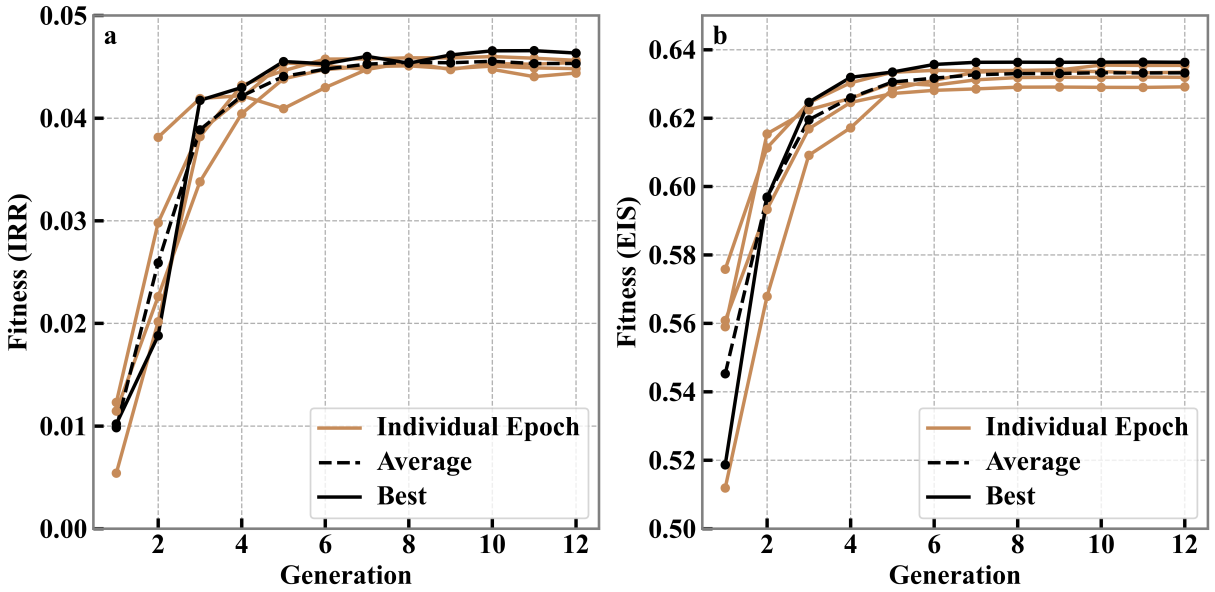


Figure 5.9: **Genetic Algorithm Convergence: Fitness Across Generations and Epochs** (a) IRR (b) EIS

encompasses regression analysis, marginal analysis, and radar plot assessments to enable objective decision-making. Initially, the application of a fourth-degree polynomial regression, illustrated in Figure 5.10 (b), models the relationship between energy independence and financial performance. Equation 5.19 provides the mathematical expression of the regression model, allowing decision-makers to estimate energy independence level of the RES based on their budgetary constraints. Further, derivative of the regression curve indicates the marginal cost of increasing the EIS per unit of initial investment, highlighting a critical threshold at an EIS of 0.4. Before the identified threshold, minimal financial sacrifices are shown to result in significant gains in energy independence. The marginal information supports optimizing the techno-economic performance within RES by balancing budget allocations against the expected environmental returns.

$$x = 43414.09y^4 - 53911.59y^3 + 25515.87y^2 - 4590.18y + 350.77 \tag{5.19}$$

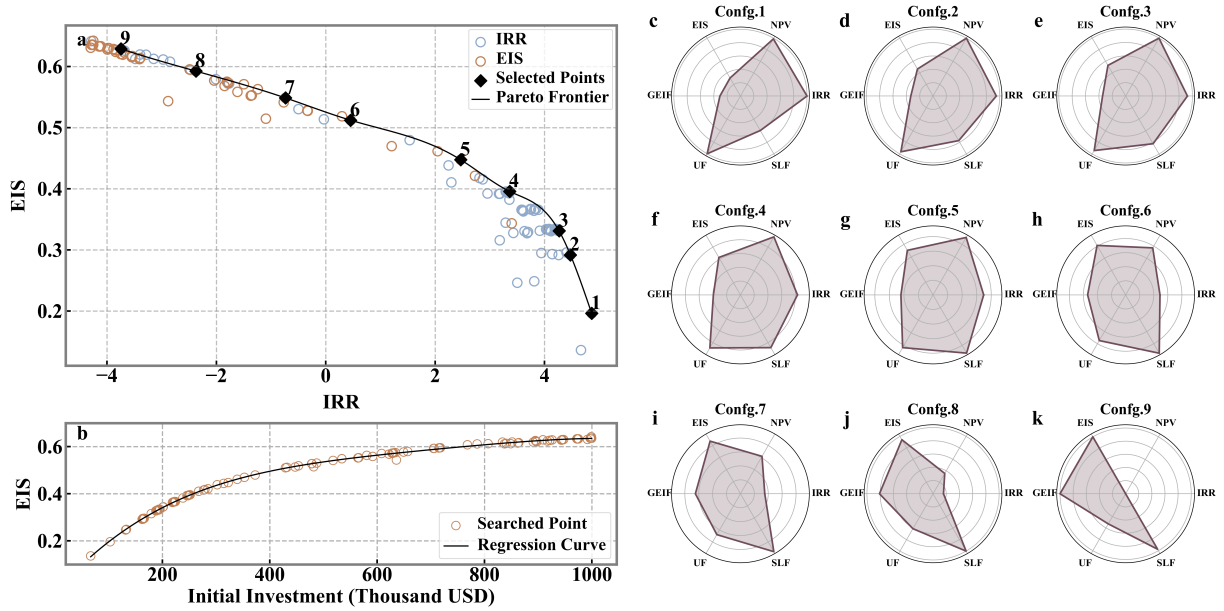


Figure 5.10: **Results of Optimization: Pareto Front and Additional Indicators** (a) Epsilon Pareto Front in Multi-objective Optimization (b) Achievable EIS by Given Initial Investment (c)–(k) Comparative Evaluation of Normalized Performance Metrics for Configurations Selected from Pareto Front, Indexed 1–9.

Moreover, the analysis along the Pareto Frontier, as shown in Figure 5.10 (a), reveals that the marginal cost of improving EIS via the IRR surges by 839%, evident to the strategic importance of identifying the optimal balance between energy independence and economic returns during decision making progress. This balance is crucial to avoid disproportionate sacrifices when increasing energy independence in RES. Furthermore, the integration of radar plot analysis into this framework provides a multifaceted view of RES performance, emphasizing systems' efficiency, cost-effectiveness, and environmental impact. Notably, the shape and balance of a radar plot provide insights into a system's all-around performance, highlighting well-rounded options that meet a broad spectrum of project goals.

In addition to providing a comprehensive performance overview, radar plot analysis uniquely supports aligning RES choices with specific preferences and unique strategic ob-

Table 5.3: **Pareto-optimal Configurations and Indicators**

Confg.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Cap_{PV}	75	120	140	180	225	310	385	515	660
Cap_{BESS}	35	60	80	110	145	175	225	245	340
IRR (%)	4.86	4.43	4.22	3.39	2.25	0.46	-0.73	-2.38	-3.75
Diff(%)	-	-8.95	-4.73	-19.55	-33.73	-79.45	-258.15	-226.09	-57.15
EIS	0.20	0.29	0.33	0.39	0.44	0.51	0.55	0.59	0.63
Diff(%)	-	22.39	18.61	7.46	5.75	3.12	3.23	2.65	2.67
GEIF	0.89	0.97	1.00	1.14	1.28	1.57	1.88	2.35	2.90
Diff(%)	-	16.70	15.87	16.98	19.35	12.31	26.25	28.31	40.40
UF	0.83	0.80	0.78	0.73	0.69	0.62	0.56	0.50	0.44
Diff(%)	-	-6.79	-9.24	-5.24	-6.51	-2.74	-4.89	-3.46	-4.84
SLF	0.18	0.23	0.25	0.27	0.28	0.29	0.29	0.30	0.30
Diff(%)	-	12.13	7.52	2.25	-0.50	1.33	0.11	0.72	-0.55

jectives. An unbalanced radar plot can highlight systems that outstanding in particular areas, meeting unique project requirements or priorities. This analysis, coupled with the overall area within the radar plot, quantifies system performance, facilitating objective comparisons. Such detailed assessments ensure that decision-makers can identify systems that not only meet broad strategic objectives and sustainability goals but also adhere to specific preferences or constraints.

This decision framework offers comprehensive support to human experts, enabling them to make informed and strategic final decisions. Through structured and integrated approach of the framework, decision-makers can achieve a balanced consideration of technical, economic, and environmental factors, ensuring that final decisions are both objective and aligned with unique strategic goals.

Given a scenario without subsidies for the RES, stakeholders require a positive return on investment and expect a balanced system performance. Configuration 4 is selected as the optimal choice for four reasons: First, Figure 5.10 (f) shows the largest area compared

to others in Figures 5.10 (c)-(k), indicating the most balanced techno-economic performance among the nine configurations. Second, configuration 4 is close to the threshold where the marginal cost of EIS significantly increases. Third, it yields a positive IRR and NPV, generating profit while maintaining a decent EIS. Fourth, the initial investment is reasonable, considering its capacity to supply half the energy demands of five commercial buildings for 18 years.

5.4.3 Energy Management

Figure 5.11 (a)-(d) demonstrated the detailed energy management simulation outcomes for configuration 4, examining the temporal dynamics of energy distribution. Figure 5.11(a) shows an hourly view of the energy supply mix from PV, BESS, and the grid, while Figure 5.11(b) offers an annual overview of the energy mix. Weekly variations in energy sourcing are described in Figure 5.11(c), and Figure 5.11(d) offers monthly variations of the energy mix.

From Figure 5.11, it is evident that the PV system serves as the primary energy supplier during the day. In the early night, the BESS plays a crucial role in supplying energy by discharging the stored energy, which is particularly cost-effective during periods of high electricity prices in the evening. On weekends, the combined capacity of these two systems is generally sufficient to meet the lower energy demands of commercial buildings. Seasonal variations also significantly influence the system's reliance, with a greater dependency on solar energy in summer due to abundant sunlight and a shift towards grid dependency in winter.

Figure 5.12 shows the RE utilization within a RES, including interactions with BESS, load, and the grid. Figure 5.12(a) shows a detailed hourly RE distribution, focusing on immediate load requirements during daylight, surplus energy storage, and grid exportation, with a noted decrease at night. On weekends, due to decreased load demand, a large proportion of RE is stored at BESS and exported to the grid. This exportation is feasible while adhering to regulatory limits on grid sell-back, balanced by the system's PV and BESS capacities. Figure 5.12(b) provides weekly data, showing successful grid exportation

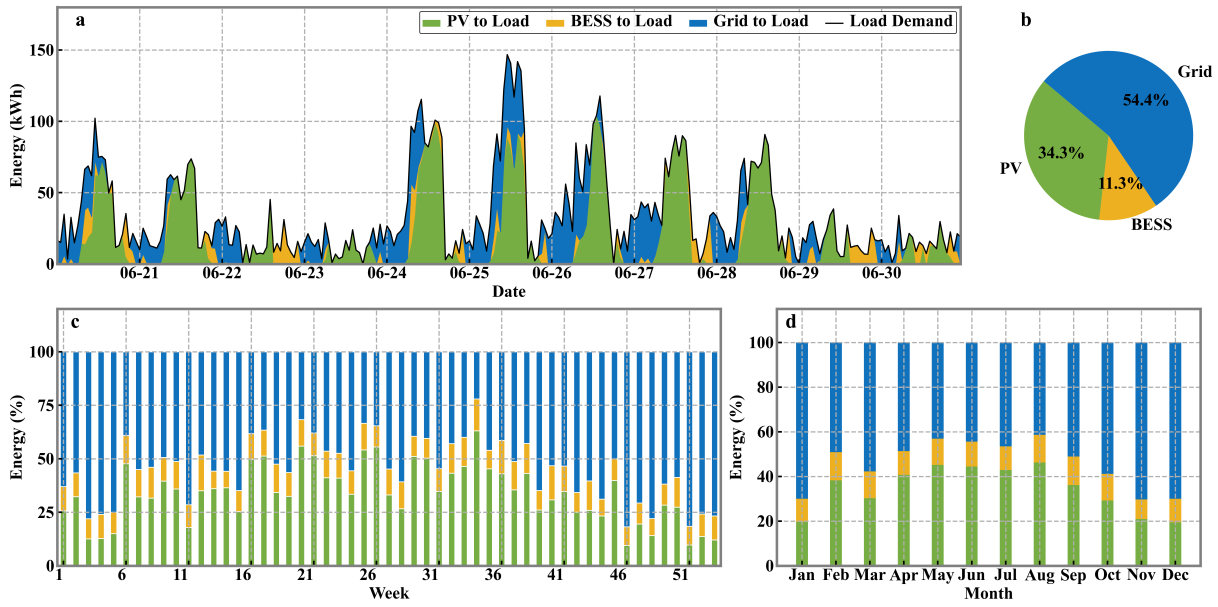


Figure 5.11: **Simulation Results for Configuration 4: Load Demand Energy Supply Breakdown** (a)Hourly Energy Supply Dynamics: PV, BESS, Grid (Zoomed View) (b)Yearly Energy Supply Percentages (c)Weekly Energy Supply Percentages (d)Monthly Energy Supply Percentages

for most of the year, while Figure 5.12(c) highlights monthly variations in energy production and grid dependency.

Observations from Figure 5.12(a) indicate that during the weekdays, the energy generated by PV panels is primarily used to meet the energy demands of buildings, followed by charging the battery system and being sold to the grid. On weekends, despite a reduction in energy demand, the generation of RE remains significant, leading to the majority of this power being sold to the grid. Figure 5.12(b) shows that the proportion of RE sold to the grid remains relatively stable, most of the time ranging between 20% and 25%. This finding underscores the critical role of the battery system as a buffer between the solar panels and the grid. However, Figure 5.12(c) shows that approximately 14% of the electricity is directed to the battery system, which differs from the previously analyzed

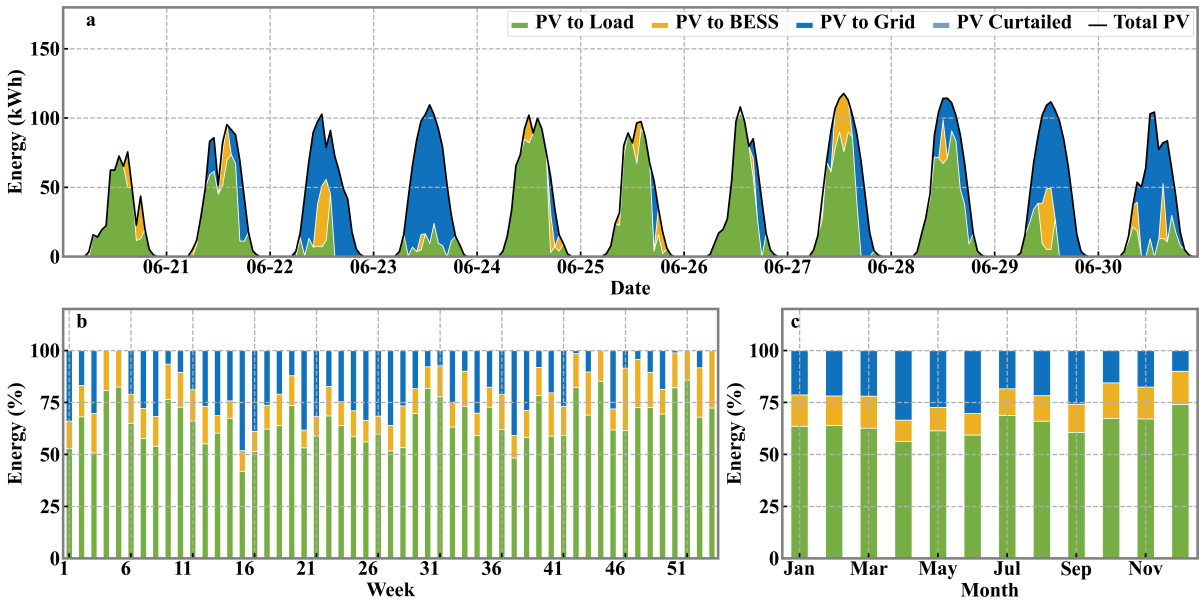


Figure 5.12: **Energy Management Simulation Results for Configurations 4: Renewable Energy Utilization** (a)Hourly Renewable Energy Utilization Dynamics: PV, BESS, Grid (Zoomed View) (b)Weekly Utilization Percentages (c)Monthly Utilization Percentages.

figure of 12%. This discrepancy is primarily due to efficiency losses incurred during the energy migration to the BESS. This observation aligns with the initial conclusion, suggesting that the generated RE is typically utilized immediately, thereby providing additional evidence to this conclusion. The preference for directly utilizing RE to avoid efficiency losses is evident, yet the decision to store or export energy to the grid depends on current and future energy prices, as well as problems related to charging BESS such as efficiency loss and degradation. Consequently, when energy prices are low, the system tends to store energy in the BESS; conversely, when prices are high, it prefers exporting the energy to the grid. This observation once again highlights the importance of optimizing PV and BESS capacity and effective energy management.

5.4.4 Comparative Analysis

The comparative analysis in Figure 5.13 indicates that expanding the capacity of the system does not proportionally enhance energy efficiency, and the primary goal of the system is self-sufficiency rather than selling electricity. This conclusion is drawn from the following observations: as the system size expands, the utilization of RE sources and dependence on the grid increase logarithmically, as shown in Figure 5.13 (a), while the proportion of RE stored in the BESS remains constant and the supply to buildings decreases linearly, as shown in Figure 5.13 (b). These observations offer another piece of evidence to support the idea that RES should prioritize supplying RE to users over profiting from the sale of surplus energy.

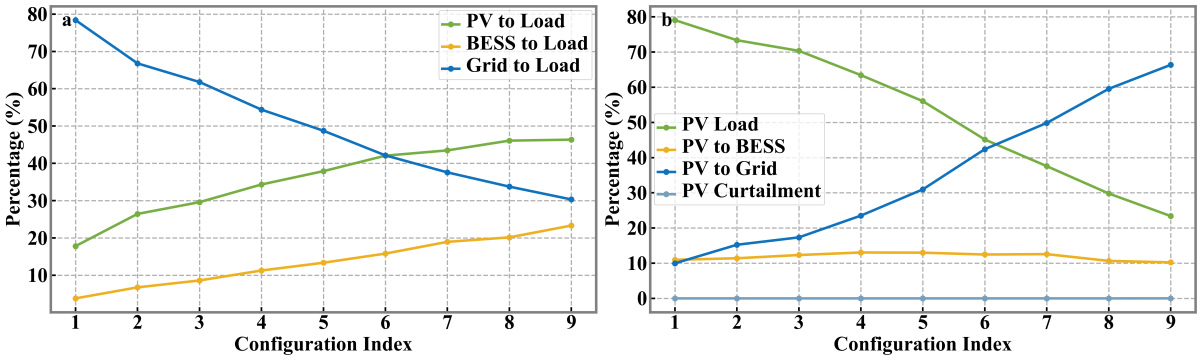


Figure 5.13: **Comparison of Technical Performance of Pareto-optimal Configurations** (a) Load Demand Energy Supply (b) RE Utilization

Figures 5.14 (a) and (b) present the average monthly utilization rate of the BESS and the SOH of the BESS, respectively. In smaller systems, where the PV capacity is limited, the BESS shows lower utilization rates, as demonstrated in Figure 5.14(a). This underutilization can be attributed to the minimal surplus energy generation, particularly under conditions such as an increased cloud cover that further reduces output. This operational characteristic is beneficial from a lifecycle perspective of the battery system, as lesser cycling of the BESS extends its SOH, as indicated in Figure 5.14(b). Conversely, larger systems expedite BESS degradation due to increased charge and discharge cycles

necessitated by higher energy outputs. Such insights emphasize the need for a balanced approach in system design, prioritizing direct energy consumption and minimal battery usage to optimize both economic and environmental benefits.

In contrast, BESSs in smaller PV systems are often underutilized. Storing and retrieving energy leads to losses, making direct RE consumption more efficient in systems with small PV capacity. Minimizing battery use in such systems extends battery life and aligns with economic considerations, especially when energy imports from the grid are affordable. This strategy highlights the need for system design optimization, balancing energy storage, direct usage, and overall efficiency.

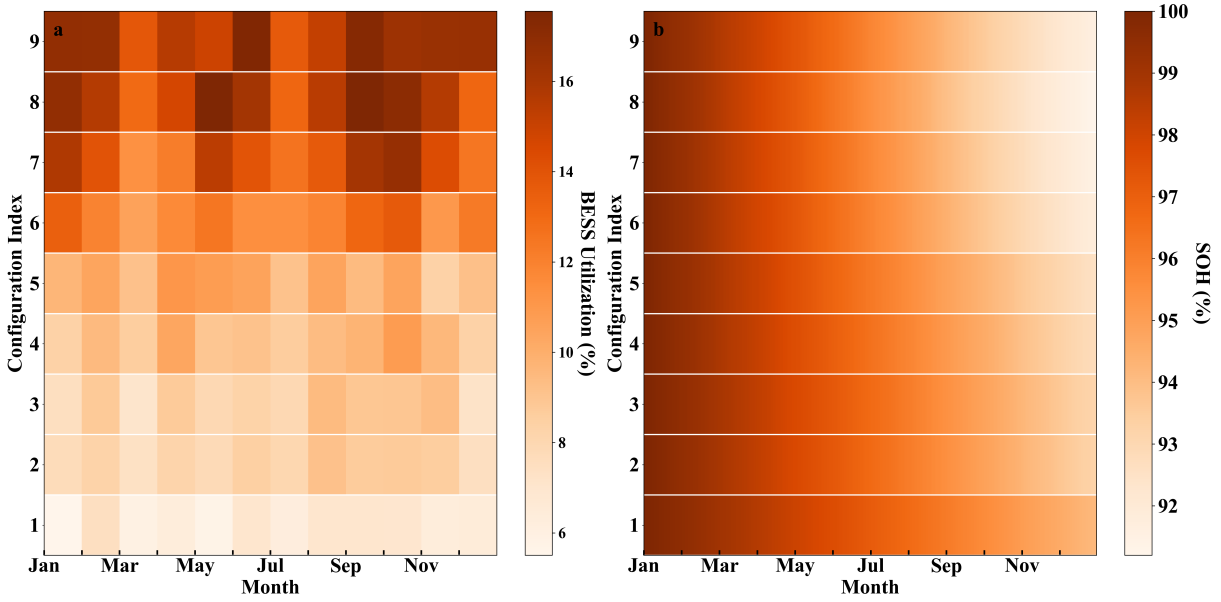


Figure 5.14: **Energy Management Simulation Results for Pareto-optimal Configurations: BESS Utilization** (a) Monthly BESS Utilization Rate for Optimized System Configuration (b) Monthly BESS SOH for Optimized System Configuration

Figure 5.15 evaluates the optimized RES's effectiveness, showing how the capacity of system components affects grid interaction and RE utilization. An increase in PV capacity changes the energy distribution among the grid, BESS, and load. Figure 5.15 (a) and (b)

display the variability of the GEIF with different system sizes. Figure 5.15(c) and (d) show that smaller systems use generated energy more efficiently. Additionally, the SLF in Figure 5.15 (e) and (f) increases with system size, suggesting larger systems benefit from a larger BESS for stable RE distribution.

Figure 5.15 (a) and (b) demonstrate the variability of the GEIF across different system configurations, underscoring that larger systems, despite their increased interaction with the grid, are less cost-effective compared to smaller systems. The primary reason for this is that in configurations 1 to 4, the majority of the generated electricity is utilized to meet the energy demands of buildings, resulting in lower GEIF values because a smaller proportion is exported to the grid. Conversely, larger systems produce more surplus electricity, which leads to greater interactions with the grid. Nevertheless, these larger configurations require substantial expansion in BESS capacity. Constructing such systems primarily for the purpose of selling excess electricity to the grid without subsidy proves to be economically infeasible. Nonetheless, these findings still support the initial conclusions that the generated RE tends to consume immediately and provide further evidence. For instance, in comparison to configuration 1, configuration 9 exhibits an 880% increase in PV capacity, yet the corresponding rise in GEIF is less than 300%. This highlights that while larger systems are capable of supplying more electricity, the negative IRR indicates they are economically unfeasible without subsidies, particularly when their primary role is to sell electricity.

Figure 5.15 (c) and (d) illustrate that smaller systems have a higher UF, indicating a tendency to directly consume electricity due to their lower RE generation. Direct consumption of RE can avoid efficiency losses typically associated with energy storage or conversion processes, thereby enhancing both energy efficiency and economic viability. Conversely, larger systems face challenges in managing surplus energy. When this surplus cannot be sold to the grid, options like curtailment or storage in the BESS bring additional costs, which consequently reduce the economic attractiveness of large-scale systems.

The SLF across various configurations, as shown in Figure 5.15 (e) and (f), demonstrates a notable increase from configuration 1 to 4, yet this trend reaches a plateau between

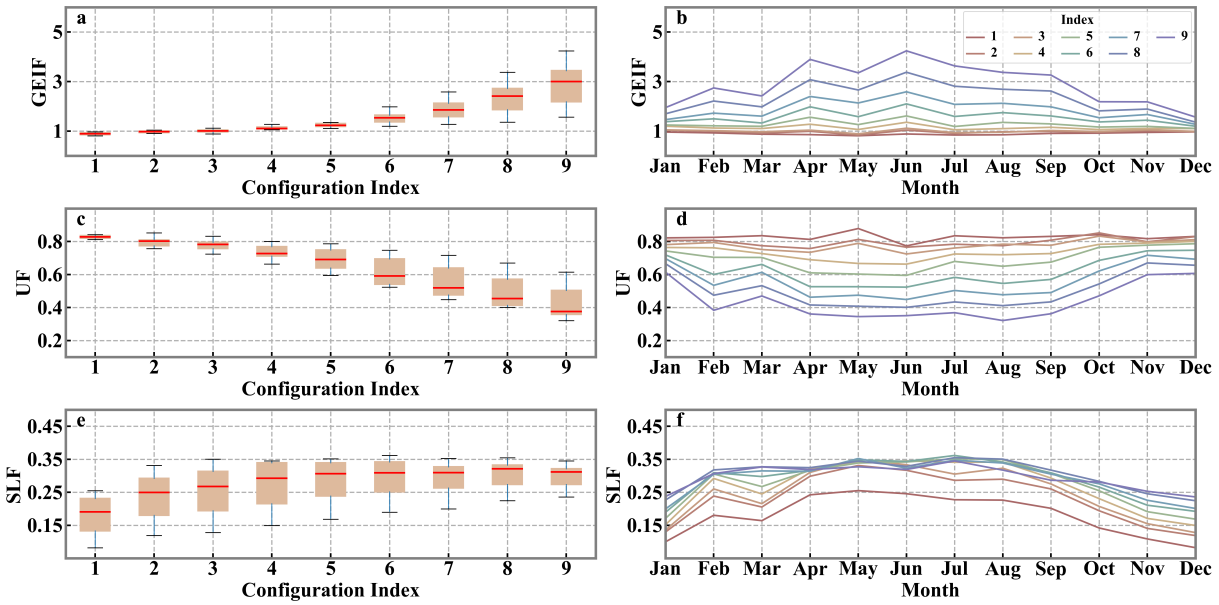


Figure 5.15: **Technical Performance Analysis of System Configurations Determined by Primary Optimization/Genetic Algorithm** (a) GEIF: Monthly Statistical Distribution (b) GEIF: Monthly Trend Analysis (c) UF: Monthly Statistical Distribution (d) UF: Monthly Trend Analysis (e) SLF: Monthly Statistical Distribution (f) SLF: Monthly Trend Analysis

configurations 5 and 9. This pattern confirms that smaller systems tend to directly meet load demands. In contrast, larger systems result in energy waste, primarily because surplus energy generation is either stored in the BESS or exported to the grid. Such practices diminish the profit of the system, underscoring the inefficiencies associated with large-scale energy production in the context of this configuration.

5.4.5 Assumptions and Limitations

This study assumes that the existing electrical grid can integrate the proposed RES without significant upgrades. An expected rise in demand for renewable energy, driven by increasing environmental consciousness and supportive policies, is anticipated to ensure

a stable market for the produced energy. Technological advancements will likely reduce costs for solar panels and batteries, making renewable energy production costs competitive with traditional fossil fuel sources. High public and stakeholder support, influenced by the recognized environmental and economic benefits, is expected to diminish social resistance to new renewable projects. The validity of applying predictive and optimization methods hinges on the stability of future user energy consumption patterns, energy prices, policies, regulations, and local meteorological conditions, which should not deviate significantly from historical trends.

The limitations of this study can be categorized into two key areas: assumptions that introduce uncertainties in our models and simplifications that may not accurately capture the dynamics of energy systems. Firstly, the assumptions about electricity prices and energy consumption patterns are based on historical trends, which are susceptible to unpredictable factors such as changes in global energy production, policy shifts, climate change, and technological advancements. These uncertainties mentioned above could lead to significant volatility in energy prices and shifts in consumption patterns. This could potentially affect the optimization results if future data do not align with the current ones. For example, an unexpected rise in energy prices might cause an underestimation of the necessary capacity of RES, while a decrease could result in overestimation.

Building on this, another notable limitation of our study arises from the modeling of the BESS charging and discharging processes, which are represented with hourly granularity. This approach simplifies the intricate dynamics of energy flow into and out of the BESS by averaging over an hour, potentially overlooking the nuances of shorter-duration fluctuations and transitions within each hour. Such simplifications may not fully capture rapid changes in energy demand or supply. This limitation could be solved by integrating another optimizer for BESS to resolve during the operation. Although this simplification introduces limitations by not fully accounting for minute variations, it's crucial to note that the primary goal of this study is not to devise a precise energy management strategy but to support capacity optimization over decades. In this extended time frame, these micro-level fluctuations become less significant. To address this limitation, integrating an additional optimizer for BESS operations could offer a more detailed resolution during op-

eration. However, given our study's focus on long-term capacity optimization, the impact of such micro-level fluctuations is deemed marginal.

Moreover, policy and regulatory frameworks are assumed to be steady, but changes here would affect the profit model and require new energy management strategies. Additionally, variables like the discount rate, based on historical data, might not align with future value due to the dynamic nature of technology, economics, and socio-political factors. Periodic reassessment of the optimization model is necessary to maintain its relevance and accuracy due to the unpredictable nature of policy and regulatory changes.

5.5 Summary

This research develops a novel methodology for optimizing the capacities of a renewable energy system consisting of photovoltaic panels and battery energy storage to achieve a critical balance between technological advancement and economic feasibility. By integrating parameters corresponding to user-specific demands, dynamic policy landscapes and volatile energy markets, this approach adopts an advanced energy management algorithm alongside stochastic modelling to enhance the optimization's accuracy, adaptability, and robustness. This study demonstrates that implementing a dynamic population decay algorithm within the non-dominated sorting genetic algorithm II framework can significantly accelerate the convergence speed (reducing the search times by 60.42 %).

A case study has been conducted on optimizing the capacities of components in a renewable energy system in the Netherlands to validate the proposed methodology. Performance analysis reveals that this optimal configuration achieves a substantial 95 % improvement in the energy independence score and a 28.1 % enhancement in the grid energy interaction factor. Meanwhile, the sustainable internal rate of return remains at 3.39 %. This configuration outperforms the reference by 216 % in the Battery Energy Storage System utilization. It is also demonstrated that a balance exists between lower capital costs and significant advancements in key performance indicators, establishing its cost-effectiveness, efficiency, and technical robustness. Compared to the rule-based non-dominated sorting

genetic algorithm, the proposed approach shows an 18.12 % improvement in the Internal Rate of Return (3.39 % vs. 2.87 %) while maintaining a similar level of energy independence (0.39 vs. 0.38). The study concludes that smaller Renewable Energy System configurations offer greater economic viability and energy efficiency through direct consumption of generated renewable energy. In contrast, larger systems encounter financial difficulties and operational inefficiencies due to the challenges of managing excess energy, often requiring subsidies and policy support. This finding underscores the importance of system capacity optimization in determining economic and environmental Renewable Energy System outcomes. This approach ensures a comprehensive system performance evaluation, aligns economic interests with sustainable energy practices, and contributes significantly to the advancement of Renewable Energy Systems. Future research will aim to identify the economic-environmental balance for Renewable Energy Systems in different locations and enhance system robustness through scenario design and pressure testing.

Chapter 6

Prediction-Based Energy Management

This chapter introduces the project for a prediction-based energy management algorithm in a RES. Section 6.1 summarizes main contributions and novelties of this project. Section 6.2 describes the target system and presents the proposed algorithm for predictive energy management. The algorithm employs an advanced machine learning algorithm for energy demand forecasting and optimizing energy distribution within the system. Section 6.3 outlines the simulation procedure, focusing on energy management optimization to ensure efficient operation and resource allocation. The results are discussed in Section 6.4, with an emphasis on machine learning techniques for energy demand prediction and the implementation of energy management strategies. This section also includes a comparative analysis of the system's performance under different configurations and discusses the assumptions and limitations that affect the results. Finally, Section 6.5 provides a summary of the key insights gained from the prediction-based energy management process.

6.1 Contribution

The main contributions of this project are as follows:

- The modelling of energy storage systems is enhanced by incorporating machine learning techniques, setting the stage for precise system simulations. Features are extracted from accurate nonlinear models, improving modelling accuracy without sacrificing computational efficiency.
- An advanced machine learning model, transformer, is employed to predict the highly volatile energy consumption profiles of users. Simulations are designed to reflect realistic operational scenarios, enhancing the reliability of the predictions and optimization.
- A novel optimization approach developed and modified based on the concept of the model predictive control. This model addresses the issue of decreased accuracy over long prediction horizons, striking a balance between optimization accuracy and strategic foresight.

6.2 Target System and Proposed Algorithm

This section introduces the target system and the proposed algorithm for predictive-based energy management. The target system focuses on optimizing the operation of the RES, ensuring efficient energy generation and storage. The predictive-based energy management algorithm, discussed in detail, uses advanced ML techniques to forecast energy demand and adjust system operations accordingly. By integrating predictive analytics into energy management, the system aims to enhance energy management by enabling data-driven adjustments that improve efficiency and operational performance without increasing costs or compromising reliability.

6.2.1 Target System

Figure 6.1 demonstrates the target RE, consisting of renewable energy, electrical energy, and hydrogen energy circuits. PV panels and WTs generate RE, which flows into the

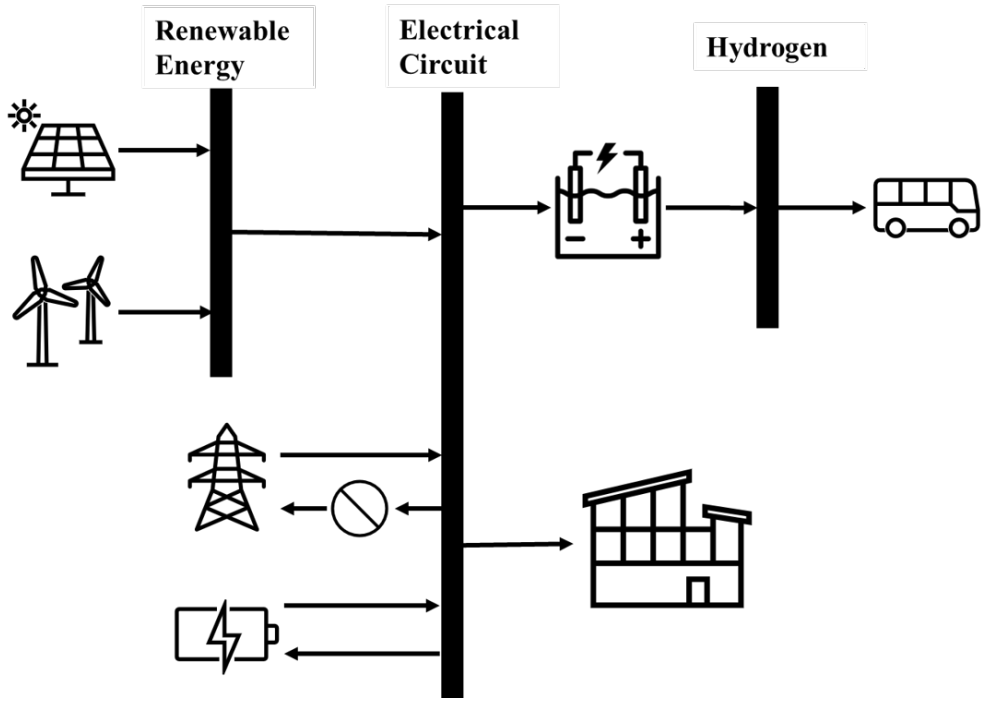


Figure 6.1: **Components Overview of Target System**

electrical circuit to meet immediate load demands, store surplus energy to ESSs, and, if permitted, sell excess energy back to the grid. When there is an energy deficit, the system draws power from the BESS and, if necessary, the electrical grid. The HESS serves two roles: it absorbs surplus electrical energy to produce hydrogen through water electrolysis, and it supplies hydrogen to the hydrogen circuit to meet the demands of the hydrogen bus. As such, the HESS serves as an ESS, and a source for fulfilling specific hydrogen energy requirements, integrating both storage and demand within the system.

Schematic in Figure 6.2 illustrates the detailed energy flow within the system, including four sections: RE, BESS, HESS, and the Grid, which are color-coded in green, orange, blue, and red, respectively. Each section features crucial variables and interactions that define the system's functionality.

The RE section utilizes solar and wind energy, where the PV systems convert solar en-

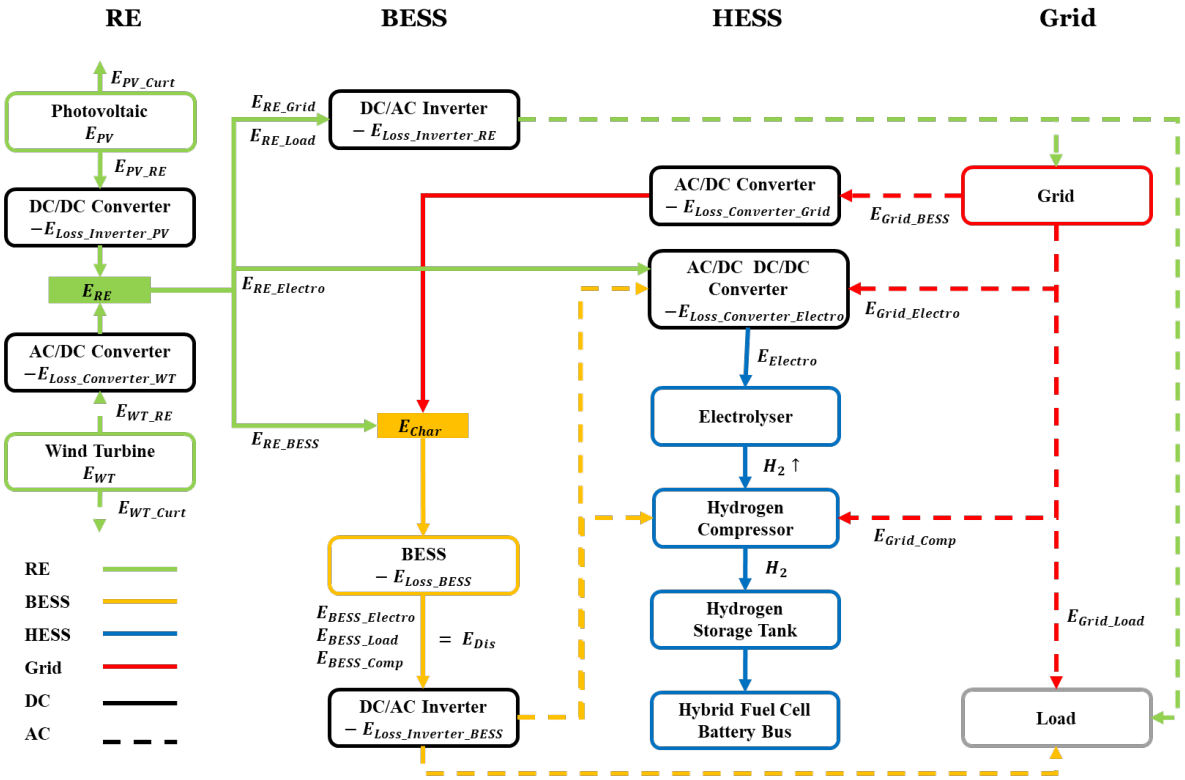


Figure 6.2: **Schematic of Target System with Key Variables**

ergy to electricity (E_{PV}). E_{PV} can either be used directly (E_{PV_RE}) after conversion losses in a DC/AC inverter ($E_{Loss\_Inverter\_PV}$), or curtailed (E_{PV_Curt}). Wind turbines generate energy (E_{WT}), similarly subject to use (E_{WT_RE}) or curtailment (E_{WT_Curt}) with conversion losses through an AC/AC converter ($E_{Loss\_Converter\_WT}$). The total renewable energy produced (E_{RE}) supports various demands including grid export (E_{RE_Grid}), building load (E_{RE_Load}), energy storage in hydrogen ($E_{RE\_Electro}$), and battery storage (E_{RE_BESS}).

The BESS section manages surplus RE and grid energy for storage (E_{RE_BESS}, E_{Grid_BESS}) and utilizes it for the electrolyzer and compressor in the HESS ($E_{BESS\_Electro}$, E_{BESS_Comp}) as well as meeting building energy needs (E_{BESS_Load}). Energy conversions here involve inherent losses (E_{Loss_BESS}), and the conversion to the necessary AC voltage occurs via a

DC/AC inverter ($E_{Loss\_Inverter\_BESS}$).

In the HESS section, the electrolyzer uses energy ($E_{Electro}$) to split water into hydrogen and oxygen, with the produced hydrogen being compressed (E_{BESS_Comp}, E_{Grid_Comp}) to the required pressure for bus fueling. This section highlights the dual energy sourcing for the compressor, enhancing reliability, especially in blackout conditions.

Finally, the Grid section plays a critical role in supplementing energy shortages within the system. It interacts with the BESS for energy storage and retrieval (E_{Grid_BESS}) and provides the necessary power for the HESS operations ($E_{Grid\_Electro}$, E_{Grid_Comp}). Additionally, it caters to direct energy consumption by buildings (E_{Grid_Load}), ensuring consistent energy availability across the system.

6.2.2 The Predictive-based Energy Management Algorithm

Figure 6.3 (a) depicts the architecture of the proposed energy management algorithm. The center of the energy management algorithm is a MILP optimizer, which has two input sources: future data and system states. The future data component includes supply data and demand data, while the system states encompass the current statuses of system components, such as the battery's SOC and the HESS's available hydrogen.

The input data originates from a data stream indicated in the green box of Figure 6.3 (a). The data's dimension is the input dimension (D_{in}) by prediction horizon (H_{Pred}). The prediction horizon is determined by a neural network output. However, due to the inverse relationship between prediction accuracy and horizon length, only the initial portion, referred to as the optimization horizon (H_{Opt}), is selected for use, which is then fed into the optimizer.

The output of the optimizer is an energy management strategy, with a dimension of output dimension (D_{out}) by prediction horizon (H_{Pred}). This strategy is designed to take into account the relationship between supply and demand over a defined optimization period. However, only a segment of this output, corresponding to the control horizon

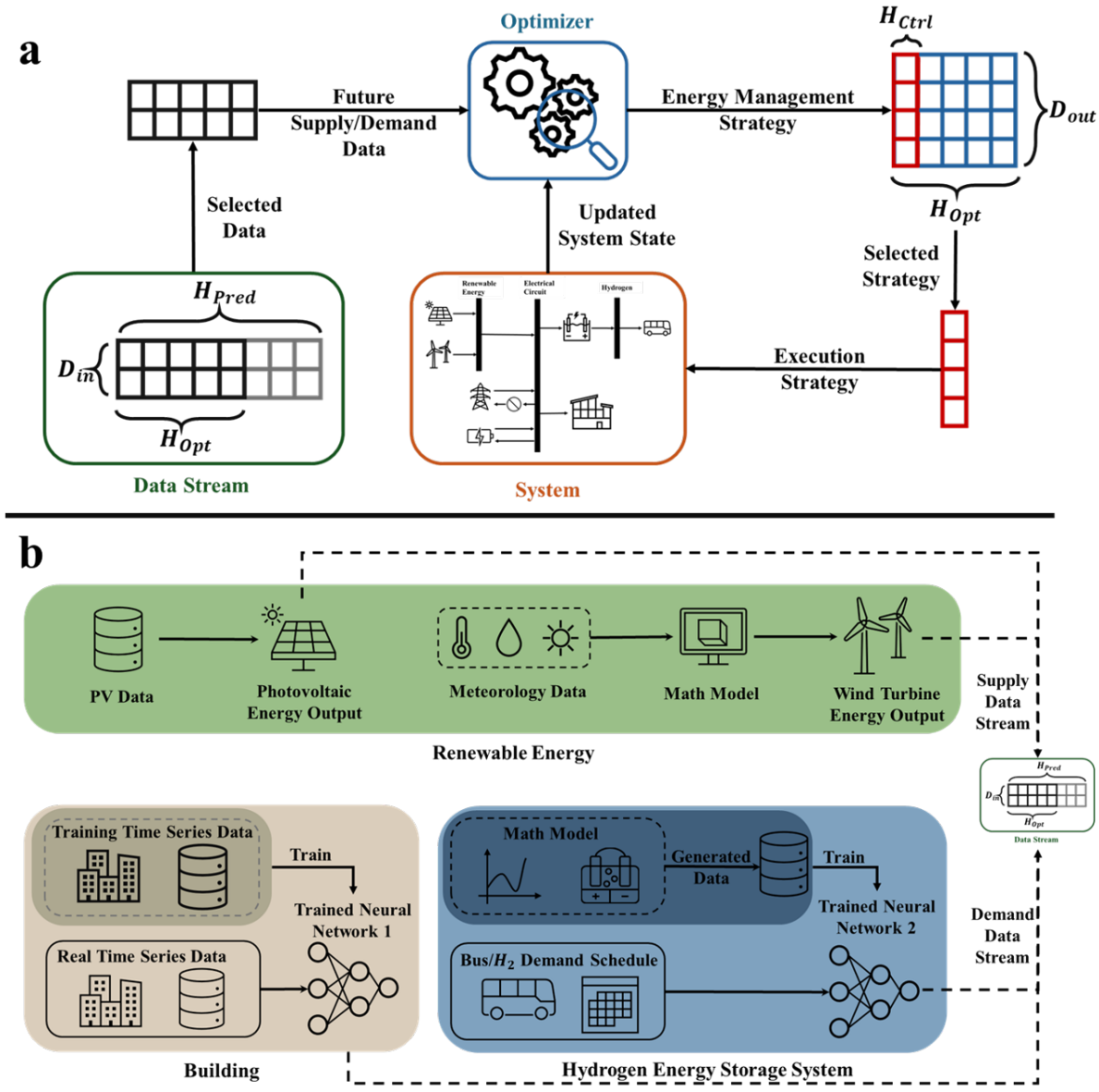


Figure 6.3: **Conceptual Framework of the Proposed Energy Management Algorithm**(a) Optimization Algorithm Structure (b) Data Stream Integration

(H_{Ctrl}), is used as the selected strategy. This portion is then executed as input to the system.

The RES makes adjustments based on the selected strategy. The system will update the system states based on the available energy allocated and the real demand data. The updated system states is then fed back into the optimizer. This iterative process allows the optimizer to refine the energy management strategy continuously, based on real-time data, ensuring adaptive and responsive management aligned with current system conditions.

Figure 6.3 (b) shows the preparation of the data stream, with supply data on the top and demand data on the bottom. The top section represents RE sources, including PV and WT segments. PV data is directly obtained from the dataset, while WT data is derived from meteorological information, gathered through simulation software TRNSYS. Since local weather stations can predict meteorological data, it is feasible to estimate renewable energy generation based on this information. Therefore, this project does not focus on predicting renewable energy generation with machine learning.

The bottom section shows energy demand data. On the lower left is the building energy demand. Initially, this dataset is used to train a neural network, with a portion set aside for testing. The test set provides time series data for the trained neural network during simulations. The neural network then predicts future building energy demand. To the right is the HESS. Here, a mathematical model generates data used to train a second neural network. This network estimates the future energy demand of the HESS, based on the hydrogen demand schedule.

6.2.3 Bus Hydrogen Demand Modelling

The technical report from Mercedes Benz [145] and Hyundai reveals that a typical 12-meter bus consumes 9 kg of hydrogen per 100 km in urban settings. According to a case study, the average daily route for a bus is 230 km [146], resulting in a daily hydrogen consumption of 21 kg. The methodology adopted from the literature [147] simulates hourly hydrogen demand using a modified Gaussian distribution shown in Equation 6.1, which varies throughout the day. Notably, there is no demand at midnight due to the cessation of bus operations. The demand peaks in the afternoon as buses refuel in preparation for night shifts. On weekends, hydrogen demand decreases to 50% of the weekday levels. Buses are

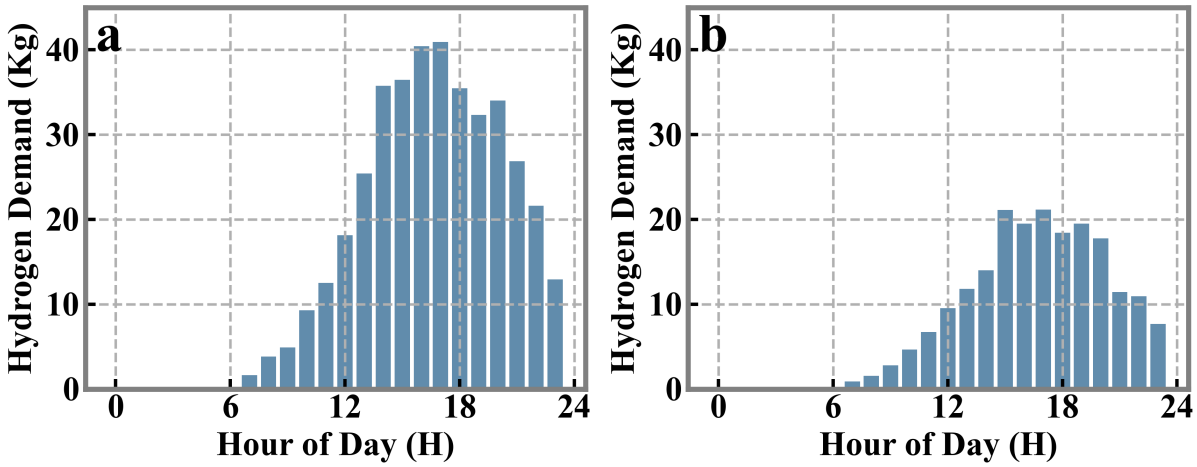


Figure 6.4: **Hydrogen Demand Profile** (a) Workday (b) Weekend and Holiday

scheduled to refuel between 7 AM and 11 PM. Figure 6.4 illustrates the hydrogen demand schedule for a typical weekday and weekend. The designed refuelling station is capable of servicing a hub of 20 buses.

$$f(t) = a \cdot e^{\frac{-(t-b)^2}{2 \cdot c^2}} \cdot (1 + Var) \tag{6.1}$$

6.3 Energy Management Optimization

The objective of the optimization Equation 6.2 is to minimize the electricity bill and reduce the shortage of the hydrogen supply among the optimization horizon T. The first two terms of the objective function are the return of trading energy and the last term is to control the importance of hydrogen shortage based on the value of coefficient M.

$$\min \sum_{t=0}^{t=T} E_{Im}^{(t)} \cdot \pi_{Im}^{(t)} - E_{Ex}^{(t)} \cdot \pi_{Ex}^{(t)} + M \cdot H_{2,\text{Shortage}}^{(t)} \tag{6.2}$$

The objective is subject to a variety of constraints detailed in the following equations.

These constraints are essential for ensuring the model adapts to the physical and operational limitations of the system, including energy generation, conversion, storage, and consumption dynamics.

Basic Rule Equation set 6.3 defines the basic rules essential for optimizing the operation of a BESS and managing its interactions with the electrical grid within a RES framework. These constraints are crucial for preventing overcharging and over-discharging, thereby safeguarding the BESS's longevity and performance. Specifically, they ensure that the BESS cannot charge and discharge simultaneously and that the system cannot import and export energy from the grid at the same time.

The variables E_{Char} and E_{Dis} represent the energy charged to and discharged from the BESS, respectively. These energies are constrained by the maximum capacity of the BESS, denoted as E_{BESS_max}. The Boolean variables B_{Char} and B_{Dis} indicate whether the BESS is in charging mode or discharging mode.

In these equations, the Boolean variables B_{Char} and B_{Dis} can take values of 0 or 1, indicating the operational state of the BESS. The inequalities ensure that when the BESS is charging ($B_{\text{Char}} = 1$), the discharging energy E_{Dis} is zero, and vice versa. The equation $B_{\text{Char}} + B_{\text{Dis}} = 1$ enforces that the BESS is always either charging or discharging, but not both simultaneously.

$$
\begin{aligned}
-E_{BESS\_max} \cdot B_{\text{Dis}} &\leq E_{\text{Dis}} \leq E_{BESS\_max} \cdot (1 - B_{\text{Char}}) \\
-E_{BESS\_max} \cdot B_{\text{Char}} &\leq E_{\text{Char}} \leq E_{BESS\_max} \cdot (1 - B_{\text{Dis}}) \\
B_{\text{Char}} + B_{\text{Dis}} &= 1 \\
-E_{Grid\_max} \cdot B_{\text{Ex}} &\leq E_{\text{Ex}} \leq E_{Grid\_max} \cdot (1 - B_{\text{Im}}) \\
-E_{Grid\_max} \cdot B_{\text{Import}} &\leq E_{\text{Im}} \leq E_{Grid\_max} \cdot (1 - B_{\text{Export}}) \\
B_{\text{Im}} + B_{\text{Ex}} &= 1
\end{aligned}
\tag{6.3}
$$

Similarly, the variables E_{Im} and E_{Ex} represent the energy imported from and exported to the grid, respectively. These energies are constrained by the grid's maximum capacity,

denoted as E_{Grid_max}. The Boolean variables B_{Im} and B_{Ex} indicate whether the system is importing from or exporting to the grid.

These constraints ensure that the system cannot import and export energy at the same time. When the system is importing energy ($B_{\text{Im}} = 1$), the exported energy E_{Ex} is zero, and when the system is exporting energy ($B_{\text{Ex}} = 1$), the imported energy E_{Im} is zero. The equation $B_{\text{Im}} + B_{\text{Ex}} = 1$ ensures that the system is always either importing or exporting energy, but not both at the same time.

Energy Balance Constraints Equation set 6.4 delineates the operational strategy for managing renewable energy sources in an integrated energy system. This framework is designed to optimize the distribution of energy harvested from renewable sources, such as PV panels and WT, by efficiently allocating energy between actual consumption and the necessary curtailment dictated by system constraints. These equations account for the inherent losses associated with energy conversion and inversion processes, ensuring that the system operates efficiently while meeting the diverse demands placed upon it.

$$\begin{aligned} E_{PV} &= E_{\text{PV\_RE}} + E_{\text{PV\_curt}} \\ E_{\text{WT}} &= E_{\text{WT\_RE}} + E_{\text{WT\_curt}} \\ E_{\text{RE}} &= E_{\text{PV\_RE}} - E_{\text{Loss\_Inverter\_PV}} + E_{\text{WT\_RE}} - E_{\text{Loss\_Converter\_WT}} \\ E_{\text{RE}} &= E_{\text{RE\_BESS}} + E_{\text{RE\_Electro}} + E_{\text{RE\_Grid}} + E_{\text{RE\_Load}} \end{aligned} \tag{6.4}$$

The first two equations represent the balance between the total energy generated by the PV panels and WT and the sum of the energy utilized and the energy curtailed due to system limitations. Here, E_{PV} and E_{WT} represent the total energy generated by the PV panels and wind turbines, respectively. The variables E_{PV_RE} and E_{WT_RE} denote the renewable energy from PV and WT that is effectively used in the system, while E_{PV_curt} and E_{WT_curt} represent the curtailed energy due to constraints such as limited storage capacity or grid limitations.

The third equation accounts for the net renewable energy available after considering losses in the inversion and conversion processes. In this equation, E_{RE} is the net renew-

able energy available for distribution. The variables $E_{Loss\_Inverter\_PV}$ and $E_{Loss\_Converter\_WT}$ represent the energy losses during the inversion process for PV energy and the conversion process for WT energy, respectively.

The final equation distributes the net renewable energy to various components of the integrated energy system. Here, $E_{\text{RE\_BESS}}$ is the portion of renewable energy directed to the BESS, $E_{\text{RE\_Electro}}$ is the energy supplied to electrolyzers, $E_{\text{RE\_Grid}}$ is the energy exported to the grid, and $E_{\text{RE\_Load}}$ is the energy consumed by the local load.

These equations ensure that the renewable energy generated is efficiently managed within the system. They account for the balance between generation and curtailment, losses in energy conversion processes, and the distribution of net renewable energy to various system demands. By optimizing these factors, the integrated energy system maximizes the utilization of renewable resources while adhering to operational constraints.

Equation set 6.5 provided model the energy dynamics within a BESS, crucial for optimizing its charging and discharging processes in the context of system efficiency and aging factors. These equations integrate the flows of energy from renewable sources and the grid into the BESS, while considering conversion losses. They also regulate the charging and discharging activities based on the SOC limits and the capacity of the aged battery system. This ensures the BESS operates within safe and efficient parameters, balancing energy storage with demand fulfillment across different system loads and contributing effectively to overall energy management and system stability.

$$
\begin{aligned}
&E_{\text{Char}} = E_{\text{RE\_BESS}} + E_{\text{Grid\_BESS}} - E_{\text{Loss\_Converter\_Grid}} - E_{\text{Loss\_Converter\_BESS}} \\
&E_{\text{Dis}} = E_{\text{BESS\_Electro}} + E_{\text{BESS\_Comp}} + E_{\text{BESS\_Load}} \\
&E_{\text{Char}}^{(t)} \leq \text{SOC}_{\max} \cdot \text{Cap}_{\text{BESS\_Aged}}^{(t)} - \text{SOC}^{(t-1)} \cdot \text{Cap}_{\text{BESS\_Aged}}^{(t-1)} \\
&E_{\text{Dis}}^{(t)} \leq \text{SOC}^{(t-1)} \cdot \text{Cap}_{\text{BESS\_Aged}}^{(t-1)} - \text{SOC}_{\min} \cdot \text{Cap}_{\text{BESS\_Aged}}^{(t)} \\
&\text{SOC}^{(t)} \cdot \text{Cap}_{\text{BESS\_Aged}}^{(t)} = \text{SOC}^{(t-1)} \cdot \text{Cap}_{\text{BESS\_Aged}}^{(t-1)} + E_{\text{Char}}^{(t)} - E_{\text{Dis}}^{(t)} - E_{\text{Loss\_BESS}}^{(t)}
\end{aligned}
\tag{6.5}
$$

Equation set 6.6 presented are integral to the management of energy and hydrogen storage within a hybrid system that includes electrical and hydrogen components. These

equations account for the transfer and conversion of energy from renewable sources, batteries, and the grid towards electrolysis processes and compression systems, incorporating efficiency factors and conversion losses. The system's dynamics also extend to hydrogen storage and supply, where stored hydrogen is balanced against supply and demand scenarios, capturing production, utilization, and any shortages. These equations facilitate the optimization of both energy and hydrogen flows, ensuring efficient operation across various components of the integrated system, crucial for advancing sustainable and resilient energy solutions.

$$
\begin{aligned}
&E_{\text{Electro}} = E_{\text{RE\_Electro}} + E_{\text{BESS\_Electro}} \cdot \eta_{Inverter} + E_{\text{Grid\_Electro}} - E_{\text{Loss\_Converter\_Electro}} \\
&E_{\text{Comp}} = E_{\text{BESS\_Comp}} \cdot \eta_{\text{Inverter}} + E_{\text{Grid\_Comp}} \\
&H^{(t)}_{2,\text{Storage}} = H^{(\text{t-1})}_{2,\text{Storage}} + H^{(t)}_{2,\text{Electro\_Storage}} - H^{(t)}_{2,\text{Supply}} \\
&H^{(t)}_{2,\text{Demand}} = H^{(t)}_{2,\text{Supply}} + H^{(t)}_{2,\text{Shortage}}
\end{aligned}
\tag{6.6}
$$

Equation set 6.7 outline energy distribution and grid interactions in a power system. They describe the total energy imported from the grid for various uses, the energy exported to the grid from renewable sources, and the total energy consumed by the load, incorporating contributions from renewables, battery discharges with efficiency adjustments, and direct grid supply. This framework optimizes energy utilization, load management, and grid efficiency, crucial for maintaining balance and reliability in smart grids.

$$
\begin{aligned}
&E_{\text{Im}} = E_{\text{Grid\_BESS}} + E_{\text{Grid\_Electro}} + E_{\text{Grid\_Comp}} + E_{\text{Grid\_Load}} \\
&E_{\text{Ex}} = E_{\text{RE\_Grid}} \\
&E_{\text{Load}} = E_{\text{RE\_Load}} + E_{\text{BESS\_Load}} \cdot \eta_{\text{Inverter}} + E_{\text{Grid\_Load}}
\end{aligned}
\tag{6.7}
$$

Energy Loss Constraints Equation set 6.8 and 6.9 define the energy losses in conversion and inversion processes within a renewable energy system. They quantify the losses from wind turbines, batteries, grid interactions, and electrolysis processes due to efficiency less than perfect (denoted by $\eta_{\text{Converter}}$ and η_{Inverter}). These losses are crucial for accurately

modeling energy efficiency and optimizing system performance, ensuring realistic planning and operation in renewable energy systems.

$$
\begin{aligned}
&E_{\text{Loss\_Converter}} = E_{\text{Loss\_Converter\_WT}} + E_{\text{Loss\_Converter\_BESS}} + E_{\text{Loss\_Converter\_Grid}} \\
&E_{\text{Loss\_Converter\_WT}} = E_{\text{WT\_RE}} \cdot (1 - \eta_{\text{Converter}}) \\
&E_{\text{Loss\_Converter\_BESS}} = E_{\text{RE\_BESS}} \cdot (1 - \eta_{\text{Converter}}) \\
&E_{\text{Loss\_Converter\_Grid}} = E_{\text{Grid\_BESS}} \cdot (1 - \eta_{\text{Converter}}) \\
&E_{\text{Loss\_Converter\_Electro}} = (E_{\text{Grid\_Electro}} + E_{\text{RE\_Electro}} + E_{\text{BESS\_Electro}}) \cdot (1 - \eta_{\text{Converter}})
\end{aligned}
\tag{6.8}
$$

$$
\begin{aligned}
&E_{\text{Loss\_Inverter}} = E_{\text{Loss\_Inverter\_PV}} + E_{\text{Loss\_Inverter\_BESS}} \\
&E_{\text{Loss\_Inverter\_PV}} = E_{\text{PV\_RE}} \cdot (1 - \eta_{\text{Inverter}}) \\
&E_{\text{Loss\_Inverter\_BESS}} = E_{\text{Dis}} \cdot (1 - \eta_{\text{Inverter}})
\end{aligned}
\tag{6.9}
$$

Power Constraints Equation set 6.10 power capacity requirements to ensure that the available converters, inverters, and electrolyzers can handle the energy processed from various sources and uses. These constraints guarantee that the system's physical capabilities are not exceeded by the operational demands, crucial for maintaining stability and efficiency in energy conversion and distribution within the power system.

$$
\begin{aligned}
&P_{\text{Converter}} \cdot dt \geq E_{\text{WT\_RE}} + E_{\text{RE\_BESS}} + E_{\text{RE\_Electro}} + E_{\text{Grid\_BESS}} + E_{\text{Grid\_Electro}} \\
&P_{\text{Inverter}} \cdot dt \geq E_{\text{PV\_RE}} + E_{\text{BESS\_Electro}} + E_{\text{BESS\_Comp}} + E_{\text{BESS\_Load}} \\
&P_{\text{Electro}} \cdot dt \geq E_{\text{Electro}}
\end{aligned}
\tag{6.10}
$$

6.4 Result and Discussion

In this study, a rigorous evaluation was conducted to determine the most effective ML algorithm for time series prediction within a proposed energy management system. Different

ML algorithms were initially compared, and the Transformer was selected for its superior ability to handle complex temporal data, establishing it as the primary predictive model for subsequent simulations. The simulation began with a reference configuration, Configuration 1, serving as the baseline. Additional simulations were then conducted across various system configurations with differing settings and energy management strategies to assess performance. Parameters for all configurations are shown in Table 6.1.

Table 6.1: **Configuration Comparison**

Config.	**1**	**2**	**3**	**4**	**5**
Prediction Horizon (h)	48	48	48	48	1
Optimization Horizon (h)	24	24	48	24	1
Control Horizon (h)	1	1	1	1	1
ML Architecture	Transformer	Transformer	Transformer	LSTM	NA
HESS Storage (MWh)	10	6	10	10	10

6.4.1 Machine Learning for Energy Demand Prediction

The comparative analysis presented in Figure 6.5 evaluates the predictive capabilities of three neural network models—Transformer, Inverted Transformer, and LSTM—for forecasting building energy demand. Figure 6.5 (a) to (d) illustrate the models' performance at 3-hour intervals, providing a detailed examination of their temporal prediction accuracy. Detailed hyper-parameters of these models are listed in Table 6.2.

The Transformer model consistently outperforms the others, closely aligning with actual energy consumption data across all time frames. This superior performance underscores the Transformer's robustness in handling the complex, nonlinear patterns inherent in energy demand, and is particularly noteworthy given the common challenge of declining prediction accuracy over extended horizons in time-series forecasting.

Table 6.2: **Training Hyper-parameters**

Parameter	**Value**	**Parameter**	**Value**
Sequence Length	48	Learning Rate	0.005
Prediction Length	48	LSTM Hidden States	30
LSTM Layers	3	Encoder Layers	2
Decoder Layers	1	Output Channels	14
Encoder Input Size	14	Decoder Input Size	14

In contrast, the i-Transformer and LSTM models exhibit a gradual decrease in predictive accuracy as the forecast horizon extends. Despite the i-Transformer's structural similarity to the Transformer, its inverted architecture does not effectively enhance its ability to capture the sequential dependencies critical for accurate energy demand prediction. The LSTM model, although proficient in modeling temporal sequences, demonstrates limitations in handling long-term dependencies compared to the Transformer's self-attention mechanisms.

Figure 6.5 (e) and (f) further quantify these findings by providing monthly and overall comparisons of the models' predictive performance. The Transformer's dominance is evident across all months, indicating its stability and reliability under varying seasonal conditions and consumption patterns.

Table 6.3 presents the models' performance metrics—MSE—across different hyper-parameters of label and prediction lengths. Label length refers to the number of historical time steps provided as input, enabling the model to analyze past patterns and trends, while prediction length specifies the number of future time steps for which the model generates forecasts. The Transformer consistently achieves the lowest MSE value, confirming its efficacy. Specifically, with a prediction length of 48, the Transformer's superior accuracy demonstrates its capacity to maintain high performance.

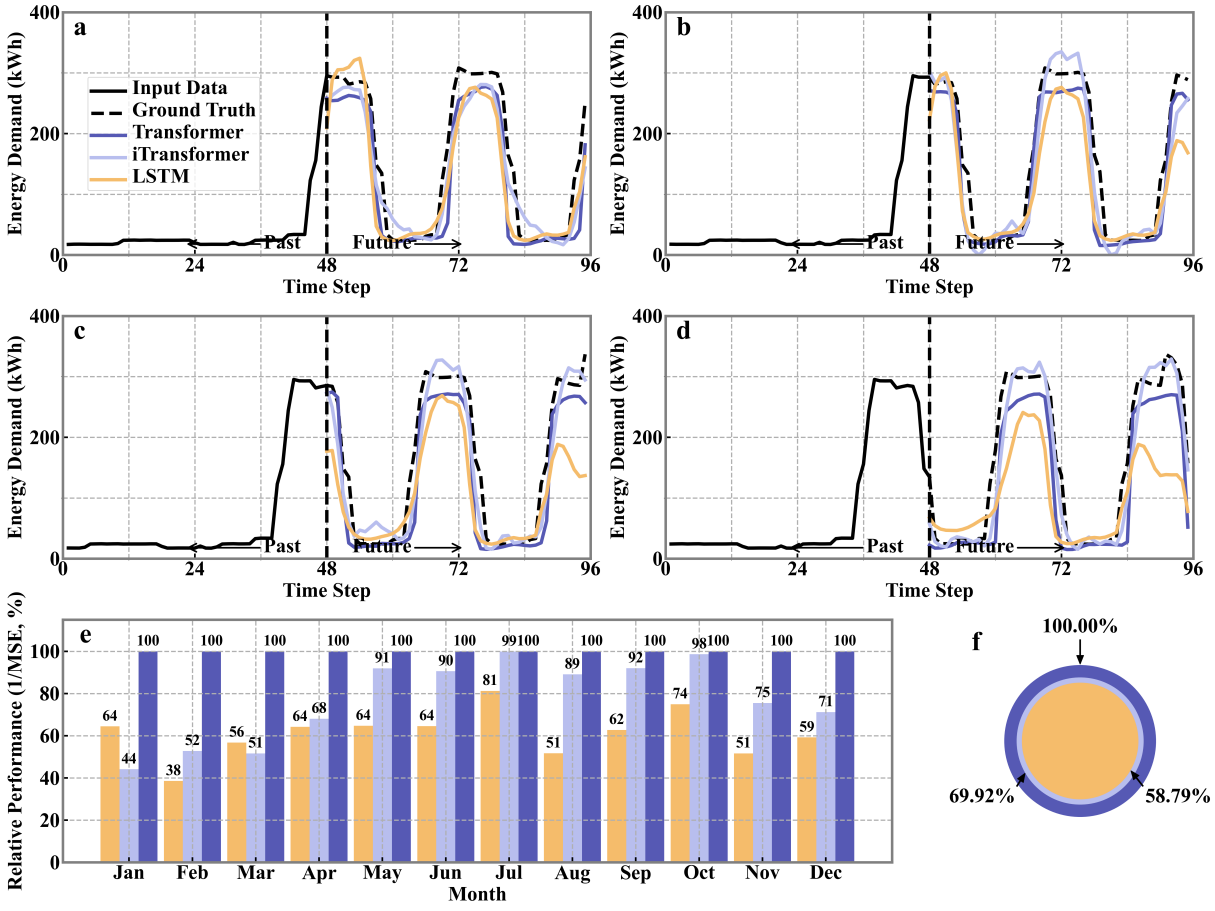


Figure 6.5: **Building Energy Demand Prediction** (a) Predicted Energy Demand at Time T (b) Predicted Energy Demand at Time T+3 Hours (c) Predicted Energy Demand at Time T+6 Hours (d) Predicted Energy Demand at Time T+9 Hours (e) Monthly Comparison of Network Prediction Performance (f) Overall Network Prediction Performance Comparison

6.4.2 Energy Management

The energy management strategies implemented in this study demonstrate the effective optimization of HESS within RES and empower green hydrogen energy production for transportation usage. Figure 6.6 illustrates the operational dynamics of the HESS during

Table 6.3: **Comparison of Prediction Performance Across Different Models**

Model	Label Len	Pred Len	MSE
Label Len = 24, Pred Len adjusted			
LSTM	24	48	0.8357
Transformer	24	48	**0.5154**
iTransformer	24	48	0.6956
LSTM	24	24	0.7188
Transformer	24	24	**0.5326**
iTransformer	24	24	0.6380
Label Len = 48, Pred Len adjusted			
LSTM	48	48	0.7381
Transformer	48	48	**0.5240**
iTransformer	48	48	0.5810
LSTM	48	96	0.8834
Transformer	48	96	**0.5670**
iTransformer	48	96	0.6072
LSTM	48	24	0.6855
Transformer	48	24	**0.4674**
iTransformer	48	24	0.5120

the simulation period, highlighting the interplay between hydrogen demand, supply, and storage levels. Specifically, Figure 6.6(a) shows the fluctuations in hydrogen demand and the corresponding supply mechanisms from randomly selected hour index 1000 to 1072. The energy management system allocates energy from RE sources, the BESS, and the grid

to the electrolyzer, ensuring that hydrogen demand is consistently met without shortages. This alignment between supply and demand as shown in 6.6(d) and 6.6(e) indicates that the system's capacity is well-matched to operational requirements, with hydrogen storage levels remaining sufficient throughout the period.

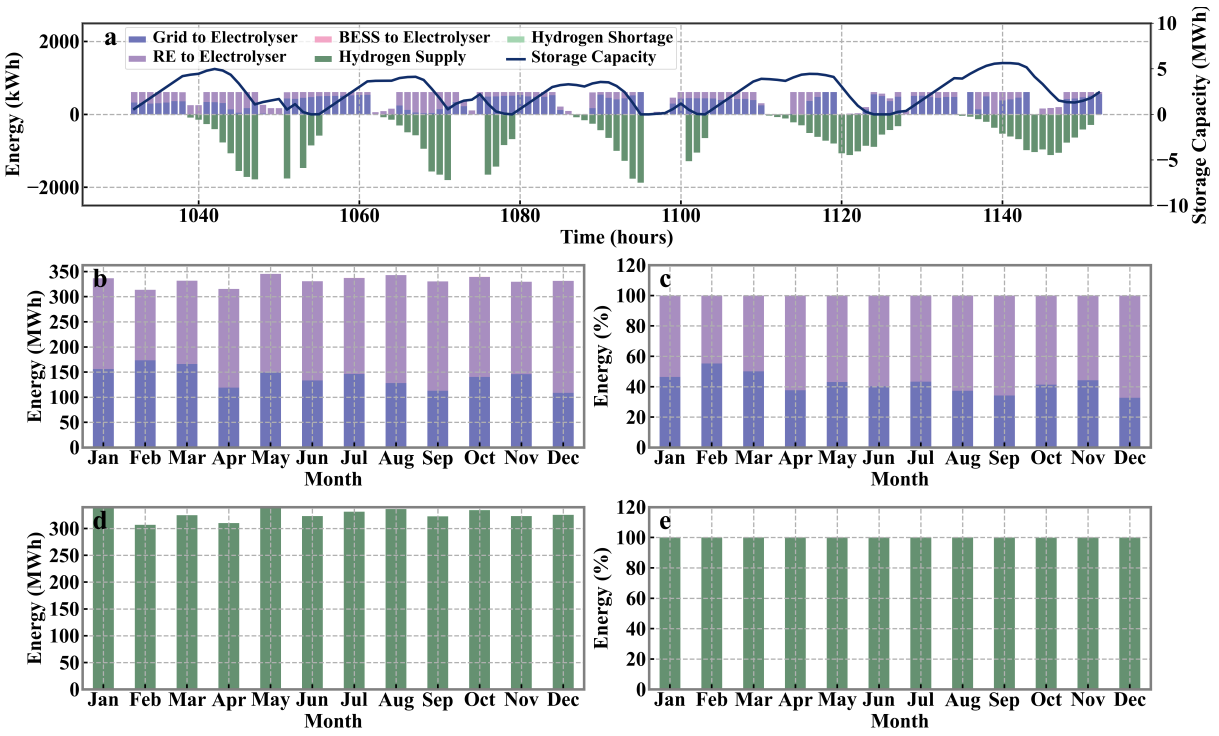


Figure 6.6: **Energy Management Simulation Results for the Hydrogen Energy Storage System** (a) Hydrogen Demand and Supply Dynamics from Hour Index 1000 to 1072 (b) Detailed Breakdown of Energy Supply by Source in MWh (c) Proportional Distribution of Energy Supply by Source (d) Detailed Breakdown of Hydrogen Supply in MWh (e) Proportional Distribution of Hydrogen Supply

Figures 6.6(b) and 6.6(c) analyze the energy sources supplying the electrolyzer. The results reveal that the GHR achieved 52.65%. The system strategically avoids utilizing energy from the BESS for hydrogen production due to inefficiencies and increased costs associated with energy conversion losses, battery aging, and DC/AC inversion processes. By

prioritizing direct renewable energy and grid energy, the system enhances overall efficiency and reduces unnecessary energy transformations.

The analysis of the building energy supply, depicted in Figure 6.7, emphasizes the importance of accurate predictive energy management in optimizing renewable energy utilization. Figures 6.7(a) to 6.7(f) illustrate the optimized energy supply at time T, based on predicted energy demand with a 5-hour rolling horizon shifting. As time advances, the future time points for which predictions were made draw nearer, and predictions become increasingly accurate, allowing for a more effective allocation of energy resources. The energy management algorithm allocates energy from renewable sources, the grid, and the BESS based on these predictions, proactively scheduling renewable energy utilization to meet the load demand. This preemptive allocation is facilitated by the rolling horizon optimization, which continuously updates the energy supply plan as new, more accurate predictions become available.

Figures 6.7(g) and 6.7(h) provide a breakdown of the weekly and monthly energy supply, revealing that renewable energy supplies the majority of the energy required by the buildings. Prioritizing renewable energy reduces reliance on grid energy and minimizes the need to draw from the BESS, enhancing the sustainability and economic efficiency of the energy system by utilizing cost-effective renewable resources.

6.4.3 Comparative Analysis

In this research, the performances of various configurations of the RES are examined through a comparative study that focuses on several key factors: the storage capacity of the HESS, the prediction horizons of the energy management algorithms, and the ML architectures employed. Each scenario specifically analyzes how variations in these factors impact the system's operational efficiency, with particular emphasis on hydrogen supply, renewable energy utilization, and energy independence.

In cases where certain configurations cannot fully meet the hydrogen demand, direct comparison of some performance indicators is challenging. To address this, adjustments

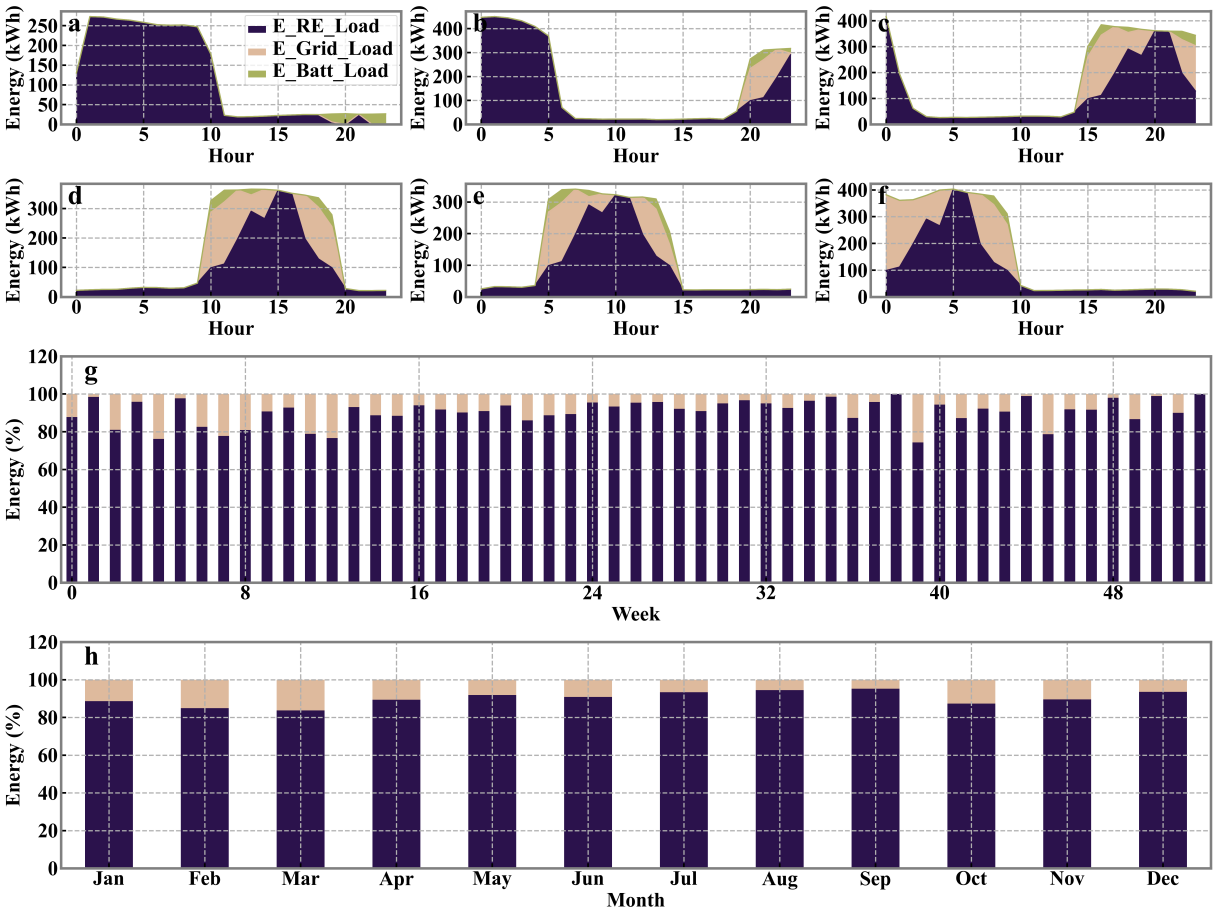


Figure 6.7: **Energy Management Simulation Results for the Building Energy Supply** (a)-(f) Optimized Energy Supply at Time T based on Predicted Energy Demand with 5-Hour Rolling Horizon Shifting (g) Weekly Energy Supply Breakdown (h) Monthly Energy Supply Breakdown

are made to these indicators to account for hydrogen supply shortages. The additional supply in these instances is sourced from the grid rather than renewable energy, allowing for consistent and equitable comparisons across configurations. These indicators, adjusted indicators, are presented in Table 6.4 and visualized in Figure 6.8.

The study first assesses the impact of altering HESS storage capacity by comparing

Table 6.4: **Summary of Indicator Values**

Configuration	**1**	**2**	**3**	**4**	**5**
EIS	0.71	0.72	0.71	0.70	0.70
Adjusted EIS	0.71	0.69	0.71	0.67	0.63
GEIF	0.44	0.46	0.42	0.48	0.57
Adjusted GEIF	0.44	0.52	0.42	0.54	0.7
UF	0.95	0.91	0.96	0.88	0.80
NSF	0.00	0.31	0.00	0.22	0.34
NES	0.00	0.044	0.00	0.159	0.40
HSSS	100	18.65	100	5.88	2.44
GHR	52.62	61.90	54.30	50.94	46.21
Adjusted GHR	52.62	47.42	54.30	43.67	29.54

Configuration 1 with a reduced capacity scenario (Configuration 2) and an extended prediction horizon scenario (Configuration 3) to examine the consequences on hydrogen supply sufficiency and green hydrogen production rate. Figure 6.9 illustrates the simulation result.

The simulation result revealed that expanding HESS storage capacity is critical for ensuring hydrogen supply sufficiency. Comparing Configuration 1 with Configuration 2, reducing the storage capacity by 40% led to a dramatic decrease in the HSSS from 100 to 18.65. This significant drop indicates the system's inability to meet hydrogen demand, as evidenced by increases in the NSF from 0.00 to 0.31 and the NES from 0.00 to 0.044.

Interestingly, the GHR increased from 52.62% to 61.90%, indicating a higher proportion of renewable energy used in hydrogen production. However, this increase was due to a reduction in the total energy supplied to the electrolyzer, stemming from the smaller storage capacity. With reduced capacity, the electrolyzer operated less frequently, leading to a higher proportion of renewable energy usage when it did operate. Despite this, the

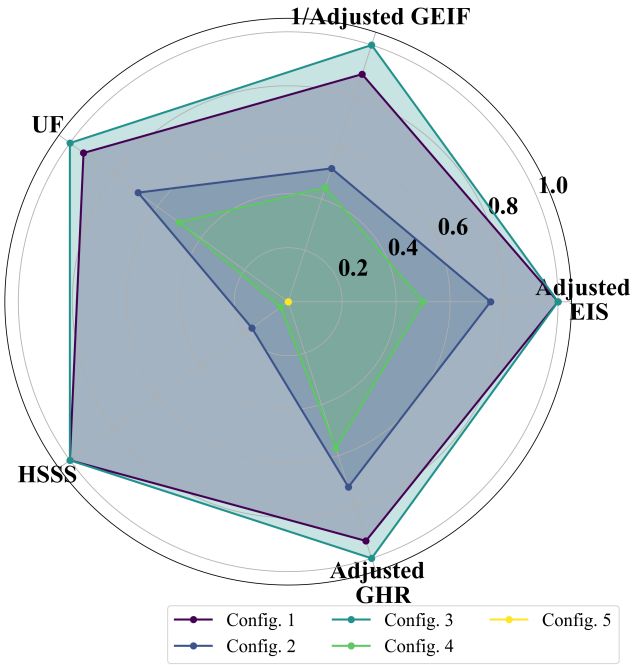


Figure 6.8: **Processed Indicator Comparison**

actual quantity of renewable energy consumed decreased, reflecting the underutilization of available resources.

This scenario highlights that higher values in indicators like GHR do not necessarily equate to better performance when total energy supplied diminishes and demand goes unmet. Adequate storage capacity is crucial not only for meeting demand but also for optimizing renewable energy utilization. The adjusted GHR further confirms this by showing a comparison of 52.62% versus 47.42%, illustrating that when demand is taken into account, the apparent benefits of a higher GHR percentage diminish, underscoring the need for sufficient storage to fully leverage renewable resources.

Extending the optimization horizon in the energy management algorithms from 24 to 48 hours, as done in Configuration 3 compared to Configuration 1, showed modest improvements in system performance. The longer horizon allowed the system to better anticipate

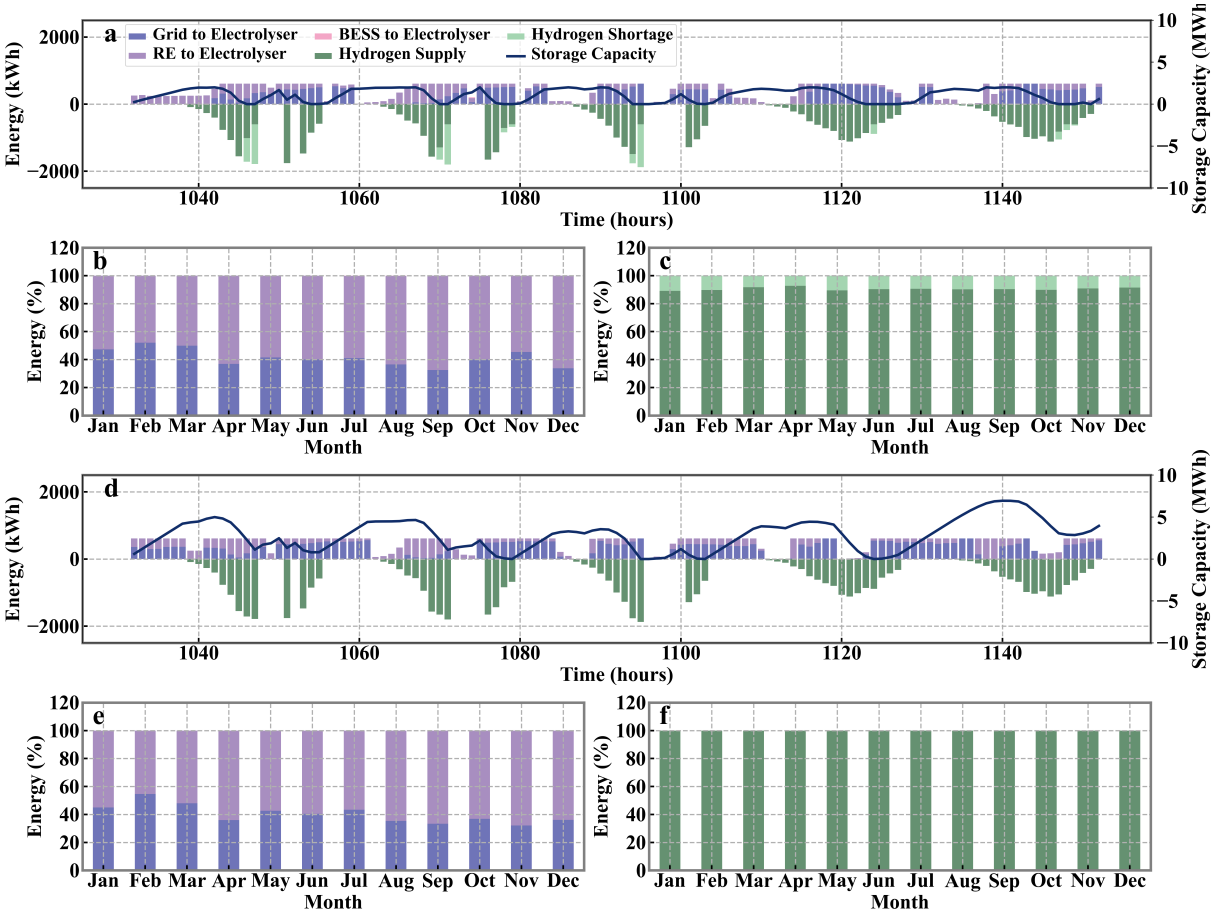


Figure 6.9: **Comparison of Energy Management Simulation Results for the Hydrogen Energy Storage System with Different Configurations** (a) 72-Hour Hydrogen Dynamics (Config. 2) (b) Energy Supply Distribution (Config. 2) (c) Hydrogen Supply Distribution (Config. 2) (d) 72-Hour Hydrogen Dynamics (Config. 3) (e) Energy Supply Distribution (Config. 3) (f) Hydrogen Supply Distribution (Config. 3)

future demand and renewable energy availability, enabling more strategic planning of energy allocation. This resulted in a slight increase in the UF from 0.95 to 0.96 and a decrease in the GEIF from 0.44 to 0.42, indicating reduced reliance on grid energy. The GHR also increased slightly from 52.62% to 54.30%. Despite these improvements, the benefits of

extending the prediction horizon appeared to have diminishing returns beyond a certain point. The HSSS remained at 100 in both configurations, suggesting that while a longer horizon enhances planning, its impact is limited without concurrent improvements in other areas like storage capacity and predictive accuracy.

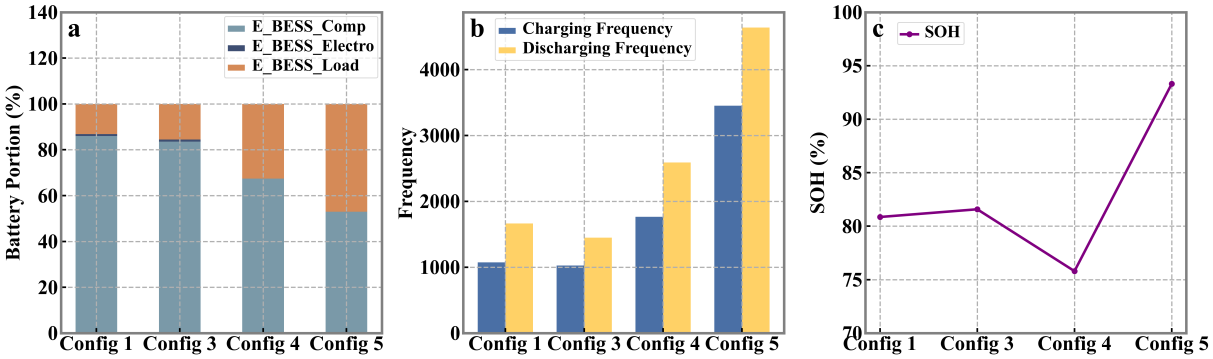


Figure 6.10: **Energy Management Simulation Results for the Battery Energy Storage System** (a) Energy Flow (b) Usage Frequency (c) State of Health

The study then evaluates the impact of algorithm performance on the BESS, focusing on how different optimization configurations affect energy allocation, battery operations, and overall BESS efficiency. The findings, illustrated in Figure 6.10, underscore the critical role of predictive capabilities in optimizing BESS performance.

As the algorithm's performance deteriorates, a larger proportion of the discharged energy from the BESS is allocated to the load, while the energy supplied to the electrolyzer and compressor decreases significantly, as demonstrated in Figure 6.10 (a). This trend is particularly evident in configurations with lower prediction accuracy or without predictive functionality. Configurations with better optimization capabilities effectively supply energy to the electrolyzer, whereas poorer-performing configurations primarily direct energy to the compressor and load. This shift indicates an imbalance in resource allocation due to the system's inability to accurately anticipate future energy demand.

Declining optimization performance also leads to a significant increase in the frequency of battery charging and discharging, as shown in Figure 6.10 (b). In scenarios with poor or

no prediction, the system frequently adjusts battery operations to manage fluctuations in energy demand. This reliance on real-time adjustments, in the absence of robust predictive capabilities, results in more frequent battery use, highlighting the system's struggle to efficiently balance supply and demand.

Furthermore, in Figure 6.10 (c), a decrease in optimization performance correlates with increased battery usage and a decline in the battery's SOH, particularly in configurations lacking predictive capabilities. Despite the higher frequency of battery operations, the overall efficiency of BESS utilization decreases. Frequent, small-scale adjustments replace more strategic and efficient battery usage, negatively impacting system efficiency and accelerating battery degradation.

These observations lead to several conclusions. Enhancing the system's predictive ability by extending the optimization horizon significantly improves its capacity to anticipate future energy demand and allocate resources efficiently. Configurations with better predictive capabilities demonstrate increased system efficiency and reduced unnecessary battery operations. Conversely, reducing prediction accuracy weakens optimization performance, causing more energy to be directed to the load instead of the compressor or electrolyzer, indicating the underutilization of resources. Eliminating predictive functionality further degrades performance, resulting in frequent battery operations, increased usage, and a decline in the battery's SOH.

Finally, the influence of different ML architectures—Transformer versus LSTM—is analyzed to determine their respective efficiencies in predictive accuracy and overall energy management. The choice of ML algorithm within the predictive-based energy management system significantly influenced operational effectiveness, as evidenced by comparing Configurations 1, 4, and 5. Configuration 1, utilizing a Transformer model, achieved optimal performance, whereas Configuration 4, which used an LSTM model, and Configuration 5, which lacked a predictive algorithm entirely, showed substantial declines in key performance indicators.

As prediction accuracy deteriorated in Configuration 4, the HSSS dropped significantly from 100 to 5.88, indicating severe hydrogen supply shortages. Correspondingly, the NSF

increased from 0.00 to 0.22, and the NES rose from 0.00 to 0.159, reflecting both more frequent and larger supply shortfalls. The UF decreased from 0.95 to 0.88, highlighting a reduction in renewable energy utilization efficiency.

The performance declined even further in the absence of predictive capability in Configuration 5, with the HSSS dropping sharply to 2.44. The NSF and NES values rose to 0.34 and 0.40, respectively, indicating both frequent and substantial hydrogen supply shortages. Additionally, the UF decreased to 0.80, while the adjusted GEIF increased to 0.7, indicating a greater dependence on grid energy and even less effective use of renewable resources. This pattern underscores the critical role of accurate predictions in sustaining renewable energy utilization and reducing grid reliance.

These findings visualized and confirmed in Figure 6.11, underscores the critical role of advanced ML architectures in enhancing predictive accuracy and, consequently, overall system performance. Without accurate predictive capabilities, the system cannot effectively balance supply and demand, leading to operational inefficiencies and unmet energy needs.

Overall, the integration of advanced predictive control techniques, particularly the use of a Transformer-based model predictive control framework with rolling horizon optimization, significantly enhanced the system's ability to effectively manage energy resources and integrate renewable energy sources. By continuously updating the energy supply plan based on accurate predictions, the system proactively scheduled renewable energy utilization to meet load demands, maximizing the use of renewable energy, reducing reliance on grid energy, and improving the reliability and sustainability of the energy supply to buildings and the HESS.

6.4.4 Assumptions and Limitations

Despite the valuable insights gained from this comparative analysis, the study has certain limitations due to its focus on a select range of HESS capacities, prediction horizons, and machine learning architectures within a simulated setting. This scope might not accurately

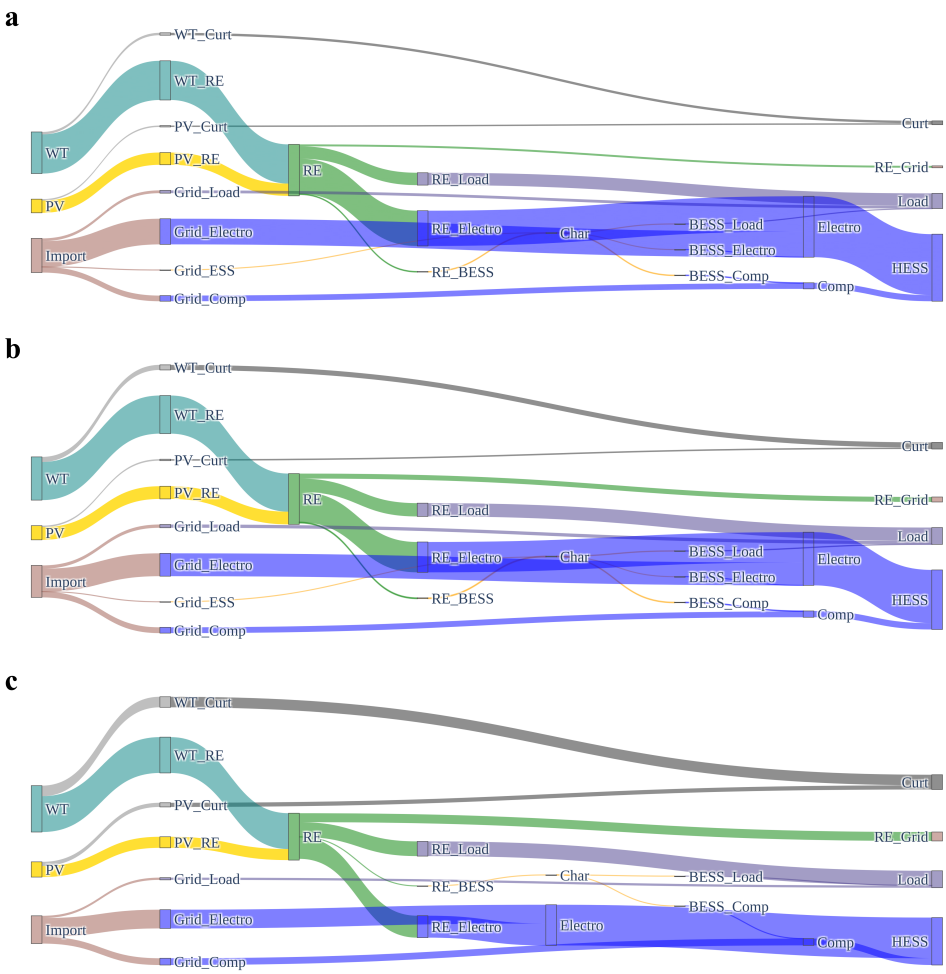


Figure 6.11: **Energy Flow within the System**(a) Configuration 1 (b) Configuration 4 (c) Configuration 5

reflect the complexities and external influences that real-world renewable energy systems encounter, such as equipment malfunctions, maintenance challenges, or fluctuations in energy generation caused by unpredictable weather conditions.

Moreover, the effectiveness of the employed predictive models, including Transformer and LSTM architectures, heavily depends on the quality and breadth of the input data.

Issues with data accuracy or representativeness could affect the performance and generalizability of the findings across different contexts or geographical areas.

Furthermore, the analysis excludes considerations of how integrating other renewable energy sources or storage technologies might influence system dynamics and performance outcomes.

Lastly, the simulation-based approach does not incorporate human factors or the potential impacts of regulatory or market dynamics, which are significant drivers of system performance. These omissions suggest that while the study's findings are instructive for understanding technical configurations and their immediate impacts, they may not be directly translatable to real-world scenarios without further consideration of these broader and more complex factors.

6.5 Summary

This study introduces a predictive-based energy management algorithm and conducts a detailed examination of how variations in hydrogen energy storage system capacity, prediction horizons, and machine learning architectures impact its performance. The proposed energy management algorithm improves electrolyzer utilization by strategically allocating energy in advance based on accurate predictions, providing a cost-effective solution for hydrogen production. Simulations demonstrate that the algorithm enables the system to operate with a smaller electrolyzer, raising the Hydrogen Supply Sufficiency Score from 18.65 to 100, representing a 436.46% improvement in hydrogen supply reliability.

Expanding the optimization horizon resulted in a modest yet notable improvement in the system's predictive accuracy and operational efficiency. Specifically, the Utilization Factor saw a marginal increase from 0.95 to 0.96, and the Adjusted Grid Energy Interaction Factor decreased from 0.44 to 0.42. The deployment of a Transformer model significantly enhanced the system's performance compared to the Long Short-Term Memory model, as evidenced by a 94.12% increase in the Hydrogen Supply Sufficiency Score. Additionally, in systems with advanced energy management capabilities, the adjusted Green Hydrogen

Ratio increased from 29.54% to 52.62%. This improvement highlights the critical role of sophisticated energy management in enhancing the sustainability of hydrogen production by maximizing renewable energy utilization. The real-world data-driven simulation demonstrates that the proposed energy management algorithm can be effectively applied to diverse systems, including those with varying scales, layouts, or geographical locations, showcasing its adaptability and robustness across different operational scenarios. Future work could explore real-time implementation under dynamic conditions and evaluate economic feasibility through cost-benefit analysis and policy alignment.

Chapter 7

Conclusion and Future Work

This chapter provides the overall conclusions of this research and offers an outlook on future work. Section 7.1 summarizes the key findings and contributions made throughout the study, highlighting the advancements in RES modeling and optimization. Section 7.2 discusses potential areas for future work, addressing how the methodologies and results from this research can be extended or applied to new challenges and opportunities in the field of RES.

7.1 Conclusion

This thesis presents several significant contributions to the field of energy management in renewable energy systems, focusing on advanced modeling, optimization, and machine learning techniques to enhance system performance and efficiency.

Firstly, a data-driven PV panel modeling and optimization framework was proposed. This approach addresses the computational complexities commonly associated with traditional mathematical modeling. By leveraging data-driven methods, the framework not only reduces computational overhead but also provides an in-depth analysis of PV panel performance, leading to improved energy system efficiency.

Secondly, a two-stage optimization approach for determining the optimal sizing of primary components in renewable energy systems was introduced. Utilizing real-world data, this method effectively balances customer return requirements and environmental impact considerations. The proposed algorithm is adaptable to individual customer needs and employs an advanced search mechanism that significantly decreases computational time and resources. This enhancement accelerates the convergence towards the optimal solution, making the optimization process more efficient and practical for real-world applications.

Finally, to further improve the efficiency of renewable energy systems, a novel energy management algorithm based on advanced machine learning techniques was developed. This algorithm significantly enhances the performance of the energy management system by increasing prediction accuracy over conventional machine learning methods. The improved accuracy contributes to a more robust and efficient energy management system, offering superior adaptability and reliability in various operating conditions.

Overall, the contributions of this thesis provide practical solutions to key challenges in renewable energy system management. By reducing computational complexity, enhancing optimization strategies, and leveraging advanced machine learning techniques, this work offers flexible and efficient tools that can be tailored to specific customer requirements. These advancements have the potential to significantly impact the renewable energy sector, facilitating more efficient and sustainable energy management practices.

7.2 Future Work

Future work can be divided into two sections: improvements within the existing projects and potential advancements beyond them. The first project focuses on photovoltaic panel modeling and optimization. Enhancing the model's adaptability to various global locations and environmental conditions would make the system more versatile and reliable in dynamic environments. By enabling automatic adjustments to different geographic settings and climates, the framework could perform optimally across a wide range of real-world conditions, significantly broadening its applicability.

The second project was on component size optimization, which was conducted prior to the development of the full system. Future research could aim to identify the economic-environmental balance for RES in different locations and enhance system robustness through scenario design and pressure testing. This approach would enable the system to account for location-specific factors, such as energy prices, weather variability, and regulatory constraints, while pressure testing would help ensure resilience under extreme conditions. In addition, a system-wide energy equation can be developed to provide a unified representation of energy flows within the entire system. This enhancement will improve the transparency and robustness of the model and promote its application to larger-scale systems or different configurations.

The third project offers a promising direction for incorporating the arge Language Model (LLM) and reinforcement learning to improve system performance. Integrating LLMs would enable the system to explain its energy management logic to users in a clear and understandable way, enhancing transparency and user trust. Additionally, LLMs could receive and interpret user feedback, allowing the system to continuously adjust and refine its energy management strategies. This feedback loop could be further strengthened through reinforcement learning, where the system learns from user interactions over time, optimizing performance based on real-world feedback. Incorporating user satisfaction levels into the optimization process would tailor the system's operations based on user preference, creating a more personalized and adaptive experience.

Additionally, several other potential advancements could significantly improve RES. One area worth exploring is the development of robust and fault-handling algorithms. These systems are often deployed in environments where failures such as energy source intermittency or grid disruptions are common. Creating algorithms that can withstand these disruptions would allow systems to maintain functionality even under challenging conditions. This could involve the integration of predictive failure detection methods and self-healing mechanisms to reduce downtime.

Another promising direction involves decentralized energy management for systems that include microgrids and distributed energy resources. As the energy system becomes more

decentralized, effectively managing energy generation and consumption at the local level is becoming increasingly critical. Future research could focus on algorithms capable of facilitating dynamic energy exchanges between multiple prosumers (producers and consumers), optimizing for factors such as grid stability, energy costs, and emissions in real-time.

Letter of Copyright Permission

MDPI, as publisher of the manuscript "Formulation and data-driven optimization for maximizing the photovoltaic power with tilt angle adjustment" fully or partly adopted in Chapter 1, Chapter 2, Chapter 3 and Chapter 4 allow the reuse of published papers in the thesis without formal permissions. Thus, the waivers of copyright from MDPI are achieved by the following statement from MDPI website:

"For all articles published in MDPI journals, copyright is retained by the authors. Articles are licensed under an open access Creative Commons CC BY 4.0 license, meaning that anyone may download and read the paper for free. In addition, the article may be reused and quoted provided that the original published version is cited."

Elsevier, as publisher of the manuscript "Component capacity optimization of a renewable energy system using data-driven two-stage algorithmic approach" fully or partly adopted in Chapter 1, Chapter 2, Chapter 3 and Chapter 5 allow the reuse of published papers in the thesis without formal permissions. Thus, the waivers of copyright from Elsevier are achieved by the following statement from Elsevier RightsLink:

"As the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source."

References

[1] J Mitali, S Dhinakaran, and AA Mohamad. Energy storage systems: A review. *Energy Storage and Saving*, 2022.

[2] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[3] Kenneth Gillingham, David Rapson, and Gernot Wagner. The rebound effect and energy efficiency policy. *Review of Environmental Economics and Policy*, 2016.

[4] Energy Information Administration. Eia projections indicate global energy consumption increases through 2050, outpacing efficiency gains and driving continued emissions growth. https://www.eia.gov/pressroom/releases/press542.php, 2023.

[5] International Energy Agency. *World energy outlook 2019*. OECD/IEA Paris, 2019.

[6] Chi-Wei Su, Li-Dong Pang, Meng Qin, Oana-Ramona Lobonţ, and Muhammad Umar. The spillover effects among fossil fuel, renewables and carbon markets: evidence under the dual dilemma of climate change and energy crises. *Energy*, 274:127304, 2023.

[7] Zahra Medghalchi and Onur Taylan. A novel hybrid optimization framework for sizing renewable energy systems integrated with energy storage systems with solar photovoltaics, wind, battery and electrolyzer-fuel cell. *Energy Conversion and Management*, 294:117594, 2023.

[8] Muhammad Khalid. Smart grids and renewable energy systems: Perspectives and grid integration challenges. *Energy Strategy Reviews*, 51:101299, 2024.

[9] Jeffrey Dankwa Ampah, Sandylove Afrane, Ephraim Bonah Agyekum, Humphrey Adun, Abdulfatah Abdu Yusuf, and Olusola Bamisile. Electric vehicles development in sub-saharan africa: Performance assessment of standalone renewable energy systems for hydrogen refuelling and electricity charging stations (hrecs). *Journal of Cleaner Production*, 376:134238, 2022.

[10] Zhaoyang Qu, Chuanfu Xu, Fang Yang, Fan Ling, and Sasan Pirouzi. Market clearing price-based energy management of grid-connected renewable energy hubs including flexible sources according to thermal, hydrogen, and compressed air storage systems. *Journal of Energy Storage*, 69:107981, 2023.

[11] Hamed H Pourasl, Reza Vatankhah Barenji, and Vahid M Khojastehnezhad. Solar energy status in the world: A comprehensive review. *Energy Reports*, 10:3474–3493, 2023.

[12] Global Wind Energy Council. Global wind report 2024, 2024. Available at: https://gwec.net/global-wind-report-2024.

[13] World Wind Energy Association. Wwea half-year report 2023: Additional momentum for windpower in 2023, 2023. Available at: https://wwindea.org/wwea-half-year-report-2023-additional-momentum-for-windpower-in-2023.

[14] Yunxia Long, Yaning Chen, Changchun Xu, Zhi Li, Yongchang Liu, and Hongyu Wang. The role of global installed wind energy in mitigating co2 emission and temperature rising. *Journal of Cleaner Production*, 423:138778, 2023.

[15] Ember Climate. Global electricity review 2023, 2023. Available at: https://ember-climate.org/global-electricity-review-2023.

[16] Nuria Martín-Chivelet, Konstantinos Kapsis, Helen Rose Wilson, Veronique Delisle, Rebecca Yang, Lorenzo Olivieri, Jesús Polo, Johannes Eisenlohr, Benjamin Roy,

Laura Maturi, et al. Building-integrated photovoltaic (bipv) products and systems: A review of energy-related behavior. *Energy and Buildings*, 262:111998, 2022.

[17] Hussein Mohammed Ridha, Chandima Gomes, Hashim Hizam, Masoud Ahmadipour, Ali Asghar Heidari, and Huiling Chen. Multi-objective optimization and multi-criteria decision-making methods for optimal design of standalone photovoltaic system: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 135:110202, 2021.

[18] Meng Yang, Lihui Zhang, Zhenli Zhao, and Liwan Wang. Comprehensive benefits analysis of electric vehicle charging station integrated photovoltaic and energy storage. *Journal of Cleaner Production*, 302:126967, 2021.

[19] Kamran Ali Khan Niazi and Marta Victoria. Comparative analysis of photovoltaic configurations for agrivoltaic systems in europe. *Progress in Photovoltaics: Research and Applications*, 31(11):1101–1113, 2023.

[20] Manoj Kumar Sharma and Jishnu Bhattacharya. Dependence of spectral factor on angle of incidence for monocrystalline silicon based photovoltaic solar panel. *Renewable Energy*, 184:820–829, 2022.

[21] Abdurashid Mavlonov, Takhir Razykov, Fazal Raziq, Jiantuo Gan, Jakapan Chantana, Yu Kawano, Takahito Nishimura, Haoming Wei, Andriy Zakutayev, Takashi Minemoto, et al. A review of sb2se3 photovoltaic absorber materials and thin-film solar cells. *Solar Energy*, 201:227–246, 2020.

[22] Kevin Beepat, Sanjay Kumar, Ankush Sharma, Davinder Pal Sharma, Dinesh Pathak, and Jean-Michel Nunzi. Perovskite materials for photovoltaics: a review. *The European Physical Journal Applied Physics*, 98:43, 2023.

[23] Liang Jiang, Sai Cui, Peng Sun, Yanyan Wang, and Chao Yang. Comparison of monocrystalline and polycrystalline solar modules. In *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, pages 341–344. IEEE, 2020.

[24] H Kang. Crystalline silicon vs. amorphous silicon: the significance of structural differences in photovoltaic applications. In *IOP Conference Series: Earth and Environmental Science*, volume 726, page 012001. IOP Publishing, 2021.

[25] Gökhan Yıldız, Büşra Çalış, Ali Etem Gürel, and İlhan Ceylan. Investigation of life cycle co2 emissions of the polycrystalline and cadmium telluride pv panels. *Environmental Nanotechnology, Monitoring & Management*, 14:100343, 2020.

[26] CB Isabela, Ricardo A Marques Lameirinhas, Joao Paulo N Torres, and Carlos AF Fernandes. Comparative study of the copper indium gallium selenide (cigs) solar cell with other solar technologies. *Sustainable Energy & Fuels*, 5(8):2273–2283, 2021.

[27] Eun-Bi Kim, M Shaheer Akhtar, Hyung-Shik Shin, Sadia Ameen, and Mohammad Khaja Nazeeruddin. A review on two-dimensional (2d) and 2d-3d multidimensional perovskite solar cells: Perovskites structures, stability, and photovoltaic performances. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, 48:100405, 2021.

[28] Rui Wang, Tianyi Huang, Jingjing Xue, Jinhui Tong, Kai Zhu, and Yang Yang. Prospects for metal halide perovskite-based tandem solar cells. *Nature Photonics*, 15(6):411–425, 2021.

[29] Yong Cui, Huifeng Yao, Ling Hong, Tao Zhang, Yabing Tang, Baojun Lin, Kaihu Xian, Bowei Gao, Cunbin An, Pengqing Bi, et al. Organic photovoltaic cell with 17% efficiency and superior processability. *National Science Review*, 7(7):1239–1246, 2020.

[30] Khaled Osmani, Mohamad Ramadan, Thierry Lemenand, Bruno Castanier, and Ahmad Haddad. Optimization of pv array tilt angle for minimum levelized cost of energy. *Computers & Electrical Engineering*, 96:107474, 2021.

[31] Ridha Ben Mansour, Meer Abdul Mateen Khan, Fahad Abdulaziz Alsulaiman, and Rached Ben Mansour. Optimizing the solar pv tilt angle to maximize the power output: A case study for saudi arabia. *IEEE Access*, 9:15914–15928, 2021.

[32] Shixun Mo, Qintao Ye, Kunping Jiang, Xiaofeng Mo, and Gengyu Shen. An improved mppt method for photovoltaic systems based on mayfly optimization algorithm. *Energy Reports*, 8:141–150, 2022.

[33] JE Hay and JA Davies. Calculations of the solar radiation incident on an inclined surface 1980: Proceedings of first canadian solar radiation data workshop. *Canadian Atmospheric Environment Service, Canada*, 5972.

[34] RSRACSR Perez, R Stewart, C Arbogast, R Seals, and J Scott. An anisotropic hourly diffuse radiation model for sloping surfaces: description, performance validation, site dependency evaluation. *Solar energy*, 36(6):481–497, 1986.

[35] Thomas M Klucher. Evaluation of models to predict insolation on tilted surfaces. *Solar energy*, 23(2):111–114, 1979.

[36] Douglas T Reindl, William A Beckman, and John A Duffie. Evaluation of hourly tilted surface radiation models. *Solar energy*, 45(1):9–17, 1990.

[37] Benjamin YH Liu and Richard C Jordan. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar energy*, 4(3):1–19, 1960.

[38] Eric Evans Osei Opoku, Alex O Acheampong, Kingsley E Dogah, and Isaac Koomson. Energy innovation investment and renewable energy in oecd countries. *Energy Strategy Reviews*, 54:101462, 2024.

[39] Ahmad Shah Irshad, Nqibullah Kargar, MH Elkholy, Gul Ahmad Ludin, Said Elias, Ahmadullah Hilali, Tomonobu Senjyu, Mahmoud M Gamil, and Gabor Pinter. Techno-economic evaluation and comparison of the optimal pv/wind and grid hybrid system with horizontal and vertical axis wind turbines. *Energy Conversion and Management: X*, page 100638, 2024.

[40] Amir R Nejad, Jonathan Keller, Yi Guo, Shawn Sheng, Henk Polinder, Simon Watson, Jianning Dong, Zian Qin, Amir Ebrahimi, Ralf Schelenz, et al. Wind turbine

drivetrains: state-of-the-art technologies and future development trends. *Wind Energy Science*, 7(1):387–411, 2022.

[41] Mehmet Bilgili and Hakan Alphan. Global growth in offshore wind turbine technology. *Clean Technologies and Environmental Policy*, 24(7):2215–2227, 2022.

[42] Yi Zhang, Hexu Sun, Jianxin Tan, Zheng Li, Weimin Hou, and Yingjun Guo. Capacity configuration optimization of multi-energy system integrating wind turbine/photovoltaic/hydrogen/battery. *Energy*, 252:124046, 2022.

[43] Abraham Alem Kebede, Theodoros Kalogiannis, Joeri Van Mierlo, and Maitane Berecibar. A comprehensive review of stationary energy storage devices for large scale renewable energy sources grid integration. *Renewable and Sustainable Energy Reviews*, 159:112213, 2022.

[44] Yi He, Su Guo, Jianxu Zhou, Jilei Ye, Jing Huang, Kun Zheng, and Xinru Du. Multi-objective planning-operation co-optimization of renewable energy system with hybrid energy storages. *Renewable Energy*, 184:776–790, 2022.

[45] Chouaib Ammari, Djamel Belatrache, Batoul Touhami, and Salim Makhloufi. Sizing, optimization, control and energy management of hybrid renewable energy system—a review. *Energy and Built Environment*, 3(4):399–411, 2022.

[46] Daniele Groppi, Antun Pfeifer, Davide Astiaso Garcia, Goran Krajačić, and Neven Duić. A review on energy storage and demand side management solutions in smart energy islands. *Renewable and Sustainable Energy Reviews*, 135:110183, 2021.

[47] J Mitali, S Dhinakaran, and AA Mohamad. Energy storage systems: A review. *Energy Storage and Saving*, 2022.

[48] Chethan Parthasarathy, Hossein Hafezi, and Hannu Laaksonen. Lithium-ion bess integration for smart grid applications-ecm modelling approach. In *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5. IEEE, 2020.

[49] Mohammad Abdul Hannan, SB Wali, Pin Jern Ker, Muhamad Safwan Abd Rahman, M Mansor, VK Ramachandaramurthy, KM Muttaqi, Teuku Meurah Indra Mahlia, and Zhao Yang Dong. Battery energy-storage system: A review of technologies, optimization objectives, constraints, approaches, and outstanding issues. *Journal of Energy Storage*, 42:103023, 2021.

[50] Shengqi Zhang, Heyu Liu, Fei Wang, Yuhui Miao, and Jingwen Dong. Dual-stage operation strategy of bess for frequency regulation considering planned peak shaving. *Energy Reports*, 6:502–508, 2020.

[51] Luis Fernando Grisales-Noreña, Brandon Cortés-Caicedo, Oscar Danilo Montoya, Rubén Iván Bolaños, and César Augusto Marín Moreno. Nonlinear programming for bess operation for the improvement of economic, technical and environmental indices by considering grid-connected and stand-alone networks: An application to the territory of colombia. *Journal of Energy Storage*, 98:112856, 2024.

[52] Yuqing Yang, Stephen Bremner, Chris Menictas, and Merlinde Kay. Modelling and optimal energy management for battery energy storage systems in renewable energy systems: A review. *Renewable and Sustainable Energy Reviews*, 167:112671, 2022.

[53] Bharat Singh and Ashwani Kumar. Optimal energy management and feasibility analysis of hybrid renewable energy sources with bess and impact of electric vehicle load with demand response program. *Energy*, 278:127867, 2023.

[54] Jianmin Ma, Yutao Li, Nicholas S Grundish, John B Goodenough, Yuhui Chen, Limin Guo, Zhangquan Peng, Xiaoqun Qi, Fengyi Yang, Long Qie, et al. The 2021 battery technology roadmap. *Journal of Physics D: Applied Physics*, 54(18):183001, 2021.

[55] Qusay Hassan, Ammar M Abdulateef, Saadoon Abdul Hafedh, Ahmed Al-samari, Jasim Abdulateef, Aws Zuhair Sameen, Hayder M Salman, Ali Khudhair Al-Jiboory, Szymon Wieteska, and Marek Jaszczur. Renewable energy-to-green hydrogen: A

review of main resources routes, processes and evaluation. *International Journal of Hydrogen Energy*, 2023.

[56] M Lopez-Juarez, T Rockstroh, R Novella, and R Vijayagopal. A methodology to develop multi-physics dynamic fuel cell system models validated with vehicle realistic drive cycle data. *Applied Energy*, 358:122568, 2024.

[57] Elsayed Abdelhameed and Hiroshi Tashima. Experimental study on the effects of methane-hydrogen jet as direct injected fuel in marine diesel engine. *Energy*, 267:126569, 2023.

[58] Adrian Lubecki, Jakub Szczurowski, and Katarzyna Zarebska. A comparative environmental life cycle assessment study of hydrogen fuel, electricity and diesel fuel for public buses. *Applied Energy*, 350:121766, 2023.

[59] Peng Liu, Tianyan Yang, Hongbin Zheng, Xiang Huang, Xuan Wang, Tian Qiu, and Shuiting Ding. Thermodynamic analysis of power generation thermal management system for heat and cold exergy utilization from liquid hydrogen-fueled turbojet engine. *Applied Energy*, 365:123290, 2024.

[60] Julian Hoelzen, Daniel Silberhorn, Thomas Zill, B Bensmann, and R Hanke-Rauschenbach. Hydrogen-powered aviation and its reliance on green hydrogen infrastructure–review and research gaps. *International Journal of Hydrogen Energy*, 47(5):3108–3130, 2022.

[61] Priyanka Saha, Faysal Ahamed Akash, Shaik Muntasir Shovon, Minhaj Uddin Monir, Mohammad Tofayal Ahmed, Mohammad Forrukh Hossain Khan, Shaheen M Sarkar, Md Kamrul Islam, Md Mehedi Hasan, Dai-Viet N Vo, et al. Grey, blue, and green hydrogen: A comprehensive review of production methods and prospects for zero-emission energy. *International Journal of Green Energy*, 21(6):1383–1397, 2024.

[62] Xiangxiang Dong, Jiang Wu, Zhanbo Xu, Kun Liu, and Xiaohong Guan. Optimal coordination of hydrogen-based integrated energy systems with combination of hydrogen and water storage. *Applied energy*, 308:118274, 2022.

[63] Jianxiao Wang, Qi An, Yue Zhao, Guangsheng Pan, Jie Song, Qinran Hu, and Chin-Woo Tan. Role of electrolytic hydrogen in smart city decarbonization in china. *Applied Energy*, 336:120699, 2023.

[64] Sara Bellocchi, Paolo Colbertaldo, Michele Manno, and Benedetto Nastasi. Assessing the effectiveness of hydrogen pathways: A techno-economic optimisation within an integrated energy system. *Energy*, 263:126017, 2023.

[65] Reza Khalili, Arian Khaledi, Mousa Marzband, Amin Foroughi Nematollahi, Behrooz Vahidi, and Pierluigi Siano. Robust multi-objective optimization for the iranian electricity market considering green hydrogen and analyzing the performance of different demand response programs. *Applied Energy*, 334:120737, 2023.

[66] Qi Li, Xukang Xiao, Yuchen Pu, Shuyu Luo, Hong Liu, and Weirong Chen. Hierarchical optimal scheduling method for regional integrated energy systems considering electricity-hydrogen shared energy. *Applied Energy*, 349:121670, 2023.

[67] Keith Scott. Introduction to electrolysis, electrolysers and hydrogen production. 2019.

[68] Mohd Nur Ikhmal Salehmin, Teuku Husaini, Jonathan Goh, and Abu Bakar Sulong. High-pressure pem water electrolyser: A review on challenges and mitigation strategies towards green and low-cost hydrogen production. *Energy Conversion and Management*, 268:115985, 2022.

[69] Sayem M Abu, MA Hannan, Pin Jern Ker, M Mansor, Sieh Kiong Tiong, and TM Indra Mahlia. Recent progress in electrolyser control technologies for hydrogen energy production: A patent landscape analysis and technology updates. *Journal of Energy Storage*, 72:108773, 2023.

[70] Yeyan Xu, Liangzhong Yao, Peng Xu, Wei Cui, Zhenan Zhang, Fangbing Liu, Beilin Mao, and Zhang Wen. Load forecasting method for building energy systems based on modified two-layer lstm. In *2021 3rd Asia Energy and Electrical Engineering Symposium (AEEES)*, pages 660–665. IEEE, 2021.

[71] Neelam Mughees, Syed Ali Mohsin, Abdullah Mughees, and Anam Mughees. Deep sequence to sequence bi-lstm neural networks for day-ahead peak load forecasting. *Expert Systems with Applications*, 175:114844, 2021.

[72] Haixiang Zang, Ruiqi Xu, Lilin Cheng, Tao Ding, Ling Liu, Zhinong Wei, and Guoqiang Sun. Residential load forecasting based on lstm fusing self-attention mechanism with pooling. *Energy*, 229:120682, 2021.

[73] M Junaid Gul, Gul Malik Urfa, Anand Paul, Jihoon Moon, Seungmin Rho, and Eenjun Hwang. Mid-term electricity load prediction using cnn and bi-lstm. *The Journal of Supercomputing*, 77:10942–10958, 2021.

[74] Sabrina J Mielke, Zaid Alyafeai, Elizabeth Salesky, Colin Raffel, Manan Dey, Matthias Gallé, Arun Raja, Chenglei Si, Wilson Y Lee, Benoît Sagot, et al. Between words and characters: A brief history of open-vocabulary modeling and tokenization in nlp. *arXiv preprint arXiv:2112.10508*, 2021.

[75] Hongtao Sheng, Qianlong Zhu, Jun Tao, Huaying Zhang, and Feixiang Peng. Distribution network reconfiguration and photovoltaic optimal allocation considering harmonic interaction between photovoltaic and distribution network. *Journal of Electrical Engineering & Technology*, 19(1):17–30, 2024.

[76] AL-Rousan Nadia and AL-Najjar Hazem. Optimized deep neural network to estimate orientation angles for solar photovoltaics intelligent systems. *Cleaner Engineering and Technology*, 20:100754, 2024.

[77] Ruimiao Liu, Zhongbing Liu, Wei Xiong, Ling Zhang, Chengliang Zhao, and Yingde Yin. Performance simulation and optimization of building façade photovoltaic systems under different urban building layouts. *Energy*, 288:129708, 2024.

[78] Zixuan Shangguan, Hao Li, Bowen Yang, Zhongkai Zhao, Tiantian Wang, Liming Jin, and Cunman Zhang. Optimization of alkaline electrolyzer operation in renewable

energy power systems: A universal modeling approach for enhanced hydrogen production efficiency and cost-effectiveness. *International Journal of Hydrogen Energy*, 49:943–954, 2024.

[79] Joshi Sukhdev Nirbheram, Aeidapu Mahesh, and Ambati Bhimaraju. Techno-economic optimization of standalone photovoltaic-wind turbine-battery energy storage system hybrid energy system considering the degradation of the components. *Renewable Energy*, 222:119918, 2024.

[80] Lyu-Guang Hua, Ghulam Hafeez, Baheej Alghamdi, Hisham Alghamdi, Farrukh Aslam Khan, and Safeer Ullah. Demand response with pricing schemes and consumers mode constraints for energy balancing in smart grids. *Applied Energy*, 377:124324, 2025.

[81] Yavuz Eren and İbrahim Küçükdemiral. A comprehensive review on deep learning approaches for short-term load forecasting. *Renewable and Sustainable Energy Reviews*, 189:114031, 2024.

[82] Abhishek Saxena, Ravi Shankar, Ehab El-Saadany, Manish Kumar, Omar Al Zaabi, Khalifa Al Hosani, and Utkal Ranjan Muduli. Intelligent load forecasting and renewable energy integration for enhanced grid reliability. *IEEE Transactions on Industry Applications*, 2024.

[83] Yanis Prunier, David Chuet, Sylvain Nicolay, Gwenaëlle Hamon, and Maxime Darnon. Optimization of photovoltaic panel tilt angle for short periods of time or multiple reorientations. *Energy Conversion and Management: X*, 20:100417, 2023.

[84] Gautam Raina, Rohit Vijay, and Sunanda Sinha. Study on the optimum orientation of bifacial photovoltaic module. *International Journal of Energy Research*, 46(4):4247–4266, 2022.

[85] Khaled Hosny Ibrahim, Eslam Mohamed Ahmed, and Saber Mohamed Saleh. Minimum cost-based design of isolated pv-wind hybrid system considering the pv tilt

angle and wind turbine hub height as design parameters using genetic algorithm. *International Journal of Energy Research*, 45(9):13149–13162, 2021.

[86] Praful Kumar Ganti, Hrushikesh Naik, and Mohanty Kanungo Barada. Hybrid tsa-rbfnn based approach for mppt of the solar pv panel under the effects of tilt angles variations and environmental effects. *International Journal of Energy Research*, 45(14):20104–20131, 2021.

[87] Ashutosh Sharma, Mehmet Ali Kallioğlu, Anchal Awasthi, Ranchan Chauhan, Gusztáv Fekete, and Tej Singh. Correlation formulation for optimum tilt angle for maximizing the solar radiation on solar collector in the western himalayan region. *Case Studies in Thermal Engineering*, 26:101185, 2021.

[88] Ali S Alghamdi. Performance enhancement of roof-mounted photovoltaic system: Artificial neural network optimization of ground coverage ratio. *Energies*, 14(6):1537, 2021.

[89] Ridha Ben Mansour, Meer Abdul Mateen Khan, Fahad Abdulaziz Alsulaiman, and Rached Ben Mansour. Optimizing the solar pv tilt angle to maximize the power output: A case study for saudi arabia. *IEEE Access*, 9:15914–15928, 2021.

[90] Qusay Hassan, Majid K Abbas, Ammar M Abdulateef, Jasim Abdulateef, and Abdulmajeed Mohamad. Assessment the potential solar energy with the models for optimum tilt angles of maximum solar irradiance for iraq. *Case Studies in Chemical and Environmental Engineering*, 4:100140, 2021.

[91] Saeed A Asiri, Elias M Salilih, Khaled M Alfawaz, Ageel F Alogla, Nidal H Abu-Hamdeh, and Osama K Nusier. Improve the performance of renewable energy conversion and storage via ann in a system of solar water heater with variable speed photovoltaic circulating pump. *International Journal of Energy Research*, 2022.

[92] Olusola Bamisile, Ariyo Oluwasanmi, Chukwuebuka Ejiyi, Nasser Yimen, Sandra Obiora, and Qi Huang. Comparison of machine learning and deep learning algorithms

for hourly global/diffuse solar radiation predictions. *International Journal of Energy Research*, 2021.

[93] Sina Akhavan Shams and Rouhollah Ahmadi. Dynamic optimization of solar-wind hybrid system connected to electrical battery or hydrogen as an energy storage system. *International Journal of Energy Research*, 45(7):10630–10654, 2021.

[94] Bashar Shboul, AL-Arfi Ismail, Stavros Michailos, Derek Ingham, Lin Ma, Kevin J Hughes, and Mohamed Pourkashanian. A new ann model for hourly solar radiation and wind speed prediction: A case study over the north & south of the arabian peninsula. *Sustainable Energy Technologies and Assessments*, 46:101248, 2021.

[95] Ümit Ağbulut, Ali Etem Gürel, and Yunus Biçen. Prediction of daily global solar radiation using different machine learning algorithms: Evaluation and comparison. *Renewable and Sustainable Energy Reviews*, 135:110114, 2021.

[96] César G Villegas-Mier, Juvenal Rodriguez-Resendiz, José M Álvarez-Alvarado, Hugo Rodriguez-Resendiz, Ana Marcela Herrera-Navarro, and Omar Rodríguez-Abreo. Artificial neural networks in mppt algorithms for optimization of photovoltaic power systems: A review. *Micromachines*, 12(10):1260, 2021.

[97] Nawin Ra, Aritra Ghosh, and Ankur Bhattacharjee. Iot-based smart energy management for solar vanadium redox flow battery powered switchable building glazing satisfying the hvac system of ev charging stations. *Energy Conversion and Management*, 281:116851, 2023.

[98] Mohammad Ali Lasemi, Ahmad Arabkoohsar, Amin Hajizadeh, and Behnam Mohammadi-ivatloo. A comprehensive review on optimization challenges of smart energy hubs under uncertainty factors. *Renewable and Sustainable Energy Reviews*, 160:112320, 2022.

[99] Aakash Hassan, Octavian Bass, Yasir M Al-Abdeli, Martin Masek, and Mohammad AS Masoum. A novel approach for optimal sizing of stand-alone solar pv sys-

tems with power quality considerations. *International Journal of Electrical Power & Energy Systems*, 144:108597, 2023.

[100] Mim Mashrur Ahmed, Barun K Das, Pronob Das, Md Sanowar Hossain, and Md Golam Kibria. Energy management and sizing of a stand-alone hybrid renewable energy system for community electricity, fresh water, and cooking gas demands of a remote island. *Energy Conversion and Management*, 299:117865, 2024.

[101] Fayza S. Mahmoud, Ahmed A. Zaki Diab, Ziad M. Ali, Abou-Hashema M. El-Sayed, Thamer Alquthami, Mahrous Ahmed, and Husam A. Ramadan. Optimal sizing of smart hybrid renewable energy system using different optimization algorithms. *Energy Reports*, 8:4935–4956, 2022.

[102] Mohamed A Mohamed, Ali M Eltamaly, and Abdulrahman I Alolah. Swarm intelligence-based optimization of grid-dependent hybrid renewable energy systems. *Renewable and Sustainable Energy Reviews*, 77:515–524, 2017.

[103] Ali M Eltamaly, Mohamed A Mohamed, MS Al-Saud, and Abdulrahman I Alolah. Load management as a smart grid concept for sizing and designing of hybrid renewable energy systems. *Engineering Optimization*, 49(10):1813–1828, 2017.

[104] P. E. Campana, L. Cioccolanti, B. François, J. Jurasz, Y. Zhang, M. Varini, and J. ... Yan. Li-ion batteries for peak shaving, price arbitrage, and photovoltaic self-consumption in commercial buildings: A monte carlo analysis. *Energy Conversion and Management*, 234:113889, 2021.

[105] Ali M Eltamaly and Mohamed A Mohamed. A novel software for design and optimization of hybrid power systems. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 38:1299–1315, 2016.

[106] A. Hassan, Y. M. Al-Abdeli, M. Masek, and O. Bass. Optimal sizing and energy scheduling of grid-supplemented solar pv systems with battery storage: Sensitivity of reliability and financial constraints. *Energy*, 238:121780, 2022.

[107] Y. Wu, Z. Liu, J. Liu, H. Xiao, R. Liu, and L. Zhang. Optimal battery capacity of grid-connected pv-battery systems considering battery degradation. *Renewable Energy*, 181:10–23, 2022.

[108] Tianye Liu, Zhen Yang, Yuanyuan Duan, and Shuozhuo Hu. Techno-economic assessment of hydrogen integrated into electrical/thermal energy storage in pv+ wind system devoting to high reliability. *Energy Conversion and Management*, 268:116067, 2022.

[109] Yi He, Su Guo, Jianxu Zhou, Feng Wu, Jing Huang, and Huanjin Pei. The quantitative techno-economic comparisons and multi-objective capacity optimization of wind-photovoltaic hybrid power system considering different energy storage technologies. *Energy Conversion and Management*, 229:113779, 2021.

[110] Jingze Yang, Zhen Yang, and Yuanyuan Duan. Capacity optimization and feasibility assessment of solar-wind hybrid renewable energy systems in china. *Journal of Cleaner Production*, 368:133139, 2022.

[111] A Al-Quraan and B Al-Mhairat. Sizing and energy management of standalone hybrid renewable energy systems based on economic predictive control. *Energy Conversion and Management*, 300:117948, 2024.

[112] Mohammad Hadi Shaban Boloukat and Asghar Akbari Foroud. Stochastic-based resource expansion planning for a grid-connected microgrid using interval linear programming. *Energy*, 113:776–787, 2016.

[113] A Al-Quraan and B Al-Mhairat. Economic predictive control-based sizing and energy management for grid-connected hybrid renewable energy systems. *Energy*, page 131795, 2024.

[114] Jiandong Jia, Haiqiao Li, Di Wu, Jiacheng Guo, Leilei Jiang, and Zeming Fan. Multi-objective optimization study of regional integrated energy systems coupled with renewable energy, energy storage, and inter-station energy sharing. *Renewable Energy*, 225:120328, 2024.

[115] R Praveen Kumar and G Karthikeyan. A multi-objective optimization solution for distributed generation energy management in microgrids with hybrid energy sources and battery storage system. *Journal of Energy Storage*, 75:109702, 2024.

[116] R Sathishkumar, M Venkateswaran, P Deepamangai, et al. An efficient power management control strategy for grid-independent hybrid renewable energy systems with hybrid energy storage: Hybrid approach. *Journal of Energy Storage*, 96:112685, 2024.

[117] Yanting Zhou, Zhongjing Ma, Jinhui Zhang, and Suli Zou. Data-driven stochastic energy management of multi energy system using deep reinforcement learning. *Energy*, 261:125187, 2022.

[118] Reihaneh Haji Mahdizadeh Zargar and Mohammad Hossein Yaghmaee Moghaddam. Development of a markov-chain-based solar generation model for smart microgrid energy management system. *IEEE Transactions on Sustainable Energy*, 11(2):736–745, 2019.

[119] Salvatore Vergine, César Álvarez-Arroyo, Guglielmo D'Amico, Juan Manuel Escaño, and Lázaro Alvarado-Barrios. Optimal management of a hybrid and isolated microgrid in a random setting. *Energy Reports*, 8:9402–9419, 2022.

[120] César Álvarez-Arroyo, Salvatore Vergine, Agustín Sánchez de la Nieta, Lázaro Alvarado-Barrios, and Guglielmo D'Amico. Optimising microgrid energy management: Leveraging flexible storage systems and full integration of renewable energy sources. *Renewable Energy*, page 120701, 2024.

[121] Leyi Yao, Zeyuan Liu, Weiguang Chang, and Qiang Yang. Multi-level model predictive control based multi-objective optimal energy management of integrated energy systems considering uncertainty. *Renewable Energy*, 212:523–537, 2023.

[122] Sina Akhavan Shams and Rouhollah Ahmadi. Dynamic optimization of solar-wind hybrid system connected to electrical battery or hydrogen as an energy storage system. *International Journal of Energy Research*, 45(7):10630–10654, 2021.

[123] How HOMER Calculates the PV Cell Temperature. *HOMER Energy*.

[124] J. A. Duffie, William A. Beckman, and W. M. Worek. Solar Engineering of Thermal Processes, 2nd ed. *Journal of Solar Energy Engineering*, 116(1):67–68, 02 1994.

[125] Abbas Fotouhi, Daniel J. Auger, Karsten Propp, Stefano Longo, and Mark Wild. A review on electric vehicle battery modelling: From lithium-ion toward lithium–sulphur. *Renewable and Sustainable Energy Reviews*, 56:1008 – 1021, 2016.

[126] A.; Janzou S.; Nelson A.; Lundstrom B. DiOrio, N.; Dobos. Technoeconomic modeling of battery energy storage in sam. *National Renewable Energy Laboratory (NREL)*.

[127] Holger C. Hesse, Michael Schimpe, Daniel Kucevic, and Andreas Jossen. Lithium-ion battery storage for the grid—a review of stationary battery storage system design tailored for applications in modern power grids. *Energies*, 10(12), 2017.

[128] Georg Angenendt, Bahram Ashrafinia, Sebastian Zurmühlen, Kevin Jacqué, Julia Badeda, and Dirk Uwe Sauer. Influence of the battery voltage level on the efficiency and cost of a pv battery energy storage system. In *32. Symposium Photovoltaische Solareergie*, 2017.

[129] Holger C Hesse, Rodrigo Martins, Petr Musilek, Maik Naumann, Cong Nam Truong, and Andreas Jossen. Economic optimization of component sizing for residential battery storage systems. *Energies*, 10(7):835, 2017.

[130] R.F.P. Bentvelsen. Modeling and scheduling of a controllable electrolyser in an industrial grid. Master's thesis, Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, 2019.

[131] Boyd M, Chen T, Dougherty B (2017) NIST Campus Photovoltaic (PV) Arrays and Weather Station Data Sets. National Institute of Standards and Technology. U.S. Department of Commerce, Washington, D.C.

[132] Sebastian Bader, Xinyu Ma, and Bengt Oelmann. One-diode photovoltaic model parameters at indoor illumination levels – a comparison. *Solar Energy*, 180:707–716, 2019.

[133] Mussad Alzahrani, Katie Shanks, and Tapas K Mallick. Advances and limitations of increasing solar irradiance for concentrating photovoltaics thermal system. *Renewable and Sustainable Energy Reviews*, 138:110517, 2021.

[134] Kunal Sandip Garud, Simon Jayaraj, and Moo-Yeon Lee. A review on modeling of solar photovoltaic systems using artificial neural networks, fuzzy logic, genetic algorithm and hybrid models. *International Journal of Energy Research*, 45(1):6–35, 2021.

[135] Elaff ALfadil Ismail and Samah Mohammed Hashim. An economic evaluation of grid connected photovoltaic system for a residential house in khartoum. In *2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)*, pages 1–6. IEEE, 2018.

[136] Yang Zhang, Anders Lundblad, Pietro Elia Campana, F. Benavente, and Jinyue Yan. Battery sizing and rule-based operation of grid-connected photovoltaic-battery system: A case study in sweden. *Energy Conversion and Management*, 133:249–263, 2017.

[137] Manbir Sodhi, Lennart Banaszek, Chris Magee, and Mercedes Rivero-Hudec. Economic lifetimes of solar panels. *Procedia CIRP*, 105:782–787, 2022.

[138] Bloomberg New Energy Finance. Top 10 energy storage trends in 2023, 2023.

[139] Why does the federal reserve aim for inflation of 2 percent over the longer run? *Board of Governors of the Federal Reserve System*, 2020. Available at https://www.federalreserve.gov/faqs/economy_14400.htm.

[140] Ying Xu, Huan Zhang, Lei Huang, Rong Qu, and Yusuke Nojima. A pareto front grid guided multi-objective evolutionary algorithm. *Applied Soft Computing*, 136:110095, 2023.

[141] A. Pena-Bello, E. Barbour, M. C. Gonzalez, M. K. Patel, and D. Parra. Optimized pv-coupled battery systems for combining applications: Impact of battery technology and geography. *Renewable and Sustainable Energy Reviews*, 112:978–990, 2019.

[142] Ember. European wholesale electricity price data. Data sourced from ENTSO-e and cleaned, October 2023. Available at https://ember-climate.org/data-catalogue/european-wholesale-electricity-price-data/.

[143] Waqas Khan, Juo Yu Liao, Shalika Walker, and Wim Zeiler. Impact assessment of varied data granularities from commercial buildings on exploration and learning mechanism. *Applied Energy*, 319:119281, 2022.

[144] Wenrui Ye, Münür Sacit Herdem, Joey Z Li, Jatin Nathwani, and John Z Wen. Formulation and data-driven optimization for maximizing the photovoltaic power with tilt angle adjustment. *Energies*, 15(22):8578, 2022.

[145] Mercedes-Benz. E-citaro fuel cell, 2024. Accessed: 2024-08-15.

[146] B Reuter, M Faltenbacher, O Schuller, N Whitehouse, and S Whitehouse. New bus refuelling for european hydrogen bus depots. *FCH JU*, 2017.

[147] L Bartolucci, S Cordiner, V Mulone, C Tatangelo, M Antonelli, and S Romagnuolo. Multi-hub hydrogen refueling station with on-site and centralized production. *International Journal of Hydrogen Energy*, 48(54):20861–20874, 2023.

Appendix A

Optimization Result Validation

The project presented in Chapter 5 has been validated using RETScreen. The validation process included the evaluation of the following key indicators to assess the simulation results: the EIS and the IRR. In RETScreen, the EIS is the fraction of load delivered, and the IRR is the pre-tax IRR.

For the simulation, the Config. 3 from Table 6.4 was selected. This configuration specifies a PV capacity of 140 kW and a battery storage capacity of 80 kWh. The climate data location was set to Amsterdam, Netherlands. The PV panel efficiency was assumed to be 15%, with mono-silicon as the chosen technology.

The simulation results show that the PV system can deliver 102,899 kWh annually, with a load fraction of 28.2%, which is equivalent to EIS of 0.282. The pre-tax IRR was calculated to be 3.6%. Comparing these results to the outcomes from Chapter 5, where the EIS was 0.33 and the IRR was 4.22%, no significant difference was observed.

These findings confirm the environmental and financial viability of the project, demonstrating its potential for meaningful contributions to energy sustainability and efficiency.